JRC TECHNICAL REPORT

A quality approach to real-time smartphone and citizen-driven food market price data

The case of Food Price Crowdsourcing Africa (FPCA) in Nigeria

Solano-Hermosilla, G., Adewopo, J., Peter, H., Barreiro-Hurle, J., Arbia, G., Nardelli, V., Gorrín-González, C., Micale, F. & Ceccarelli, T.

2020

Joint Research Centre

EUR 30027 EN

This publication is a Technical report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication. For information on the methodology and quality underlying the data used in this publication for which the source is neither Eurostat nor other Commission services, users should contact the referenced source. The designations employed and the presentation of material on the maps do not imply the expression of any opinion whatsoever on the part of the European Union concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Contact information
Name: Gloria Solano-Hermosilla
Address: Edificio Expo – Inca Garcilaso, 3 E-41092 Seville – SPAIN
Email: Gloria.SOLANO-HERMOSILLA@ec.europa.eu
Tel.: +34 954 48 8475

EU Science Hub
https://ec.europa.eu/jrc

JRC119273

EUR 30027 EN

PDF	*ISBN 978-92-76-14305-5*	*ISSN 1831-9424*	*doi:10.2760/366116*
Print	*ISBN 978-92-76-20599-9*	*ISSN 1018-5593*	*doi:10.2760/120193*

Luxembourg: Publications Office of the European Union, 2020

© European Union, 2020

The reuse policy of the European Commission is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Except otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (https://creativecommons.org/licenses/by/4.0/). This means that reuse is allowed provided appropriate credit is given and any changes are indicated. For any use or reproduction of photos or other material that is not owned by the EU, permission must be sought directly from the copyright holders.

All content © European Union, 2020, except: cover page, © IITA - Julius Adewopo, 2019.

How to cite this report: Solano-Hermosilla, G., Adewopo, J., Peter, H., Barreiro-Hurle, J., Arbia, G., Nardelli, V., Gorrín-González, C., Micale, F.& Ceccarelli, T., A quality approach to real-time smartphone and citizen-driven food market price data: The case of Food Price Crowdsourcing Africa (FPCA) in Nigeria, EUR 30027 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-14305-5, doi:10.2760/366116, JRC119273.

Access to the interactive dashboard to analyse and explore the data (scan with the mobile phone the QR code or click on the link below):

https://datam.jrc.ec.europa.eu/datam/mashup/FP_NGA

Contents

Acknowledgements

The authors would like to thank Giampiero Genovese, Head of the Economics of Agriculture Unit, Pavel Ciaian, Berta Sánchez, Federica Di Marcantonio, Dimitrios Kremmydas and Arnaldo Caivano from the European Commission's Joint Research Centre for their continuing support and additional input to the production of this study.

The authors would like to extend their gratitude to Yakob Seid (FAO) for sharing the experience and data about the AMIS-FAO crowdsourcing initiative in Nigeria and to the attendees of the two stakeholder and expert workshops, in Brussels, Belgium (2015) and in Abuja, Nigeria (2019), and to all the citizen volunteers who have participated in this project and made this innovative collection of data from citizens for citizens possible. The project was part-financed by the European Commission's Directorate-General for International Cooperation and Development (DG DEVCO).

Authors

Gloria Solano-Hermosilla, Jesús Barreiro-Hurle, Celso Gorrín González and Fabio Micale from the European Commission, Joint Research Centre (JRC), Economics of Agriculture Unit, Seville, Spain;

Julius Adewopo and Helen Peter from the International Institute of Tropical Agriculture (IITA-CGIAR), Nigeria;

Giuseppe Arbia and Vincenzo Nardelli from the Università Cattolica del Sacro Cuore, Milano;

Tomaso Ceccarelli from Wageningen University & Research (WUR).

Abstract

Timely and reliable monitoring of commodity food prices is an essential requirement for the assessment of market and food security risks and the establishment of early warning systems, especially in developing economies. However, data from regional or national systems for tracking changes of food prices in sub-Saharan Africa lacks the temporal or spatial richness and is often insufficient to inform targeted interventions. In addition to limited opportunity for [near-]real-time assessment of food prices, various stages in the commodity supply chain are mostly unrepresented, thereby limiting insights on stage-related price evolution. Yet, governments and market stakeholders rely on commodity price data to make decisions on appropriate interventions or commodity-focused investments. Recent rapid technological development indicates that digital devices and connectivity services are becoming affordable for many, including in remote areas of developing economies. This offers a great opportunity both for the harvesting of price data (via new data collection methodologies, such as crowdsourcing/crowdsensing — i.e. citizen-generated data — using mobile apps/devices), and for disseminating it (via web dashboards or other means) to provide real-time data that can support decisions at various levels and related policy-making processes. However, market information that aims at improving the functioning of markets and supply chains requires a continuous data flow as well as quality, accessibility and trust. More data does not necessarily translate into better information. Citizen-based data-generation systems are often confronted by challenges related to data quality and citizen participation, which may be further complicated by the volume of data generated compared to traditional approaches. Following the food price hikes during the first noughties of the 21st century, the European Commission's Directorate General for International Cooperation and Development (DG DEVCO) started collaborating with the European Commission's Joint Research Centre (JRC) on innovative methodologies for real-time food price data collection and analysis in developing countries. The work carried out so far includes a pilot initiative to crowdsource data from selected markets across several African countries, two workshops (with relevant stakeholders and experts), and the development of a spatial statistical quality methodology to facilitate the best possible exploitation of geo-located data. Based on the latter, the JRC designed the Food Price Crowdsourcing Africa (FPCA) project and implemented it within two states in Northern Nigeria. The FPCA is a credible methodology, based on the voluntary provision of data by a crowd (people living in urban, suburban, and rural areas) using a mobile app, leveraging monetary and non-monetary incentives to enhance contribution, which makes it possible to collect, analyse and validate, and disseminate staple food price data in real time across market segments. The granularity and high frequency of the crowdsourcing data open the door to real-time space-time analysis, which can be essential for policy and decision making and rapid response on specific geographic regions.

1 INTRODUCTION

Timely and reliable monitoring of commodity food prices is an essential requirement when assessing market and food security risks, and building up early warning systems, especially in developing economies. This is due to several reasons, including – i. relevance as an indicator of the balance between agricultural supply and market demand, and ii. being a measure of market performance as well as markets' influence on households' food affordability and livelihood. This is especially relevant in sub-Saharan countries, where poor households spend more than 60% of their budget on food (Lozano-Gracia & Young, 2014), and agriculture is still the primary source of households' income (Davis, Di Giuseppe, & Zezza, 2017). Amidst an array of other food security indicators, prices can be observed more frequently and at less cost and they generally reflect expectations and sentiments of market actors (Kalkuhl, von Braun, & Torero, 2016). Farmers and agri-food businesses make decisions according to their price expectations, based on available and accessible information. Timely and reliable price data can help farmers and businesses to form better market expectations, optimise their production or commercial decisions (Haile, Kalkuhl, & Usman, 2015), while informing governments (and other organisations) to take timely and right decisions and policies on market-level interventions, particularly in times of crisis.

One of the reasons agriculture in Africa lags behind its potential (AGRA, 2017) is the poor transfer of market signals to farmers, which prevents them from making informed decisions to maximise their income and investment capacity (Short, Barreiro-Hurle, & Balié, 2014). Generally, there is a lack of regional or national systems for high-frequency tracking of food price changes in sub-Saharan Africa (Zeug, Zeug, Bielski, Solano-Hermosilla, & M'barek, 2017). Yet, governments and market stakeholders rely on commodity price data to make decisions regarding appropriate market-level interventions or commodity-focused investments. Predominantly, individuals and institutions rely on snapshot surveys of markets carried out by national or international institutions, or through non-transparent proxy methods. This data often lacks timeliness and sufficient spatial granularity, nor does it cover all stages in the supply chain to provide relevant insights on price evolution (Green et al., 2013; Kalkuhl et al., 2016). Therefore, the quality and usefulness of available commodity price data are limited because governments, stakeholders and market actors cannot fully understand market dynamics and are rarely proactive in identifying opportunities for market-level interventions to bolster food security. This indicates the imperative need to identify alternative and efficient data collection methods to generate timely and reliable data for decision-support at various levels, including governments, private institutions, and supply chain actors.

The rapid development and access to technology and digital tools have been inclusive of remote areas in developing economies, where digital devices and connectivity services are becoming accessible and affordable for many. This offers excellent opportunities for harvesting price data through new data collection methodologies, such as crowdsourcing/crowdsensing (citizen-generated data) on mobile apps/sensors. This information can be disseminated, e.g. via web dashboards ([1]) providing access to real-time ([2]) geo-located data to support decision- and policy-making processes. Since 2012, various ICT and crowdsourcing approaches have been implemented in Africa (Zeug et al., 2017). These approaches vary in their degree of openness (groups/crowd open to new members or closed groups/crowd), type of crowd (general public or limited to experts), and duration (spontaneous vs. longer-term continued contributions). Also, approaches vary in terms of the type of rewards provided to the crowd (e.g. monetary payment, mobile phone airtime or none), and in the ICT tools (e.g. web platform, SMS- or smartphone-based solutions). Considering the growing adoption of smartphones, exploring the use of smartphone-based solutions further is highly recommended (Seid & Fonteneau, 2017). These, compared to SMS-based solutions are more user-friendly, less prone to errors and have the potential of collecting data faster with higher volume and enhanced spatial richness (e.g. geolocation) that would be otherwise not possible. Crowdsourcing allows people living in urban, suburban, and rural areas to report and to receive relevant data about their immediate reality (Blaschke et al., 2013), and empowering citizens both as providers and users of information.

However, any market information system that aims at improving the functioning of markets and supply chains requires quality data, accessibility to the data and trust in the system. Therefore, to be reliable, crowdsourcing methods need an effective quality assurance system (Pedersen et al., 2013). More data does not necessarily translate into better information. Compared to classical price-data generation approaches which are based on probability sampling, citizen-based systems rely on non-probabilistic sampling (e.g.

([1]) Web dashboards make it possible to visualise relevant, collectively produced information while contribute to reducing search costs and information overload of data users (Matheus et al, 2018).

([2]) In this context, real-time refers to delays in publication of less than one or two days.

convenience sampling) must be adjusted, among others, for sampling bias. This poses challenges for data quality, especially considering the volume and velocity of data generated. Therefore, new statistical methods, rigorous algorithms, and smart tools are needed for timely data processing with quality controls to ensure reliability and usability of insights generated from such data (DGINS, 2013, 2018).

Further, citizen-based approaches add a new dimension to the generation of data. As data provision is voluntary, getting the crowd engaged and maintaining their engagement is essential. Therefore, in designing such initiatives, it is crucial to consider appropriate incentive systems, possibly combining extrinsic (e.g. monetary rewards) and intrinsic (e.g. personal satisfaction) factors, with the inclusion of an effective feedback mechanism (Pedersen et al., 2013). Evidence from most of the crowdsourcing initiatives suggests that it is impossible to sustain voluntary contribution from the crowd once rewards are phased out. Behavioural sciences provide an additional tool to help sustain crowd contributions by mobilizing behavioural factors via 'nudges' (i.e. information that suggests or gently pushes people to behave in a certain way while maintaining the individual's freedom of action) to strengthen the engagement of individuals with the crowdsourcing platform (Sunstein, 2014a). Similarly, data protection and privacy need to be included as part of the design of crowdsourcing, to avoid reluctance of the citizens to participate in crowdsourcing, out of concerns for their privacy and identity (Ziegler et al., 2017).

Statistical authorities are encouraged to explore innovative methods for collecting and processing data and to find ways to integrate this new data as complementary sources in the statistical process (ESS, 2019). This has already been the case for several countries and institutions. As an example, as of 2017 Eurostat has a new section on its website dedicated to experimental statistics (https://ec.europa.eu/eurostat/web/experimental-statistics), where examples of uses of new data sources and methods to better respond to data users' needs can be found. Another example is that of the Organisation for Economic Cooperation and Development (OECD) with its Observatory of Public Sector Innovation (OPSI). There we find examples of how Artificial Intelligence (AI) in the form of machine learning algorithms can be used to unlock the potential of crowdsourcing for decision-making (OECD-OPSI, 2019). In African countries, some examples of the use of crowdsourcing for collecting data on food prices are provided by the World Bank (WB) (Hamadeh, Rissanen, & Yamanaka, 2013), the Mobile Vulnerability Analysis & Mapping (mVAM) of the World Food Programme (WFP) (WFP, 2016), the European Commission's Joint Research Centre (JRC) crowdsourcing pilot run in cooperation with the African Development Bank (AfDB) (Donmez et al., 2017) and the AMIS-FAO crowdsourcing initiative in Nigeria (Seid & Fonteneau, 2017).

Following the food price hikes and volatility events during the first noughties of the 21st century, the European Commission's Directorate-General for International Cooperation and Development (DG DEVCO) started collaborating with the JRC on innovative methodologies for real-time food price data collection and analysis in developing countries. The work carried out so far includes a crowdsourcing pilot in several African countries (Donmez et al., 2017), two stakeholder and expert workshops in Brussels (2015) and Abuja ([3]) (2019), a methodological study (Zeug et al, 2017), the development of a spatial statistical quality methodology to facilitate the best possible exploitation of geo-located crowdsourced data on food prices (Arbia, Solano-Hermosilla, Micale, Nardelli, & Genovese, 2018), and the preliminary validation of this methodology using AMIS-FAO data from Nigeria (Seid & Fonteneau, 2017). Based on the developed methodology, JRC designed a food price collection system based on crowdsourcing to validate the method as an alternative way to generate reliable real-time information on food prices, namely the Food Price Crowdsourcing in Africa (FPCA). This report provides detailed information about the implementation of the FPCA in two Nigerian states (Kano and Katsina) during 2018 and 2019 by JRC in collaboration with the International Institute for Tropical Agriculture (IITA) and Wageningen University and Research (WUR), and the main results achieved. Nigeria was considered a suitable country to develop and test methodology due to the size of its economy and population, degree of smartphone penetration (GSMA, 2018) and the expansion of mobile broadband (GSMA, 2019), its regional influence, and the diversity of commodities traded. Furthermore, the Nigerian government has recently shown a strong interest in exploring strategies/approaches for smart national-scale data collection, which can improve the quality and timeliness of the Agricultural Performance Survey (APS) data. The latter initiative was generated by the National Agricultural Extension, Research, and Liaison Services (NAERLS) and disseminated through the National Bureau of Statistics (NBS).

The FPCA is presented here as a credible methodology, based on the voluntary provision of data by a crowd (people living in urban, suburban, and rural areas) using a mobile app and a web dashboard to facilitate supply and use of data from the crowd, leveraging monetary and non-monetary (behavioural) incentives to

([3]) https://ec.europa.eu/jrc/en/publication/evidence-food-price-crowdsourcing-africa-fpca-project-nigeria enhance contribution. It applies a sound statistical quality methodology, which allows staple food price data to be collected along the food chain, analysed and validated, and disseminated in real time. This involves three steps:

1. gathering timely, accurate, cost-efficient, high-frequency and highly detailed spatial coverage of food price data at several stages of the food chain from citizen contributions using a mobile app, whose participation is encouraged through monetary and non-monetary incentives (i.e. nudges);
2. processing of data with a formal quality methodology to produce accurate and reliable price estimates at the regional and local level in real time; and
3. disseminating the validated crowdsourced information in real time through an open-source web dashboard.

Finally, we developed an indicator framework for assessing the quality of the FPCA, which we believe can be used to evaluate the reliability of crowdsourced data.

The rest of the report is structured as follows. Section 2 introduces the definition of crowdsourcing and key concepts. Section 3 provides an analysis of the results of a survey conducted to understand better the needs of potential data users and stakeholders. Section 4 describes the setting up of the crowdsourcing system, including the steps involved in building the crowd, and configuring the IT platform and the smartphone app. Section 5 and 6 describe the quality methodology used to produce timely and reliable price estimates and related quality indicators and the web dashboard used to disseminate the crowdsourced prices in real time. Section 7 then combines the results of the data collection and the quality indicators to assess the quality performance of the system. Section 8 presents the analysis and results of the monetary and non-monetary incentives. Section 9 includes some data insights. Finally, Section 10 presents the challenges and conclusions.

2 DEFINITIONS AND KEY CONCEPTS

In this section, definitions and key concepts for the Food Price Crowdsourcing Africa (FPCA) project are presented. As described in the introduction, this study focuses on using smartphone technology, a linked internet platform, and a web dashboard to implement a *crowdsourcing* approach to gathering and disseminating quality data on food market prices in real time, with the data received through voluntary contributions from citizens. This initiative, therefore, explores a possible way to improve *market transparency* in the agri-food sector through ICT and the participation of citizens. However, the focus is not just on technical requirements. Instead, there is particular emphasis on the engagement and motivation of citizens and the quality of the data/information obtained. Accordingly, the concept of *crowdsourcing* is introduced first, along with key elements and examples. Next, differences between the crowdsourcing approach taken in this study and other types of crowdsourcing applied in the same field are noted. Then the concepts of *citizen science and crowdsourcing as a subset of it* and that of *open data are* explained. Finally, some terms related to the crowdsourcing task of this project are explained.

Crowdsourcing is an online, participatory method for getting information or input for a task (e.g. data collection) from a number of people by distributing the task to a pool of people who are usually not employees (Brabham, 2013). The term *crowdsourcing*, coined by Jeff Howe (2006), is a combination of 'crowd' (people) and outsourcing (externalisation). Now, the interconnection of digital devices via the internet, enables the use of *crowdsourcing* to send and receive geo-located inputs by/from citizens. More specifically, *crowdsourcing* can be defined as a sourcing model that consists of *"a type of participative online activity in which an individual, an institution, a non-profit organisation, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task [..] always entails mutual benefit"* (Estellés-Arolas, Enrique, González-Ladrón-De-Guevara, 2012). The crowd participant also referred to as *volunteer,* obtains a reward, be it economical (i.e. *extrinsic motivation* or external rewards*)*, social recognition, self-esteem, fun, or skills improvement (i.e. *intrinsic motivation* or personal satisfaction or accomplishment). Crowdsourcing can, therefore, be either paid or unpaid, but examining the types of motivation is essential for understanding how to promote crowd participation best. In developing countries, the motivation to participate in crowdsourcing may be lacking if there is no economic incentive (Bott & Young, 2012). Yet, behavioural sciences provide an additional tool to help sustain crowd contributions by activating behavioural factors, which strengthen the engagement of individuals with the platform. These factors can be included in the design of crowdsourcing tools to maximize the quantity of data gathered. We use the term 'behavioural factors' synonymously with psychological factors— i.e. the cognitive, emotional, personal and social processes or stimuli underlying human behaviour (APA, 2018). In particular, nudges or information provided to promote a specific behaviour have been useful in multiple interventions, e.g. from promoting tax payments to encouraging citizens to provide development information in a crowdsourcing initiative in Uganda (Blaschke et al., 2013). The latter is one of the few examples of the use of nudges in crowdsourcing.

Concerning configuration, a crowdsourcing system revolves around four pillars (Hosseini, Phalp, Taylor, & Ali, 2019):

- The *crowdsourcer* is the task launcher. The crowdsourcer (which may be an individual or an organisation) launches the (open) call, manages the budget and provides the incentives (financial, social or entertainment-based), provides feedback and data to participants, and takes care of data privacy and confidentiality.
- The *crowd* is the group of volunteers that do the task.
- The *crowdsourcing task* is the piece of work to be done. The task must be suitable for crowdsourcing, i.e. it must be possible to deliver it online and with little guidance.
- The *crowdsourcing platform* is the system used to facilitate interaction between the crowd and the crowdsourcer. The platform must make it possible the enrolment of the crowd, its interaction with the crowdsourcer, carrying out the task, its validation, privacy and confidentiality, the exchange of feedback and the management of payments.

Furthermore, the literature on crowdsourcing distinguishes approaches by the nature of the crowd (Estellés-Arolas, Enrique, González-Ladrón-De-Guevara, 2012). It identifies three types of call: (i) an open call in which anyone can participate, (ii) a call limited to a community with the required knowledge and expertise, and (iii) a mixed call in which an open call is made, but those who can participate are controlled. Moreover, *crowdsourcing* relies on the *wisdom of the crowd* to find solutions. It is based on the idea that large groups of people are collectively smarter than the individuals in the group (Surowiecki, 2005). In this regard, crowdsourcing approaches can be categorised by the nature of the crowd's contributions (Soliman &

Tuunainen, 2015). They distinguish first, *integrative crowdsourcing*, i.e. the contributions are complementary, as their value resides not in the individual contribution but the aggregation of many of them. Second, *selective crowdsourcing*, i.e. the contributions are competitive as it is expected that one provides the optimal solution. The FPCA approach is integrative as other approaches implemented to collect food market prices. Indeed, in crowdsourcing for numerical tasks (i.e. surveying prices), the truth is usually inferred based on the aggregation of multiple observations from multiple contributors for the same task, e.g. collecting the price of 1 kg of rice in a market. If sufficient observations are available, averaging can eliminate random errors that affect each participant's data submissions differently, but not systematic errors which affect them all in the same way. A crowdsourcing initiative must be contextualised according to the reality of the region, especially in a developing context despite the growing spread of the internet and mobile phones, the digital divide and literacy may hinder the potential of the *wisdom of the crowd* if, for example, the crowd is urban biased (Bott & Young, 2012).

The FPCA differs from other *crowdsourcing* efforts that have been implemented in African countries to collect food market prices in several ways. For instance, many of them deviate from the strict concept of open call as they rely on a limited group of trusted and trained individuals or experts (Donmez et al., 2017; Seid & Fonteneau, 2017). In contrast, The FPCA project relies on an open call with minimal control of *crowd participants.* Besides, different to other approaches that require from participants to visit specific markets, in the FPCA project *crowd participants* are not required to go to specific markets at certain times, but to report prices during their routine visits to the market whenever they wish to do so. This approach can be referred to as *spontaneous crowdsourcing* or *opportunity crowdsourcing*.

In contrast to other *crowdsourcing* initiatives for food market prices in Africa that have used SMS-based technology to exchange data on prices, FPCA relies on the internet and smartphones that allow for flexibility in collecting and quality-checking of data by georeferencing each submission. At the moment, the internet is the most used medium in *crowdsourcing* due to the opportunities for citizen collaboration that it offers. Especially, smartphones with GPS technology allow for the collection of geo-located data in [near-]real time. According to literature (e.g. Goodchild, 2007), spatially-referenced crowdsourced data have been also defined as *Volunteered Geographic Information* (VGI). Much of the VGI research in the past has focused on OpenStreetMap (OSM) which is one of the most successful examples of VGI to date, a citizen-driven initiative to create an open map of the world, which rivals many official data sets in richness (Mooney & Minghini, 2017). But a large class of crowdsourcing is represented by geo-data collected through smartphones, to measure phenomena that are otherwise difficult to quantify precisely and in a time-efficient way. Globally, this approach is becoming more popular in several sectors, including Agriculture and Health (Buchbinder, 2017). Recent research suggests that data crowdsourced by the civil society should be integrated into the international process for monitoring Sustainable Development Goals (SDGs), and for this, it is essential to address potential data quality issues (Flückiger & Seth, 2016). In the agricultural sector in the developing world, large amounts of geo-located citizen-generated data can be put in the context of big data[4] and high-speed communication to support the provision of extension services for optimised input use (e.g. optimised irrigation and pesticide and fertiliser use), pest control and early warning systems, but also to improve market transparency, leading to greater arbitrage opportunities, to reduce information asymmetries and to improve farmers' bargaining power (Deichmann, Goyal, & Mishra, 2016). In this respect, market transparency is defined as '… the availability of relevant market information to market participants. [This includes] prices, weather, production … and stocks' (AMTF, 2016; Solano-Hermosilla, Ciaian and Kathage, 2019). In this project, it is expected that a crowdsourcing approach to monitoring food prices will be a convenient, inexpensive, and relatively quick way of compiling a large dataset that will be useful for markets actors, policy-makers and institutions.

*Crowdsourcin*g is a form of *citizen science* when it enables open collaboration in which citizens participate in the process of generating scientific outputs (Wiggins & Crowston, 2011). Applying *crowdsourcing* to *data science* initiatives, via open data portals, allows large amounts of valuable data to be acquired at a potentially low cost. Crowdsourced data is a form of *open data* if it can be freely used and reused. This is, it is made *accessible* in such a manner that data can be readily found and used, is intelligible and can be assessed with respect to reliability (The Royal Society, 2012). Open data is expected to foster citizens' collaboration (Hofmokl, 2010). With this in mind, the FPCA project disseminates validated data in close to real time through a web dashboard embedded in an open data portal. The *validation process* and automating it is, therefore, a

(4) We refer to the broad concept of 'big data' as advanced by Beręsewicz et al. (2018), which classifies 'big data' types in (i) data generated by people and stored in a digitalised format (e.g. from mobile apps, twitter), (ii) data produced automatically by people when interacting with IT systems (e.g. scanner data) and (iii) machine-generated data usually captured by sensors (Arbia, Solano-Hermosilla, Micale, Nardelli, & Genovese, 2020, forthcoming).

key aspect of a *crowdsourcing system* before data is used and/or shared with the public for further use in analysis or decision-making. Finally, concerning terminology, in FPCA the digital contribution from citizens consists of data submissions by volunteers, each of which gives rise to a *data record* containing the price of one or more food products, their characteristics, the time, and geo-location expressed in geographic coordinates. A data point contains the same information as a data record for one single food product.

3 Horizon-scanning and stakeholder perceptions

Data on food prices gathered collaboratively by citizens using the presence of Internet-connected geo-located devices allows large amounts of data to be collected, which makes crowdsourcing a way of improving market price information systems. A better price information system that enhances market transparency can help market actors to make better informed decisions on commercial, production and investment transactions, but also support policy-makers, international organisations and the general public (FAO, 2018). By definition, market transparency implies that relevant information is available to all data users. A sound information system with a well-designed dissemination tool should help data users by reducing search processes and information overload. For this reason, to be relevant, any information system must include in its design the needs and views of the potential data users and stakeholders.

For this purpose, an online survey was sent to a number of stakeholders possibly interested in information on food prices in Nigeria through a questionnaire. This was sent to 46 people representing, among others, government, research, development cooperation and farmers' organisations. There were 14 respondents to the survey. Many of the non-respondents belong to farmer's crop associations in Kano and Katsina. Despite the small number of responses, these have made it possible to create a number of profiles or "personas" from respondents, associating them with the typical answers received.

3.1 THE SURVEY

The survey, entitled "Questionnaire on price and market information systems in the agri-food food chain in Nigeria" was designed using the *EU survey tool* and consisted of several sections about:

1. The business or organisation represented by the respondent
2. The current agricultural/food price and market information systems: availability and use in the business or organisation.
3. Gaps, challenges and potential for improving market information systems
4. Crowdsourcing or a citizen engagement approach for voluntary data collection to inform on market food price developments
5. Optional information including the name and references of the respondent and her/his organisation, awareness of other similar initiatives, knowledge of other organisations and people potentially interested.

The survey was composed of 30 questions in total (questionnaire available at https://ec.europa.eu/eusurvey/runner/MarketInformationNigeria).

3.2 THE RESPONDENTS' PROFILE

As anticipated, a number of profiles or "personas" (a character representing a particular type of user of the information which can incorporate one or more respondents) have been identified and associated with the most frequently received answers.

The following "personas" were identified:

1. The Trader. Includes importers and exporters of agricultural goods
2. The agro-processor
3. The government manager, at Federal Ministry of Agriculture and Rural Development (FMARD-Nigeria) and in the extension services
4. The technical assistance advisor that operates in development cooperation, including bilateral and multilateral organisations; private consultants are also included
5. The researcher, which includes data analysts and researchers from academia.

3.2.1 The Trader

The Trader's main commodities of interest are rice, maize, other cereals and beans. He only uses sources of price information from within Nigeria, i.e. the Commodity Exchange (NCX) and National Bureau of Statistics (NBS), with weekly to monthly coverage and reporting at the state level. However, the Trader would require daily information to address his organisation's needs, and especially management decisions and market price dissemination.

He sees improved price and market information systems in Nigeria as a means to (most importantly) reduce uncertainty for operators (including price volatility), increase opportunities for risk management (futures markets, insurance options, access to credit, improved contracts, etc.), increase competition, and improve investment decisions of farmers (in the long term).

How can we improve market price information systems? According to the Trader, this will be achieved by improving the quality of existing information (harmonisation, independent verification, etc.,) and by collecting data at farm gate and wholesale trade levels (most importantly), and also retail and input suppliers level. A commodity list and geographic coverage are the most important aspects to him, followed by timeliness and dissemination methods.

The Trader is familiar with price data collection activities and had some previous experience with crowdsourcing methods. He believes that funding for the set-up and ongoing costs of the system and lack of trust are the main barriers.

Obtaining timely and reliable market food price data and analysis from crowdsourced contributions would add an incentive to participate as a data contributor. However, an additional reward system may be needed. He is keen on a follow-up and receiving more detailed information.

3.2.2 The Agro-Processor

The Agro-Processor's (only) commodity of interest is tubers. No sources of price information are used, as the organisation relies on fixed supplier contracts.

In any case, improved price and market information systems are seen as a means to improve farmer's production decisions (in the short term); they would also help operators to identify opportunities (better product offer/better market) within their country, and contribute to research and the generation of knowledge in the agri-food supply chain.

Better use of existing information and dissemination methods would improve market information systems. Levels to be addressed should include input suppliers, the first processing stage and retail. He has little or no knowledge of price data collection systems and crowdsourcing methods, and no opinion as to whether crowdsourced food price data may be an incentive to participate. Also no interest in a follow-up, nor in receiving more detailed info.

3.2.3 The Government Manager

He is interested in a very ample range of food commodities: rice, maize, other cereals, beans, tubers, banana (Plantain), Gari, eggs and chicken, beef, fish, milk and vegetables.

He uses a variety of sources of price information, mostly from within Nigeria: National Agricultural Extension Research Liaison Services (NAERLS), the Strategic Grains Reserve (SGR), NCX, NBS; he also uses the Famine Early Warning Systems Network (FEWSNET) Nigeria price bulletin. Given the wide variety of sources, the frequency is weekly, monthly, and yearly, and reporting is at market, municipality, and district levels. The main organisational need is for management decisions.

Improving information systems would reduce uncertainty for operators, increase opportunities for risk management, and improve farmers' investment decisions (in the long term).

How can we improve market price information systems? According to the Government Manager, this should be through a large number of actions, i.e. mandatory reporting by operators, voluntary reporting by operators, better use of existing information, improved quality of existing information, and facilitation of dialogue among stakeholders.

At what level should the information be collected? Almost all, according to him: input suppliers; farm gate, first processing stage, subsequent processing stages, wholesale trade, retail, and consumers.

Timeliness, geographic coverage, commodity list, and dissemination methods are all aspects which should be improved. He is familiar with price data collection activities and had some previous experience with voluntary unpaid crowdsourcing.

Funding for set-up and ongoing costs is the constraint identified. Obtaining timely and reliable market food price data and analysis out of the crowdsourced contributions would add an incentive to participate as a data contributor. He is interested in a follow-up and receiving more detailed information.

3.2.4 The Technical Assistance Advisor

His interests are very specific: vegetables, tomatoes, ginger and chilli. He is currently not making use of any information system. This is because of insufficient commodity coverage, lack of timeliness, and lack of reliability in the data provided by current systems. Improving information systems would lead to increased competition and "level the playing field for farmers" (access to market information, bargaining power, mutual trust).

How can we improve market price information systems? According to the Technical Assistance Advisor, this would be mostly done through mandatory reporting by operators, better use of existing information and improved quality of existing information.

At what level should the information be collected? Input suppliers are most important, followed by farm-gate and wholesale trade.

Timeliness is the most important aspect to be improved, followed by commodity list and dissemination methods.

He is not familiar with price data collection activities, nor with crowdsourcing. He believes that the main constraints are funding, lack of incentives for continuous participation, technology and legal issues.

The constraints identified are lack of trust, lack of adequate statistical methods (including validation) to deal with this type of data and funding for setting up and current costs.

For him, obtaining timely & reliable market food price data and analysis would be an incentive to participate as a data contributor. He is interested in a follow-up and receiving more detailed information.

3.2.5 The Researcher

She is interested mostly in staple food, i.e. rice, maize, other cereals, beans and tubers, but also in Gari, bread, oil and vegetables.

The National Bureau of Statistics (NBS) and FEWSNET Nigeria price bulletin are her main sources of information. The frequency of the data available is monthly to yearly and reporting is at district and state levels (for NBS data) but also at market level (for FEWSNET data). Food security, other sectoral policies and research are indicated as the main organisation needs.

Improving information systems would lead to a range of positive effects, most importantly, due to increased opportunities for risk management, reduced uncertainty for operators and improved farmer investment decisions (in the short term).

How should we improve market price information systems? According to the Researcher, this would be by improving the quality of existing information, voluntary reporting by operators, and better use of existing information.

At what level should the information be collected? She indicates a wide range, from farm gate to wholesale trade, retail, consumers and input suppliers.

Geographic coverage and dissemination methods followed by timeliness and commodity list are the aspects which should be improved. She is familiar with price data collection activities and has some previous experience with voluntary rewarded crowdsourcing.

The constraints identified are lack of trust, lack of adequate statistical methods (including validation) to deal with this type of data and funding for set-up and ongoing costs.

Obtaining timely & reliable market food price data and analysis would be an incentive to participate as a data contributor. She is interested in a follow-up and receiving more detailed information.

4 Setting up the system

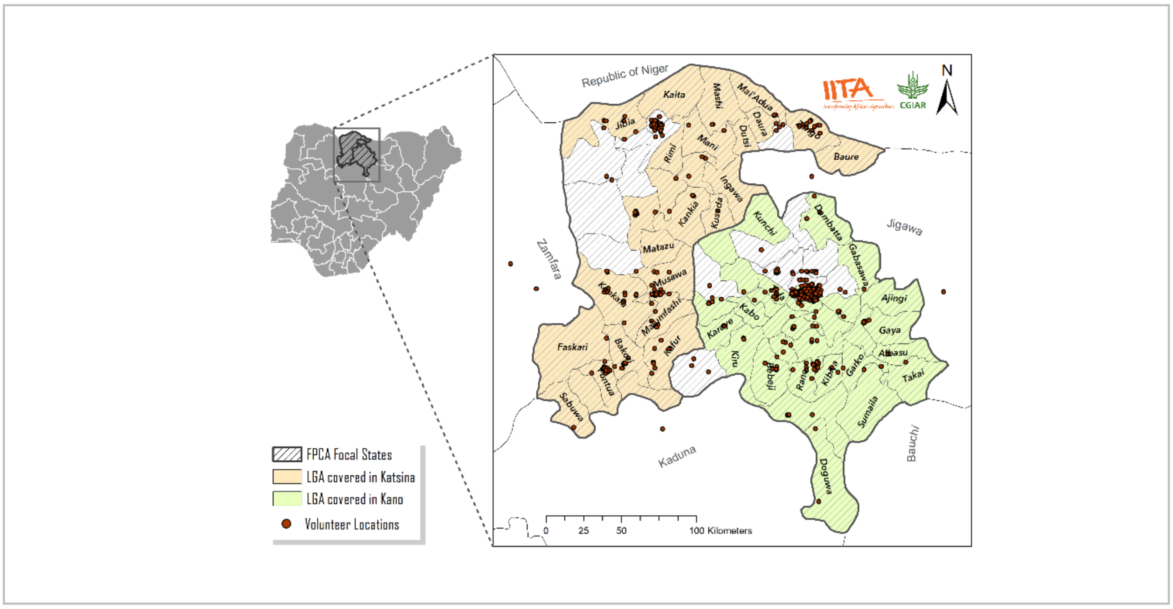
The FPCA project aims at implementing a crowdsourcing method to collaboratively build a reliable, timely, high-frequency and cost-efficient dataset of food prices in the areas of interest in Nigeria based on a smartphone application, use of the internet, and the participation of citizens. This section describes the implementation of the crowdsourcing approach for an on-the-ground data collection on food prices, with a focus both on the technology and on the voluntary participants (the crowd) as individuals contributing data to the crowdsourcing system, and at the same time being potential users of the collective output of the crowdsourcing system.

4.1 SCOPE AND APPROACH

4.1.1 Geography

The project was implemented in Nigeria, which is administratively subdivided into 36 States and the Federal Capital Territory (FCT) of Abuja, 774 Local Government Areas (LGAs), by order of hierarchy (HDX, 2019). The third level of administrative unit is the wards. The project mainly focused on the states of Kano and Katsina, in the Northern region of the country (Figure 1). These two states are among the largest cereal and legume producers in Nigeria (Table 2). They have major accessible rural and urban markets, which are spatially dispersed in different parts of each state. Also, the presence of large cities with a medium-density population is considered a favourable factor for assessing intra- and inter-urban market variations. For instance, according to the most recent census figure, Kano was ranked as the most populous state in Nigeria with a moderately high population density of ~450 people/km^2. Similarly, Katsina was ranked in the 4th position, thus indicating high potential for engaging citizens as volunteers of the crowd to conduct the price data collection. Broadband mobile technology is readily accessible in most parts of these states, and the majority of the people living within and outside city limits currently use and access mobile services for everyday communication or business. Generally, food security issues are common in some areas of both states (Ovaga, 2012). Food insecurity is related to the impact of the massive displacement of farmers to camps due to terror-related threats to their lives and property ([5]) (Babagana et al., 2018). Despite the gradual return to normality, revitalising the food production system in affected zones is a tall order, and this is complicated by population growth trends, low levels of literacy in rural areas, and poor access to basic social infrastructure.

Figure 1. Map showing the focal states of the FPCA project, Local government areas (LGAs) where flyers were distributed, and the geo-referenced locations of the volunteers as of the registration time



([5]) The situation has greatly improved over the past two-three years, and more progressively during the implementation period of the FPCA project.

4.1.2 Commodities

The price data collection was focused on four major grain-legume commodities, namely maize, rice, beans, and soybeans, either sold in locally measured quantities or packaged in standard quantities. However, considering the dominant differentiation of these commodities in the market, with price differing based on the source and type, it was relevant to sub-categorise some of the commodities. Therefore, maize was sub-categorised as yellow and white maize, rice was sub-categorised as Thailand, Indian, and Local Grade (1-5) class, while beans were sub-categorised into white and red beans (with grades ranging from 1-5).

4.1.3 Market/outlet types and food chain stages

Timely and reliable price data can help participants in agri-food value chains to make commercial decisions at each stage of the chain and support the decisions of governments and other stakeholders (FAO, 2018). Due to the progressive evolution of commodity prices from the farm to end consumers, the crowdsourcing tool was designed to ensure full coverage of stages along the commodity value chain. This describes the flow of activities/services from the primary producer to the final consumer (i.e. "from field to fork") and thus includes all levels of market transaction (stages) and actors that play a role in the production, distribution and transformation of the commodity. The most common food chain stages at which prices can be observed are: farm-gate (transaction occurs near the farm), wholesale (where traders usually sell to other traders, normally in large volumes (e.g. >50 kg bags) and retail (where food commodities are mostly sold to end users, i.e. consumers, and transaction volumes are usually small) (Ismail, 2010).

In practice, we let participants identify the type of market/outlet (e.g. open market, supermarket and convenience shop) from which prices are submitted. This information is then used to derive the transaction level/food chain stage (Table 1). This approach appears to be more suitable given that volunteers are spontaneous, untrained and not necessarily familiar with the different levels of commodity transaction (i.e. farm gate, wholesale, and retail). The final standardised definition of market/outlet types that were covered (Table 1) was subject to contextual descriptions, in line with known categories of markets/outlets for African markets as described by the International Comparison Program (ICP) (World Bank, 2015).

Table 1. Mapping typical categories of outlet to food chain segments.

Market/outlet Types	Description	Chain level
Directly from farmer	Located at or close to the farm	Farm gate
Bulk and discount stores	Wholesale stores, discount stores	Wholesale
Markets (city market, village market, open air or covered market)	Open markets, covered markets, wet markets	Wholesale (> 50 kg bag)
		Retail (< 50 kg bag)
Street outlets	Mobile shops, street vendors	Wholesale (> 50 kg bag)
		Retail (< 50 kg bag)
Medium and small shops	Mini-markets, kiosks, neighbourhood shops, grocery stores, convenience stores	Retail
Large shops	Supermarkets, hypermarkets, department stores	Retail

Market/outlet Types	Description	Chain level
Specialised stores	Mid-sized supermarkets, raw food stores, commodity-specific marketss/supermarket	Retail

Source: Based on classification of outlet types by the International Comparison Program (ICP) (World Bank, 2015) and Ismail (2010)

Note: (1) Coverage of aggregating points (where smaller quantities from farmers or small scale traders are accumulated for further sales, to reduce marketing costs) and processing points was considered to be beyond the scope of this data collection, so they were excluded.

(2) During the pilot phase, the option 'Other market' was left open to capture other all potential types of store/outlet from which prices can be reported. After the pilot, the list of market types was expanded and the open category for 'Other markets' was excluded to minimise entry errors.

(3) Although the transaction stage is usually associated with the actual outlet/market, retail and wholesale sales may occur in the same market. Therefore, the transaction volume is used as a further qualifier. Typically, volumes per transaction in wholesale nodes tend to be larger, e.g. multiple 50 kg bags or more (Ismail, 2010).

4.1.4 Prices

Transaction prices. The goal of this crowdsourcing price survey is to collect the prices that purchasers pay to sellers to acquire the food commodities on the list. Therefore, the volunteers (sellers or buyers) were asked to submit actual transaction prices; however, the survey tool was flexible enough to accommodate price submission when the volunteers assume the role of a price observer ([6]).

Type of prices. Based on the definitions of food chain levels (Section 4.1.3), we considered the three types of prices: farm gate price (price received by farmers), retail price (price paid by final consumers) and wholesale price (price between farm and retail).

Currency. Prices were submitted in the local currency, Nigerian Naira ([7]) (₦).

Packaging unit. a pre-defined set of common packaging units was defined for the commodities, relative to conventional and recognised standards. The packaging units were based either on actual standard weights (100kg bag, 50kg bag, 25kg bag, 10kg bag, 5kg bag, 1kg bag) or on the most popular volume-based measure (Mudu/Kwano/Tiyya), which roughly equates to 2.6 kg per measure (Kormawa & Ogundapo, 2004) ([8]). It should be noted that the conversion of Mudu to kg varies among the Nigerian states.

4.1.5 Crowdsourcing model and sampling

The design of the FPCA crowdsourcing model is as follows:

Crowdsourcer: The JRC is the launcher and funder of the FPCA project, which was implemented on-the-ground in collaboration with IITA and Wageningen University and Research.

Crowd: FPCA centred on engaging citizen volunteers (the crowd) from different backgrounds and geographical areas. A participant may be a seller, a buyer, or a price observer that participates voluntarily, i.e. convenience sampling (a type of non-probability sampling approach). In the implemented approach, the crowd is initially addressed via an open call through a media campaign based on flyer distribution and radio ads. This was then enhanced by the participation of extension agents and by word of mouth, thus resembling partly a snowball sampling approach (a type of non-probability sampling). The requirements for volunteers to register is to own a smartphone with GPS functionality and be able to follow online instructions, which, in particular, in a developing country context can produce a certain selection bias related to the need for internet and smartphone access and a degree of literacy.

Crowdsourcing task: The FPCA asks volunteers through a survey/questionnaire in the app to submit prices of certain foods, anonymously, voluntarily and spontaneously during their routine visit to the market if they wish to do so. Spontaneity refers to the fact that the volunteers are not requested or

([6]) See the questionnaire in Annex 1 for the closed list of transaction type options implemented in the app during the roll-out phase from which participants had to choose.

([7]) During the time of the FPCA implementation, 1 Naira corresponds 0.002867 EURO (https://www.cbn.gov.ng/rates/exrate.asp).

([8]) Automatic conversion to standardized values per kilogram occurs during the data processing phase of the project (Section 5.4.1) for comparison across commodities and market segments.

committed to visiting the market regularly for data submission, seeking to reduce self-selection bias. It was expected that a diverse and robust crowd size would suffice to generate a daily data flow of commodity food prices with extensive spatial coverage, without any absolute commitment from any volunteer to submit data.

Crowdsourcing platform: A pre-configured open source server (ONA) and a mobile-based tool (ODK) is used in this project. Communication with the volunteers is carried out via an information website, SMS messages and a web dashboard ([9]) to disseminate the crowdsourcing outcome, where data confidentiality and privacy is fully ensured. A system for paying incentives and motivating participants was developed and implemented. Finally, a series of algorithms to retrieve, validate and aggregate the individual contributions to produce reliable price estimates were programmed in R software (R Core Team, 2020) and are available upon request by the authors.

4.1.6 Period and phases of the project

Based on the expected seasonality of commodity price trends, the data collection was planned to cover the period around the onset of the harvest season (i.e. late September), up to the period around the commencement of planting (i.e. late June), in the region. The overall project implementation was further sub-divided into three (3) phases, which include:

- The ideation phase (end of July 2018 to mid-September 2018)
- The pilot phase, six weeks (mid-September Sep 2018 to end of October 2018), and
- Full roll-out, 35 weeks (End October 2018 to end of June 2019).

After the roll-out phase was completed, the platform was kept open for a period for voluntary contributions from citizens without the payment of monetary or other incentives, which has made it possible to explore the sustainability of the voluntary contribution on this type of platform.

Table 2. Basic description of pilot focal states for food price crowdsourcing in northern Nigeria.

State	Demographics	Popular agricultural produce	Mobile broadband and cell phone access	Major opportunities	Potential limitations
Kano	Population = 9.5m Density = 470 people/km^2 Male: 5m; Female: 4.5m	Sesame, soybeans, garlic, chilli peppers, millet, cowpeas, sorghum, maize, and rice	MTN, Airtel, 9 mobile, Globacom Overall mobile phone access: ~40%	True to its appellation as the centre of commerce, Kano is home to major markets in northern Nigeria and a confluence of trade routes originating from other parts of the country and other sub-Saharan countries such as Niger and Chad. Therefore, commodity prices in Kano have economic implications for other neighbouring countries.	- Language barrier (this can be mitigated by preparing publicity materials in dual languages – i.e. English and local languages)
Katsina	Population = 5.8m Density = 323 people/km^2 Male: 3.0m; Female: 2.8m	Millet, guinea corn, sesame, cowpeas, groundnuts, maize, rice, wheat	MTN, Airtel, 9 mobile, Globacom Overall mobile phone access: ~40%	Katsina shares its northern boundary with the republic of Niger. The state is mostly characterised by dryland agriculture due to semi-arid conditions. The last 3-5 years, food production has been threatened in many areas of the state due to several instances of terrorist attacks. As the farming	Language barrier (this can be mitigated by preparing publicity materials in dual languages – i.e. English and local languages)

([9]) The dashboard is published at https://datam.jrc.ec.europa.eu/datam/mashup/FP_NGA

State	Demographics	Popular agricultural produce	Mobile broadband and cell phone access	Major opportunities	Potential limitations
				communities resettle, there are new opportunities to assess the dynamics of food pricing within the state.	- Security: fairly recent spate of terrorist activities may limit access to markets or farms in the northern parts of the state. However, we anticipate that the ongoing improvement in security outlook will continue.

4.2 IMPLEMENTATION APPROACH (A-I)

The overall implementation followed an A-I stepwise approach, which covers the required elements for successful implementation of a price-monitoring crowdsourcing system, ranging from initial planning to system set-up, and interaction between collaborators. The overall implementation was divided into three phases: ideation, pilot, and roll-out.

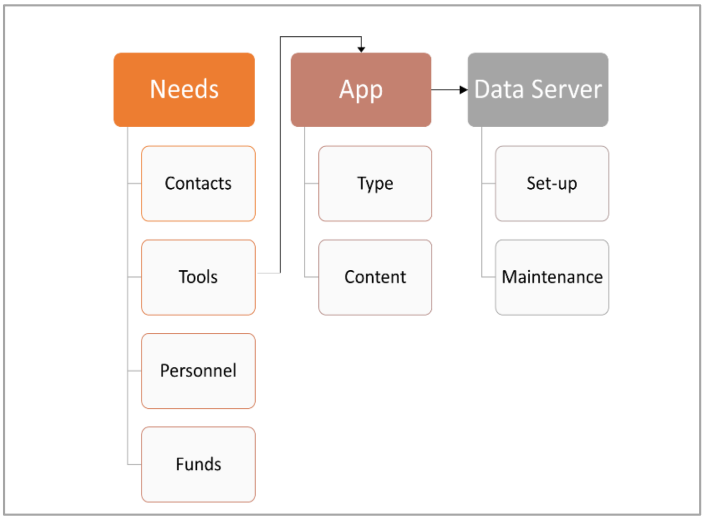
The ideation phase included the inception, initial planning, and set-up of the data collection system through careful assessment of needs and tweaking of the overall method.

The pilot phase, which lasted for four weeks after the ideation phase, involved the initial set-up of the crowdsourcing system and testing with a limited number of volunteers (200). The data collection continued from the pilot phase into the full roll-out phase, which lasted for around eight months (following the pilot phase) and involved engaging a larger number of volunteers of the crowd (737 volunteers) and incorporating insights from the pilot phase. Details of the stepwise (A-I) approach, which covers these three phases, are provided below.

4.2.1 A - Assessing needs and designing platforms

At the preparatory stage, an initial scoping of needs for the crowdsourcing task was conducted (Figure 2). This step is considered important before the commencement of the pilot phase, for strategic coordination and timely delivery of targets. The major considerations were guided by requirements for crowd independence, data visibility and accessibility, and flexibility to implement changes within the system set-up, when and where necessary. Therefore, the initial needs identified include:

Figure 2. Initial needs identified for implementation of crowdsourcing



i. Contacts: These were based on the notion that successful implementation of the crowdsourcing initiative requires proper understanding of contextual realities and opportunities to engage a large number of volunteers and sustain dataflow. Therefore, potential local contacts were considered and consulted to provide insights and support critical aspects of the project implementation, mainly publicity for crowd enlistment and reward incentive management. Moreover, considering that initial (direct or indirect) contact with prospective volunteers is indispensable for crowd activation, appropriate publicity options were considered relative to cost, penetration potential, and timeliness. Publicity options considered include "word of mouth", flyers, phone calls, website, radio, and TV advertising as potential means to initiate contact with the public and engage the interest of prospective volunteers.

ii. Tools: Because this project should provide a proof of concept for crowdsourced food price monitoring within a fairly short time frame (less than a year), it was imperative to implement it within the framework of an existing technology/system that is secure, flexible, and accessible. For that, the relevant tools and available technology options to operationalise the major components of the crowdsourcing system were assessed, including the front-end crowdsourcing tool with the crowd (app) which must seamlessly interact with the back-end database and visualisation platform (the system) in real time. The major criteria for the final selection of the tool include (1) Readiness for use as an existing resource/technology, (2) Ease of deployment and (3) Minimal or no downtime risk for steady operation. Based on previous crowdsourcing experience in IITA ([10]) and evaluation of the available tool options, the smartphone-based open data kit (ODK) was selected as the most compatible front-end tool for the envisioned crowdsourcing system. The ODK tool works on major smartphone operating systems. It can synchronise with cloud-based servers, which are configured to assimilate data from the ODK-based survey tool in real time, the multi-user volunteers submit it. For the back-end system, the robust and reputable cloud-based server ONA (www.ona.io) was selected and configured to provide functionality for visualising, downloading and editing submitted data (Figure 3 to Figure 5). Although ONA provides free accounts for public use, this project was set up using a paid enterprise (super-user) account to ensure secure access and protect any personal data associated with volunteer submissions. In the overall crowdsourcing set-up (Figure 3), the central administrator prepares the survey questions (app questionnaire or *Data Submission Form*) in .xls format (Annex 2) and administers it by uploading it into the ONA server (Figure 4). Each volunteer (observer, buyer, or seller) installs the ODK app on their phone and accesses the administered form (in mobile-compatible format) by specifying relevant configurations, such as the server URL and account details (Figure 5). Subsequently, at each instance of data submission, the volunteers complete the relevant questionnaire, either by providing actual price at the point of transaction or by asking vendors (farmers/traders/sellers) about the price of each commodity. The ODK tool has the advantage of functioning both online and offline; however actual data upload (into the server) from each volunteer only occurs when the volunteer is connected to the internet. As each volunteer submits a new record, the data becomes immediately accessible on the ONA server dashboard and can be formatted and downloaded as maps, charts, and tables by the administrator (Figure 4).

iii. Personnel: Although the crowdsourcing system was envisioned as a fully automated system, we assessed personnel needs for tasks such as distributing flyers, handling enquiries, managing the reward system, carrying out initial reviews of data and other related tasks. In summary, one full-time locally recruited staff member was considered necessary to support office-based tasks (managing enquiries and the reward system, monitoring data flow, and handling administrative issues). Besides, two ad-hoc locally recruited staff were periodically engaged to support both out-of-office and in-office tasks (including preparing and disseminating communication materials, verifying volunteer profiles, and occasionally checking data before processing rewards).

iv. Funds: The costs of various tasks, sub-tasks, and reward incentives were calculated so that we could budget accordingly and guide cost-effective decision-making regarding major needs such as the choice of publicity approach (and vendor), and thresholds for crowd reward incentive. For instance, by breaking down the cost of implementation (including the staff costs) relative to available budget, it became clear that the reward system must be managed in-house (with the use of an existing mobile payment platform). The reason for that is that major banks or telecom companies that can fill this role required robust contracts for partnership. Similarly, due to the cost implication, and perceived low value to the project objectives, an initial proposal for periodic contact sessions for volunteers of the crowd within their respective localities was abandoned. In contrast, after receiving cost estimates for flyer design and production, more flyers were produced beyond the initial target (5000 copies), to improve the likelihood of reaching more prospective volunteers.

([10]) IITA previously implemented non-spontaneous crowdsourcing of agro-dealers' information within similar geography in Nigeria.

All these needs were considered with the overall intention of deploying a fully functional and nearly contactless system that minimises the costs of management and engagement with prospective and current volunteer crowd members, and that makes real-time tracking of/access to data possible.

Figure 3. Schematic representation of crowdsourcing as a model for efficient collection of robust data

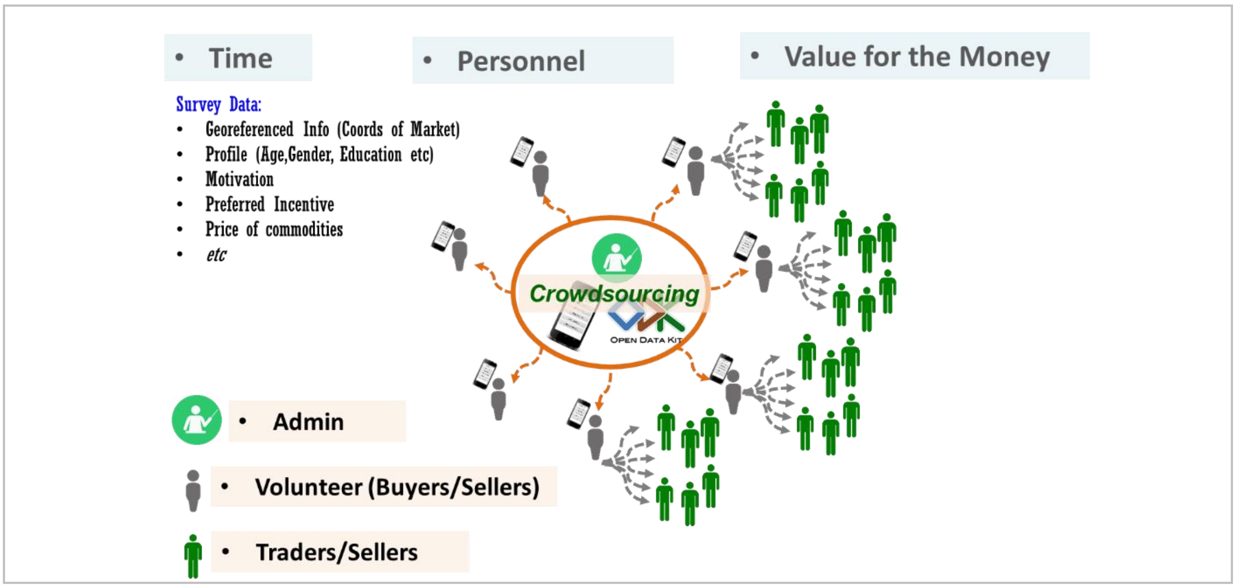


Figure 4. Schematic representation of overall set-up for mobile-based data collection and data output, starting with the preparation of form in Excel (.xls) format, upload onto the server, and data submission

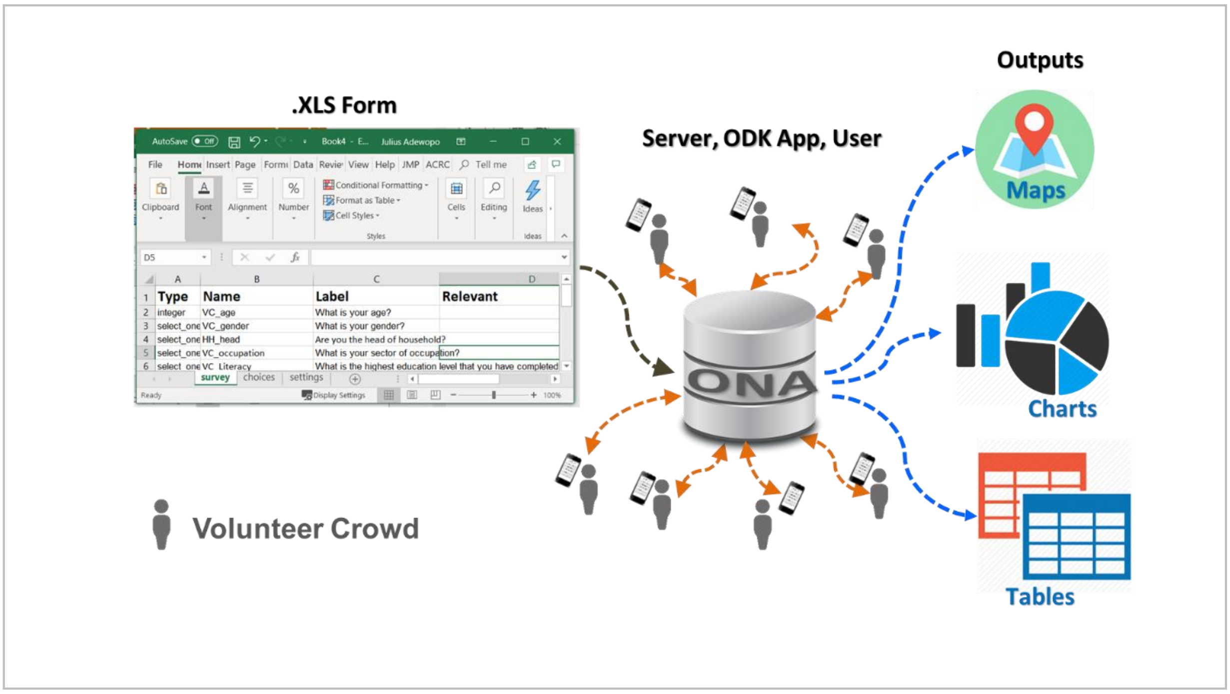
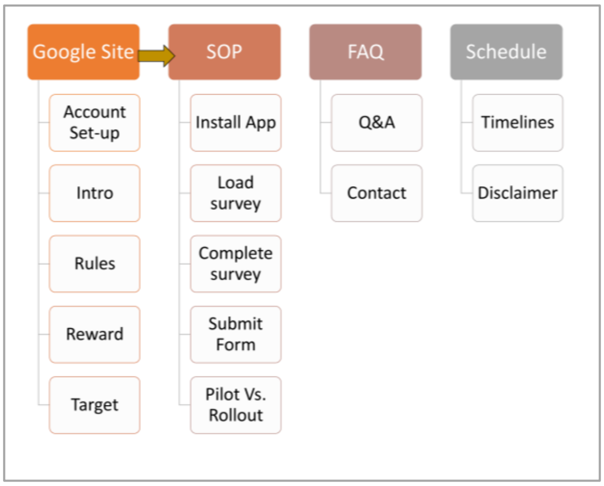


Figure 5. User interface steps for FPCA data collection using the smartphone-based ODK app

4.2.2 B – Build Google site with contents

Based on the identified need for a "go-to" information platform that allows prospective and current volunteer crowd members to access information contactless, a miniature Google-hosted website (https://sites.google.com/view/foodprice; Figure 6) was developed. The website provided basic information about the purpose and approach of the project, how to participate, rules and guidelines, and frequently asked questions (FAQs).

Figure 6. Major elements of the Google site as a simple one-stop platform for volunteer access to information

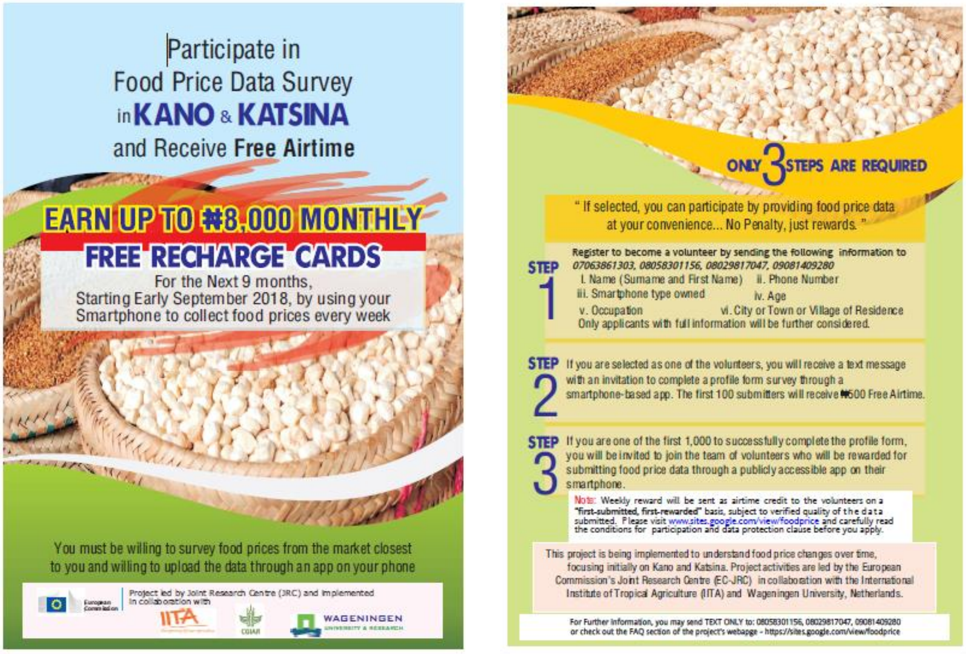


4.2.3 C – Compose and disseminate publicity (radio broadcast and flyers)

Based on the identified need for effective publicity to initiate the crowdsourcing campaign, flyers and radio media were selected as the most promising means to reach the prospective volunteer crowd in the region. The flyers were designed to provide sufficiently detailed, yet succinct information to attract the attention of the crowd and refer them to the earlier mentioned Google site or prompt interested prospective volunteers to contact the local implementation team through a dedicated phone line (Figure 7). Eight thousand copies of the flyers were printed and distributed through various means. Agricultural Extension Agents (EAs) who often work within rural areas were also contacted to support the wide distribution of the flyers. Out of the 44 local government areas (LGAs) in Kano, the flyers reached 35, while 27 LGAs were reached (out of 33 LGAs) in Katsina State (Figure 1). Additionally, we made attempts to reach out to graduates who are participating in the mandatory 1-year National Youth Service Corps (NYSC) program (https://portal.nysc.org.ng/). The NYSC members are typically posted to every Local Government Area (LGA) within each State. Within each LGA, a local government inspector (LI) has oversight functions on the activities and engagement of the NYSC entities. The flyers were partially distributed to these LIs with the expectation that this may broaden the geographical representation of the volunteers of the crowd (i.e. targeting coverage of all LGAs). A digital version of the flyer was posted on platforms such as WhatsApp and Facebook groups for the NYSC.

Additionally, two prominent radio media houses (Cool FM and Arewa FM) in the two States were contracted to develop a 1-minute-long advert based on a drafted transcript to capture quick information on the project and invite the listeners to participate by taking further steps. The advert([11]) was broadcast at the start of the project over a week, and it encouraged prospective volunteers to visit the project's website and complete a profile form or send a text to register their interest. Those who sent in texts were invited to visit the project website, follow the protocol and submit the profile form.

Figure 7. Flyers which were designed and distributed to publicise the crowdsourcing initiative across the focal States in Nigeria



4.2.4 D – Develop data reporting templates

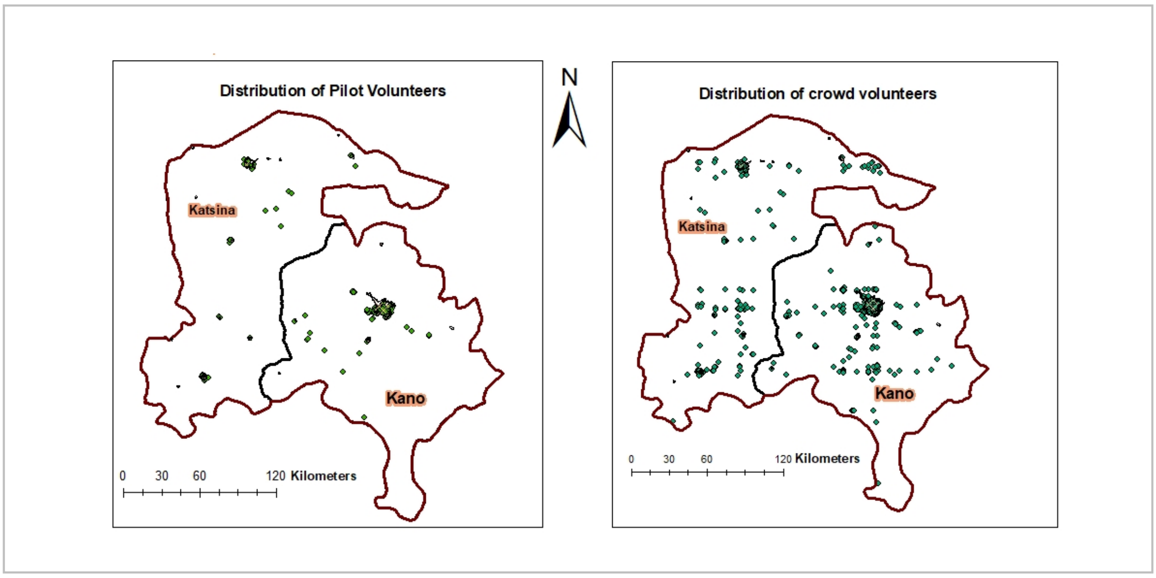
The raw daily prices submitted by the volunteers are categorised as "Level 0 Data" (the raw crowdsourced data). The subsequent revised versions (e.g. cleaned versions with "reward status" column included) were sequentially categorised as "Level 1 Data" and "Level 2 Data". Each step corresponds to a different quality level based on the data validation and quality control procedure applied (see Section 5.4). The Level 0 data was readily accessible and directly downloadable from ONA platform in various formats (.xls, .xml, .csv, .pdf). The Level 1 data was initially prepared manually and periodically uploaded into JRC's ownCloud service known as JRC Box facility for data management. Then, as part of the development of the validation and quality control procedure, an automatic routine was implemented in parallel to the citizen-driven data generation process to extract the data from ONA via their API, filter out the noise in the data and create a useful dataset for sharing with the public ([12]). Then the data transfer was ultimately stopped because JRC was able to pull the data directly from the ONA server.To standardise data reporting and ensure that variables were properly documented, a .xls data reporting template was developed with four tabs for metadata, variables, data, and data history.

([11]) https://www.dropbox.com/s/8uhzwmk5qeimg5k/FPCA_Jingle_v2.mpeg?dl=0
([12]) The automated script which was developed pulls, edits and filters raw data directly from the ONA server platform via their API. However, the feedback loop to the reward system could not be implemented during the duration of the project, so this step had to be done manually but could be considered for full automation in future implementations.

4.2.5 E - Engage prospective volunteers and tally qualified candidates:

Before onboarding the prospective volunteer, it was considered necessary to verify candidates' competency, to increase the expected integrity and reliability of the crowdsourced data. This was designed as a self-evaluation process that allows volunteers to demonstrate their ability to use the smartphone-based app, answer basic structured questions (in English), and send timely data. Therefore, all prospective volunteers were invited to complete the "*Data Profile Form*" (Annex 2) and submit it through the mobile-based ODK app. This initial task was useful in validating the competency of the volunteers of the crowd and prompting timeliness of profile data submission, before inviting each volunteer to commence actual data submission. The profile form submission was active during and after the pilot phase for access by new prospective volunteers, who were eventually enlisted in the full roll-out (Figure 8).

Figure 8. Distribution of volunteers who successfully registered during the pilot and full roll-out phase of FP CA project. The green dots indicate volunteers whose location was geo-referenced at the point of registration



4.2.6 F - Finalise onboarding of candidates and deploy survey form

After verifying the profile data for each member of the volunteer crowd (VC), the first set of candidates who completed the profile form were manually assigned unique IDs (VC_ID) and invited to participate in the pilot survey. Each volunteer was required to provide their assigned VC_ID (as a mandatory field) before proceeding to fill out the survey or submit data. The VC_ID is also vital to link each piece of data submitted to a volunteer profile while maintaining the desired anonymity of the participating citizens. In the roll-out phase, each new profile form submitted by prospective volunteers was manually checked against existing records to avoid intentional or unintentional duplication of registration ([13]).

4.2.7 G - Gamify rewards for valid submissions and maintain records

The reward system was intentionally designed to be "***non-committal***" and yet "***promising***" for the crowd (i.e. somewhat "gamified"). For this, the prospective volunteers of the crowd were informed, through initial publicity, that each valid submission stands a chance of being compensated for N2 000 (~€4), and a volunteer **may** receive up to a maximum reward of N8 000 (~€16) per month. Based on the need to implement viable controls against fraud or hacking of the system, including multiple redundant submissions by crowd members, maximum compensation thresholds were pre-set for individuals and the entire crowd per day, per week, and per month (Table 3). These thresholds, along with warnings about being blocked , were expected to be an effective strategy to deter any malicious or fraudulent manipulation of the crowdsourcing system. Subject to the pre-set thresholds, crowd members whose data submission(s) were marked as "valid"

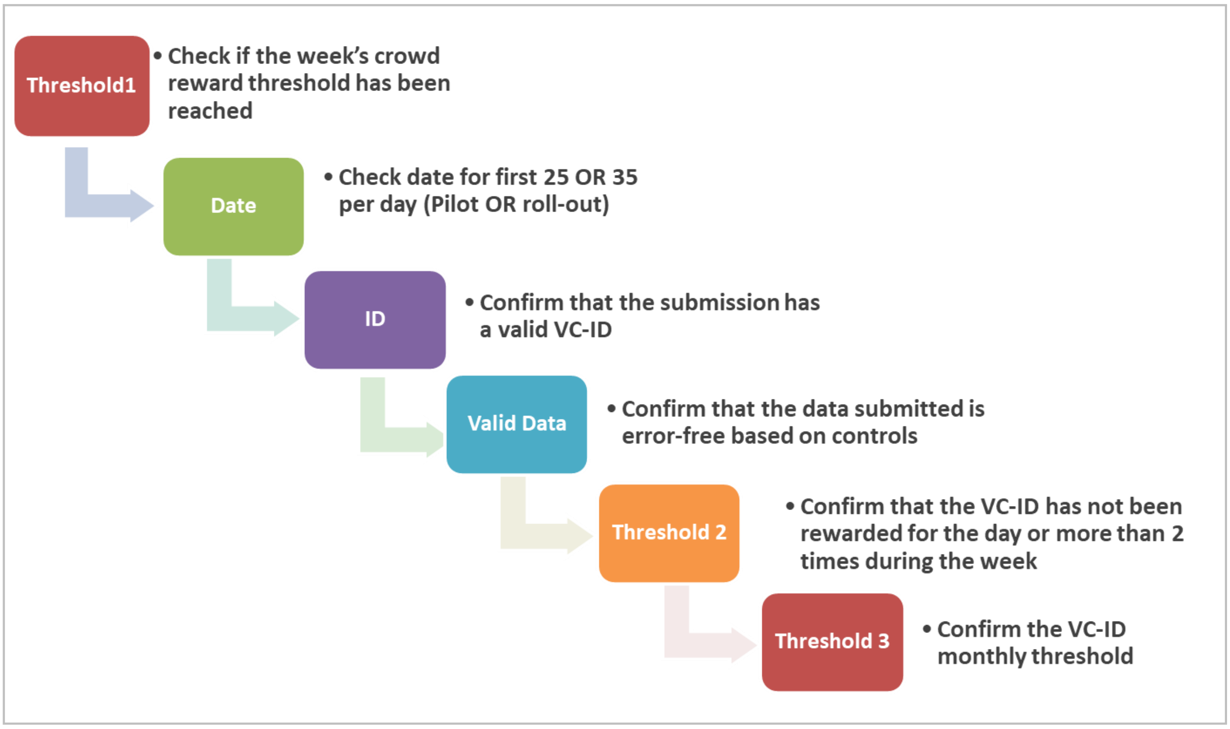
([13]) Future improvements should consider automating this step.

were rewarded on a "first-submit, first-rewarded" basis ([14]). Therefore, each data record submission was carefully checked and screened step-wise before being marked for reward, subject to the periodic thresholds (Figure 9).

Table 3. Pre-set reward ceilings for the number of submissions for each phase of the crowdsourcing.

Frequency	Pilot Phase		Roll-out Phase	
	Crowd	Individual	Crowd	Individual
Daily	25	1	35	1
Weekly	175	2	145	2
Monthly	700	4	1000	4

Figure 9. Stepwise process for checking and confirming the eligibility of each submission for rewards



4.2.8 H – Handle enquiries in real time and monitor quality controls

Dedicated phone lines and a project email account were maintained to receive and respond to enquiries from volunteers of the crowd. Responsiveness to VC's enquiries at the pilot stage through multiple communication channels helped to address/clarify various issues that can impact the quality of data submitted by the volunteers. As expected, at the onset of the project (first two weeks), the number of enquiries per day was high (~20 calls per week) compared to later stages (after ~6 weeks) where about two enquiries per week were received ([15]). Generally, detailed information was provided for the onboarding process and navigation of the ODK app (through the Google-based website), and some of the steps are intuitive, so most of the

([14]) 'Validity check' here refers to quick review of each data record submitted for any apparent and likely intentional error and "incompleteness" in the data submitted.

([15]) Actual records of enquiries were not kept, so these figures are based on memory recall.

volunteers clearly understood the requirements for successful participation. Necessary quality control measures were implemented on the ground for crowd and data management. These include the association of valid VC_ID with each data submission, pre-setting attribute domain for each data record in the app, requiring geo-location for each piece of data submitted, and tracking reward incentive thresholds.

4.2.9 I - Interaction between IITA, JRC and WUR collaborators

Cohesive interaction was maintained between the collaborating researchers from JRC, WUR and IITA (as the local implementer of on-the-ground data collection) to address needs and implement tweaks in the methodology/approach. Major periodic consultations/meetings were held to agree on the content of survey questions, the scope of data collection, standard operating procedures, and to support timely decisions regarding crowd engagement, publicity, rewards, data management. The "unconventional" approaches that were implemented emerged from the collective insights and aspirations of the collaborators to implement a credible and replicable crowdsourcing methodology that transcended the status-quo, at a minimal cost.

5 Data quality methodology for crowdsourced/citizen data

In an increasingly connected world with a growing number of citizens using smartphone and tablet technologies, crowdsourcing data, i.e. digital contributions from citizens, can be of great help due to ease of access (even from remote and difficult-to-access areas), extraordinary speed, volume and variety, and details (e.g. geolocation).

But collecting many data does not necessarily translate into having more useful information to produce better knowledge and better decisions. Instead, data must be timely, accessible, accurate, consistent and relevant for the purpose. Also, processing a large amount of citizen-generated data can be very time-consuming, which may prevent its use (OECD-OPSI, 2019). This poses challenges for decision-makers and crowdsourcers on *how to ensure trust in crowdsourced data and thus maximise its value and usefulness*. Appropriate steps and statistical methods must be found to process and ensure the quality of a large amount of raw crowdsourced data of a priori unknown quality and collected without following traditional, official data collection methodologies before including it in any analysis or decision-making process (DGINS, 2013, 2018).

Establishing a digital data collection platform is thus not enough to ensure the quality of digital contributions and related information. Crowdsourced data may be characterised by severe non-sampling and sampling errors caused by problems in the data collection process. These are usually: over- or under-coverage, measurement errors, possible fraudulent activities, non-response or participation bias. Indeed sampling bias remains a critical concern in crowdsourcing in obtaining representative samples. In our field of interest, the question remains of how to infer the 'true' prices of different food items in time and space using all digital contributions from citizens (Donmez et al., 2017).

Therefore, an important goal of this project has been to develop and implement a quality methodology for crowdsourced/citizen-generated food prices consisting of:

(i) making use of spatial statistical methods to produce a series of algorithms to automatically validate and aggregate raw crowdsourced data, to allow for dissemination in real time of accurate and reliable estimates of food prices at the sub-regional and regional level, and

(ii) an indicator framework to assess the quality of crowdsourced food prices. The latter can be very helpful, both for the crowdsourcer to monitor the quality and performance of the crowdsourcing system, and for the data users to understand the quality of the data, which is essential for data to be useful in decision making (Chengalur-Smith, Ballou, & Pazer, 1999).

5.1 WHAT IS DATA QUALITY?

Making sound decisions depends on using high-quality data, which is timely, reliable and accurate, and relevant to the task at hand. In the terminology of Juran and Godfrey (1999) and from a users' perspective, quality data is 'fit for use'.

In data quality frameworks for official statistics, data quality is referred to as a multidimensional concept usually defined using three pillars: the institutional environment, the statistical processes and the statistical output (ESS, 2019). Statistical processes are expected to be *sound and appropriate, not to imply excessive burden on respondents, and to be cost-effective*, while the final data output of a statistical process is expected to be *relevant, accurate and reliable, timely and punctual, coherent and comparable, and accessible*. Finally, the institutional environment pillar includes important quality aspects such as the use of adequate resources and ensuring statistical confidentiality of data, among others (ESS, 2019; UN, 2015).

5.2 WHY QUALITY CONTROL IN CROWDSOURCING/CITIZEN DATA?

Ensuring data quality is always a challenge, especially when working with large datasets that contain hundreds or thousands of pieces of geo-located data provided by citizens. Effective data validation and management practices and as much automation as possible are necessary for crowdsourced/citizen data to be made open for use in a timely, accurate, consistent and accessible manner while maintaining cost efficiency and data confidentiality. Indeed crowdsourcing data may be characterised by a plethora of errors and imperfections that seriously undermine the possibility of using it to describe spatial phenomena. Although they have many distinctive features, these errors can be classified into the two traditional broad statistical categories of sampling and non-sampling errors.

5.2.1 Non-sampling error

Crowdsourcing data may be affected by measurement errors due to wrong interpretations of the volunteers of the phenomenon to observe, by location errors due to mistakes in recording the coordinates, by the non-independence of collectors and by possible fraudulent activities if there is a reward associated with collection (see e. g. Arbia et al., 2018).

Although the sources of non-sampling error may be very different, their effect is very similar in that they tend to manifest themselves in the form of outliers. However, a standard analysis of outliers (e. g. by examining ranges greater than a certain proportion of the standard deviation) maybe not the best way of identifying them in the case in hand because:

- the distribution of food prices tends to be positively skewed;
- the spatial distribution of food price may better reveal geographical anomalies.

For example, crowdsourced information on food prices that is geographically located can be verified with other crowdsourced data that is available for the same location. Increasing amounts of data improve consistency and can provide an adequate data validation mechanism in crowdsourcing.

The non-sampling error relates to errors in observations at the point level. Analysing outliers to remove non-sampling errors constitutes the phase of pre-processing, which is explained later in Section 5.4.1.

5.2.2 Sampling error

Although non-sampling error may dramatically affect the quality of price surveys, even in the absence of non-sampling error, the most severe drawback related to crowdsourcing for drawing inferences on food prices is the sampling bias. This is because data is gathered *as-it-comes*, without following any formal sample design. This makes it impossible to carry out any classical probabilistic inference properly. In crowdsourcing, participation is voluntary, leading to self-selection of the data collectors. The digital divide can be a problem in crowdsourcing because the internet can provide the infrastructures and solutions for enabling the development of a crowd. However, the diversity of contributions may be limited by technological inequality (Winner, 1984). In crowdsourcing, there is no rigorous planning of selection of the individuals. This situation is not new in statistics, where it is traditionally described as "convenience sampling" (also known as availability, haphazard, accidental, grab, or opportunity sampling), in which members of the population are chosen only based on their relative ease of access. Convenience sampling belongs to the broader category of the so-called non-probabilistic sampling. When data is gathered on a convenience basis, the probabilities of inclusion cannot be accurately determined. As a consequence, all inferential statistical methods developed to produce a sound inductive inference cannot be employed because, in general, all the optimal properties of the estimators are lost (Hansen, Hurwitz, & Madow, 1953).

Sampling error relates to errors in the dataset at the aggregated level. The corrections implemented to mitigate sampling errors constitute the phase called post-sampling, which is explained later in Section 5.4.2.

Despite the sampling and non-sampling errors that may affect crowdsourced data, if they are treated with the appropriate quality control procedure, it can be a valuable source of additional information that can complement official statistics, rather than replace. Besides, adequate methods and tools for quality in crowdsourcing may encourage the emergence of innovative applications to improve a public good such as market transparency to support a better decision-making process for market actors (from farmers, through traders to consumers), and also governmental actors.

5.3 LINKS BETWEEN THE QUALITY CONTROL PROCEDURE AND DATA QUALITY

A data quality control procedure directly tackles several of the dimensions of data quality frameworks, such as those related to the structure of the data, namely: *accuracy and reliability, coherence and comparability, and accessibility* (ESS, 2018). However, it is essential to see how a quality control procedure relates to all quality dimensions/criteria directly or indirectly (ESS, 2018; Eurostat, 2014). We analyse this in the context of crowdsourcing:

> *Relevance* is a characteristic of statistics/data output measuring the degree to which statistical information meets the current and potential needs of the users. Only information that is timely, accessible and accurate is relevant. This way, relevance is indirectly related to the quality control procedure, as this must be able to produce the levels of accuracy and timeliness needed by potential data users.

Accuracy refers to the degree to which the data match the phenomena they are designed to measure. The gap is explained by potential sampling and non-sampling errors, as explained above. The quality control procedure must take both types of errors into account.

Coherence and comparability refer to the fact that data must be internally consistent, consistent over time and comparable across regions and countries, and with external data sources. The quality control procedures must ensure the consistency of the checking rules over time and the use of common standards where possible to allow for regional and external comparisons. This quality aspect is thus a direct part of the quality control procedure.

Accessibility and clarity refer to the need to ensure that the data output of the quality control procedure can be read and inputted automatically into an IT dissemination tool (such as a web dashboard), without any misunderstanding. It is, therefore, a quality aspect that is a direct part of the quality control procedure. Also, the dissemination tool must be designed in such a way that data output is presented in a clear and understandable form and an adequate dissemination format. The dissemination tool is not a direct part of the quality control procedure but indirectly is related to this quality dimension. On the one hand, it affects the data format, and on the other hand, the dissemination tool subjects the data to public scrutiny, which can help to detect quality problems that are not evident.

Timeliness refers to the length of the time gap between actual data collection and its dissemination/publication. This quality aspect may be a conditioning factor in the design of the procedure. For example, to meet the required timing of publication, quality checks must be automated and may need to be designed in a less restrictive manner. Yet software availability and the development of adequate algorithms can contribute to reducing the time gap.

Also, quality aspects of the statistical process can be related to the quality control procedure in crowdsourcing. Process quality consists of two broad aspects: effectiveness in obtaining outputs of high quality and efficiency to produce them at minimal cost to the data producer, and to the provider of the original data (Eurostat, 2014). The quality dimensions of the process and their relation to the quality control procedure applied to crowdsourced data are as follows:

Sound methodology refers to using adequate tools and procedures from data collection, through data processing to data dissemination. It is not directly addressed in the quality control procedures, but it is related to them. For example, the existence of adequate guidelines for data collectors, (e.g. including pictures), can contribute to improved data at entry point reducing so the burden of the quality control procedure. Also, the use of common information standards in the design of questionnaire embedded in the smartphone app with closed option lists can minimise the number of manual entries and possible related errors.

Appropriate statistical procedures from data collection, through data processing to disseminating quality statistics. For example, questionnaires implemented in the smartphone app (the survey tool) must be tested before data collection. The way questionnaires are developed impacts on the quality control procedure. Also, controls at the point of entry and *ex post* must be adequate to identify outliers/errors in the data and minimise the possibility of disseminating inaccurate data.

Non-excessive burden on respondents refers to the fact that the data request should be proportionate to user needs and the burden (e.g. time devoted) on respondents is monitored and kept to a minimum. This is indirectly related to quality control procedures because these need sufficient crowdsourcing contributions from citizens to be effective, and the excessive burden could discourage citizen participation.

Cost-effectiveness refers to the efficient use of resources. This quality dimension may condition the quality control procedure, as meeting budgetary restrictions may imply limitations for processes.

Finally, the quality aspects of the institutional environment can also be related to the quality control procedure.

Adequacy of resources refers to whether the resources available (human, financial, technical) are adequate to meet the quality requirements. This can have an impact on quality control procedures, since it can affect the number and quality of contributions. For example, the absence of financial resources to sustain citizen contributions through monetary rewards or publicity campaigns may limit the number of voluntary data contributions and make the quality control procedure ineffective. Similarly, the available human and technical resources may also condition the design of the quality control procedure.

Statistical confidentiality in crowdsourcing refers to securing the privacy of volunteers who submit data. This can be directly addressed in the quality control procedure by setting anonymisation rules as required, along with the different steps of the process.

5.4 A QUALITY CONTROL PROCEDURE FOR CROWDSOURCED/CITIZEN PRICE DATA (16)

A data quality control procedure consists of a system of routine technical/statistical steps applied to raw data (i.e. raw crowdsourced prices), to measure and control their quality. A quality control procedure aims to ensure a certain level of quality of the final data concerning the various quality characteristics or dimensions. It is an essential part of the production-use data cycle (see Figure 28 in Section 6.3). This cycle can be described in four main phases:

1. data collection/generation,
2. data extraction, processing (editing, automatic quality checks and management) and storage,
3. data sharing (anonymisation, publication and visualisation),
4. data usage (statistical analysis, machine learning, visualisation, decision-making).

The quality control procedure implemented in this project includes steps such as automated retrieval of the mass of citizen-generated data. Its standardisation and transformation, geo-location, accuracy checks such as detecting and flagging observations with missing mandatory information data outside the target area and outliers (i.e. out of scale data). It makes use of sound spatial statistical methodologies for calculations in order to produce reliable price estimates at the local and regional level.

All steps from the quality control procedure are integrated in algorithms programmed in R code (R Core Team, 2020) to reduce processing effort and time. The R code is available upon request to the authors. This way, citizen contributions can be more efficiently processed to quality to facilitate their use in analysis and decision making.

Based on the refinement of previous work of Arbia et al. (2018), we propose and apply a stepwise quality control procedure to FPCA data in two main phases, fully integrated into the sequence of algorithms:

Phase 1: extract raw crowdsourced data on food prices from the digital platform, edit and validate it in real time. Data validation is the decisional procedure that ends with the acceptance or refusal of data as acceptable (pre-processing phase).

Phase 2: make accurate and reliable food price estimates accessible and actionable in real time at the desired regional level (post-sampling phase).

The different steps in the quality procedure lead to different datasets with different levels of quality.

Quality level 0 – Raw Crowdsourced Data is the individual prices submitted by volunteers without any editing or another form of processing than transforming the JSON semi-structured data into structured data.

Quality level 1 – Primary Processed Crowdsourced Data is the Raw Crowdsourced Data after some preliminary processing has taken place. Mainly related to editing, automatic conversion to standard packaging units, geo-location to administrative divisions, i.e. State, Local Government Area (17) (LGA) and Ward-urban/rural.

Quality level 2 – Processed Crowdsourced Data is the Primary Processed Crowdsourced Data after it has gone through the validation process based on spatial methods and outlier detection techniques. The quality level 2 output feeds the online indicator dashboard.

Quality level 3 – Aggregated Crowdsourced Data is the Processed Crowdsourced data after aggregation at the desired temporal (e.g. daily, weekly, monthly) and geographic (e.g. State) level through the post-

(16) Arbia et al. (2018, 2020, forthcoming).

(17) The LGAs are the geographic units that correspond to the Second Administrative Level Boundaries as developed by the United Nations (UN, 2001) to support the availability of reliable geospatial information for sustainable development (i.e. policy making, programming, and operations) as well as knowledge- and information-sharing.

sampling process (averaging using post-sampling weights). The quality level 3 output feeds the online indicator dashboard.

5.4.1 The pre-processing phase

This phase consists of four steps related to data and quality management:

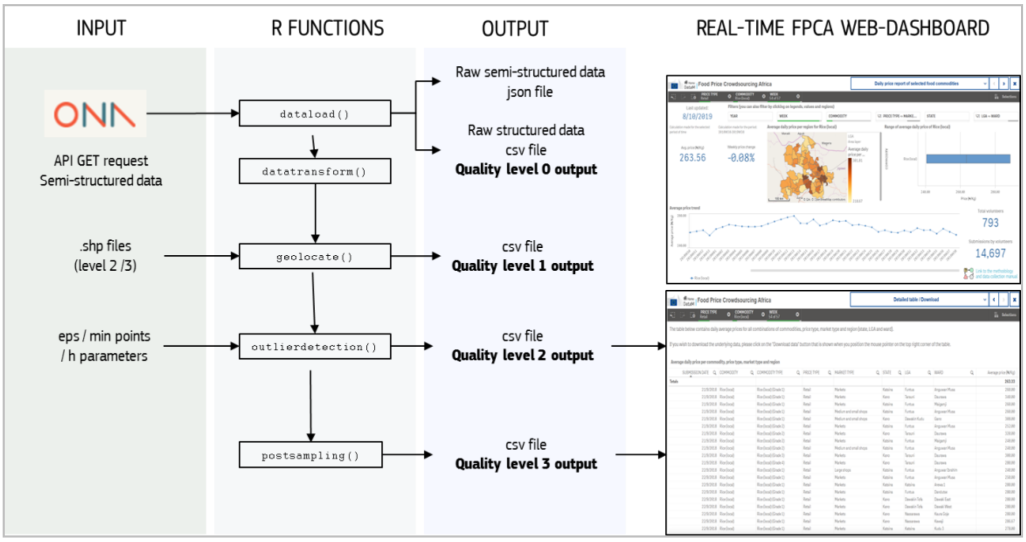
1. Automatic retrieval of data from the digital platform, transforming the JSON semi-structured data into structured data. The output of this step is a .csv and a .rds file containing the Quality level 0 data. This step is incorporated in R function `dataload.R`.
2. Data transformation mainly comprises the standardisation of measurement units to one single measure, and classification of the different market/outlet types available in the *Data Submission Form* of the smartphone app, in line with known categories of markets/outlets for African markets, as described by the International Comparison Program ([18]) (ICP) (World Bank, 2015). The output of this step is stored in a .csv file and a .rds file containing the Quality level 1 data. This step is incorporated in R function `datatransform.R`.
3. Data geo-location comprises the allocation of each observation to the different levels of administrative subdivision ([19]) (i.e. State, Local Government Area (LGA) and Ward-urban/rural) based on the coordinates at the point of data collection. The output of this step is stored in a .csv file and a .rds file containing the Quality level 1 data. This step is incorporated in R function `geolocate.R`
4. Outlier detection. Once data is retrieved, transformed and geo-located, the pre-processing phase focuses on detecting outliers and removing them from the dataset before analysis. This operation is performed separately for each commodity and each price type. The output of this step is stored in a .csv file and a .rds file containing the Quality level 2 data. This step is incorporated in R function `outlier_detection.R`. Especially this part of the quality control procedure aims to:
   - detect missing mandatory information
   - detect information outside the geographic target area
   - detect outliers (e.g. out of scale data)
   - attach a quality flag to each price observation, to avoid modifying the observed data points

Figure 10 shows a representation of the whole quality control procedure up to dissemination through the web dashboard.

([18]) The market typology is based on the classification of outlet types used in the International Comparison Programme of the World Bank (World Bank, 2015).

([19]) The administrative boundaries (Admin 0 – 2) are based on the Common Operational Data (COD) for administrative boundaries of Nigeria. For each administrative unit there is a p-code and a name. Admin COD datasets (Admin 0 – 2) for Nigeria are endorsed by the Office of the Surveyor General of the Federal Republic of Nigeria (OSGOF) and the IMWG (Feb 2017). Admin 1 (name and pcode) indicates the State name and code. The country is divided into 36 States and Abuja, which is the Federal Capital Territory (FCT); admin 2 (name and pcode) indicates the name of the Local Government Areas. The country is divided into 774 LGAs that aggregate into 36 States and the FCT. Admin 3 (name and pcode) indicates the ward (HDX, 2019).

Figure 10. Stepwise description of the quality control procedure, R functions, data output and real-time dissemination



5.4.1.1 Outlier detection

Especially three methods are suggested for detecting and removing outliers and are implemented in the R codes, leaving the user free to choose one or more of them.

— The first method consists in classical removal of values that exceed k times the standard deviation from the mean. If we define P_l as the price of a commodity in the local area/market l, $m(P)$, as the overall mean price of the commodity observed over all the data collectors at a given moment and $sd(P)$ their standard deviation, an outlier can be identified as the price P_l for which we have:

$$P_l > m(P) + h \times sd(P) \; or P_l < m(P) - h \times sd(P)$$

The parameter h can be customised by running the R codes. In practice, reasonable values are h = 2 or h = 3.

This first method may fail in identifying outliers in the left tail due to the typical right-skewed distribution of prices. To overcome this, we propose a second method.

— A second method for identifying outliers uses the median instead of the mean, and the interquartile range (IQR) instead of the standard deviation. As it is well known, median and IQR are more robust measures that are less affected by extreme values, and may mitigate the problems associated with skewed distributions. The method then consists in removing values that exceed h times the interquartile range from the median price. If we now define $M(P)$ as the median price, $Q1(P)$ and $Q3(P)$ respectively as the first and the third quartile of the distribution and $IQR(P)$ as its interquartile range with $IQR(P) = Q3(P) - Q1(P)$, an outlier can be detected as the price P_l for which we have:

$$P_l > Q3(P) + h \, IQR(P) \; or \; P_l < Q1 - h \, IQR(P)$$

Similarly to the previous case, the parameter h can be customised by running the R codes. In practice, reasonable values are h = 1.5 or h = 2.

Although more robust than the first method, this second strategy does not completely rectify the problem of missing the left tail outliers in right-skewed distributions.

The statistical literature on outlier detection for non-symmetrical, non-normal distribution is inconclusive because there is no solution which can be recommended in all cases and, conversely, solutions should be tailored to any specific case. The general suggestion in the statistical literature is to try to identify the form of the distribution (e. g. lognormal, exponential, Weibull) and then carry out a transformation on the data to reduce them to normality. For instance, if the data approximately follow a lognormal distribution, data can be transformed to normality by taking the logarithms of the data. Once the data is transformed, the outliers can be identified using one of the methods described above.

A third method relies on the idea that, in practice, local competition between selling points means that prices are distributed over space without significant discontinuities. If this happens, unusual data (possibly generated by non-sampling errors) can be detected by looking at the price values for the commodity in question in the neighbourhood. A spatial outlier is thus defined as a value that departs dramatically from the values observed in the spatial neighbourhood. We say that a point is a 'neighbour' to a data point if it is among the first, say *c*, closest neighbours to the given data point (the value of *c* is a parameter in our R codes, but a reasonable value is c=5) and if it is at a distance (e.g. Euclidean, travel or time distance) less than an arbitrary threshold (say d*) which is also parameterised in the R code. Essential to the method is the possibility of calculating inter-point distances. In the FPCA case where each observation has attached its geo-coordinates, this task can be automatically accomplished using Google Maps. Travel distances (e.g. km) or time (e.g. hours) for a matrix of origins and destinations, based on recommended routes from a start to an endpoint, can be similarly obtained through Google Maps Distance Matrix API.

According to this third definition, a spatial outlier is thus intuitively defined as the value which exceeds *r* times the standard deviation from the mean price in the neighbourhood:

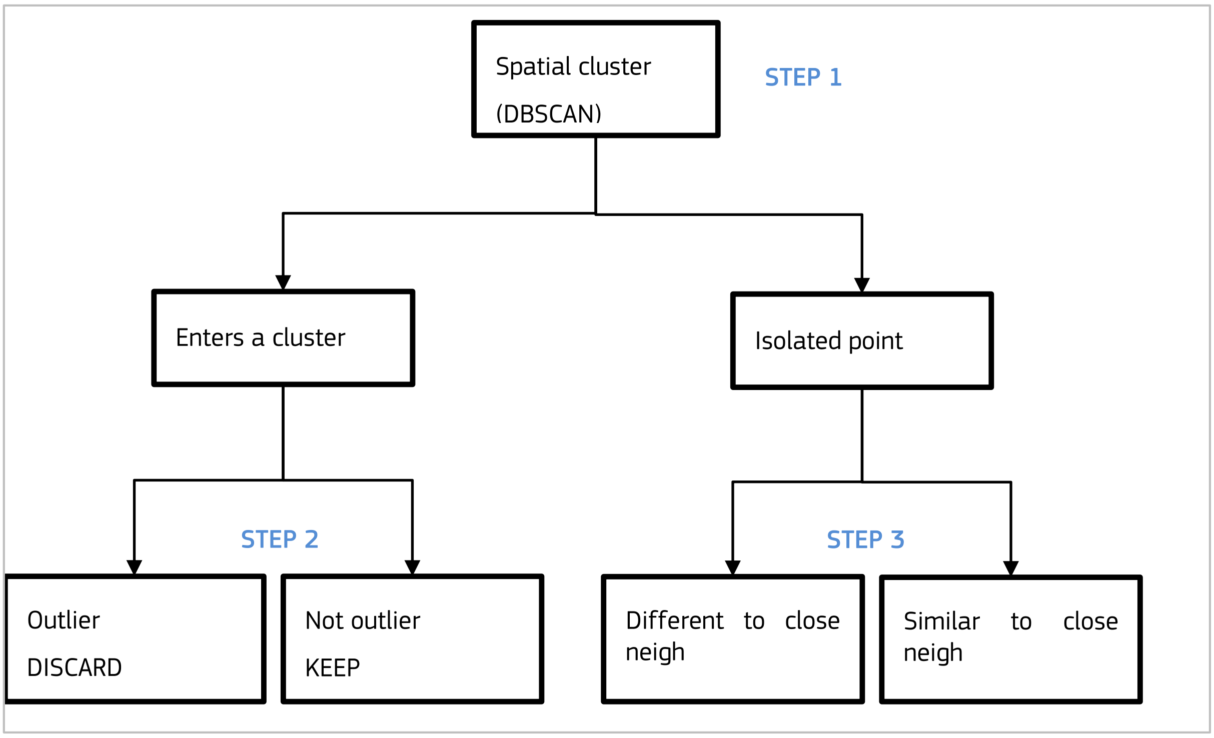
$$P_j > lag(P_j) + r\ sd(P_j)\ or\ P_j < lag(P_j) - r\ sd(P_j)$$

$$Where\ lag(P_i) = \sum_{i=1}^{n} w_{ij} P_j\ and\ w_{ij} = \begin{cases} 1 \text{ if } i \text{ and } j \text{ are neighbours} \\ 0 \text{ otherwise} \end{cases}$$

The spatial lag is thus a weighted sum of the values observed at neighbouring locations, since the non-neighbours are not included (those i for which $w_{ij} = 0$).

Now we have defined the three outliers detection strategies, the full methodology proposed in the current project (and incorporated in the R function `outlier_detection.R`) can be summarised in Figure 11.

Figure 11. Diagrammatic representation of the three steps of outlier detection



The full strategy is described in detail here below.

STEP 1 – Spatial cluster (DBSCAN)

The procedure firstly requires first of the detection of outliers based only on spatial proximity (outliers in geographic space, not value, therefore also called isolated points). The method we choose for cluster detection is the so-called Density-based spatial clustering of applications with noise (or DBSCAN).

DBSCAN is a density-based clustering non-parametric algorithm (Ester, Kriegel, Sander, & Xu, 1996) which, given a set of points distributed in space, groups together those that are close together (points with many points nearby) and identifies as outliers points that stand alone in low-density regions.

We know that due to local competition between selling points, the same commodity will be traded at similar prices at the same place and time. DBSCAN can be then used to cluster data points in geographic space and time to define *spatio-temporal markets*. Since within each spatio-temporal market, the price for the same commodity and market type is expected to be similar, an outlier detection algorithm as described above can help to identify erroneous data (i.e. out of scale). In crowdsourcing, this is making use of the 'wisdom of the crowd' by comparing multiple contributions for the same task. The DBSCAN method needs two parameters as input: a threshold distance between points (eps) and the minimum number of points (minPts) needed to form a cluster.

To briefly present the method, consider a set of points distributed in a geographical space, and define ε as a parameter specifying the radius of a circular neighbourhood.

The method identifies four typologies of points:

— A *core point* (say p) is defined as a point with at least a pre-specified number of points (say $minPts$) within distance ε.

— A *directly reachable point* (say q) is defined as a point that falls within a distance ε from a core point p.

— A point q is *reachable* from p if there is a path $p1,\ldots,pn$ with $p1 = p$ and $pn = q$, where each $p_{i+1}$is directly reachable from p_i. Note that this implies that all points on the path must be core points, with the possible exception of q.

— All points that are not reachable from any other point are classified as *isolated points.*

Notice that reachability is not a symmetrical relationship, in that, by definition, no point may be reachable from a non-core point, regardless of distance (a non-core point may be reachable, but nothing can be reached from it). Therefore, it is necessary to define a further typology of points:

— Two points p and q are *density-connected* if there is a point o such that both p and q are reachable from o.

This new concept of *density-connectedness* is now symmetrical.

A clustering system like this satisfies three properties that are desirable in general and extremely relevant in our case:

1. All points within the cluster are mutually density-connected.
2. If a point is density-reachable from any point of the cluster, it is part of the cluster as well.
3. There is a geographically based criterion for identifying outliers.

STEP 2 – Outlier detection

If after the DBSCAN classification, a point enters a cluster, then the price data in that location/cluster are used to detect price outliers (using the standard deviation or IQR criterion). The outliers thus identified are discarded. All other values are kept for the analysis contained the next steps.

STEP 3 – Relocation and outlier detections

If after the DBSCAN classification, a point does not enter any cluster, technically speaking, it is classified as an isolated point, but, before discarding it, further analysis is performed. Indeed, if the value observed in that point is similar to the mean of some cluster in the neighbourhood, the point is associated with that cluster.

If, conversely, the isolated point is very different from the mean of any other existing cluster, then the point is discarded.

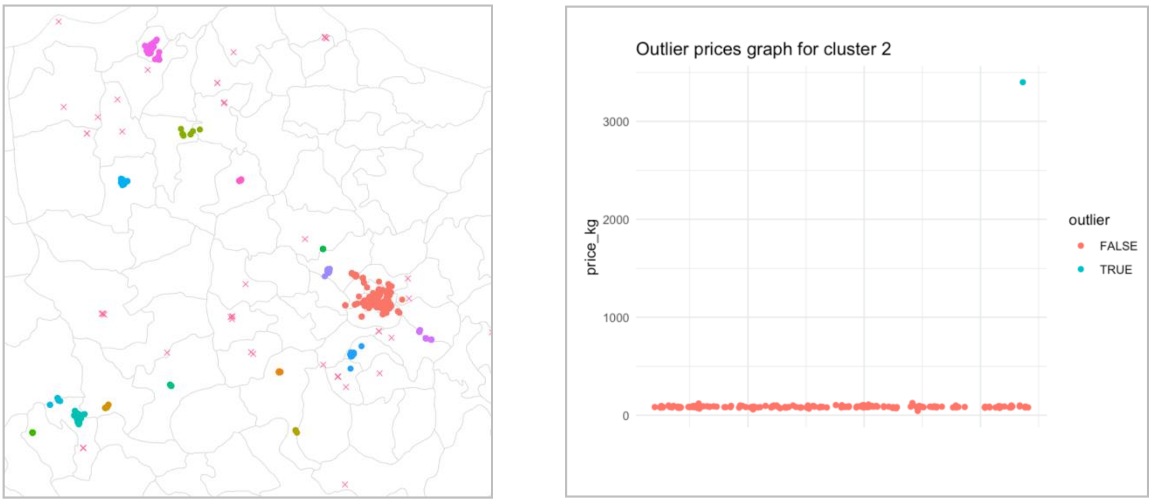
The output of this procedure reports some measures that are useful in assessing the quality of the crowdsourcing data collection approach, namely:

- The percentage of standard outliers detected
- The percentage of spatial outliers detected
- The percentage of DBSCAN isolated points detected

5.4.1.2 Example of the outlier detection procedure

For the sake of exemplifying the outlier detection procedure of the pre-processing phase, Figure 12 reports the process of detection of outliers in the retail *prices of local rice that were crowdsourced in Kano and Katsina* LGAs during the period going from 1 *April to 30 April 30 2019*. The graph reports the spatial distribution of 3282 observations. In Step 1, the DBSCAN procedure described above identifies 12 clusters and 17 (or 0.52%) isolated points, which are reported as red crosses on the graph (Figure 12, left). In Step 2, outliers inside the clusters are identified with the 2-standard deviations criterion, resulting in an outlier rate of 7% (Figure 12, right).

Figure 12. Results of DBSCAN procedure (left) and outlier detection within a DBSCAN cluster (right) for retail prices of maize, white in the period 3 March – 10 April 2019



5.4.2 The post-sampling phase

As described above, raw crowdsourced data, by definition, does not obey any formal spatial sampling plan since it is collected voluntarily. Then sample bias is a concern. As such, it cannot be directly used to draw reliable statistical inferences on food price variation and food crisis detection.

The strategy employed in this project to tackle this problem consists in subjecting the crowdsourced data to a process (which we refer to as *post-sampling*), before using them in subsequent analysis and inference. The procedure develops through the several steps which are incorporated into the R function `post-sampling.R`.

Following this basic idea, Arbia et al. (2018) suggested transforming crowdsourced datasets by discarding information in such a way that they resemble a formal sample design.

In general, we can imagine three forms of post-sampling:

- A subset of the units is drawn from the dataset according to some design (*hard-core* post-sampling). This implies a reduction in the number of observations available.
- The dataset is corrected to resemble a formal design (*flexible* post-sampling). This implies a more moderated reduction in the number of observations available.

— No observation is discarded, but data is re-weighted, taking into account the requirements of a formal sampling design (*weighted* post-sampling). See Arbia et al. (2018).

In a nutshell, the post-sampling method can be described as follows. Suppose that a set of, say, *N* food prices are collected without a proper statistical sampling design on a set of (say *L*) given geographical locations (e.g. LGAs in Nigeria). To implement the strategy, we can then compare the location of the observed data with that of a set of points selected following a reference formal sample design of equivalent sample size.

As a reference sample design, we can use, for example, any of the following:

— A simple random scheme

— A stratified random sample with geographical stratification based on an auxiliary variable such as population

— An optimal spatial sample design (Arbia, 1993; Arbia et al., 2018)

In each of the L subareas considered, the N observations can be then re-weighted to resemble the formal sampling scheme. To illustrate the method, suppose that a variable X is observed in the $L = (l = 1, \ldots, L)$ locations, and we also obtain n_l crowdsourced observations in location l, the total number of observations being equal to $N = \sum_{l=1}^{L} n_l$. A formal sampling design is then defined to select, from a list of possible locations in the study area, the same number of observations as those empirically available. We call m_l the number of observations required by the formal design in each location with $N = \sum_{l=1}^{L} m_l$. To adopt a weighted post-sampling procedure, we then calculate, in each location, a *post-sampling ratio*, defined as the ratio between the number of observations required by the reference sampling design and those available in each area, $PS_l = m_l/n_l$. In Arbia et al. (2018), an estimate of the mean of X is then obtained as a weighted average of the observations in each location using the post-sampling ratio as weights. Thus, if $PS_l = 1$, the number of observations available in location l is precisely that required by the sampling plan, and no adjustment is needed. The available observations are then over-weighted if $PS_l > 1$ and, on the contrary, down-weighted when $PS_l < 1$. In summary:

First of all, a set of, say, *N* observations on food price is obtained through crowdsourcing in a set of given collection points.

Secondly, the map of the observed data points is compared with a map of the points selected according to a formal sample design with a sample size which equals that achieved with crowdsourcing.

Before being used for drawing inference on the food price, the observations are then re-weighted in such a way that they should resemble the formal spatial sampling scheme. Especially observations are down-weighted if the actual crowdsourced observations are more than those required based on the formal sampling plan. They are conversely over-weighted if the actual crowdsourced observations are less than those required by the formal sampling plan.

Below we describe the procedural steps of the post-sampling applied to the FPCA crowdsourced food price data.

The application of the method requires two different levels of geographic units. For example, prices for the same commodity and price type can be averaged at location level (generally a city, town or village) and aggregated or post-sampled at a higher geographic level, e.g. LGA or State level. The FPCA application considers as the first geographic level the clusters of points artificially generated through the clustering procedure, referred to as *locations* and as the second geographical level the State level. We use the population numbers available for each LGA, which are proportionally distributed among the clusters, as an auxiliary variable for the formal sample design. In any case, the second geographic level requires information on the auxiliary variable at a lower geographic level. Alternatively, the user can choose to use the LGAs as the first geographic level and post-sample at State level.

The average prices are calculated at a regular time interval (e. g. week. month) whenever a reasonable number of observations are collected (e. g. more than 40) through the crowdsourcing procedure. A global price estimate for each commodity and price type of is obtained as a weighted average of the prices collected in each location.

The procedure develops through the following steps.

STEP 1 – Count of observations at the location level

We start considering the price $P_{m,l}^{t}$ as the average price observed at time t in market m which is located in location l.

The data prices $P_{m,l}^{t}$ are considered to be already pre-processed as described in Section 5.4.1; that is, we have cleaned them from the presence of outliers and so discarded potential data entry errors.

We assume that there are L locations in the study area and that we have observed a total number of n_1 prices in the markets observed in location 1, n_2 in location 2, … , and n_L in location L, with:

$$n_l = \sum_{m} n_{m,l}$$

The total number of observations in the whole crowdsourced exercise is $= \sum_{l=1}^{L} n_l$.

STEP 2 - Aggregation of observations at the location level

Prices are averaged at the location level with a simple unweighted mean through the expression:

$$P_l^t = \frac{\sum_m P_{m,l}^t}{n_l}$$

STEP 3 – Count the number of data points selected by the sampling procedure at the location level

Using a random stratified sample with geographical stratification based on population size as a reference, we select a sample of data collection points precisely equal in number to those observed (that is N). The inputs of this procedure are thus the list of all locations in the study area, their coordinates, the distance between them (route distance or time), their population and the total sample size N.

We count the number of data points selected by the procedure in each location l and we call them m_l.

STEP 4 – Post-sampling ratio

We build up the *Post-sampling ratio,* defined as the ratio between the number of observations required by the random stratified sampling plan and the number of observations available (i.e. crowdsourced) in each location.

So, in location l, we have:

$$PS_l = \frac{m_l}{n_l}$$

STEP 5 - Aggregating location prices at the targeted study area

Location prices within the targeted administrative unit (e.g. State) are then aggregated.

The average price for the study area subdivision is then obtained as a weighted average of the prices in each location using the post-sampling ratio as weights.

Formally, we have:

$$P^t = \frac{\sum_{l=1}^{L} PS_l * P_l^t}{\sum_{l=1}^{L} PS_l}$$

Thus, if in location i $PS_i = 1$ then the number of observations available in location 1 is precisely that required by the reference sampling plan, and no adjustment is needed. Conversely, if in location i $PS_i > 1$, then the number of observations available in location i is less than that required by the reference sampling plan, and the observations are over-weighted. Finally, if in location i $PS_i < 1$, then the number of observations available in location i is more than that required by the reference sampling plan, and the observations are down-weighted in the average process.

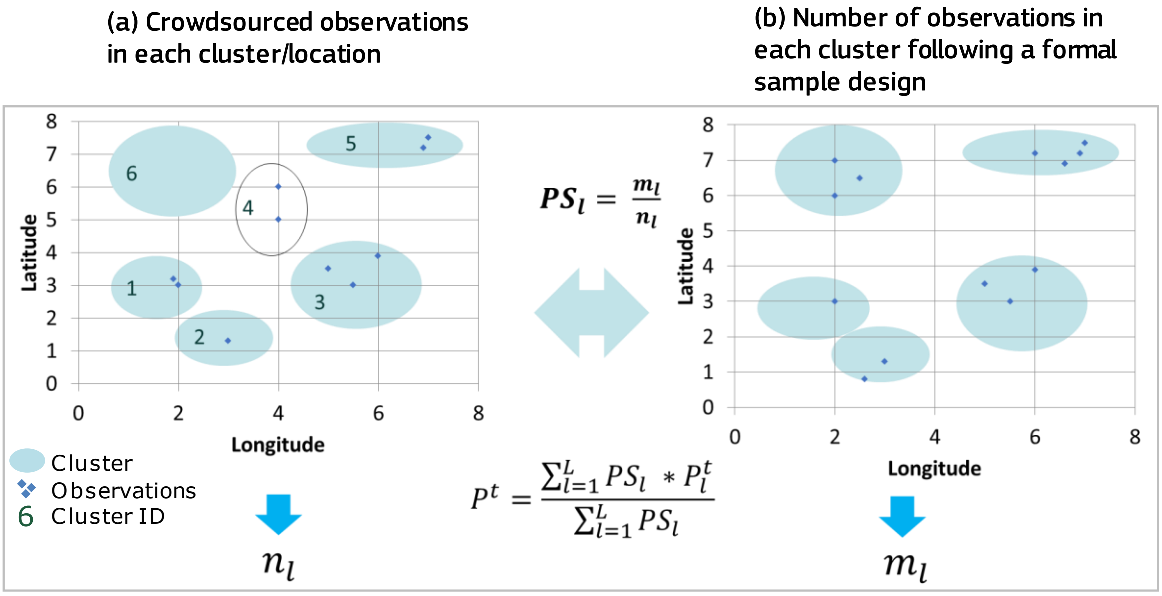
If no observations are available in location i ($n_i = 0$), then the location is not considered in the averaging process. If no observations are required in location i ($m_i = 0$), then the observations collected in location i will also not contribute to the calculation of the global price.

5.4.2.1 Example of the post-sampling procedure

Figure 13 graphically describes the post-sampling process. The figure reports the case of six hypothetical locations (e.g. LGAs) (reported in the graph as Venn diagrams). Figure 13a shows the location of crowdsourced observations in the six areas (e.g. LGAs). There are n_l crowdsourced observations in the l-th area ($l = 1, \dots, 6$).

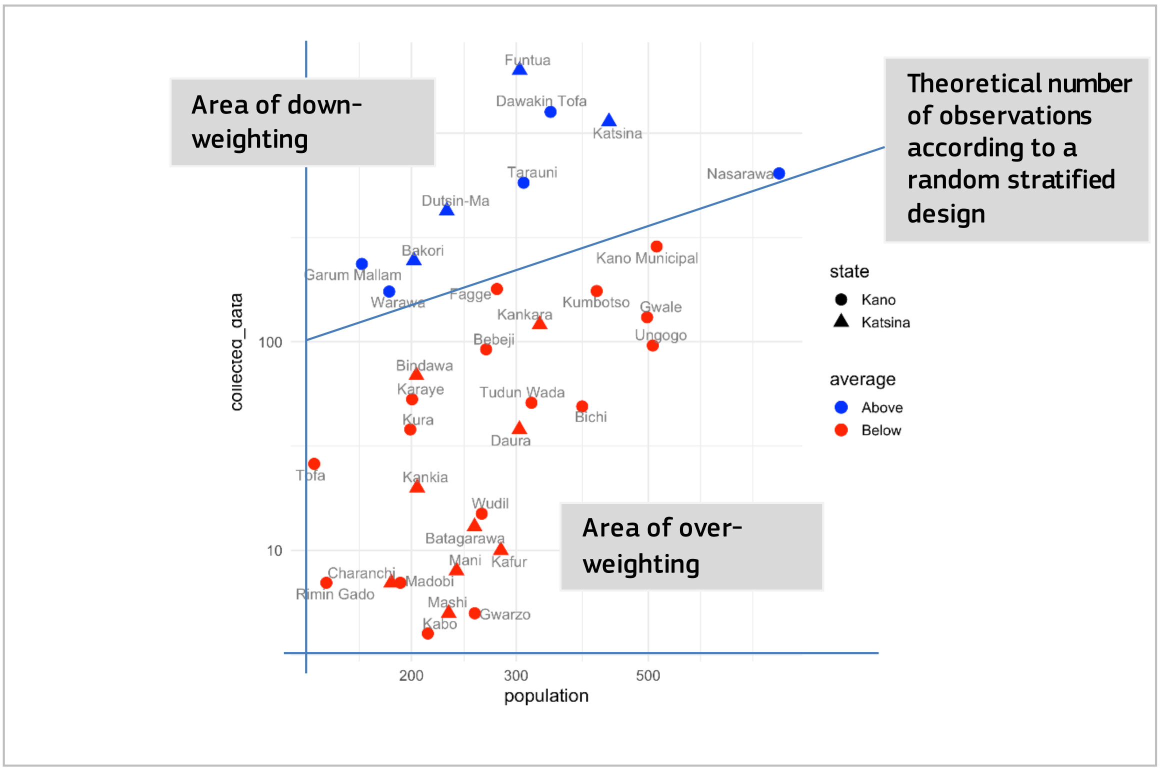
In contrast, Figure 13b displays the location of a sample with the same sample size which conforms to a formal sample design (e.g. stratified random with the area's (e.g LGA's) population as a stratification variable). There are m_l sample observations in the l-th area ($l = 1, \dots, 5$) in the formal sampling design. The ratio of these two quantities defines the post-sampling ratio (PS) in each area (e.g. LGA). The final average price P^t is obtained as a weighted average of the price in each of the 6 areas, using the post-sampling ratio as weights.

Figure 13. Location of the crowdsourced observations (a) and the theoretical observations following a formal sample design (b) in 6 hypothetical zones. The sample design requires no points in area number 4.



To visualise the effect of post-sampling re-weighting, Figure 14 reports, as an example, the number of crowdsourced observations collected in the LGAs of Kano and Katsina States (during the FPCA crowdsourcing data collection exercise) as a function of the LGAs' population size. In a geographically stratified sample design, the sample size should be proportional to the population, so in Figure 14, we should observe all points arranged on a straight line. Points that lie above the line represent LGAs where the number of crowdsourced observations is higher than that required by the sample design and which therefore require a down-weighting. Conversely, points that lie below the line represent LGAs where the number of crowdsourced observations is smaller than that requested by the sample design and which therefore require overweighting.

Figure 14. Number of crowdsourced observations collected in the LGAs of Kano and Katsina States in September 2018 as a function of the LGAs' population size. Blue points fall above the line and require down-weighting in the weighted average. Red points fall below the line and require over-weighting in the weighted average.



5.4.2.2 Difference between post-sampling and post-stratification

Although there are some similarities between the post-sampling procedure suggested and the traditional post-stratification methods, there are some differences.

Post-stratification is a well-known method introduced to reduce bias when the sample fails to include some people for technical reasons. Indeed, since some groups of individuals have different response propensities, they may be under- or over-represented. As a consequence, the process of sample selection may not represent the characteristics of the population of interest well. For instance, some groups of individuals may be intentionally oversampled to ensure that enough individuals are present with a specific characteristic and so to obtain a target value for statistical power. For one reason or another, differences between sample and population characteristics can lead to biased estimates. To mitigate this potential bias, survey researchers post-stratify the probability of selection, so that sample characteristics match population control totals (Little, 1994). Typical post-stratification variables are gender or age class.

Therefore, in post-stratification, we assume that there is one or more stratification variable.

In contrast, the post-sampling procedure that we propose only needs a stratification variable to build up the formal design to contrast crowdsourcing with and create the weighting scheme. Post-sampling is thus a way of reweighting the sample considering the difference between the actual collected data and that which ideally should have been collected, according to some specific sample design.

Therefore, post-stratification and post-sampling share the idea of re-weighting observations to correct for under- over-representation, but they differ substantially in the way the weights are derived.

5.4.2.3 Quality indictor of the reliability of crowdsourced prices

In this section, we present a measure to quantify the reliability of the crowdsourcing data. In general, if the crowdsourcing data collection resembles a formal design, it will be of high quality. Conversely, if the data collected by the voluntary collectors is very different from that that we would have gathered following a formal design, the quality will be low.

From the post-sampling procedure, n_l was defined as the size of the crowdsourced data and m_l was defined as the target sample size according to a formal reference sample design.

Moreover, $\sum_{l=1}^{L} m_l = \sum_{l=1}^{L} m_l$ was defined as the total number of crowdsourced observations.

The reliability of crowdsourced data can be measured using the formula:

$$CRI = 1 - \frac{\sum_{l=1}^{L}(m_l - n_l)^2}{\sum_{l=1}^{L} n_l^2 - 2N \min_l(n_l) + N^2}$$

This index ranges between 0 and 1:

$$0 \leq CRI \leq 1$$

Indeed, the case of CRI = 0 represents the case OF LOW RELIABILITY, where all crowdsourced data is concentrated in one single spatial unit, whereas the reference sample design requires a minimum number of points. Formally, we have:

$$n_l = 0, \forall l \neq \min_l(m_l)$$

And

$$n_l = N, if\ l = \min_l(m_l)$$

Conversely, the case of CRI = 1 represents the case OF PERFECT RELIABILITY where the crowdsourced data and the formal design perfectly coincide $(m_l = n_l), \forall l)$ so that $\sum_{l=1}^{L}(m_l - n_l)^2 = 0$.

The CRI index described above can be calculated in two different ways, firstly including only the sub-areas where data is collected or, alternatively, including all sub-areas present in the study area. We can refer to the first as a *local* and the second as a *global* index. If all observed data points are concentrated in one sub-area of the State, the local CRI indicator still delivers good results, although actual data would not cover most of the State. Conversely, if data is collected in all sub-areas, but have only few observations in each area (in extreme cases only one), then local CRI performs poorly even if data is reasonably distributed.

In the R codes, we accounted for the calculation of both the global and local index.

5.5 QUALITY CONTROL INDICATORS IN CROWDSOURCING/CITIZEN DATA

Once the set of practices and algorithms developed to ensure the quality of crowdsourced food prices are implemented (the quality control procedure) through the proposal of a set of quality control indicators, the quality control methodology offers a tool for quantifying quality aspects of the data, and thus provides valuable information on the adequacy of the data for food price monitoring/tracking and different types of decision-making (e.g. short-term commercial or policy decisions, long-term investments)

Crowdsourcing and citizen science practitioners collecting data need to consider the extent to which accurate, timely, consistent and relevant data can be collected/accessed at lower costs while ensuring confidentiality and security of data. For this purpose, once the quality control procedure has been described, the second part of the crowdsourcing quality control methodology provides a selection of quality indicators that measure different aspects of data quality.

By analysing the selected quality indicators, the launcher of a crowdsourcing/citizen data initiative can quantify the quality of the data (Vetrò et al., 2016). If quality information is attached to data (e.g. quality labels), this provides valuable information on the adequacy of the crowdsourced/citizen-generated data for use in analysis and decision-making (Chengalur-Smith et al., 1999).

The ESS handbook for quality reports (Eurostat, 2014) provides guidelines for preparing detailed data quality reports. It suggests organising them by statistical output and process quality criteria or dimensions, as presented in Box 1.

Box 1. ESS handbook - Guidelines for preparing detailed quality reports

Part II of the ESS handbook for quality reports provides guidelines for preparing detailed quality reports. It suggests organising them by statistical output and process quality criteria or components, with the primary section headings being:

1. Synthesis of the quality report, introduction to the statistical process and its outputs – an overview to provide the context of the report;

2. Relevance, assessment of user needs and perceptions – an output quality component;

3. Accuracy and reliability - an output quality component;

4. Timeliness and punctuality - output quality components;

5. Accessibility and clarity - output quality components;

6. Coherence and comparability - output quality components;

7. Cost and burden – process quality components;

8. Confidentiality – a process quality component;

9. Statistical processing

Source: Eurostat, 2014

The proposed quality indicator framework follows this guidance. Then, given the set of data quality dimensions, the two necessary elements to measure and report on data quality are (i) quantifiable measures/indicators, which require refinement of quality criteria and metrics. And (ii) measurement of indicators built up in a sound information system ([20]),to determine the degree of concordance with data quality standards (Eurostat, 2014).

5.5.1 Crowdsourcing quality indicators framework

Table 4 presents a non-exhaustive list of 38 indicators and their descriptions which have been proposed to measure the different quality dimensions. The formulas used to compute them, their measurement and interpretation are provided later in Section 7, when the performance analysis of the FPCA project is presented. Sub-section 5.5.2 presents some examples of quality indicators and empirical results. Indicators can be measured for the entire dataset or subsets of data (Vetrò et al., 2016), for example, by commodity item or geographic region. Similarly, they can be calculated for different periods (e.g. daily, weekly, monthly) depending on target frequency of quality monitoring.

Besides quantitative indicators (metrics) qualitative indicators (e.g. description in words, graphs) may also be included, as well as and simple yes/no indicators to assess whether something has happened or not (e.g. existence of a manual for data collection) (FAO, 2018).

([20]) For this project, Microsoft Excel and Stata (StataCorp., 2017) has been used to produce the quality metrics when the appropriate data is entered. Further work to incorporate the data quality metrics into the Qlik Sense dashboard tool used to disseminate data was outside the scope of this project but could help to automate the data extraction, production and dissemination of quality metrics.

Table 4. Description of indicators for assessing selected quality dimensions of crowdsourced/citizen-driven data.

Quality dimension	**#**	**Indicator**	**Level**	**Freq. (1)**	**Type (2)**	**Description**
Relevance (output)	1	Track involvement of data users and stakeholders	Project	R	D	Indicates the different types and dates of surveys conducted to involve potential data users and stakeholders.
	2	Track feedback from data users and stakeholders	Project	R	D	Indicates the different types and dates of surveys conducted to track feedback from data users and stakeholders.
Accuracy and reliability (output)	3	Valid data index	Dataset, region, commodity, price type, volunteer	D, W, M, Y	Q	Indicates the share of valid observations (data points) of the total number of observations. Valid observations are complete, (without missing information) belong to the target geographic area, and are not identified as outliers or as isolated points during the validation procedure.
	4	Outlier rate	Dataset, region, commodity, price type, volunteer	D, W, M, Y	Q	Indicates the number of outlier observations (data points) as a proportion of the total number of observations.
	5	Isolated points rate	Dataset, region, commodity, price type, volunteer	D, W, M, Y	Q	Indicates the number of isolated points as a proportion of the total number of observations.
	6	Time series completeness rate	Dataset, region, commodity, price type, volunteer	R	Q	Indicates the number of days (or weeks, or months) for which valid observations are available as a proportion of the total number of days (or weeks, or months) during a target period.
	7	Spatial completeness rate	Dataset, region, commodity, price type	W, M, Y	Q	Indicates the number of administrative areas (e.g. LGA, ward) for which valid observations are available as a proportion of the total number of administrative areas in the target region.
	8	Track number and percentage of sub-regions with completeness rate above a given threshold	Dataset, region, commodity, price type	W, M, Y	Q	Indicates the number of administrative areas (e.g. LGA, ward) with time series completeness rates above a given threshold (e.g. >75%), and the share of administrative areas with time series completeness

Quality dimension	#	Indicator	Level	Freq. (1)	Type (2)	Description
						rates above the threshold of the total number of administrative areas.
	9	Crowdsourcing reliability	Region, commodity		Q	Indicates the extent to which the map of crowdsourced data points in space reflects that of a formal sample design.
Timeliness and punctuality (output)	10	Time for publication	Dataset	D, W, M, Y	Q	Indicates the average time gap between data submission by volunteers and data publication.
	11	Up-to-dateness	Dataset, region, commodity, price type	W, M, Y	Q	Indicates the time gap between today (current week, month) and the last day (week, month) with available information.
Accessibility and clarity	12	Track machine readability of data	Dataset	R	D	Track availability of machine-readable data, that is, data that does not need any manual steps before being uploaded in the dissemination tool.
	13	Track availability of an open data dissemination tool	Dataset	R	D	Track availability of open data, updates, ease of access and user-friendly dissemination tools.
	14	Track availability of metadata	Dataset	R	D	Describes the information (metadata) accompanying the statistics (documentation, explanations, etc.).
	15	Track conditions of accessibility	Dataset	R	D	Describes the conditions for access to data: means/tools, open/private access.
	16	Track user feedback on accessibility, clarity and means of dissemination	Dataset	R	D	Describes feedback from users on accessibility, clarity and dissemination format.
	17	User visualisation rate	Dataset	D, W, M	Q	Indicates the number of visualisations of the whole dataset.
Coherence and consistency (output)	18	Share of common classifications and standards used	Dataset	R	Q	Indicates the share of fields in the *app Data Submission Form* that are associated with existing common classifications and standards (e.g. geographic information, measurement units, and commodity definitions).
	19	External consistency ratio	Commodity, region	D, W, M	Q	Indicates the relationship between the crowdsourced price and a reference price.
Cost and burden (process)	20	Time to register in the app	Dataset	R	Q	Indicates the time that volunteers require to register in the smartphone app.

Quality dimension	#	Indicator	Level	Freq. (1)	Type (2)	Description
	21	Time to send data through the app	Dataset	R	Q	Indicates the time that volunteers require to submit data through the smartphone app.
	22	Reward efficiency	Dataset	W, M	Q	Indicates the number of data submissions (data records) rewarded as a proportion of overall data submissions.
	23	Cost per data point	Dataset	W, M	Q	Indicates the share of the total monetary costs for rewarding the volunteers for each submitted data point (rewarded or not).
	24	Track time of execution of validation script	Dataset	R	Q	Indicates the time the script needs to run.
Confidentiality	25	Track data anonymisation	Dataset	R	D	Tracks the application of a process to ensure that data has been anonymised and is published with no reference to personal data
Statistical processing	26	Track availability of data collection manual	Dataset	R	D	Indicates the availability or not of a manual associated with the process of data collection.
	27	Track updates to data collection manual	Dataset	R	D	Tracks updates to a manual associated with the process of data collection.
	28	Track availability of an operating manual to guide data management.	Dataset	R	D	Tracks availability of an operating manual to guide data processing, validating, compilation and dissemination.
	29	Track updates to the operating manual to guide data management.	Dataset	R	D	Tracks updates to a manual associated with the process of processing, validating, compilation and dissemination.
	30	Track number of questions from volunteers on the data submission app	Dataset	R	Q	Indicates the number of questions from volunteers on use and participation through the data submission app.
	31	Effectiveness of publicity channels	Dataset, regions	R	Q	Indicates the share of volunteers engaged per publicity channel.
	32	Crowd size index	Dataset, regions	R	Q	Indicates the changes (increases) in the total number of registered volunteers from start point.
	33	Crowd engagement index	Dataset, regions	W, M	Q	Indicates the share of active volunteers (i.e. those submitting data) from the volunteer pool for each target

Quality dimension	#	Indicator	Level	Freq. (1)	Type (2)	Description
						period (weekly or monthly).
	34	Data submissions	Dataset, regions, commodity, price type, volunteer	W, M, Y	Q	Indicates the number of data observations submitted.
	35	Week-day Bias Index	Dataset, regions, commodity	W	Q	The ratio of the maximum and the minimum number of daily data submissions to the total number of submissions in a week.
	36	Market-type Bias Index	Dataset, regions, commodity	W, M, Y	Q	Indicates the degree of concentration of price submissions in few market/outlet types vs. various.
	37	Track feedback to volunteers	Dataset	R	Q	Number of SMS messages sent to volunteers
	38	Track availability of an automated control dashboard of quality indicators	Dataset	R	D	Indicates the availability of a dashboard that provides a regular overview of quantitative (calculated automatically based on the dataset) and qualitative quality indicators

Note: (1) In Frequency D: daily, W: weekly, M: monthly, Y: yearly and O: occasionally; (2) in Type D: Descriptive, Q: quantitative.

5.5.2 Quality assessment

In this section, some results are presented concerning the quality assessment of the FPCA data collection that took place in Kano and Katsina States in the North of Nigeria in the period from September 2018 to June 2019. During this time a financial incentive was granted to the participants according to specific rules and several motivational tools were deployed, which helped sustain the data flow.

From July 2019 to October 2019, the IITA has kept the platform and app open for data collection without any type of incentive for participants, and in November it has again established a series of economic and non-economic incentives, after which the system has continued to operate without incentives.

Further analysis of the quality performance of this crowdsourcing project is presented in Section 7.

5.5.2.1 Analysis of the number of price data submissions

The first aspect of data quality to be analysed is the number of crowdsourced observations, which reflect the soundness of the data collection method and thus the quality dimension relating to statistical processing (see Table 4). Concerning this, Figure 15 reports the weekly number of crowdsourced observations (or price data submissions) received in the FPCA project in the 41 crowdsourcing weeks from September 2018 to June 2019 and in the subsequent period.

Especially Figure 15 shows the absolute number of weekly observations, which ranges from 1 000 to 8 000 in the observational period. The figure shows that in the pilot phase (6 weeks) the amount of data collected through crowdsourcing steadily increases from 1 000 to 3 000. The roll-out produces an immediate jump to a level of about 8 000 observations a week, followed by a decrease until it reaches a stable value of around 4 000 data points collected per week. About potential regional differences, Figure 15 and Table 5 below indicate that both States - Kano and Katsina - show similar trends in the number of price observations submitted throughout the implementation period. The figure also shows that once incentives to participate are removed in week 42, the amount of data collected decreases to close to zero until a new incentive is introduced in week 59 (post-project end).

Figure 15. Amount of crowdsourced data by week from the FPCA data collection in the 41 crowdsourcing project weeks from September 2018 to June 2019 and beyond

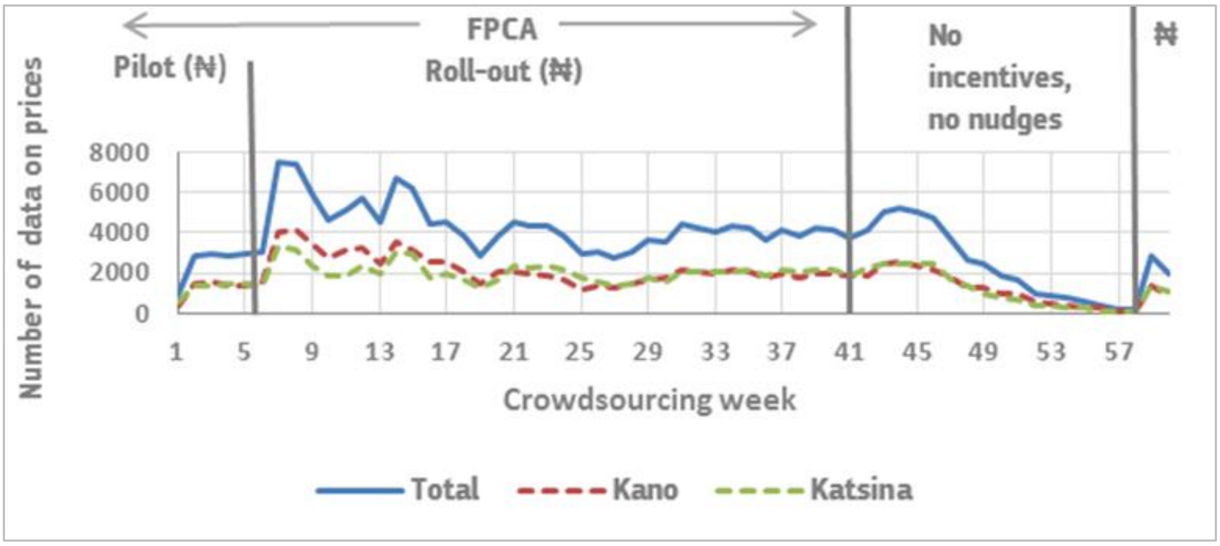


Table 5. Average weekly data submission rate in the pilot, roll-out phase and beyond, total and by region.

Region	**Pilot**	**Roll-out**	**Post**		
	week 1-6	**week 7-41**	**week 42-51**	**week 52-58**	**week 59-61**
Total	**2614**	**4412**	**3672**	**329**	**2415**
Kano	1276	2257	1771	329	1121
Katsina	1297	2087	1759	229	1229

Box 2. Interpretation of the number of data submissions

The interpretation of this indicator suggests good progress in the trend in data submissions during the project period, a period during which monetary incentives and behavioural tools were in place to promote voluntary participation by the crowd. The trend stabilizes around week 21.

The indicator also suggests that once monetary payments and other behavioural tools are removed (from week 42 on), the data submissions cannot be sustained over time. Yet the introduction and communication of a new monetary incentive from week 59 served to reactivate the number of data submissions provided by the volunteer crowd.

5.5.2.2 Analysis of reliability and accuracy

In terms of the accuracy and reliability of the crowdsourced data, the analysis is done separately for each food commodity. Table 6 and Table 7 report the monthly calculation in the States of Kano and Katsina for four quality indicators for local rice and white beans, respectively:

1. the CRI (global indicator) monthly index, as described in Section 5.4.2.
2. details on the number of observations in each State and each period
3. the number of outliers in the mean of the reported values identified with the procedure illustrated in Section 5.4.1, and
4. the percentage of outliers detected in each State and in each period.

The analysis period stretches from the start in September 2018 to the end of June 2019. This corresponds to 41 crowdsourcing weeks in the FPCA project.

In the analysis of local rice (retail), CRI values are relatively high and stable in the observation period, with systematically higher values for reliability in Kano (average 0.93) than in Katsina (average 0.80).

Table 6. Summary of accuracy indicators calculated in Kano and Katsina in the period Sep 2018 – Jun 2019, local rice, retail prices.

Month	State	CRI	# of observations	Outliers	Percentage of outliers
Sep 18	Kano	0.93	340	14	4%
Oct 18	Kano	0.93	1 504	41	3%
Nov 18	Kano	0.94	3 132	72	2%
Dec 18	Kano	0.90	2 906	205	7%
Jan 19	Kano	0.95	2 167	136	6%
Feb 19	Kano	0.95	1 805	98	5%
Mar 19	Kano	0.92	1 466	46	3%
Apr 19	Kano	0.94	2 015	81	4%
May 19	Kano	0.93	2 097	113	5%
Jun 19	Kano	0.91	1 979	76	4%
Averages		**0.93**	**1 941**	**88**	**4%**
Sep 18	Katsina	0.76	339	35	10%
Oct 18	Katsina	0.78	1 201	97	8%
Nov 18	Katsina	0.71	1 418	55	4%
Dec 18	Katsina	0.76	1 656	82	5%
Jan 19	Katsina	0.80	1 322	75	6%
Feb 19	Katsina	0.84	1 396	48	3%
Mar 19	Katsina	0.85	1 185	73	6%
Apr 19	Katsina	0.81	1 267	148	12%
May 19	Katsina	0.81	1 334	141	11%
Jun.19	Katsina	0.85	1 317	103	8%
Averages		**0.80**	**1 244**	**86**	**7%**

The monthly outlier percentage for the retail price of local rice varies from 2% to 7%, with an average of 4% in Kano and from 3% to 11%, with an average of 7% in Katsina.

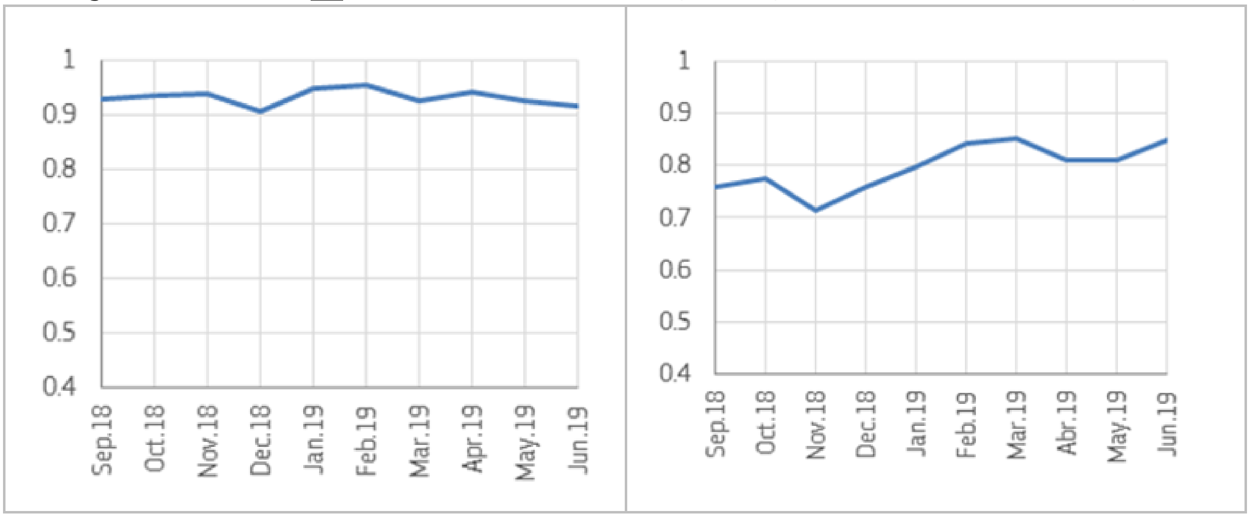
Outliers in the value of prices may indicate problems in data quality or changes in patterns of prices, or both. It is important to further analyse the case of outliers, for example, to check the impact that different commodity grades and market types may have in the outlier detection procedure. If price differences between grades are high, those prices which correspond to the highest or lowest commodity quality levels may be erroneously identified as outliers.

The same information reported in Table 6 is also displayed in graphical form in Figure 16, Figure 17 and

Some shared features that emerge from inspecting the graphs are first that quality tends to remain stable in Kano and slightly increases in Katsina in terms of representativeness as measured by the CRI. Second that the number of monthly observations has stabilised since April 2019 around 2000 in Kano and since January 2019 below 1500 in Katsina, and third, that the percentage of outliers tends to oscillate in Kano and to increase in Katsina over time.

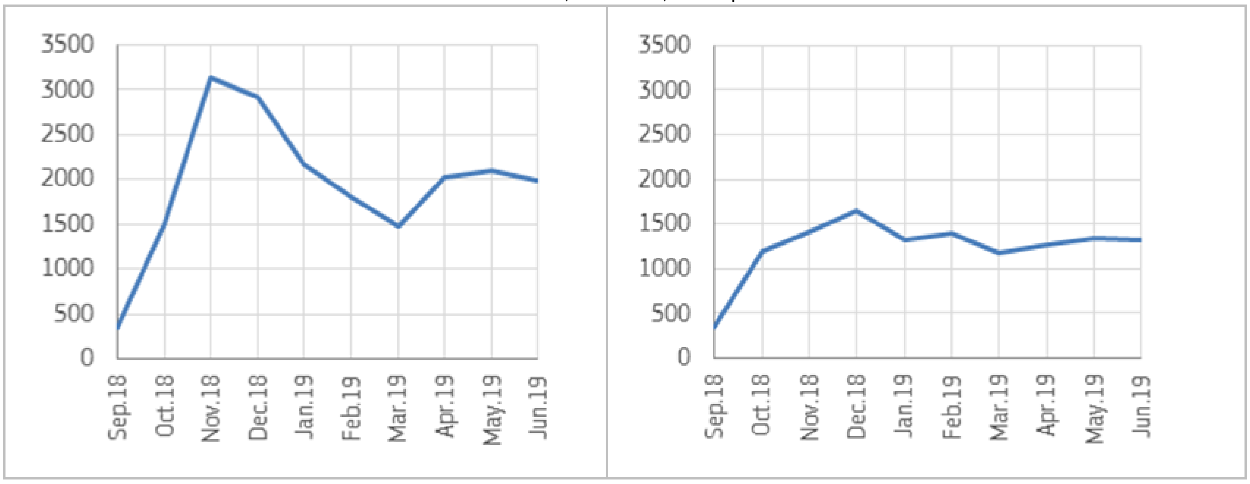
A joint examination of the various quality indicators in the two States leads to the general conclusion that the crowdsourcing price survey is comparatively more accurate in Kano State than in Katsina.

Figure 16. Evolution of CRI in (a) Kano and (b) Katsina in the period Sep 2018 – Jun 2019, local rice, retail price



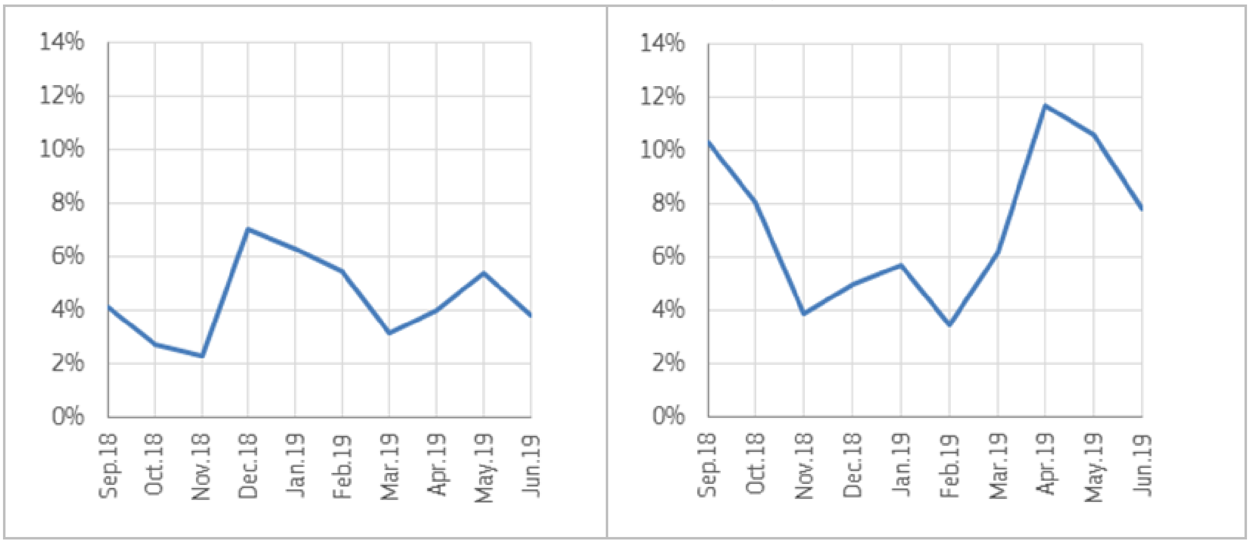
(a) (b)

Figure 17. Evolution of the number of observations crowdsourced in (a) Kano and (b) Katsina in the period Sep 2018 – Jun 2019, local rice, retail price



(a) (b)

Figure 18. Evolution of the percentage of outliers detected and eliminated in (a) Kano and (b) Katsina in the period Sep 2018 – Jun 2019, local rice, retail price



(a) (b)

Below, Table 7 shows that the CRI values also demonstrate stable slightly increasing trends in the case of white beans, with Katsina also showing systematically lower levels of CRI.

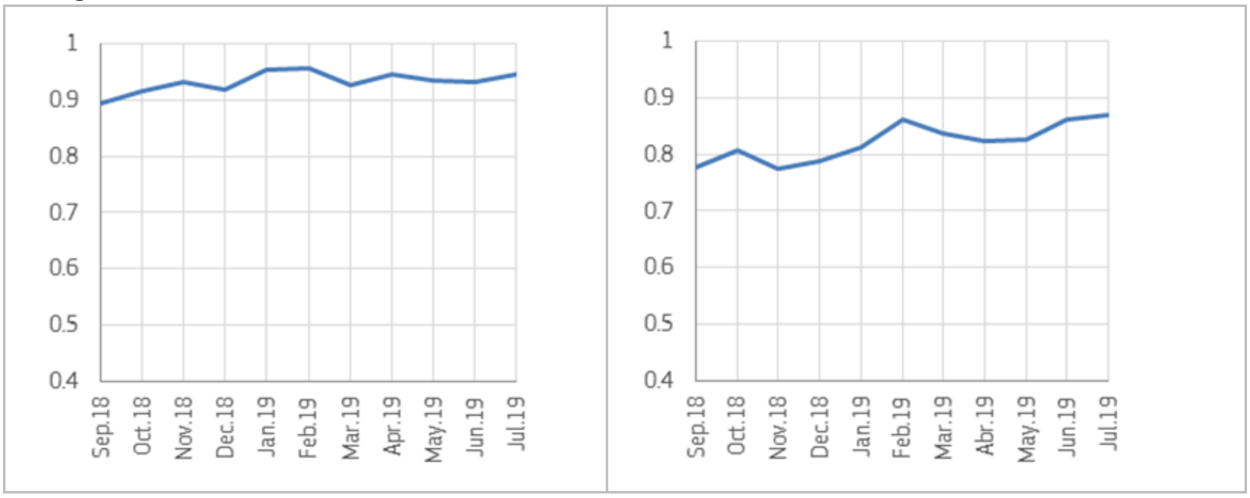
Table 7. Summary of quality indicators calculated in Kano and Katsina in the period Sep 2018 – Jun 2019, white beans, retail prices.

Month	State	CRI	N_obs	Outlier	Percentage of outliers
Sep.18	Kano	0.89	257	18	7%
Oct.18	Kano	0.91	1 219	59	5%
Nov.18	Kano	0.93	2 434	102	4%
Dec.18	Kano	0.92	2 316	82	4%
Jan.19	Kano	0.95	1 806	69	4%
Feb.19	Kano	0.96	1 475	109	7%
Mar.19	Kano	0.93	1 125	93	8%
Apr.19	Kano	0.94	1 693	72	4%
May.19	Kano	0.93	1 842	129	7%
Jun.19	Kano	0.93	1 702	126	7%
averages		**0.93**	**1 587**	**86**	**6%**
Sep.18	Katsina	0.78	261	15	6%
Oct.18	Katsina	0.81	1 040	90	9%
Nov.18	Katsina	0.77	1 277	87	7%
Dec.18	Katsina	0.79	1 407	130	9%
Jan.19	Katsina	0.81	1 048	124	12%
Feb.19	Katsina	0.86	1 173	156	13%
Mar.19	Katsina	0.84	950	78	8%
Abr.19	Katsina	0.82	969	68	7%
May.19	Katsina	0.83	943	67	7%
Jun.19	Katsina	0.86	1 078	112	10%
averages		**0.82**	**1 015**	**93**	**9%**

The outlier percentage for the retail price of white beans varies from 4% to 8%, with an average of 6% in Kano and from 6% to 13%, with an average of 9% in Katsina.

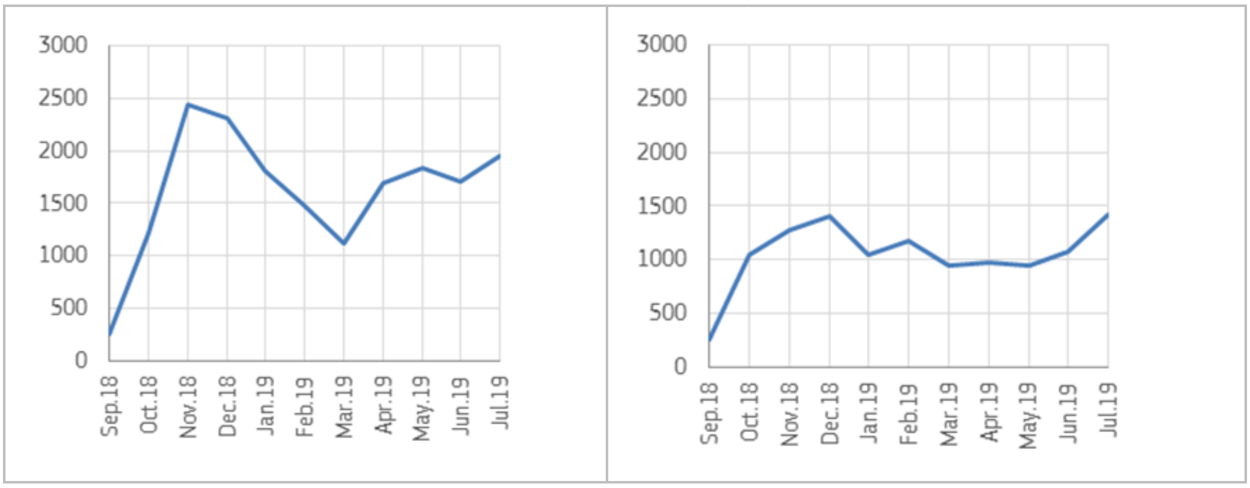
The visualisations of the quality indicators in Figure 19, Figure 20 and Figure 21 provide similar results in the CRI trends, with lower levels in Katsina. The number of price submissions increases in both Kano and Katsina, with lower levels in Katsina. Concerning the percentage of outliers, this follows a similar trend in both States, with higher percentages in Katsina.

Figure 19. Evolution of CRI in (a) Kano and (b) Katsina in the period Sep 2018 – Jun 2019, white beans, retail price



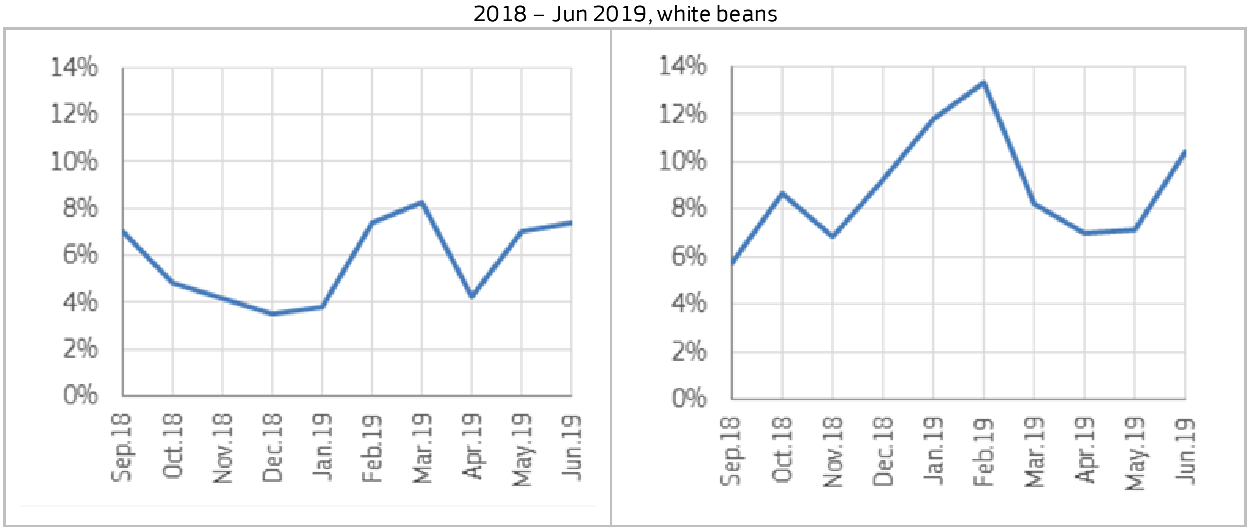
(a) (b)

Figure 20. Evolution of the <u>number of observations</u> crowdsourced in (a) Kano and (b) Katsina in the period Sep 2018 – Jun 2019, white beans, retail price.



(a) (b)

Figure 21. Evolution of the <u>percentage of outliers</u> detected and eliminated in (a) Kano and (b) Katsina in the period Sep 2018 – Jun 2019, white beans



(a) (b)

In Table 8, it can be observed that CRS values are similar across commodities within each price type, with farm gates prices scoring systematically lower in CSR than retail and wholesale prices. The number of crowdsourced observations is considerably higher for retail prices across all commodities than for other price types. Yet the percentage of outliers is not so different across price types for the local commodities as for the imported rice types.

Table 8. Summary of accuracy indicators (monthly) calculated in Kano and Katsina in the period Sep 2018 – Jun 2019 by commodity and price type.

	CRI			# of crowdsourced observations			% outliers		
	Retail	Farm gate	Wholesale	Retail	Farm gate	Wholesale	Retail	Farm gate	Wholesale
White maize	0.86	0.75	0.87	693	37	159	12%	13%	14%
Yellow maize	0.86	0.73	0.87	622	33	139	11%	17%	14%
Local rice	0.86	0.77	0.84	1 592	75	302	6%	8%	9%
Indian rice	0.85		0.86	316		88	10%		16%
Thailand rice	0.87		0.88	538		277	11%		21%
Red beans	0.84	0.69	0.82	670	35	151	12%	6%	10%
White beans	0.87	0.75	0.85	1 301	70	313	7%	9%	6%
Soybeans	0.86	0.75	0.86	710	36	145	13%	17%	12%

Box 3. Interpretation of reliability and accuracy

Good representativeness of data in terms of CRIs with slightly increasing trends over time.

Farm-gate, and to some extent wholesale, prices appear to be less represented in this crowdsourcing data collection than retail prices. Yet, these numbers must be related to the representativeness of each price type in the total population. However, new ways to involve farmers and wholesalers in submitting prices could be investigated.

There are consistently enough price contributions in both States for the commodities of local rice and white beans, with low percentages of outliers between 4 and 13% at the retail level.

However, the consistently worst values for accuracy indicators in Katsina compared to Kano could be further investigated.

5.5.2.3 Analysis of comparison with an external data source

Another important aspect of the quality of crowdsourced data is the consistency with other data sources, even if reference data is only available at lower temporal and geographic frequencies. Table 9 and Table 10 present the ratio of FPCA prices to prices published monthly by the National Bureau of Statistics (NBS) Nigeria for local rice (retail) and white beans (retail), respectively. For this purpose, we compare our crowdsourced prices aggregated at monthly level for local rice and white beans in both the States of Kano and Katsina. We compare both the post-sampled and simple averaged FPCA prices with the NBS prices.

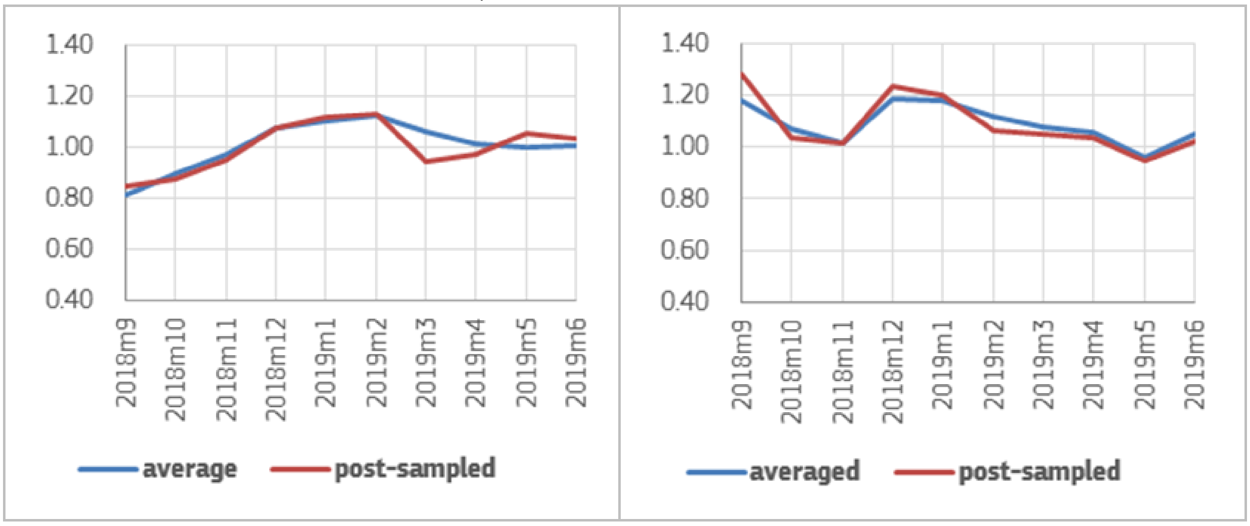
Table 9. Consistency of FPCA data with data from the National Bureau of Statistics Nigeria, rice local, retail price.

Month	State	Av price	Ps price	Ratio av (FPCA avg price/NBS price)	Ratio ps (FPCA ps price/NBS price)
Sep 18	Kano	268.68	279.28	0.82	0.85
Oct 18	Kano	264.66	260.38	0.89	0.88

Month	State	Av price	Ps price	Ratio av (FPCA avg price/NBS price)	Ratio ps (FPCA ps price/NBS price)
Nov 18	Kano	258.22	253.09	0.97	0.95
Dec 18	Kano	265.54	264.59	1.08	1.07
Jan 19	Kano	271.84	275.60	1.10	1.12
Feb 19	Kano	269.45	270.84	1.12	1.13
Mar 19	Kano	273.65	243.73	1.06	0.94
Apr 19	Kano	278.86	267.62	1.01	0.97
May 19	Kano	284.96	300.73	1.00	1.06
Jun 19	Kano	276.60	284.20	1.00	1.03
Averages		**271.25**	**270.01**	**1.01**	**1.00**
Sep 18	Katsina	269.408	291.933	1.18	1.28
Oct 18	Katsina	245.6171	238.4435	1.07	1.04
Nov 18	Katsina	253.0259	252.5965	1.02	1.01
Dec 18	Katsina	258.7735	269.7557	1.19	1.24
Jan 19	Katsina	256.9011	262.4927	1.18	1.20
Feb 19	Katsina	251.3961	239.5169	1.12	1.06
Mar 19	Katsina	255.9218	249.5775	1.07	1.05
Apr 19	Katsina	255.6256	251.3056	1.05	1.04
May 19	Katsina	258.5641	254.918	0.96	0.95
Jun 19	Katsina	254.0076	247.4056	1.05	1.02
Averages		**255.92**	**255.79**	**1.09**	**1.09**

Figure 22 shows relatively stable trends for the ratios of local rice prices in both States Kano and Katsina since the start of the roll-out phase and converging towards 1 since February 2019. An additional comparison with data available from the World Food Programme (WFP) delivers an average ratio of 1.03.

Figure 22. Evolution of the consistency ratio of FPCA data vs data from the National Bureau of Statistics Nigeria, for local rice, retail price in (a) Kano and (b) Katsina (below)



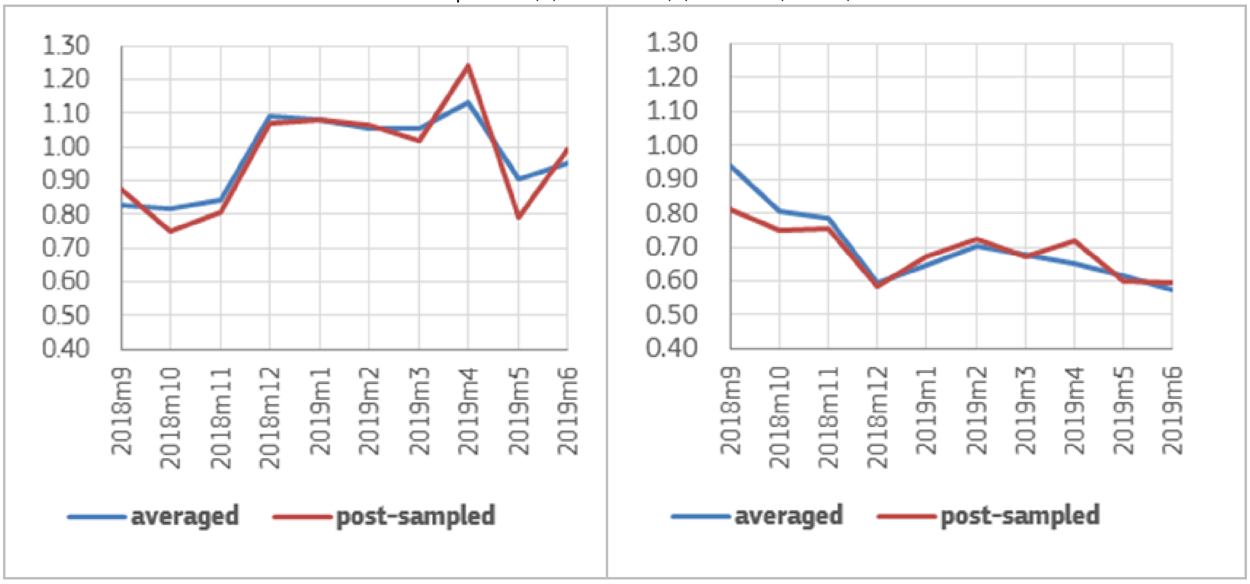
(a) (b)

Table 10. Consistency of FPCA data with data from the National Bureau of Statistics Nigeria, white beans, retail.

Month	State	Av price	Ps price	Ratio av (FPCA avg price/NBS price)	Ratio ps (FPCA avg price/NBS price)
Sep 18	Kano	270.18	284.12	0.83	0.87
Oct 18	Kano	252.34	232.37	0.82	0.75
Nov 18	Kano	248.53	237.07	0.85	0.81
Dec 18	Kano	245.88	241.45	1.09	1.07
Jan 19	Kano	243.95	243.09	1.08	1.08
Feb 19	Kano	240.22	242.52	1.05	1.06
Mar 19	Kano	234.83	226.34	1.06	1.02
Apr 19	Kano	233.39	255.10	1.13	1.24
May 19	Kano	219.12	191.54	0.90	0.79
Jun 19	Kano	196.36	203.98	0.95	0.99
Averages		**238.48**	**235.76**	**0.98**	**0.97**
Sep 18	Katsina	219.54	189.41	0.94	0.81
Oct 18	Katsina	210.06	195.49	0.81	0.75
Nov 18	Katsina	207.02	198.70	0.79	0.75
Dec 18	Katsina	192.96	188.88	0.60	0.58
Jan 19	Katsina	206.74	213.80	0.65	0.67
Feb 19	Katsina	209.57	216.43	0.70	0.72
Mar 19	Katsina	206.22	203.94	0.68	0.67
Apr 19	Katsina	196.92	217.78	0.65	0.72
May 19	Katsina	174.90	170.89	0.62	0.60
Jun 19	Katsina	162.25	168.07	0.57	0.59
Averages		**198.62**	**196.34**	**0.70**	**0.69**

Figure 23 shows a stable trend in Kano, with a ratio slightly above 1, between December 2018 and March 2019, while since then the relationship has oscillated around 1. The case of Katsina is different; FPCA prices seem to be systematically lower than those of the NBS. Nevertheless, when comparing FPCA prices with the prices published by the WFP, the relation seems to be more consistent with an average value of the ratio between October 2018 and April 2019 (WFP data on prices after April 2019 currently not available) of 0.97, with the ratio moving between 0.92 and 1.04. It is worth noting that the NBS refers to a particular variety of white beans.

Figure 23. Consistency of FPCA data with data from the National Bureau of Statistics Nigeria over time, white beans, retail price in (a) Kano and (b) Katsina (below)



(a) (b)

Box 4. Interpretation of consistency of FPCA data with NBS data

The comparison between FPCA monthly price estimates at State level with data published monthly by NBS shows proper levels of agreement relatively stable over time for local rice (retail) in both States.

However, retail prices for white beans are consistently lower over time than those reported by NBS Nigeria in Katsina, and the relationship in Kano oscillates between a long period with a ratio of above 1, followed by a period scoring below 1. Yet the comparison with data on prices published by the WFP for white beans provides a more stable relationship in both States. It is worth noting that the NBS refers to a particular variety of white beans.

Discrepancy ratios may be due to differences in the commodity variety. For example, the NBS data refer to a particular variety of white beans, the black-eyed bean, while the FPCA does not. Other reasons for discrepancies such as the market type could be further investigated.

5.5.3 Spatial analysis of the quality of price submissions

In this section, we explore the spatial aspects of the data, especially analysis of local spatial autocorrelation statistics that can help to discover hotspots (e.g. high food prices) and cold spots (e.g. low food prices) in the data, but also spatial outliers.

5.5.3.1 Location of crowdsourcing observations in space

From the spatial perspective, Figure 24 and Figure 25 report the location of crowdsourced observations in Kano and Katsina limited to the period January 2019 – June 2019, and to the commodities of local rice and white beans at retail level.

In the graphs, the population levels are also reported for comparison between actual data collection and the ideal one based on a stratified random sample. Indeed, if data were obeying a geographically stratified random sample with stratification at the LGA level, we should expect to observe a higher concentration of observed points in the most highly populated LGAs.

Box 5. Interpretation of the map of data points compared to population numbers.

Inspecting the maps shows that it is undoubtedly true that data points have been mostly collected near the capital cities of Kano and Katsina, and other cities. However, there are several LGAs that are relatively highly populated in which no data was collected in the FPCA crowdsourcing exercise, thus reducing the quality of the data collection process.

Figure 24. Crowdsourced observations for retail prices of local rice in Kano State (left) and Katsina State (right) in the period January 2019 – June 2019. Population levels are also reported in the graph for comparison between the actual data collection and ideal data collection based on a stratified random sample

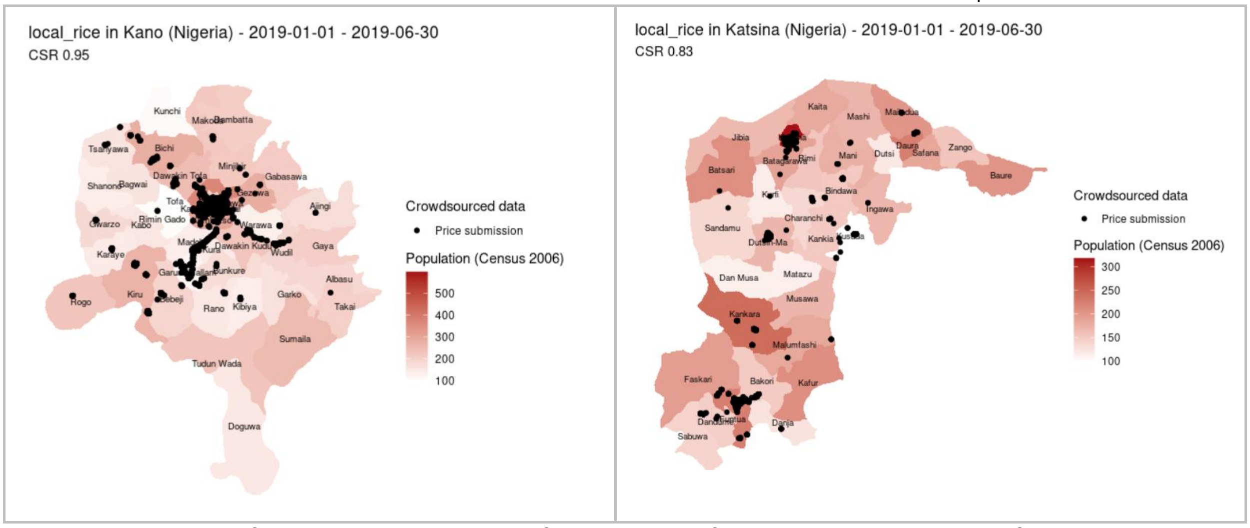
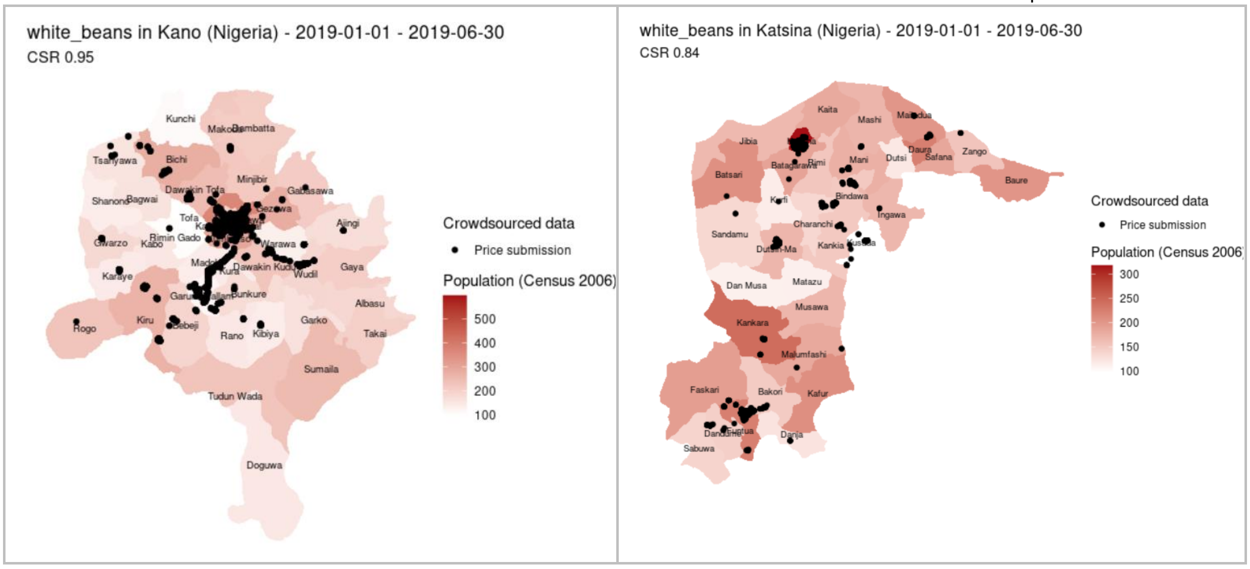


Figure 25. Location of crowdsourced observations for retail prices of white beans in Kano State (left) and Katsina State (right) in the period January 2019 – June 2019. Population levels are also reported in the graph for comparison between the actual data collection and ideal data collection based on a stratified random sample



6 Dissemination in real time: the web dashboard

High-quality data is needed to make sound decisions. Beyond accuracy, two important dimensions of data quality are '*timeliness and punctuality*' and '*accessibility*' (ESS, 2019; UN, 2015). Online dashboards can be used to disseminate crowdsourced and validated food price information to food chain market actors in a timely manner, to help them improve their production and consumption decisions, but also to other decision-makers, such as governments, donors, international organisations and to the general public. Citizen-driven dashboards may help to generate trustworthiness, advance knowledge and enhance citizens' participation in and interaction with the data generation process, and in use of data for decision-making (Matheus, Janssen, & Maheshwari, 2018).

To disseminate the data among the volunteers (and, at a later stage, the general public), an online interactive dashboard was implemented ([21]), using cutting-edge business intelligence frameworks. This tool shows daily price changes and trends, and users navigate through the data by filtering by region, product and several other fields.

6.1 The dashboard

The tool comprises three sheets:

— **Daily price report on selected food commodities:** a traditional dashboard that gives an overview of prices. It includes charts with price trends over time, average daily price per commodity, per region ([22]), market type ([23]), etc. Also, this sheet incorporates several key performance indicators (KPIs) that give an overall idea of how prices are changing, the total number of volunteers and price submissions. This sheet is addressed to consumers and users that do not need to get all the data at once, but might only care about certain regions or products.

— **Detailed table/Download:** a very simple sheet where users can download the data with a high level of granularity, i.e., average daily prices grouped by almost all available dimensions (e.g., date, commodity, price type, market type, State, LGA and ward).

— **Post-sampled table/Download:** This sheet contains a table with "post-sampled" (i.e. adequately weighted) weekly price averages per commodity, State and food chain segment (i.e. retail, wholesale and farm gate).

After accessing the dashboard, users can navigate between sheets using the sheet navigation dropdown menu in the top right (see Figure 26).

This second and third sheets are aimed at more advanced users who would like to use the data for their own calculations or projects.

6.1.1 Structure of data

The underlying data for the dashboard is the quality level 2 data produced by the R pre-processing/post-sampling script (see Section 5.4.1). Only price submissions from the States of Kano and Katsina that have not been flagged as outliers and have been assigned a cluster number (reliable price observations) are loaded into the dashboard.

In addition, quality level 3 data is data that has been post-sampled (see Section 5.4.2), and is reliable at State level, and so is loaded onto the dashboard.

([21]) The dashboard is published at https://datam.jrc.ec.europa.eu/datam/mashup/FP_NGA

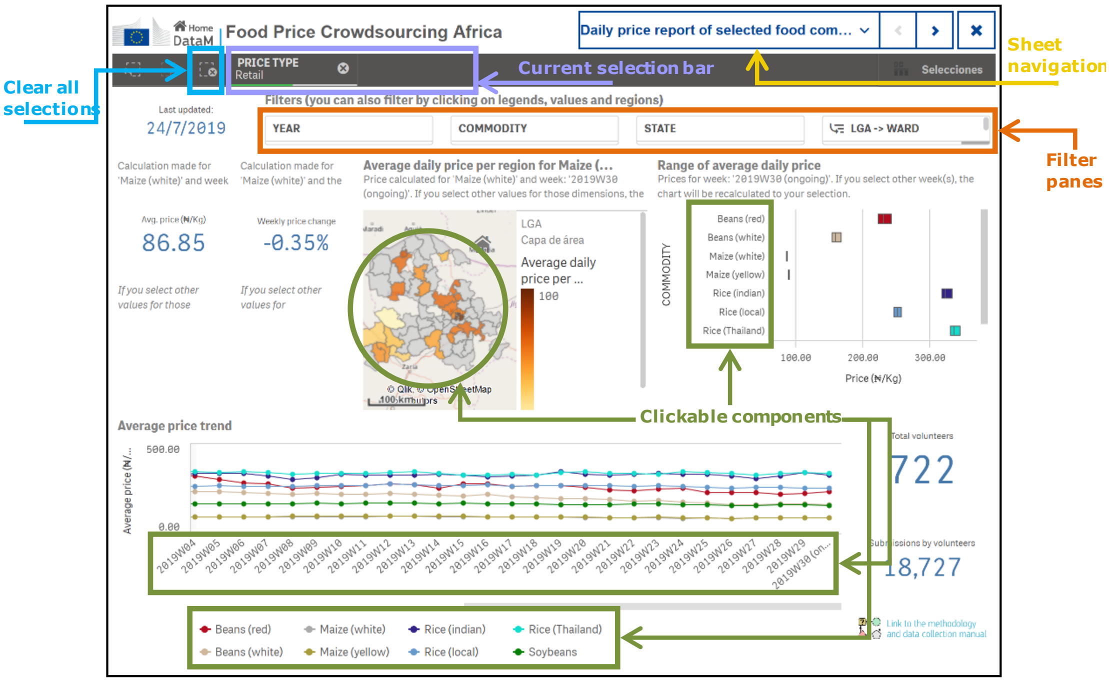
([22]) The administrative boundaries (Admin 0 – 2) are based on the Common Operational Data (COD) for administrative boundaries of Nigeria. For each administrative unit, there is a p-code and a name. Admin COD datasets (Admin 0 – 2) for Nigeria are endorsed by the Office of the Surveyor General of the Federal Republic of Nigeria (OSGOF) and the IMWG (Feb 2017). Admin 1 (name and pcode) indicates the State name and code. The country is divided into 36 States, and Abuja which is the Federal Capital Territory (FCT); admin 2 (name and pcode) indicates the name of the Local Government Areas. The country is divided into 774 LGAs, twhich aggregate into 36 States and the FCT. Admin 3 (name and pcode) indicates the ward (HDX, 2019).

([23]) The market typology is based on the classification of outlet types of the International Comparison Programme of the World Bank (World Bank, 2015).

6.1.2 Interactivity

The dashboard is fully interactive. This allows users to navigate through the data to reach the portion or "slice" that is more meaningful or important for them.

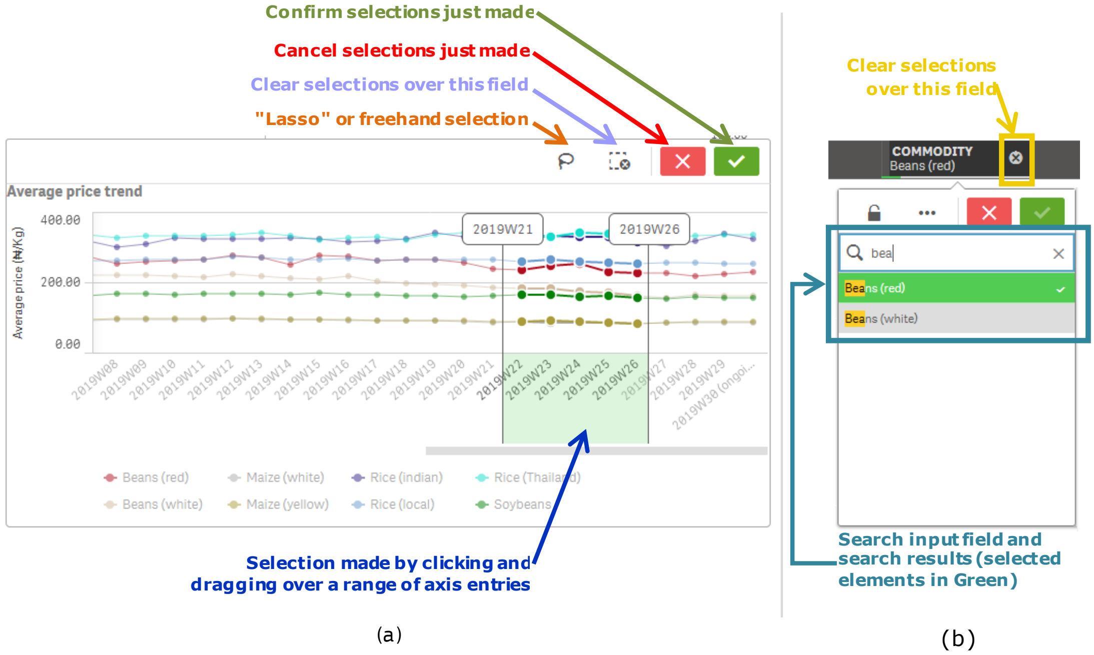
Figure 26. Initial state of dashboard



6.1.2.1 Making selections

After being presented with the dashboard in its initial state, users can choose to make additional selections in order to filter out the data that is not relevant or useful for them.

Figure 27. Making selections by interacting with the widgets



By default, a selection is made for the "price type" dimension (see "current selections" on Figure 26) because prices of different types ([24]) (farm gate/wholesale/retail) cannot be aggregated together. By selecting "retail" by default, all visualisations can be properly calculated and rendered. This selection, however, can be changed afterwards.

Selections can be made or changed by clicking on filter panes, legend entries, chart bars, chart lines, or axis entries; the data is filtered to use only the selected values. Then, the whole dashboard is recalculated over the selection, allowing the user to perform a more detailed analysis of a specific portion of the data.

Once the interaction with the component/visualisation starts, it enters "selection mode" (see Figure 27.a), i.e. all selections made are applied to the rest of the dashboard on a provisional basis. After all selections are made, they can either be confirmed (using the green button or by clicking outside of the component) or cancelled (by clicking on the red button).

Some selection components, like filter panes and fields in the "current selection" bar, allow users to run searches over the set of selectable values (see Figure 27.b). Whenever a list of elements is shown, if there are selections over that field, selected elements will be presented with green background, while selectable elements will be displayed in grey. If an element is shown in dark grey, it means selecting this element is not compatible with the current selection, and selecting it will clear all other selections first.

6.2 Methodology

6.2.1 Source

All data shown in the dashboard comes from original volunteer-submitted in-the-field price observations, which are stored in the ONA platform (see 4.2). Afterwards, the R pre-processing/post-sampling script (see Section 5.4.1) automatically scans the data, searching for outliers and invalid observations (which is known as quality level 2 of the script). As a last step, the script also calculates post-sampled weekly price averages (known as quality level 3 data).

The only input data used by the dashboard is quality level 2 and quality level 3 data coming from the aforementioned script.

([24]) See Section 4.1.3 for definitions.

6.2.2 Indicators

Prices of different types or different commodities cannot be aggregated. This is why none of the price-related visualisations are shown if more than one price type is selected, and most visualisations either group prices by commodity or choose "Maize (white)" as the default value when more than one commodity is selected.

The average price for a given day is calculated as the mean of all the submitted prices for the selected observations ([25]), using the following formula:

$$avgPrice_{day} = \frac{\sum_{observation \in selectedObservations_{day}} submittedPrice_{observation}}{N_{observation \in selectedObservations_{day}}}$$

When prices for more than one day are to be aggregated (e.g., a week), the average is calculated as the mean value of the daily average prices, as expressed in the following formula:

$$avgDailyPrice_{daySelection} = \frac{\sum_{day \in daySelection} avgPrice_{day}}{N_{day \in daySelection}}$$

6.2.2.1 Weekly prices

There are two types of weekly prices in the dashboard: simple averages and post-sampled or weighted averages.

As a general rule, unless explicitly stated, all visualisations use the simple average, which is calculated by using the formula defined above.

The post-sampled weekly price comes from quality level 3 of the R pre-processing/post-sampling script, where observations are weighted according to their representativeness. This indicator, due to data restrictions, can only be calculated at State level and on a weekly basis.

6.2.2.2 Weekly price changes

The weekly price change, computed as a compound weekly growth rate, indicates the percentage change in prices over a certain period of time, expressed as a weekly percentage. Prices usually change over a given period, but often at an uneven rate. The compound weekly growth rate provides one rate for the period in weekly terms. In other words, this indicates the percentage by which prices would have had to increase every week in order to go from the initial value to the final one (assuming constant weekly change). This indicator is calculated using the following formula:

$$weeklyPriceChange_{minWeek,maxWeek} = \left(\frac{avgDailyPrice_{maxWeek}}{avgDailyPrice_{minWeek}}\right)^{\left(\frac{1}{\#\ of\ weeks_{minWeek,maxWeek}}\right)} - 1$$

6.3 The process

For near-real-time dissemination to work, it requires a data stream which is fed by volunteers and produces valuable information after a short time. The dashboard is updated twice a day: very early in the morning, loading data up to the previous day, and in the afternoon, loading data that was submitted in the morning.

As depicted in Figure 28, the workflow is a cycle where volunteers submit their price observations, which are subsequently quality-checked to produce a reliable set of data that is then served to the public through an open dashboard. Volunteers can come back to the system to obtain larger-scale data based also on

([25]) In the interactive dashboard, observations can be filtered out in many ways (by selecting a region, a type of price, etc.). The formula for the aggregation is always calculated over all selected observations and, depending on the visualisation, it is grouped by commodity, time, region, etc.

submissions made by their peers. However, the dashboard may be of interest not only to volunteers, but also to decision makers, people in the agri-food business, the general public or scientists interested in food prices in Nigeria.

In order to protect volunteers' identities, all price observations undergo an anonymisation process, where all data that can be linked to a particular person is either removed or securely hashed so there is no way to trace an observation back to the person who submitted it.

Figure 28. Data collection and dissemination workflow

Source: JRC elaboration, loosely based on the workflow proposed by (Matheus et al., 2018) for Big Data-related projects.

Following the recommendation made by Matheus, Janssen and Maheshwari (2018), the dashboard was not only used as a tool for disseminating data, but also for fostering interaction with volunteers, such as motivating volunteers to continue to participate in the project by submitting price observations every day.

The FPCA collective knowledge shared in an open online dashboard is expected to represent a public good that uses data to provide feedback, which in turn may trigger further data collection. A more detailed analysis of this aspect is provided in Section 8 Sustaining the System.

6.3.1 Technologies

As explained above, this dashboard uses data produced by the R pre-processing/post-sampling script (specifically from quality level 2 and 3), which is programmed using the R language (R Core Team, 2020). The output of this script is written as a CSV file, which then is read by Qlik Sense (the self-analysis and business intelligence framework behind the dashboard).

Then, the Qlik Sense dashboard is embedded in the JRC's Data portal of agro-economics Modelling (DataM)([26]), implemented using HTML, Javascript and Java Platform Enterprise Edition (Java EE) technologies. Using DataM allowed us to save time since it was an existing solution, which was already online and offered the possibility to embed Qlik Sense dashboards with minimum effort. The JRC's Data portal of agro-economic Modelling (DataM) is integrated with the JRC's Data Catalogue ([27]), which is also integrated with the EU Open Data portal ([28]).

([26]) Data portal of agro-economics Modelling: https://datam.jrc.ec.europa.eu/
([27]) https://data.jrc.ec.europa.eu/dataset/36a2ac99-87db-4069-9426-995482100a6b
([28]) http://data.europa.eu/89h/36a2ac99-87db-4069-9426-995482100a6b

7 System performance

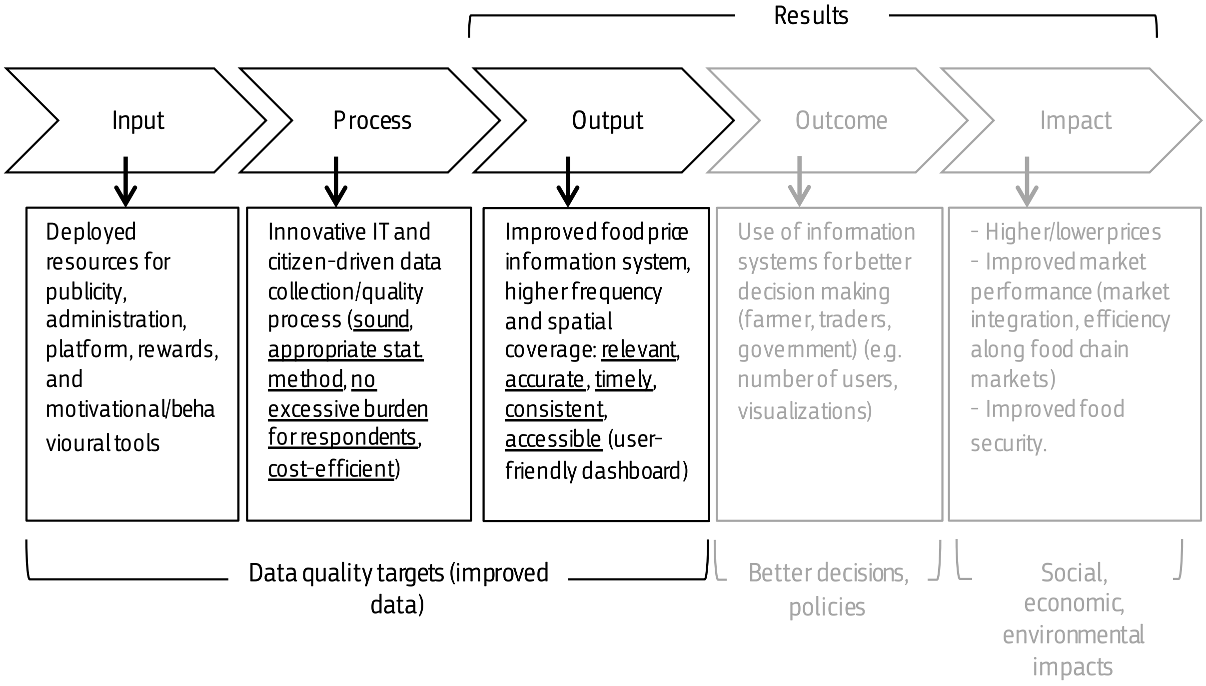
A unique element of this initiative is the aspiration to implement a mobile-app-based crowdsourcing approach to get on-the-ground insights with minimal bias (i.e. “spontaneous crowdsourcing”), at minimal cost, with continuous data flow, and with a minimal error rate associated with the data submitted.

For a crowdsourcing initiative to be effective and sustainable, we need to understand how the crowd is generated and how it contributes to creating a quality output. In this section, we try to provide quantitative and qualitative measures of how these two aspects performed for the FPCA initiative. Note that some measurements of the accuracy and reliability of the crowdsourced data have already been presented in Section 5.5.2.

Performance measurement is a key tool to assess how well an approach achieves its objectives. In this sense, to monitor and improve the performance of a crowdsourcing system in collecting and disseminating data, there is a need to have a set of measurable and reliable indicators built on a sound information system, and clear objectives and priorities.

For measuring the different levels of results (i.e. output, outcome, impact) against strategic objectives, one can use a results-oriented indicator framework (Figure 29) (EC, 2015, 2017). In this framework our analysis refers to results at output level (reliable/quality data) and the process for reaching them. Besides, through participation in citizen science projects, participants may gain better awareness and knowledge of the subject, which can be integrated into their decision making process (Brossard, Lewenstein, & Bonney, 2005). Yet, the evaluation of outcomes and impacts is beyond the scope of this report.

Figure 29. Results-oriented indicator framework applied to the quality targets of a Quality Assurance Framework for a citizen-driven data collection and dissemination system



Source: Authors' elaboration based on EC (2015, 2017) and ESS (2019)
Note: The shaded area is beyond the scope of this report and should be addressed in a future study.

To be useful, outputs from crowdsourced volunteering activities have to meet quality criteria related to relevance, accuracy and reliability, timeliness, coherence and consistency, and accessibility, both for the statistical process and its output. Internationally recognised Quality Assurance Frameworks (QAF) for data statistics (e.g. UN, 2015; ESS, 2019) rely on this multi-dimensional concept of quality. This builds around the three pillars of the institutional environment, statistical process and output (ESS, 2019) and provides a perfect set of goals and targets around which to build the performance indicators for the final product of our citizen-driven information system, as described in Section 5.5.

We propose a set of 38 indicators (see Section 5.5.1) to evaluate the output performance of the crowdsourcing system, relative to a benchmark for desirable performance, along the quality dimensions and targets of a QAF, such as the European Statistical System ([29]) (ESS). As indicated in Section 5.5, these indicators can be applied at different levels of aggregation (e.g. dataset, commodity, region, participant), and different time scales. When focusing on the full period of the FPCA activity, they serve as a *general performance index* (**GP**) to evaluate the performance of the crowdsourcing system. Also, they can be calculated for weekly or monthly periods to serve as a *weekly performance index* (**WP**) or as a *monthly performance index* (**MP**) which allows monitoring the progression and stability of the crowdsourcing system. The choice to calculate the indices as GP, WP or MP indicators depends on the specific objectives of the system administrator. While GP can be used to assess the performance of the system as a whole, WP and MP can also evaluate temporal improvement or decline in the performance of the crowdsourcing system, and allow system administrators to initiate corrective actions if required.

The following sections describe the calculation and interpretation of the non-exhaustive list of quality indicators developed in Section 5.5 along the different quality criteria for data output, and the process and environment required to reach that output.

7.1 RELEVANCE

This sub-section provides an assessment of the consideration of user needs and perceptions. Two indicators have been proposed for this.

7.1.1 Track involvement of data users and stakeholders

Two different types of surveys were conducted:

A questionnaire addressed to potential data users and stakeholders, conducted through the EU online survey tool and addressed to 46 individuals drawn from government, research, development cooperation and farmers' organisations. The results of the survey are thoroughly described in Section 3.

A questionnaire to participant citizens/volunteers at the moment of registration in the mobile app. The results are thoroughly described in Section 9.

Box 6. Interpretation of involvement of potential data users and stakeholders

The survey addressed to potential data users and stakeholders revealed that commodity list, geographic coverage, timeliness and dissemination method were important aspects of a price information system. Also, collecting data at different stages of the food chain (input suppliers; farm gate; processing stage; wholesale trade; retail) and data quality were considered important. The proposed methodology addresses these aspects. In particular, timeliness and accessibility are ensured through the automated quality process that allows for daily dissemination of data through an open-source web dashboard. The data collection in this project was limited to a reduced number of commodities and states, but is easily expandable.

The app questionnaire revealed that 8 out of 10 volunteers were interested in receiving food price data and that they would be willing to participate in the initiative without monetary rewards in exchange for data. SMS-based messaging was the preferred method for receiving information or updates, and this is likely due to the non-invasiveness of text messages and the fact that SMS is a well-established and well-known technology. At least 6 out of 10 volunteers indicated that they would prefer to receive food price data by SMS.

7.1.2 Track feedback from data users and stakeholders

Feedback from data users and stakeholders has not been gathered to date.

7.2 ACCURACY AND RELIABILITY

The assessment of some indicators of accuracy and reliability for the crowdsourced data has already been presented in Section 5.5.2, especially the percentage of outliers as identified in the pre-processing phase

([29]) The ESS follows and aligns with the European Statistics Code of Practice (ES CoP), which builds on the UN's Fundamental principles for Statistics (UN, 2015).

(Section 5.4.1) and the reliability of the crowdsourced data as indicated by the CRI value (Section 5.4.2). Here, three further indicators are presented.

7.2.1 Valid Data Index

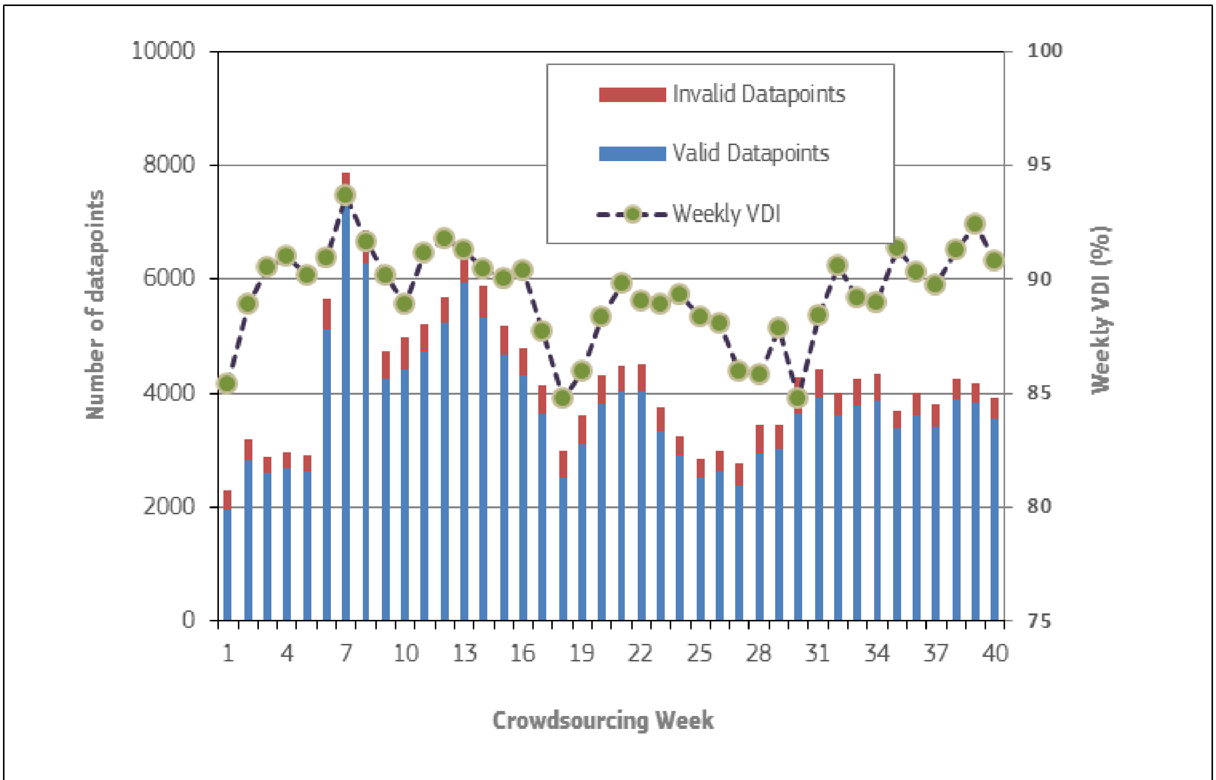
This is an index that shows the quality of data submissions by volunteers of the crowd within the period of interest. A reliable crowdsourcing system should provide useful datasets that are free from fake or fraudulent data points which can undermine the quality and usability of the data for real-life applications. The VDI is calculated as the percentage of total data point submissions at the desired time step or timeframe that are considered valid (as defined in the pre-processing phase in Section 5.4.1).

$$VDI_t = \frac{number\ of\ price\ submissions_{valid,t}}{numer\ of\ price\ submissions_{total,t}}$$

Where $t(= 1, \dots, T)$ are the target sub-periods (e.g. days, weeks, months).

Figure 30 presents the weekly calculation of this indicator for the whole set of crowdsourced prices during the FPCA implementation period.

Figure 30. Weekly frequency of valid and invalid crowdsourced data point submissions and percentage of valid data point submissions during the FPCA project implementation period



Box 7. Interpretation of the Valid Data Index

We observe a good level of data validity throughout the project, with an overall VDI of 90% and weekly variations between 85% and 94% valid data.

As shown by the trend in the VDI during the later stages of data collection (i.e. week 30-40), a higher number of valid submissions, together with a slightly declining number of invalid observations, resulted in lower percentages of invalid data.

Further analysis could track if invalid data is consistently attributable to the same volunteers, or test the validity of data provided by registered and non-registered volunteers.

7.2.2 Time series completeness

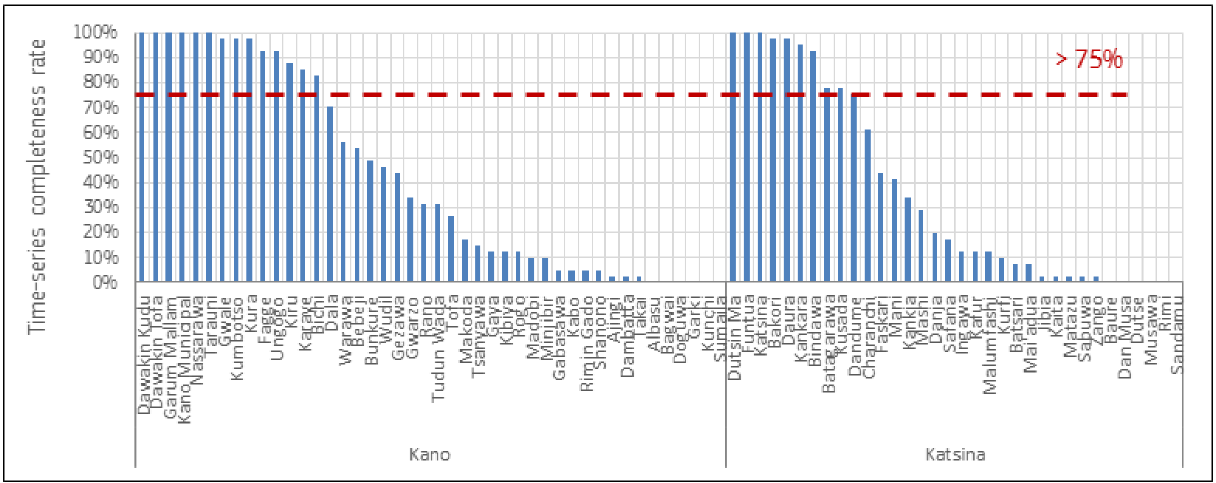
An essential aspect of quality is the completeness of reported data. Figure 31 shows the time series completeness rate in each LGA, measured as the share of weeks for which valid price information is submitted, out of the total 41 weeks of the FPCA project. The data completeness rate is calculated as follows:

$$Time\text{-}series\ completeness = \frac{number\ of\ subperiods\ with\ data\ available}{number\ of\ total\ subperiods} \times 100$$

Sub-periods can be, for example, days, weeks, months.

Figure 31 shows the calculation of the completeness of data for each LGA at weekly level. Consdering 75% as a minimum completeness to allow for informed decision making we can see that in 14 LGAs in Kano and 10 LGs in Katsina the time-series completeness rate is above 75%.

Figure 31. Time series completeness rate by LGA in the period Sep 2018 – Jun 2019.



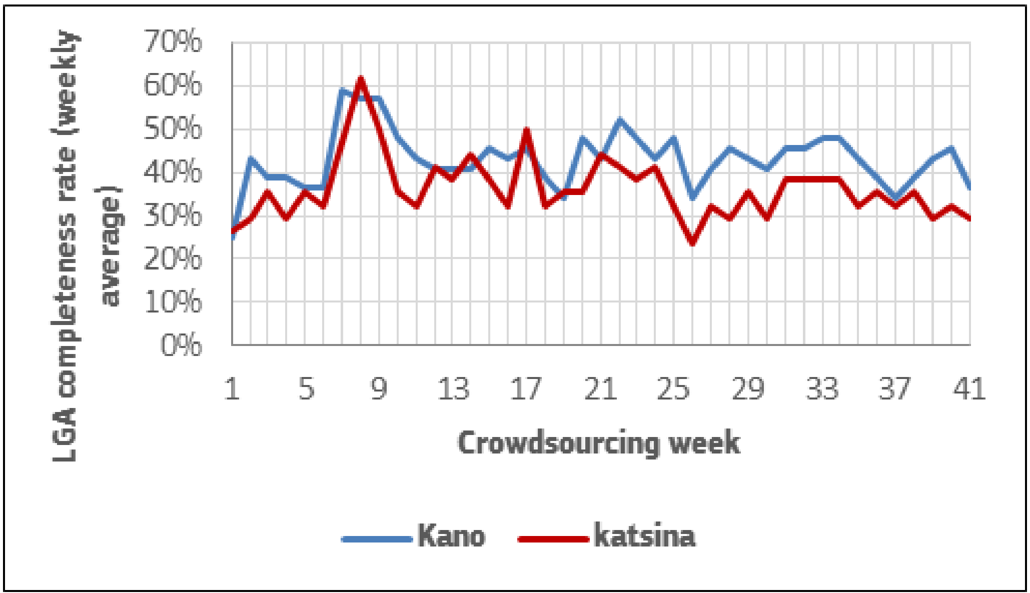
7.2.3 Spatial completeness

From a geographical perspective, it is interesting to understand the regional and sub-regional coverage.

$$Spatial\ completeness = \frac{number\ of\ administrative\ units\ with\ data\ available}{number\ of\ total\ administrative\ units}$$

Figure 32 shows by State the share of LGAs for which valid data is submitted in each week of the total number of LGAs.

Figure 32. Weekly evolution of sub-regional (LGA) completeness rate over the period Sep 2018 – Jun 2019



Both Figure 31 and Figure 32 suggest that Kano shows better data completeness rates, with a larger proportion of LGAs with a rate of completeness above 75% and better sub-regional coverage sustained over time. This is further confirmed by the summary of completeness indicators in Table 11 and was already pointed to by the CRI (Crowdsourcing reliability) indicator calculations presented in Section 5.5.2.

Table 11. Summary of completeness indicators.

	Kano	Katsina	Total
LGA completeness rate (weekly average)	43%	36%	40%
LGA completeness rate (total period)	86%	82%	85%
Number of LGAs with weekly completeness rate >75%	14 (of 44)	10 (of 34)	24 (of 78)
In % of total number of LGAs	32%	29%	31%
of which	64% urban/36% rural areas	41% urban/59% rural areas	53% urban/41%rural areas
LGA with no data submitted at all	Albasu, Bagwai, Doguwa, Garko, Kunchi, Nasarawa, Sumaila	Baure, Dan Musa, Dutsi, Dutsin-Ma, Musawa, Rimi, Sandamu	

Box 8. Interpretation of completeness indicators

About half of LGAs in Kano and about one third in Katsina show a time series completeness above 75% over the project period.

There are high levels of submission completeness for LGAs over the whole crowdsourcing period, with a ratio of 86% in Kano and 82% in Katsina.

However, the weekly submission completeness ratio for LGAs drops to an average of 43% in Kano and 36% in Katsina with slightly decreasing trends towards the end of the period.

As for the comparison between urban and rural areas, urban areas are better covered in Kano while in Katsina it is rural areas that are best covered.

LGAs without submitted data and those with a completeness rate of less than 75% could be further investigated.

7.3 TIMELINESS AND PUNCTUALITY

7.3.1 Time for publication

Data collected in this crowdsourcing project is uploaded to the web dashboard twice a day, after going through the algorithms of the quality process. When launched daily, the algorithms need 3 to 5 minutes to run.

7.3.2 Up-to-dateness

This indicates the time gap between today (current week, current month) and the last day (week, month) with available information.

Table 12 presents the calculation of the average number of days between two releases for each commodity and price type over the period of the project.

Table 12. The average gap of days between two releases for each commodity and food chain state over the FPCA project from Sep 2018 to Jun 2019.

	Kano State - days (lag)			Katsina State - days (lag)		
commodity	**retail**	**wholesale**	**farm gate**	**retail**	**wholesale**	**farm gate**
Indian rice	1.004	1.286		1.203	1.000	1.000
Local rice	1.004	1.259		1.004	1.033	1.083
Maize white	1.004	1.000		1.004	1.116	2.895
Maize yellow	1.004			1.004	1.233	5.400
Red beans	1.004	1.000		1.004	1.271	
soybean	1.004	1.000		1.007	1.089	4.231
Thailand rice	1.004	3.211		1.004	1.026	4.231
White beans	1.004	1.000		1.004	1.007	

Box 9. Interpretation of timeliness and up-to-dateness indicators

Timeliness, measured by the time gap between data collection and publication, is ensured by the dissemination of the validated data through the web dashboard, which is updated twice a day.

The up-to-dateness indicators show that retail prices are daily updated for every commodity in both States. Wholesale prices are also daily updated except in Kano State where Thailand rice presents an average lag of 3 days and yellow maize with no data available. Farm gate prices are not available in Kano State and show in Katsina lags in availability of between 1 and 5 days, except for beans (white and red) for which farm gate prices are not available.

7.4 COHERENCE AND CONSISTENCY

While triangulation with external data sources such as the NBS and the WFP is conducted in Section 5.5.2, here we present an indicator that measures to what extent data is collected using common classifications and standards that allow for comparability within the dataset and with external sources.

7.4.1 Share of common standards used

This indicates the share of fields in the app Data Submission Form that are associated with existing common classifications and standards (e.g. geographic information, measurement units, commodity definitions, etc.).

$$Number\ of\ standardised\ data = \frac{number\ of\ standard\ fields}{number\ of\ total\ fields}$$

This is a GP index with a score of 0.8.

Box 10. Interpretation of the share of standardised data

A score in this indicator of 0.8 suggests that the developed information system applies known classifications and standards, which enhances comparability and consistency of output, and especially at the point of data entry may prevent interpretation errors. The rest of the fields contain for example information on the volunteer number, whether they are making a purchase or sale transaction or are market observers.

7.5 COST AND BURDEN

Cost efficiency and limited burden on volunteers and data administrators are essential quality aspects of a crowdsourcing initiative.

Since the volunteer participants are not requested to go to the market to collect data but to provide the data during routine visits to the market, the time required to register in the app (once) and the time devoted to submit data must be kept to a minimum to prevent volunteers from giving up. These times are a function of the number of fields requested, and for that reason, only essential information should be required. It also makes a difference if the fields are open or the respondent can select from fixed options, which speeds up the response time.

The next two indicators track the time required to register and the time required to submit data through the app.

7.5.1 Time required to register in the app

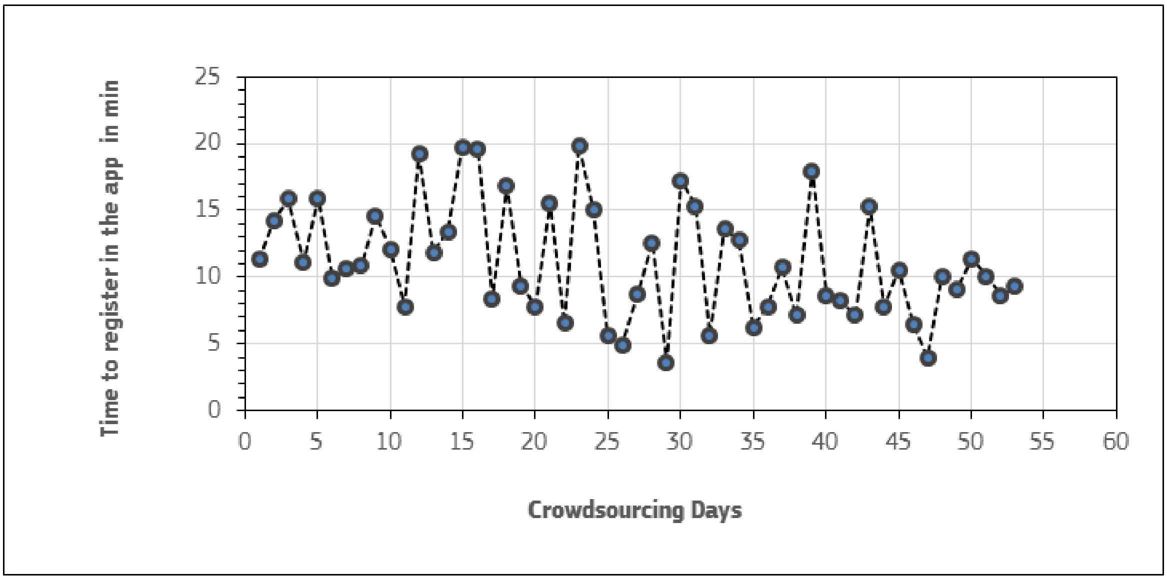
This indicates the time that volunteers require to register in the app. The indicator is measured as the median time across registered volunteers from starting to fill in the app profile form to submit. We use the median time to avoid including the effect of those rare cases where volunteers start filling the form and finish hours or even days later.

$$Time\ to\ register\ in\ the\ app_t = \underset{i}{\mathrm{median}}_t(start\ time_i - \ end\ time\ _i)$$

Where $i(= 1, \dots, I)$ are the volunteer citizens and $t(= 1, \dots, T)$ is the target period.

Figure 33 shows the time trend for this indicator over the registration period. Volunteer citizens usually take between 5 and 20 minutes to register, which situates this indicator within the 20-minute benchmark for answering an online questionnaire (7.5 seconds per online survey question) (https://verstaresearch.com/newsletters/how-to-estimate-the-length-of-a-survey/).

Figure 33. Time devoted to registration (filling in the app profile form) in the app by volunteers during the FPCA project implementation period



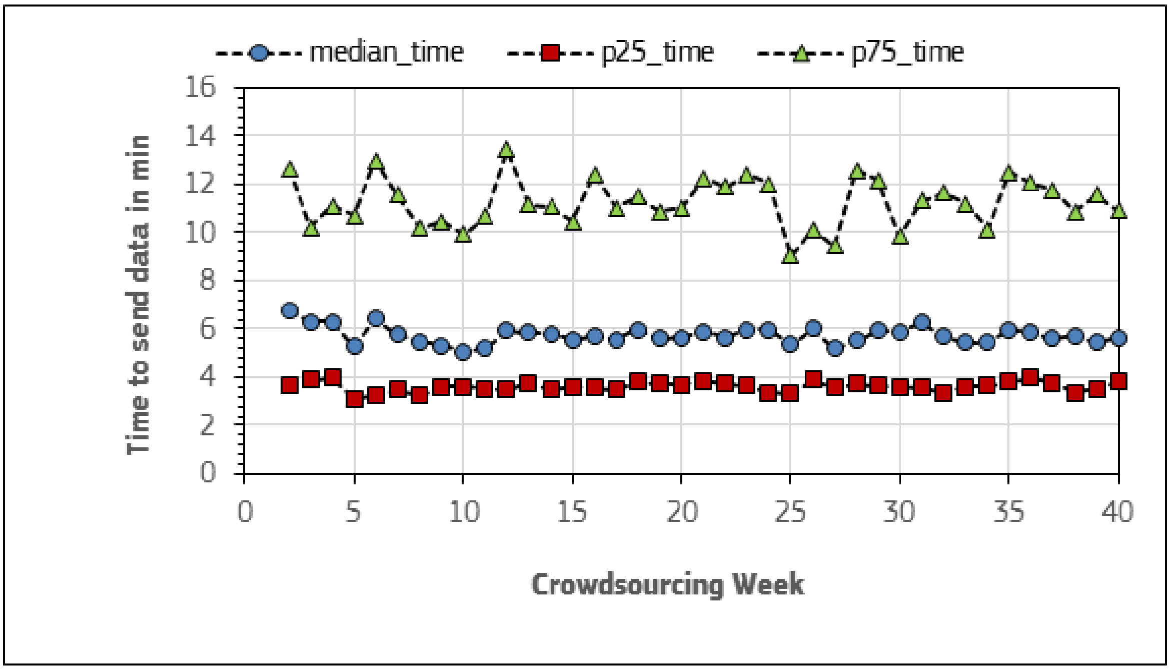
7.5.2 Time required to send data through the app

This indicates the time that volunteers require to submit data through the app. It is measured as the median time across registered volunteers from starting filling in the data form to submission. Again we use the median time to disregard the effect of participants who start data collection at the market but finish the submission at a later time (e.g. from home) since the app allows them to work off-line.

$$Time\ to\ send\ data\ through\ the\ app_t = \underset{i}{median}_t(start\ time_i - end\ time_i)$$

Figure 34 presents the calculation of this indicator at weekly level depicting the median value as well as the Q75 and Q25. While the median time is 6 minutes, we can see that at least 25% of the sample can upload data in less than 4 minutes and another 25% of the sample needs between 10 and 14 min. However, we must highlight that no individual exceeds the benchmark of 20 minutes.

Figure 34. Time devoted to data submission (filling in the app data submission form) in the app by volunteers during the FPCA project implementation period



To some extent, the differences might be due to the number of commodities reported in each data record, which can vary from 2 to 10 (including commodity varieties).

7.5.3 Reward efficiency

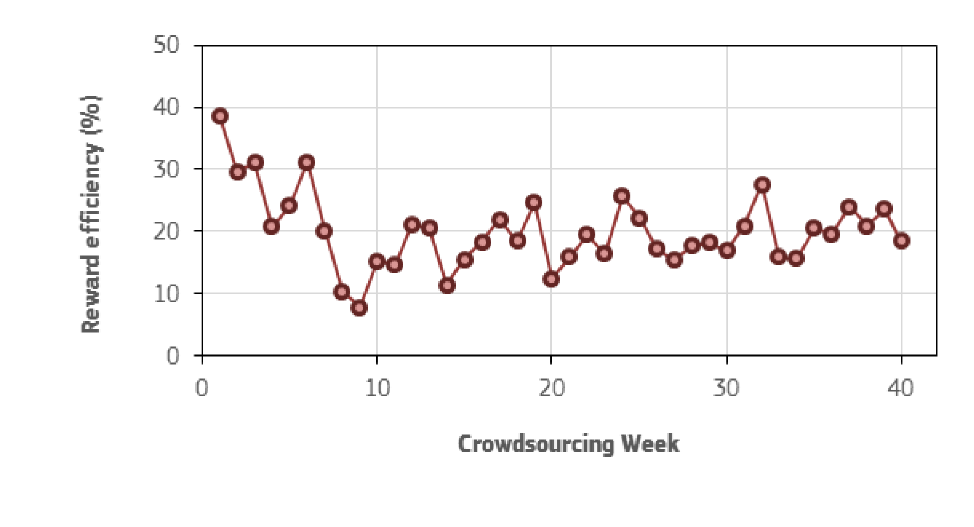
Under ideal conditions, the cost of crowdsourced data submissions should be zero, as citizens could be expected to have intrisic motivations to provide data as a contribution to a common good, mainly market transparency in our case or simply for fun, especially if it does not cost them considerable time and money. However, within the context of developing countries, motivation to participate in data submission may be lacking if there is no economic incentive (Bott & Young, 2012). Initially, we needed to activate/attract the crowd with economic rewards as part of the system set-up. However, ideally, the crowdsourcing system should function efficiently over time with or without economic rewards.

The indicator is a measure of this efficiency and it is calculated as a percentage of data submissions rewarded (rewarded data records) relative to overall valid data submitted (valid submitted data records) within each week. In an unsustainable and inefficient crowdsourcing system, this indicator is very high (>80%), while it is very low (<30%) in an efficient system. Progressively lower values indicate that the rate of data submitted by the volunteers is becoming less dependent on reward (a sort of weaning-off), vice-versa for higher values. Within the implementation of the FPCA project, the best score (~9%) was achieved after the full roll-out stage and was consistently lower than 30% till the end of the project, with weekly variations (Figure 35).

$$Reward\ efficiency_t = \frac{number\ of\ rewarded\ records_t}{total\ number\ of\ valid\ records_t} \times 100$$

Where t are the periods of data collection included in the calculation (e.g. weeks, months).

Figure 35. Weekly evolution of the reward efficiency indicator



7.5.4 Cost per data point

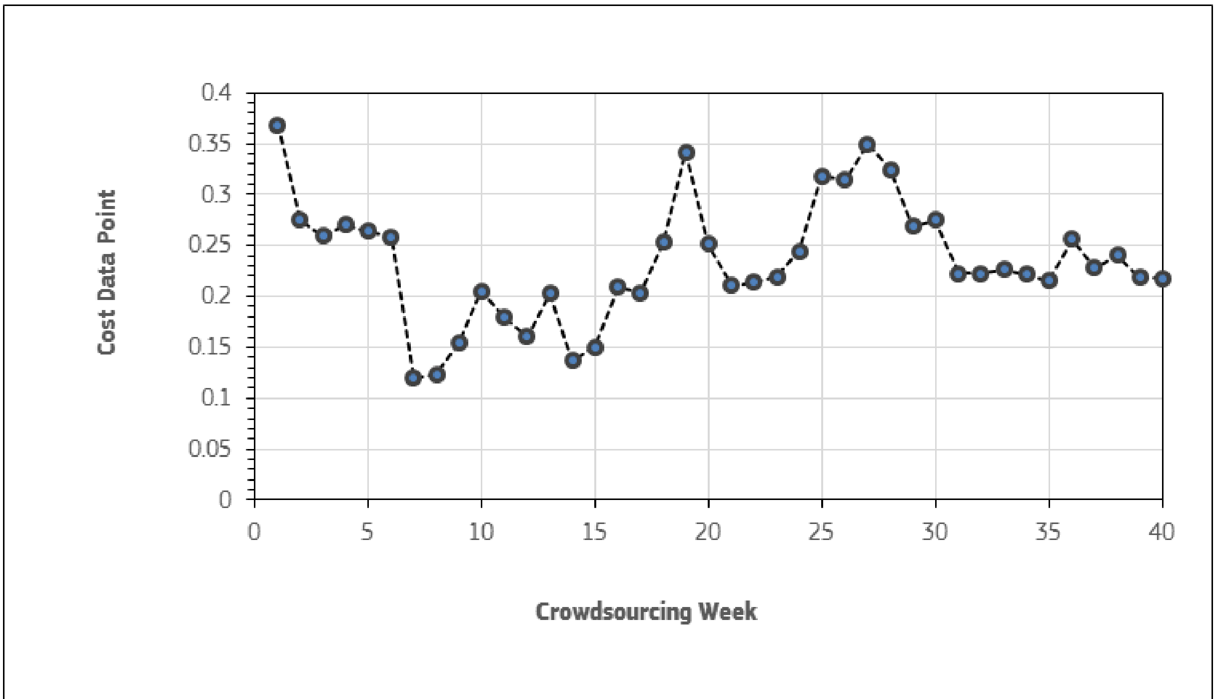
A major advantage of a crowdsourcing approach to data collection is the possibility of collecting a large volume of high-resolution data over space and time at a possible low cost. This indicator is calculated as the monetary cost of motivating the crowd (i.e. through financial compensation) divided by the total number of data points generated by the crowd. The indicator is expressed in monetary terms for direct comparison with other data crowdsourcing efforts.

$$Cost\ per\ data\ point_t = \frac{number\ of\ rewarded\ records_t \times r}{number\ of\ valid\ data\ points_t}$$

Where $t(= 1, \dots, T)$ is the period for calculation (e.g. weeks, months) and r is the amount of the monetary incentive per submission (data record).

Figure 36 shows the evolution of this indicator over the project implementation period.

Figure 36. Cost per data point during the FPCA project implementation period.



Ideally, the cost per data point should tend towards zero (0) as data points increase and the reward efficiency indicator (above) decreases, however, stabilising the cost of crowdsourced data at €1 per data point (or less) is arguably a reasonable benchmark for sustainability. Based on the actual costs of the reward incentive offered to volunteers during data collection, notwithstanding institutional/administrative costs ([30]), the average cost per data record was €0.85. The average cost per data point was €0.21 (assuming an average of four commodities reported per valid data record) (Figure 36). It is noteworthy that, in a practical sense, governments/businesses cannot acquire commodity data for free, and often must make huge investments (e.g. in staff or consultancy projects) to collect data and to obtain data of the quality of the dataset so far collected.

7.5.5 Track time to execute the validation/post-sampling script

The R script for data validation and aggregation loads at each time (twice a day) the data from the current week, which is appended dynamically to a historical dataset. By doing so, the full set of codes run for about five minutes.

Box 11. Interpretation of cost and burden

The burden to volunteers as shown by the time to register and time to submit indicators remains relatively low as for the majority of participants it only takes around 10 min to register and around 6 min to submit data. Volunteers are not requested to go to any specific market but can submit data in their routine visits to the market, therefore no transport costs can be accounted for here.

The burden (excluding set up) to the data administrator is also low once the system is set up since all processes run automatically and data is validated by the algorithms included in the script and uploaded to the web dashboard within minutes.

Furthermore, concerning running costs, the rewarded submissions account for around 30% of total submissions, which indicates that the crowdsourcing system is efficient.

The cost per data point stays at around $0.20, which could be further improved if an increasing amount of data were contributed with no monetary reward. The use of behavioural tools has proven useful for this (see Section 8).

7.6 CONFIDENTIALITY

Ensuring the privacy of personal information is a crucial aspect of the quality of the data in a crowdsourcing data collection effort, and this must be part of the design of the initiative.

In this project, to protect volunteers' identities, the data submissions undergo an anonymisation process..

Box 12. Application of confidentiality

Confidentiality and privacy of personal data are ensured by design in the crowdsourcing initiative. All price observations undergo an anonymisation process, where all data that can be linked to a particular person is either removed or securely hashed. Hence, there is no way to trace an observation back to the person who submitted it.

7.7 STATISTICAL PROCESSING

Statistical processing covers all activities and procedures from data collection, through data validation to data dissemination.

7.7.1 Crowd size index

This indicates the evolution of the total number of registered volunteers over time:

([30]) The associated administrative costs are not included because this is mainly associated with data review rather than actual data submission from the crowd.

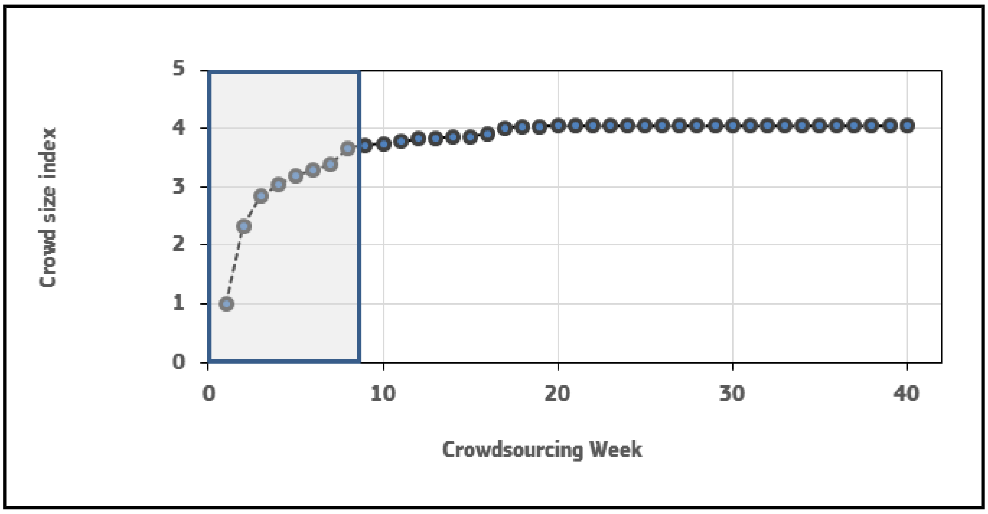
$$Crowd\ size\ index_t = \frac{crowd\ size_t}{crowd\ size\ _0}$$

Where $t(= 1, \ldots, T)$ is the target period of evaluation and 0 is the start time.

Increasing the amount of contributed data can be achieved at the intensive (encouraging the participation of registered volunteers) or the extensive margin (engaging new volunteers). Therefore, an index measuring the registration of new volunteers can be useful in evaluating the performance of the system. If communication channels are used at a specific time, the change of index before and after use of a communication channel can also be assessed using this indicator.

The number of new volunteers more than doubled in the second week (~2.3), while it achieved a score of 3.5 at the 8th week. The index stabilised after the 17th week at around 4.0 (Figure 37). This evolution is consistent with the timing of the awareness campaign (see Section 4.2) that was carried out during the ideation and pilot phases (~8 weeks) (Figure 37, shaded area) and that was discontinued once a sufficient number of volunteers was reached. This development suggests that engaging new volunteers may require additional and sustained awareness activities.

Figure 37. Crowd size Index (CSI) during the period of implementation of the FPCA project in Kano and Katsina States



Note: shaded area corresponds to the period in which different communication channels were used.

7.7.2 Crowd engagement index

This index is calculated as the relationship between the number of active volunteers (i.e. those submitting data) and the volunteer pool during each target period or time step (e.g. weekly or monthly).

$$Crowd\ engagement\ index_t = \frac{active\ crowd\ size_t}{crowd\ size_t}$$

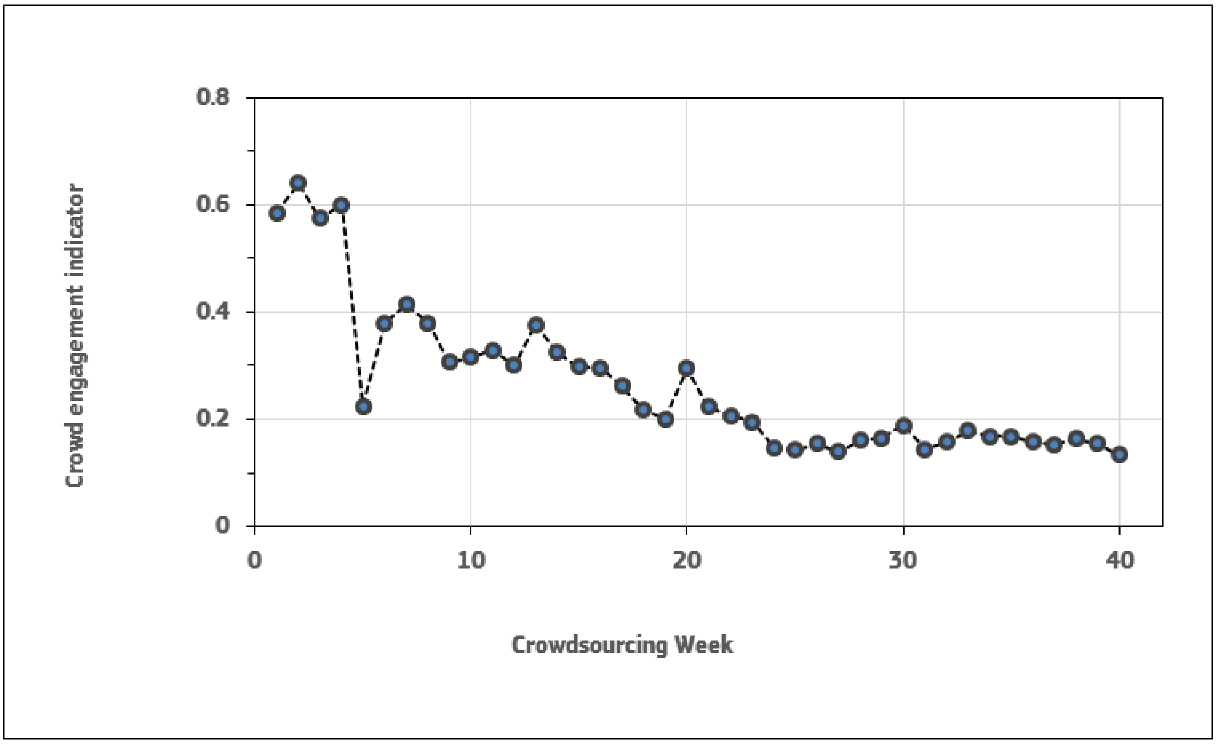
$t(= 1, \ldots, T)$ is the target period.

The index is based on the underlying assumption volunteers are considered active if they submit data at least once in the period (e.g. week or month)([31]). Therefore, a perfect value for this indicator (i.e. 1) is attained when at least one data point is associated with each unique volunteer ID. In contrast, a value of 0 implies that volunteers are disengaged for the period assessed. Taking into account the scope of the project and contextual realities within the region, a value of 0.1 can be considered as a benchmark for success for the period of the crowdsourcing. We achieved an indicator range of 0.14 – 0.64, with the maximum score attained

([31]) For the purposes of this project, the crowd engagement indicator was assessed as a WP index.

during the second week (the pilot phase). The indicator tended to stabilise around 0.15 after week 24 (Figure 38).

Figure 38. Crowd engagement index (CEI) during the period of implementation of the FP CA project in Kano and Katsina States



Different strategies can be suggested to improve crowd engagement. These strategies should activate participants' extrinsic motivation (e.g. by tailoring rewards) or intrinsic motivation (e.g. by sharing purpose, sharing the 'social norm' or how other individuals perform, gamifying the task, improving features of the platform) (Daniel, Kucherbaev, Cappiello, Benatallah, & Allahbakhsh, 2018).

During project roll-out, several strategies were implemented to keep the crowd engaged. These included both monetary rewards and behavioural incentives or 'nudges'. The assessment of the performance of these strategies is discussed in Section 8. The behavioural factors were implemented based on an experimental design that allowed for rigorous testing of their effectiveness; however, no experimental test was implemented for monetary reward. Despite the non-testing of the effect of monetary reward (due to time and budgetary constraints), the set-up of a reward system provided a basis for gamifying the crowdsourcing task, with the expectation that it would have a positive impact on the intrinsic motivation of the volunteers of the crowd (see Section 8).

7.7.3 Week-day bias index (WBI)

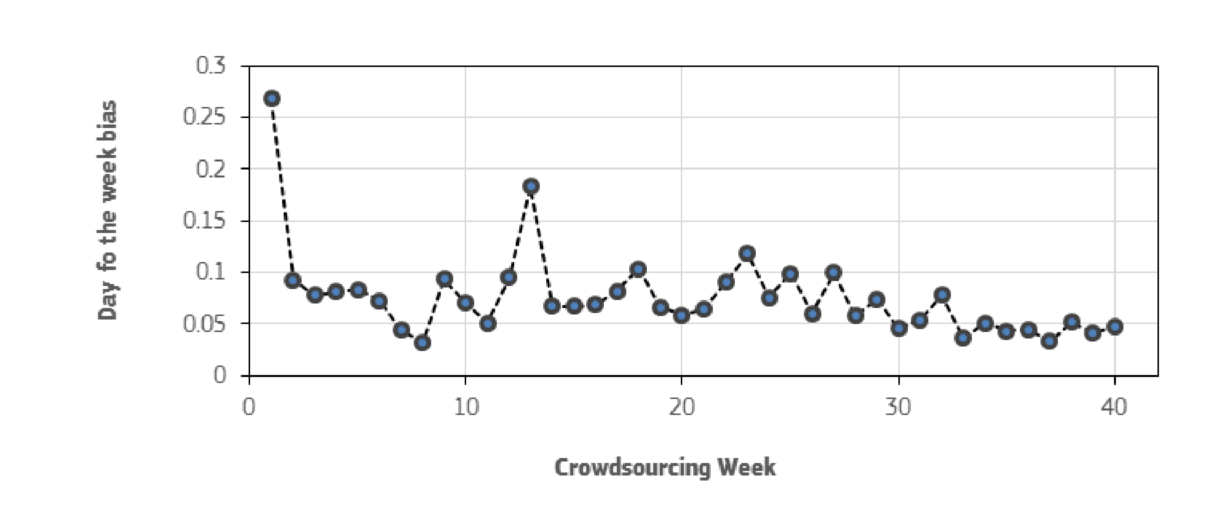
This is an index for the range of total daily data submission frequencies within the specific period of interest and can reveal "day-of-the-week" bias in the crowdsourcing system.

$$Week\text{-}day\ bias\ indicator_t = \frac{\max\limits_d s_{d,t} - \min\limits_d s_{d,t}}{\sum_{d=1}^{d=7} s_d}$$

Where $s_{d,t}$ is the total number of daily price submissions in week $t(=1,\dots,T)$ in a given day of the week $d(=1,\dots,7)$. An ideal food price crowdsourcing system should generate more-or-less constant daily submission of food price data, irrespective of the day of the week. If markets are only open during a sub-set of the week, the period of analysis should exclude those dates when markets are closed. The indicator is calculated as the difference between the maximum and the minimum number of daily data submissions within the period of interest, divided by the total number of submissions during the same period. An unbiased crowdsourcing system would be characterised by a low value (closer to 0), while a biased system would exhibit a high value of 1.

During the FPCA implementation, the indicator stabilised at a value of ~0.04, indicating that the system is not biased towards any particular day of the week (see Figure 39).

Figure 39. Week-day bias Index during the period of implementation of the FPCA project in Kano and Katsina States



7.7.4 Market type Bias Index (MBI)

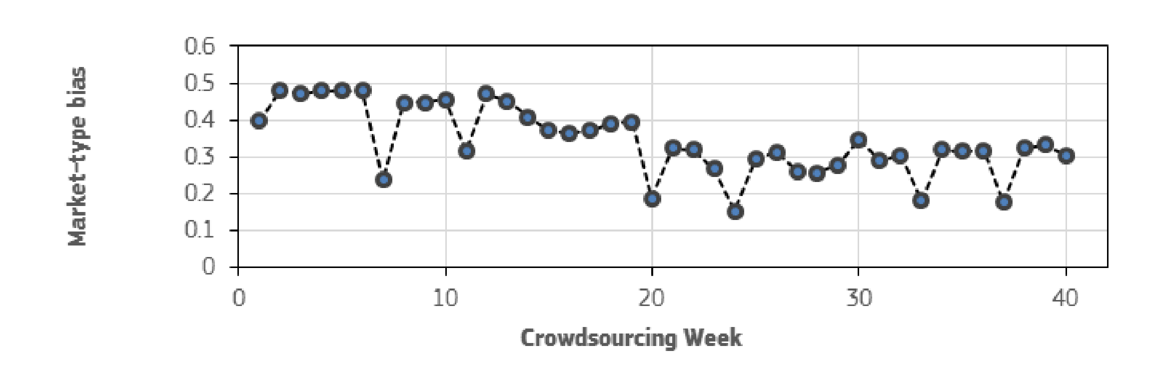
Measures the number of price submissions from each market/outlet type in relationship to overall price submissions and indicates the bias towards single market type. A Herfindahl type index has been selected for this measure.

$$Market\text{-}type\ bias\ indicator_t = \sum_j \left(\frac{s_{t,j}}{s_t}\right)^2$$

$s_{t,j}$ is the number of price submissions in period $t(= 1, \ldots, T)$ in market type $j(= 1, \ldots, 6)$ and s_t is the total number of price submissions in t.

Given that volunteers can submit food price data sourced from different types of markets (e.g. market, supermarket, kiosk, farm gate), it is relevant to ensure that the crowdsourcing system is designed to minimise bias towards a particular market type. This index compares the submissions from each market type relative to the full set of submissions. Although there was no a priori benchmark for representation across the market types, the use of this indicator is focused on its temporal change, as the main categories of market outlets are assumed to remain constant over the project's timeframe. An unbiased crowdsourcing system would be characterised by a low value (closer to 0) while a biased system would exhibit a high value of 1 (maximal concentration toward one type of market). This index stabilised after the 20th week at around 0.25, which suggests minimal bias in data submission across market types within the crowdsourcing system (Figure 40).

Figure 40. Market Bias Index during the period of implementation of the FPCA project in Kano and Katsina States



7.7.5 Track availability of a manual associated with the process of data collection

An online information platform (https://sites.google.com/view/foodprice) was developed to ensure contactless information access to potential and engaged volunteers about the purpose and approach of the project, how to participate, rules and guidelines and frequently asked questions (FAQs).

7.7.6 Track availability of an operating manual to guide data processing, validation, compilation and dissemination

A spatial statistical methodological approach has been developed and documented to deal with potential non-sampling and sampling error inherent to crowdsourcing approaches. This allows different aspects of the submitted data to be validated in close to real-timeto filter out potential erroneous observations and to aggregate information in a way that allows for statistical inference. All the steps and algorithms are programmed in R code and are well documented.

Box 13. Interpretation of the quality of the statistical processing

Statistical processing refers to the whole set of activities and processes from data collection, through validation to dissemination.

An important quality aspect of a crowdsourcing dataset is the number of voluntary contributions which can be achieved by engaging new volunteers or by promoting the regular participation of engaged volunteers. The crowd size index reflects that the awareness campaign run during a limited number of weeks was efficient to engage a sufficient number of interested volunteers (737). Yet the crowd engagement index reflects that about a 10% of them regularly provide data, which is considered as an acceptable value.

Moreover, since a citizen-generated dataset usually does not follow any sample design it is important to measure the potential biases. Both indices day-week bias and market-type bias show that data is collected to some extent homogenously from all days of the week and a variety of market types.

A manual associated with the process of data collection offered online ensures that data collectors are correctly informed on methodological issues to promote correctness at data entry.

Finally, for a crowdsourcing initiative to perform well, it is essential to use sound methodological statistical approaches for data processing, validation, compilation and dissemination which is ensured by the automation of all steps and algorithms through the R script, which is well documented.

7.8 Crowdsourcing/citizen datasets quality labels

The measurement of all quality indicators as proposed in Section 5.5.1 provides a useful tool for monitoring the performance of this crowdsourcing initiative, while it may also be of great help in designing new initiatives. Yet, from a users' perspective, it is essential to accompany a crowdsourced dataset with a quality measurement (e.g. in the form of a label). This provides an idea of the general quality of the dataset and each of the dimensions. There might be trade-offs between dimensions (e.g. completeness for accuracy). For this purpose, a process of normalisation and weighting of indicators should follow, but this is beyond the scope of this report.

8 Sustaining the system ([32])

The previous sections have described the design of the citizen-driven data generation process, its performance in terms of quality assurance and its dissemination via the web dashboard. One of the main challenges of citizen science approaches in general and of crowdsourcing initiatives, in particular, is how to assure contributions from citizens while maintaining the costs at a reasonable level and assuring a minimum level of submission quality. Thus, once the system is up and running and the crowd has been engaged via dissemination campaigns (see Section 4.2.3), the next challenge is to assure that the crowd provides a steady flow of (quality) data contributions. As part of the FPCA implementation, we tried to mobilise the crowd using two types of incentives.

First, we focused on a classic monetary incentive approach by including a monetary reward for contributors. Besides, we tried to mobilise behavioural leverages using several nudges to influence the behaviour and decision making of individuals through positive reinforcement and indirect suggestions (Thaler & Sunstein, 2009). The first one would aim to activate what researchers on crowd motivational factors call *extrinsic motivations,* those that come from outside the individual, providing an immediate or delayed payoff. The second would aim to activate *intrinsic motivations*. This is intangible motivations that come from inside the individual, linked to personal satisfaction or accomplishment, for example, 'fun', 'enjoyment', 'learning', 'social interaction' and 'social contribution' (Daniel et al., 2018; Kaufmann, Schulze, & Veit, 2011; Pedersen et al., 2013; Zeug et al., 2017). Intrinsic incentives motivate people to engage in behaviour because it is personally rewarding, not because of an external reward. There is no consensus as to whether intrinsic or extrinsic motivational factors are more effective. We further discuss the two strategies and evaluate their performance in relation to the objective below. A more detailed analysis can be found at Solano-Hermosilla *et al.*(forthcoming).

8.1 MONETARY REWARDS

Setting up a crowdsourcing initiative requires a significant up-front investment to cover the design and roll-out of the platform, and to publicise it to potential crowd participants. In the case of the FPCA, €44 000 out of a total budget of €122 000 was invested in financial rewards for ten months' data collection. The rest of the budget was devoted to outreach, the data platform, local staff and overall administration. These costs exclude the contribution of JRC to conceptual development and implementation. In principle, the investment cost could be integrated into the working programs of institutions devoted to data collection (DGINS, 2013, 2018), thus reducing the additional investment needed to set them up and maximising complementarity with traditional statistical operations.

However, once the system is up, the main cost is the monetary reward incentive ([33]). Monetary incentives intended to motivate crowd volunteers by providing small rewards to a sub-set of them who were selected on a "first-submit first-rewarded" basis, subject to other controls. The fact that not all data submissions were rewarded introduced a challenge for volunteers. Although the monetary reward was expected to motivate volunteers, they were unable to pre-determine if their data submission would be rewarded because the information about the daily submission sequence or frequency was not visible to the crowd volunteers. In FPCA project, as explained in Section 4.2.7, a token amount of ₦2 000 (~€4) was paid for 25 submissions (in the pilot) or 35 submissions (during the roll-out) daily, which:

- are valid according to the pre-defined price ranges, as established in the quality control procedure (unknown to the crowd); and
- have not been rewarded over a certain number of times in a specific period (maximum of twice per week and four per month).

Thus, the FPCA reward is somehow "gamified". The combination of a monetary reward and gamification constitutes a mix of extrinsic and intrinsic incentives to motivate the initial decision to participate, and then enhance participants' experience over time to sustain engagement (Ziegler et al., 2017).

Figure 41 shows how the number of weeks in which volunteers have made submissions, which were either validated (Figure 41b) or not (Figure 41a) is positively correlated with the number of weeks in which the

([32]) Solano-Hermosilla *et al.*, 2020, forthcoming.

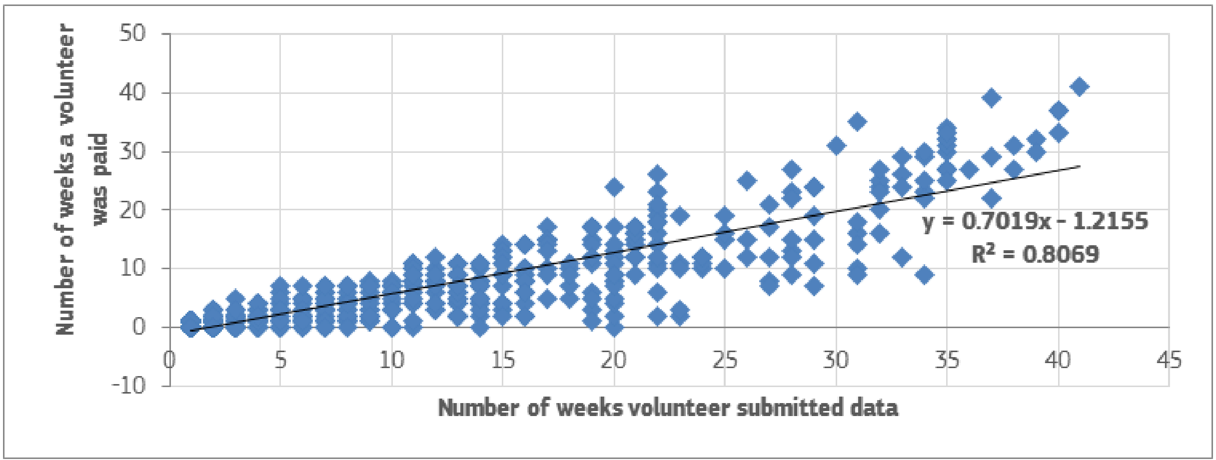
([33]) Assuming that the platform use fees are minimal and there are no maintenance costs.

volunteer received a monetary reward (correlation coefficients of 0.80 and 0.79 respectively, and *p*-values < 0.01).

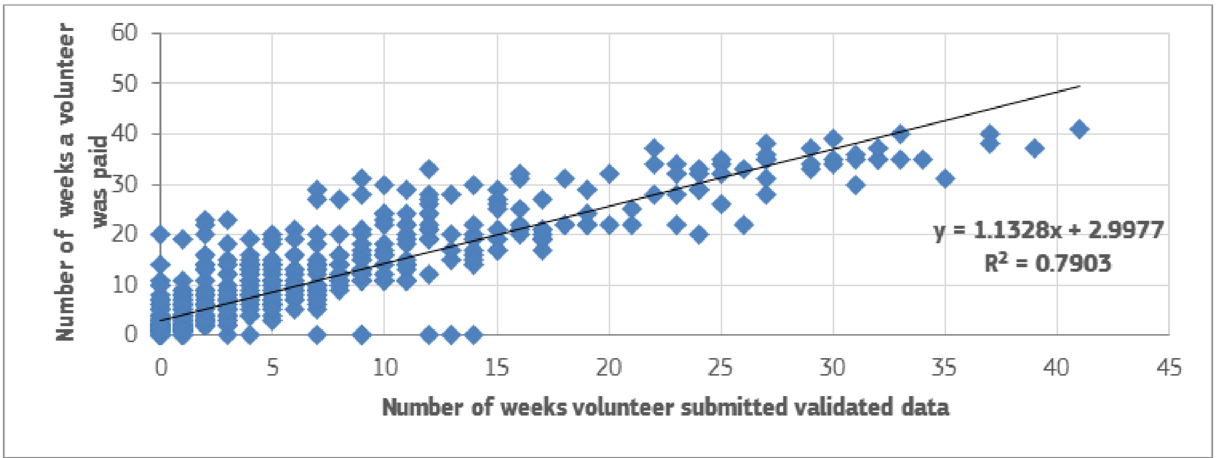
Moreover, we compared the number of weeks in which volunteers submitted data for two groups: volunteers, which had been monetarily rewarded at any moment in time, and those that had not. So we find that the rewarded group contributes significantly more prices, both raw and quality validated (*p*-value = 0.0000 for a two-sided *t*-test).

Figure 41. Correlation between the number of weeks in which volunteers submitted data which either was validated (b) or not (a) and the number of weeks in which the volunteer was monetarily rewarded during the 41 weeks of the FPCA crowdsourcing project between September 2018 and June 2019

a)

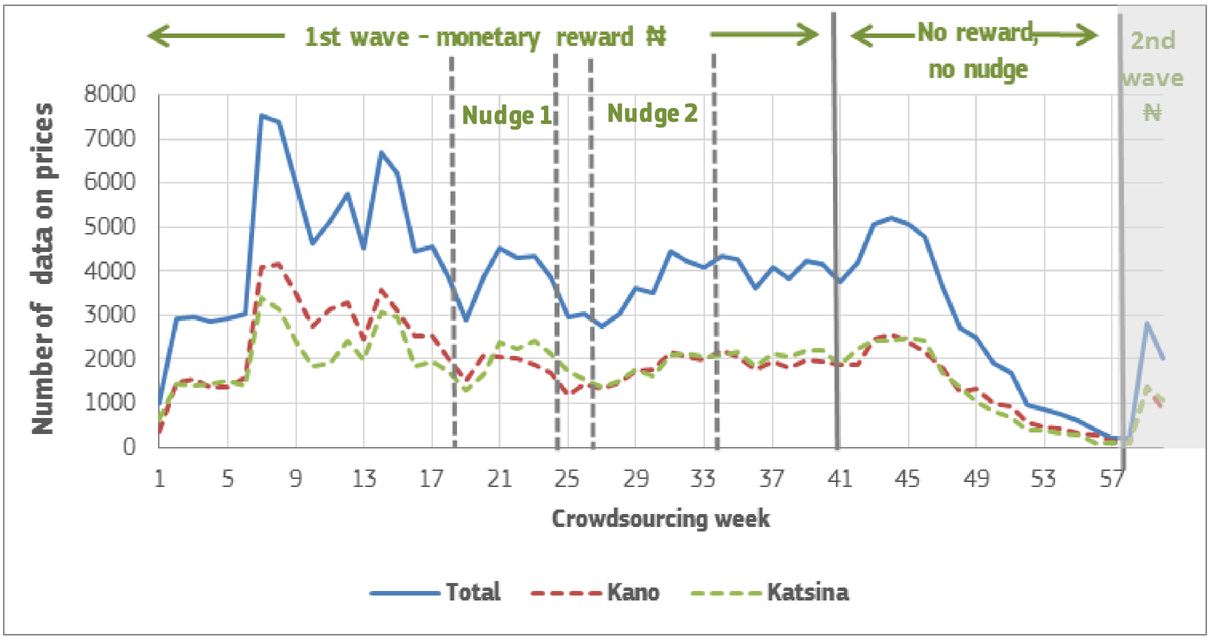


b)



The financial reward was part of the FPCA marketing campaign (both the flyers and the radio adverts mentioned “Earn up to ₦8,000 monthly”), meaning that initial participation was affected by the opportunity to receive financial compensation. However, a few weeks after the start of the roll-out phase, volunteer participation diminished. Hence, the impact of the monetary reward on participation seems to diminish (see weeks 1 to 17 in Figure 42). At this stage, we implemented new tools to sustain contributions, focusing on activating behavioural leverages as a complementary tool. The monetary reward appeared to be a necessary condition to ensure participation, because once the financial reward disappears, contributions also disappear. The data from the second wave of incentives not included in this project shows how introducing a new monetary reward reactivates immediately crowd participation (Figure 42). However, it is not a sufficient condition as the participation rate diminishes without additional incentives, as mentioned above. The next section describes the mechanisms and results of the implemented behavioural tools or nudges.

Figure 42. Number of price submissions per week during the 41 weeks of the FPCA project and after



8.2 BEHAVIOURAL NUDGES

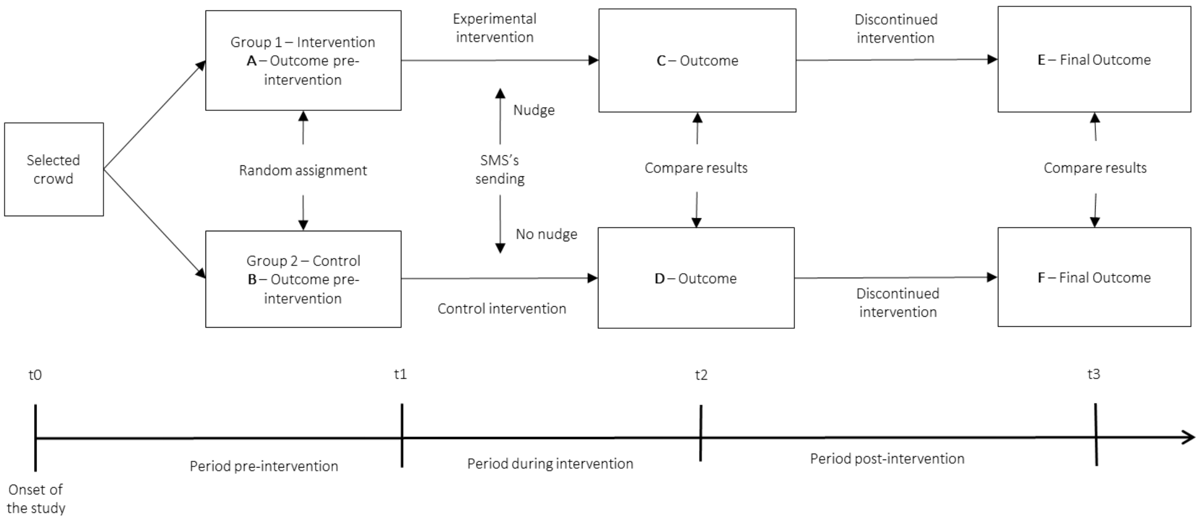
As mentioned above, humans are not pure income maximisers; there are many other factors that drive their behaviour. Insights from behavioural sciences can be applied to design additional tools to help sustain crowd contributions by mobilising well-documented behavioural leverages to foster engagement of individuals (APA, 2018). In the implementation of the FPCA, two behaviourally informed interventions were implemented in the form of nudges. Sunstein (2014b) defined nudges as 'liberty-preserving approaches that steer people in particular directions'. Nudges have proven to be successful in different policy interventions, such as, for example, promoting tax payment (Hernandez, Jamison, Korczyc, Mazar, & Sormani, 2017), child school attendance (Damgaard & Nielsen, 2018) or healthy eating (Kroese, Marchiori, & de Ridder, 2015).

The nudges were implemented following a random control trial to allow impacts on participation in the treatment group to be compared with the control group (Solano-Hermosilla *et al.*, forthcoming). Half of the registered crowd received the nudge (the treatment), and the other half did not (the control). The experimental design permits analysis based on the gold standard of intervention evaluation: the Difference-In-Difference (DID) analysis. DID analysis is used to measure differences between the treatment and control groups in terms of changes in the outcome variable (in our case number of price submissions) over time. Figure 43 presents this design graphically. We do this for two types of impacts:

a) Short term impacts: comparing the increase in the average number of submissions before and during the intervention

b) Long-term impacts: comparing the increase of the average number of submissions before and after the intervention

Figure 43. Experimental design of the implementation of the nudges



In practical terms, the nudge is implemented as SMS messages sent to the crowd with different information for the experimental and control group. SMS was used as the communication vehicle because SMS was the preferred option for communication as indicated at registration by the volunteers. To automate communication and handle it efficiently, we put a system in place that was able to generate all messages for volunteers and send them using the APIs of two online SMS platforms ([34]). This allowed us to minimise the time devoted to this task, as well as reducing the room for human error. The script was able to calculate necessary information and automatically generate the "nudge" as the text to be included in the SMS and send it with minimal human intervention. The script was programmed using the Python programming language and an SQLite database for calculating individual performance indicators (for each volunteer), generating batches of messages using SQL queries, and for storing generated messages and their status.

We use DID analysis to test the impact of two types of behavioural leverages on the contributions of the crowd, one based on 'social norms' and one based on 'information disclosure' (Sunstein, 2014a).

The DID analysis is usually implemented in a regression model as an interaction term between time and treatment group dummy variables. In this regression, time is a variable that takes a value 0 (time = 0) if the period corresponds to the time span before intervention (t < t1 in our model) and 1 (time = 1) after (t > t1 in our model) and treatment takes a value of 0 if the volunteer belongs to the control group (treatment = 0), and 1 if the volunteer belongs to the treatment group (treatment = 1). The basic form of the regression can be shown as follows, where Y_i is the outcome variable of individual $i(= 1, \ldots, I)$ (Villa, 2016):

$$Y_i = \beta_0 + \beta_1 Time + \beta_2 Treated_i + \beta_3 (Time \times Treated_i) + \varepsilon_i$$

The estimated coefficients can be interpreted as follows:

Table 13. Interpretation of the DID estimated coefficients for comparing the outcome before (baseline or pre-intervention period) and at the end of intervention.

Coefficient	Calculation	Interpretation
β_0	B	The mean outcome of the control group at the baseline
β_1	D - B	The time trend in the control group

([34]) Most of the messages were sent using a local Nigerian SMS platform (MTNbulksms.com), but in a few cases where this platform could not deliver them, we used a backup international SMS service (Fryngа.com).

Coefficient	Calculation	Interpretation
$\beta_0 + \beta_1$	B + D - B = D	The mean outcome of the control group in the intervention period
β_2	A - B	The single difference between the treatment and the control groups at the baseline (pre-intervention period)
$\beta_0 + \beta_2$	B + A – B = A	The mean outcome of the treatment group at the baseline
β_3	(C – A) – (D – B)	The DID estimator measuring the difference in changes over time of the control and intervention group
$\beta_0 + \beta_1 + \beta_2 + \beta_3$	B + D – B + A – B + (C-A) – (D – B) = C	The mean outcome of the treatment group in the follow-up (intervention period)

A similar interpretation can be made when comparing the outcome before the intervention and after the implementation period of the intervention, where C and D should be replaced by E and F.

The DID analysis also allows for the inclusion of covariates, known as controls or observable characteristics of individuals, in order to control for compositional changes. Adding those covariates may improve the precision of the DID estimate even if the intervention is independent of observed covariates (Lechner, 2011; Villa, 2016). If we denote $covariate_{k,i}$ as the *k*th covariate, the regression can be expressed as follows:

$$Y_i = \beta_0 + \beta_1 Time + \beta_2 Treated_i + \beta_3 (Time \times Treated_i) + \beta_k covariate_{k,i} + \varepsilon_i$$

The inclusion of covariates allows for a balancing *t*-test for difference in the means of the specified covariates between the control and treatment groups at time = 0.

Yet, to estimate any causal effect in DID analysis, several assumptions must hold:

- The allocation of the intervention cannot be determined by the outcome variable, and is therefore unconnected to the outcome at the pre-intervention period (the baseline);
- The treatment group and the control group have parallel trends in outcome in the pre-intervention period (the baseline);
- The composition of the treatment/intervention group and the control group is stable over time;
- There are no spillover effects from the intervention group to the control group.

In our study, we apply a random allocation of the intervention to the individuals to ensure that the intervention is unrelated to the outcome at the pre-intervention period. The allocation remains stable over time, ensuring the stable composition of the groups. The parallel trend assumption, which is critical, means that in the absence of treatment/intervention, both groups have a constant difference over time. Visual inspection is used later to confirm the parallel trend. Spillover effects could occur through social interactions (between participants in the treatment and control group) or other externalities, and occur when actions taken by individuals in the treatment group affect the outcomes of individuals of the control group. We cannot test for this last condition, so we have to assume that there were no systematic interactions between groups. This assumption is explained in more detail below for each of the interventions.

8.2.1 The impact of sharing the “social norm" as a nudge in crowdsourcing

The first nudge focused on activating "social norms”. Research on nudges in behavioural economics finds that individual behaviour is greatly influenced by the perceived behaviour of other people (Sunstein, 2013). It may raise cooperation and competition factors that can increase intrinsic motivation in situations when the participants can compare their performance to that of others. Social norms have been successfully tested as drivers to increase tax compliance (Cullis, Jones, & Savoia, 2012) and waste recycling (Linder, Lindahl, & Borgström, 2018) among others. There are different types of social norms that can be activated: descriptive norms (based on what others do), injunctive norms (based on what should be done according to what most people approve or disapprove of), subjective norms (based on what should be done according to what individuals close to them approve or disapprove of), and personal norms (based on the person's own standards or expectations. In our case, we use a descriptive norm, defined as a norm derived from the observation of what other people do, which is considered helpful in social situations to maximise effectiveness (Bobek, Roberts, & Sweeney, 2007). Therefore, our nudge does not add any judgement value to the behaviour, just reports the behaviour of the crowd and how the individual's behaviour relates to it.

For this, from week 20 up to week 25, the crowd started receiving a weekly SMS (Table 14).

Table 14. SMS texts for the intervention and control groups, respectively.

Group	SMS text
Treatment	Thank you for submitting XX prices last week to FPCA! On average, participants submitted XX prices, that puts you among the top 10% of participants.
	Thank you for submitting XX prices last week to FPCA! On average participants submitted XX prices, at least 10% of them submitted more prices than you.
	Thank you for submitting XX prices last week to FPCA! On average participants submitted XX prices, at least 25% of them submitted more prices than you.
	Thank you for submitting XX prices last week to FPCA! On average participants submitted XX prices, at least 50% of them submitted more prices than you.
	Thank you for submitting XX prices last week to FPCA! On average participants submitted XX prices. You are at the lower end, with 75% submitting more than you.
Control	Thank you for submitting prices last week to FPCA.

One of two versions (treatment or control) of the SMS was randomly assigned to the registered crowd members that had submitted data at any time during the first 20 weeks. Half of them received a simple thank you message while the SMS sent to the other half also included information on their number of submissions, and how this compared to the average submissions of the crowd. We included a non-committal 'thank you' message for the control group to allow us to identify the impact of the social norm independently of the impact of an SMS message. After week 25, the system continued to work without any further SMSs being sent. Before any analysis can be done, the effectiveness of the randomisation process needs to be tested. Table 15 presents a description of the characteristics of the sample as a whole, and of the two sub-samples. None of the characteristics analysed was significantly different between the two sub-samples, thus allowing us to conclude that the randomisation process was effective (Table 16).

Table 15. Description of the characteristics of the sample for the social norm nudge.

Nudge 1: Social norm							
	Social norm		**Control**		**Full sample**		***p*-value**
Variable	**n**	**%**	**n**	**%**	**n**	**%**	
Total	191	100%	186	100%	377	100%	
Gender							0.36
Male	31	16%	24	13%	55	15%	
Female	160	84%	162	87%	322	85%	
Education							0.41
Primary	0	0%	1	1%	1	0%	
Secondary	24	13%	29	16%	53	14%	
Tertiary	167	87%	156	84%	323	86%	
Food chain stage							0.41
Farmer	80	42%	88	47%	168	45%	
Final consumer	73	38%	56	30%	129	34%	
Retailer & wholesaler	30	16%	20	11%	50	13%	
Other	8	4%	22	12%	30	8%	
Preferred communication channel							0.23
app	24	13%	12		36	10%	
email	33	17%	27		60	16%	
SMS	117	61%	131		248	66%	
web	5	3%	6		11	3%	
unknown	12	6%	10		22	6%	
	Mean		**Mean**		**Mean**		
Age	27.25		27.46		27.35		0.69
Avg. years smartphone use	6.3		6.6		6.4		0.43

Note: *p*-values are the results of Chi-Square tests of equality between categorical variables or from *t*-test of equality of means for continuous variables; *** $p<0.01$; ** $p<0.05$; * $p<0.1$

The outcome variable we focus on in analysing this intervention is the number of prices submitted weekly by each volunteer. Table 16 describes the variables for the treated and control groups at the pre-intervention stage.

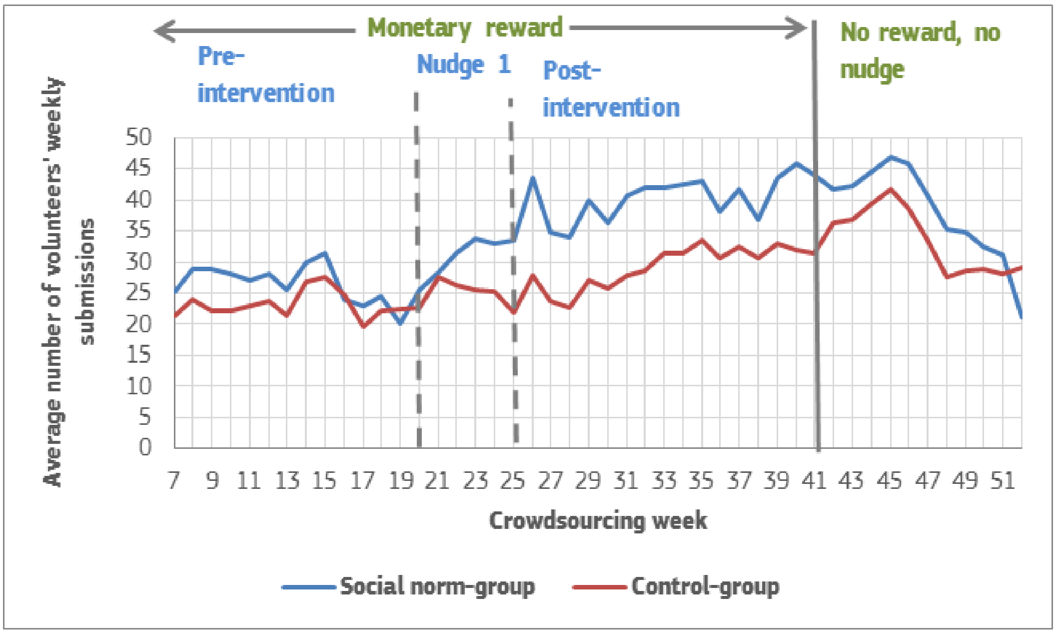
Table 16. Description of the variables and results of *t*-test at the baseline (pre-intervention) stage for the social norm nudge.

Variable(s)	Description	Mean control	Mean treated	Diff.	|t|	Pr(|T|>t)
n_weekly_obs	The number of prices submitted by a volunteer per week.	23.213	25.475	2.263	1.49	0.1365
treated	=1 if the volunteer has been allocated to the nudged group, otherwise 0	0	1	1		
time	=1 if the week corresponds to the intervention period					
lag_reward	=1 if the volunteer has been rewarded in the current week, otherwise 0	0.315	0.331	0.016	0.49	0.6262
nudge2	=1 if the volunteer receives the treatment message of the second nudge	0.519	0.536	0.017	0.47	0.6371
gender	=1 if the volunteer is male, otherwise 0.	0.826	0.863	0.037	1.43	0.1519
age	=1 if the volunteer is < 30 years old, otherwise 0	0.749	0.736	-0.012	0.39	0.6958
high_education	=1 if the volunteers has tertiary education, otherwise 0	0.878	0.865	-0.014	0.57	0.5676
farmer	=1 if the volunteer is registered as a farmer, otherwise 0.	0.478	0.482	0.004	0.12	0.9083
consumer	=1 if the volunteer is registered as a consumer, otherwise 0.	0.331	0.331	0	0.01	0.9902
trader	=1 if the volunteer is registered as a trader, otherwise 0.	0.086	0.088	0.002	0.11	0.9122
other	=1 if the volunteer is registered as other, otherwise 0.	0.105	0.099	-0.006	0.27	0.7841

*** $p<0.01$; ** $p<0.05$; * $p<0.1$

An initial inspection of trends in the number of price submissions (Figure 44) shows that before the intervention (baseline) both groups had a similar average number of contributions and trends (confirming the assumption of the parallel trend). Following the intervention, we see a spike in the number of contributions for both groups. However the increasing trend continues in the treatment group while it plateaus (and even reverses) for the control group. Only from week 51, 26 weeks after intervention and after economic incentives were discontinued, did the two groups suffer a decline in the average number of weekly contributions, and the differences between the groups cease to be significant.

Figure 44. Average number of price submissions per week by group before, during and after the implementation of the social norm nudge



The DID analysis was carried out using Stata (StataCorp., 2017). The results show that both the control and treatment groups experienced a hike in price submissions following the introduction of the SMS reminder. However, the treatment group increased their submissions more than the control group. The additional increase was on average 4.8 when comparing the pre-treatment with the treatment period, and on average 6.9 when comparing the pre-treatment and the post-treatment stages. These effects are statistically significant at 5% and 1%, respectively (Table 17). We can thus conclude that the nudge based on social norms was effective, boosting participants' engagement in both the short and medium-term.

Table 17. Difference-in difference estimation results for the implementation of the "social norm" intervention, before vs. period of intervention (left) and before vs. after period of intervention (right).

Outcome var.	**Mean**	**S. Err.**	**|t|**	**P>t**
Before				
Control	23.213			
Treated	25.475			
Diff (T-C)	2.263	1.678	1.35	0.178
After				
Control	23.441			
Treated	30.256			
Diff (T-C)	7.085	1.791	3.96	0.000***
Diff-in-Diff estimator	**4.823**	**2.454**	**1.96**	**0.050****

Outcome var.	**Mean**	**S. Err.**	**|t|**	**P>t**
Before				
Control	23.213			
Treated	25.475			
Diff (T-C)	2.263	2.357	0.96	0.337
After				
Control	31.336			
Treated	40.567			
Diff (T-C)	9.231	1.229	7.51	0.000***
Diff-in-Diff estimator	**6.969**	**2.658**	**2.62**	**0.009*****

*** $p<0.01$; ** $p<0.05$; * $p<0.1$

Even though the key covariates are balanced at the baseline (the pre-intervention period), and the differences are not statistically significant (Table 16), for robustness, we tested the DID model with covariates. The results for the comparison of the pre-intervention period and the intervention period are in line with the previous estimation. They show a significant DID coefficient of 4.2 at 10% significance level. Similarly, the results for comparison of the pre-intervention period and the post-intervention period are in line with the previous estimation. They show a significant DID coefficient of 7.0 at 1% significance level. When controlling for the receipt of financial compensation, a time-varying factor, the result of the regression shows that financial compensation increases the participation of the whole sample (positive and statistically significant coefficient). This suggesting that the effects of nudge and reward are additive.

8.2.2 The impact of sharing (disclosing) the "collective dataset" as a nudge in crowdsourcing

For the second nudge, the FPCA project draws on the utilitarian and hedonic perspectives of the FPCA crowdsourced dataset, and especially of the web dashboard. The main purpose of the FPCA project is to produce a collaborative dataset of daily food prices, disseminated in real time as open data through an IT dissemination tool (the web dashboard), to support data users in decision-making processes. Thus, the final product can be considered an Information System (IS). Information system research distinguishes between utilitarian systems, which are utility-oriented systems that intend to provide instrumental value, and hedonic systems, which are pleasure-oriented systems that intend to provide self-fulfilling value. Crowdsourcing studies find that crowdsourcing systems serve the dual purpose of being utilitarian and hedonic information systems (Soliman & Tuunainen, 2015).

From a behavioural perspective, Sunstein (2013, 2014a) claims that information disclosure can also be considered leverage and thus can be implemented as a nudge, replacing or complementing other approaches. This is the case, for example, when immediate information allows consumers and producers to make better decisions. If provided to the public, the FPCA web dashboard can be seen as a useful tool for participants and for the general public (a "common good") to support better decisions. In the case of the FPCA, it is expected that giving access to the web dashboard to participants as information users may contribute to improved user decisions. At the same time, the perceived usefulness of the tool, on the one hand, and contribution to a common good, on the other hand, may awaken both extrinsic and intrinsic motivations to increase the citizen volunteer's participation. The last effect is analysed here.

For this, from week 28, half of the sample was randomly assigned to receive a weekly SMS with a link providing access to the continuously updated web food price indicator dashboard. In contrast, the other half received a simple 'thank you' message. The intervention lasted seven weeks (week 28 to week 34). Again with this second intervention, the non-committal thank you message was sent to allow us to identify the impact of information disclosure independently of the impact of an SMS message. After week 34, the system continued to work without any further SMSs being sent. With this second nudge, the total sample consisted of all registered volunteers, who were randomly allocated to the intervention/treatment group or to the control group. We controlled whether the implemented randomisation process led to two equivalent sub-samples. We concluded that the randomisation process had indeed been successful. Only one of the analysed characteristics (age) was significantly different between the two sub-samples, at 10% significance level. Table 18 presents the description of the characteristics of the sample as a whole, and of the two sub-samples.

Table 18. Description of the characteristics of the sample for the information disclosure nudge.

Nudge 2: Disclosing collective data							
	Dashboard		Control		Full sample		*p*-value
Variable	n	%	n	%	n	%	
Total	372	100%	365	100%	737	100%	
Gender							0.84
Male	61	16%	51	14%	112	15%	
Female	311	84%	314	86%	625	85%	
Education							0.54

Nudge 2: Disclosing collective data							
	Dashboard		**Control**		**Full sample**		***p*-value**
Primary	1	0%	1	0%	2	0%	
Secondary	49	13%	52	14%	101	14%	
Tertiary	322	87%	312	85%	634	86%	
Food chain stage							0.41
Farmer	151	41%	167	46%	318	43%	
Final consumer	138	37%	119	33%	257	35%	
Retailer & wholesaler	41	11%	46	13%	87	12%	
Other	42	11%	33	9%	75	10%	
Preferred communication channel							0.37
app	31	8%	28	8%	59	8%	
email	67	18%	70	19%	137	19%	
SMS	236	63%	231	63%	467	63%	
web	13	3%	11	3%	24	3%	
unknown	25	7%	25	7%	50	7%	
	Mean		**Mean**		**Mean**		
Age	27.08		27.89		27.48		0.05
Avg. years smartphone use	6.5		6.6		6.5		0.8

Note: p-values are the results of Chi-Square tests of equality between categorical variables or from t-test of equality of means for continuous variables; *** p<0.01; ** p<0.05; * p<0.1

The outcome variable in the analysis of this intervention, like the previous nudge implementation, is the number of prices submitted weekly by each volunteer.

Table 19 describes the variables. Notice that in this analysis, we aimed to control for the case in which the volunteers had been recipients of nudge 1.

Table 19. Description of the variables and results of *t*-test at the baseline (pre-intervention) period of disclosure nudge.

Variable(s)	Description	Mean Control	Mean Treated	Diff.	|t|	Pr(|T|>t)
n_week_obs	The number of weekly prices submitted by a	29.569	28.778	-0.791	0.39	0.6945

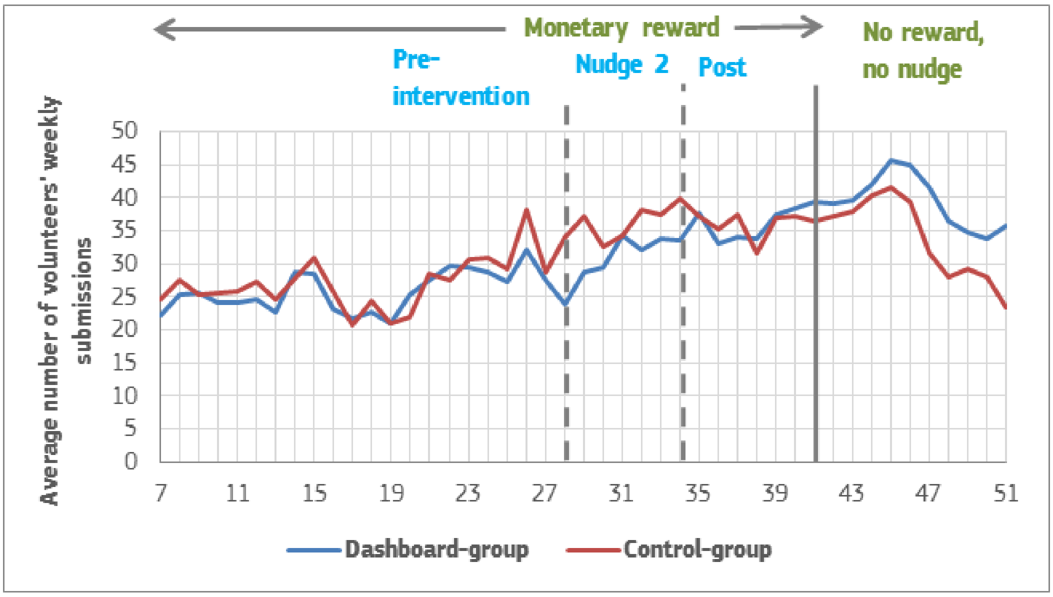
Variable(s)	Description	Mean Control	Mean Treated	Diff.	|t|	Pr(|T|>t)
	volunteer.					
%_week_valid_obs	The percentage of weekly valid prices on total prices submitted by a volunteer.	0.77	0.76	-0.01	0.58	0.5617
treated	=1 if the volunteer has been allocated to the nudged group, otherwise 0	0	1			
time	=1 if the week corresponds to the intervention period					
nudge1	=1 if the volunteers was receiver of the previous nudge	0.513	0.505	-0.008	0.23	0.8168
lag_reward	=1 if the volunteer has been rewarded in the current week, otherwise 0	0.426	0.397	-0.029	0.83	0.4043
gender	=1 if the volunteer is a male, otherwise 0.	0.954	0.888	-0.067	3.5	0.0005***
age	=1 if the volunteer is < 30 years old, otherwise 0	0.656	0.665	0.009	0.28	0.7767
high_educ	=1 if the volunteers has tertiary education, otherwise 0	0.878	0.895	0.017	0.77	0.4417
pref_comm	=1 if the volunteer prefers SMS as the communication way, otherwise 0	0.615	0.641	0.026	0.77	0.4386
farmer	=1 if the volunteer is registered as a farmer, otherwise 0.	0.607	0.51	-0.098	2.8	0.0052***
consumer	=1 if the volunteer is registered as a consumer, otherwise 0.	0.247	0.22	-0.027	0.92	0.3583
trader	=1 if the volunteer is registered as a trader, otherwise 0.	0.046	0.134	0.088	4.39	0.0000***
other	=1 if the volunteer is registered as other, otherwise 0.	0.099	0.136	0.037	1.62	0.1049
personal_motiv	=1 if the volunteers declared to have personal interest, otherwise 0.	0.459	0.409	-0.05	1.44	0.1509

Variable(s)	Description	Mean Control	Mean Treated	Diff.	|t|	Pr(|T|>t)
reward_motiv	=1 if the volunteers declared to have interest in the reward, otherwise 0.	0.38	0.344	-0.036	1.05	0.2925
data_motiv	=1 if the volunteers declared to be interested in the data and not in the rewards	0.956	0.92	-0.035	1.97	0.0488**
market_actor	=1 if the volunteer reports on actual transaction prices, otherwise 0.	0.35	0.39	0.04	1.22	0.22

*** p<0.01; ** p<0.05; * p<0.1

The trends in the number of price submissions (Figure 45) show that before the intervention (baseline), both groups had a similar average number of contributions and trends (confirming the parallel trend assumption). Following the intervention, we see an increasing trend in the number of price submissions for both groups, sustained in both groups throughout the implementation period. Once the intervention is over, the trend plateaus in both groups.

Figure 45. Average number of price submissions per week by group before, during and after the implementation of the social norm nudge



Results from DID analysis show no significant increase in the number of weekly price submissions of crowd participants in the treatment group compared to those in the control group from the period before the intervention to the period during the intervention (weeks 28 to 34), or the period after the intervention (weeks 35 to 38). Only if taking the post-intervention period starting once economic rewards stopped being paid, a higher number of weekly price submissions can be observed in the treatment group (Figure 45). However, the DID effect is not statistically significant.

Table 20. Difference-in difference estimation results for the implementation of the "information disclosure" intervention, before vs. period of intervention (left) and before vs. after period of intervention (right).

Outcome var.	Mean	S. Err.	|t|	P>t
Before				
Control	29.569			
Treated	28.778			
Diff (T-C)	-0.791	2.284	-0.35	0.729
After				
Control	36.105			
Treated	31.439			
Diff (T-C)	-4.666	2.426	1.92	0.055**
Diff-in-Diff estimator	**-3.874**	**3.332**	**1.16**	**0.245**

Outcome var.	Mean	S. Err.	|t|	P>t
Before				
Control	29.569			
Treated	28.778			
Diff (T-C)	-0.791	2.169	-0.32	0.752
After				
Control	34.735			
Treated	34.269			
Diff (T-C)	-0.466	3.06	0.15	0.879
Diff-in-Diff estimator	**0.219**	**3.751**	**0.06**	**0.953**

*** $p<0.01$; ** $p<0.05$; * $p<0.1$

For robustness, we also tested the DID estimation with covariates. As well as the reception of economic rewards (resulting in positive and statistically significant at 1%), and gender and age of the participants, we need to control for whether they were recipients of the previously implemented social norm nudge (resulting in positive and statistically significant at 1%). The DID estimator remains statistically non-significant, implying that the intervention consisting in disclosing information did not have a positive effect on the treated group compared to the control group. Some explanations for this result are the following:

1) Connectivity issues or technological barriers (35) that prevent the crowdsourcing user from accessing the web dashboard, which therefore did not help them to understand its value for themselves or for the general public. In fact, in a survey run at registration, about the initiative the majority of volunteers of the crowd indicated SMS as the best way of communicating information;

2) individuals do not know how to use the tool, or the way and type of information provided is not useful;

3) existence of spillover effects in the case of interaction between participants that have received the link and not, whereby the first group (treatment group) share the dashboard link with the second group (the control group). These, however, we assume that can be discarded since, during the intervention, we found no evidence that participants of the control group accessed the web dashboard; and

4) the period of implementation of the nudge was too short. It can be concluded that sharing the link to a web dashboard is not more effective than a 'thank you' message in promoting participation. Moreover, investigations could be carried out to find out the most appropriate format/tool and type of information to be disseminated and whether the use of the web dashboard as a price information system could have some impact on the transacted/observed prices. Indeed a final test confirmed that the volunteers providing data that are market actors (make a buying or selling transaction) and received the link to the dashboard contributed significantly more than the market observers.

8.3 SUMMARISING RESULTS OF THE MONETARY AND NON-MONETARY INCENTIVES

Despite the lack of a clear experimental design to test this, it seems that the expectation of receiving an economic reward does drive participation in the crowdsourcing data supply exercise through extrinsic or/and intrinsic (gamification) factors. Indeed the results of the DID regressions show that when including

(35) In countries with poor Internet connections or expensive data plans the use of Opera Mini browser is quite extended, which is a very compact mobile browser whose main function is to compress all traffic in order to save megabytes from data plans and for loading web pages faster. Unfortunately it is also not compatible with the latest web technologies like WebSockets, which is used by the underlying data analysis solution behind the dashboard (Qlik Sense). Also accesses to the web dashboard using Virtual Private Networks (VPNs) depend on their quality, which can affect the user experience.

behavioural nudges, the economic reward continues to have a positive influence (and statistically significant at 1%) on volunteer data submission, hence, complementing the effect of the nudge.

Moreover, our results show that the well-known 'social norm' nudge has a significant positive effect on the number of price submissions. The effect occurs both during implementation, as well as after the removal of the nudge. By contrast, the 'information disclosure' nudge appears not to have an effect. This nudge can be effective only if the information disclosed is perceived as relevant by the participants for their own use or for the community. This might require more time and additional awareness campaigns as well as interaction with the crowd about the type of information and dissemination formats. Besides, internet connectivity problems, lack of knowledge of how to use the tool or lack of interest in the disseminated data may have prevented participants from opening and using the dashboard and may explain the ineffectiveness of the disclosure nudge.

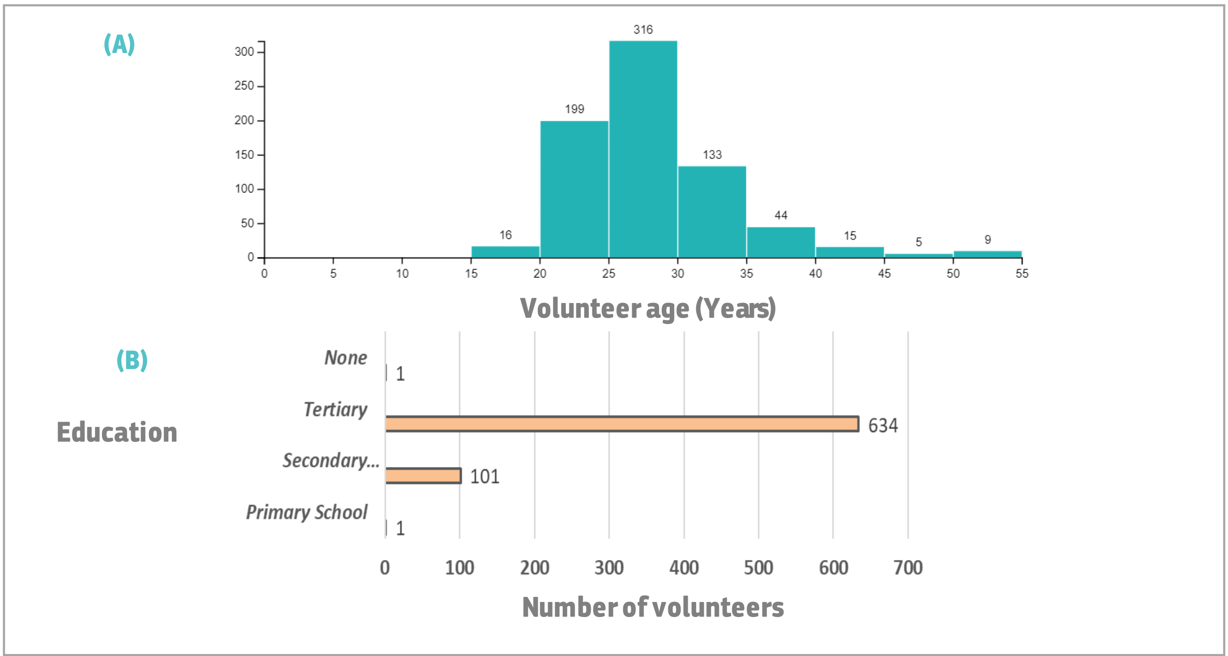
9 Data insights

The crowdsourcing system was successfully set up by enlisting volunteers who submitted price data for the target commodities daily. The actual enlisting of volunteers commenced in the second week of September 2018, and continued until the end of January 2019 ([36]), after ~1 300 prospective applicants had submitted their profile information ([37]). During the initial pilot phase, the first 200 volunteers who completed the profile form were fully onboarded, and an additional 537 were onboarded during the full roll-out phase, all bearing unique and verified information. Based on the implementation structure, actual data submission commenced during the last week of September 2018 and continued until June 2019. As of the last day of June, ~19 900 data records had been submitted, with each data recordcontaining food price data for all the target commodities. In total, over 160 000 individual food price points had been submitted by the volunteer crowds over the course of months.

9.1 CROWD PROFILE

The overall enlisted crowd was composed of 625 (85%) male and 112 (15%) female volunteers. Most (77%) of the volunteers were between the ages of 20-35 years, and nearly all have attained tertiary or secondary school education (Figure 46). Breakdown of the crowd by profession shows that majority (77%) of the volunteers are either students or directly engaged in the agricultural sector (Figure 47). Generally, 9 out of every 10 volunteers has acquired more than four years' experience with using smartphones. These metrics are indicative of the demographics of the region. For instance, the gender composition (and disparity) of the crowd may have been influenced by the cultural male dominance and lower female literacy levels in the region. The dominance of the crowd by students and workers in agricultural professions is likely due to their understanding of the relevance of the initiative. This fairly mirrors national demographics, which indicate that the largest proportion of the national labour force (~30%) are employed in agriculture and related sectors (NBS, 2010). On a general note, the success of crowdsourcing can depend on the significance of the crowdsourcing theme to the potential crowd. Therefore, the dominant representation of students and professionals in agriculture may be an indication that the participants are partly or fully motivated by the relevance of the crowdsourcing theme to their professional/personal interest (sell or buy at the best prices). It may also be that in the case of students, the monetary reward is an essential source of motivation.

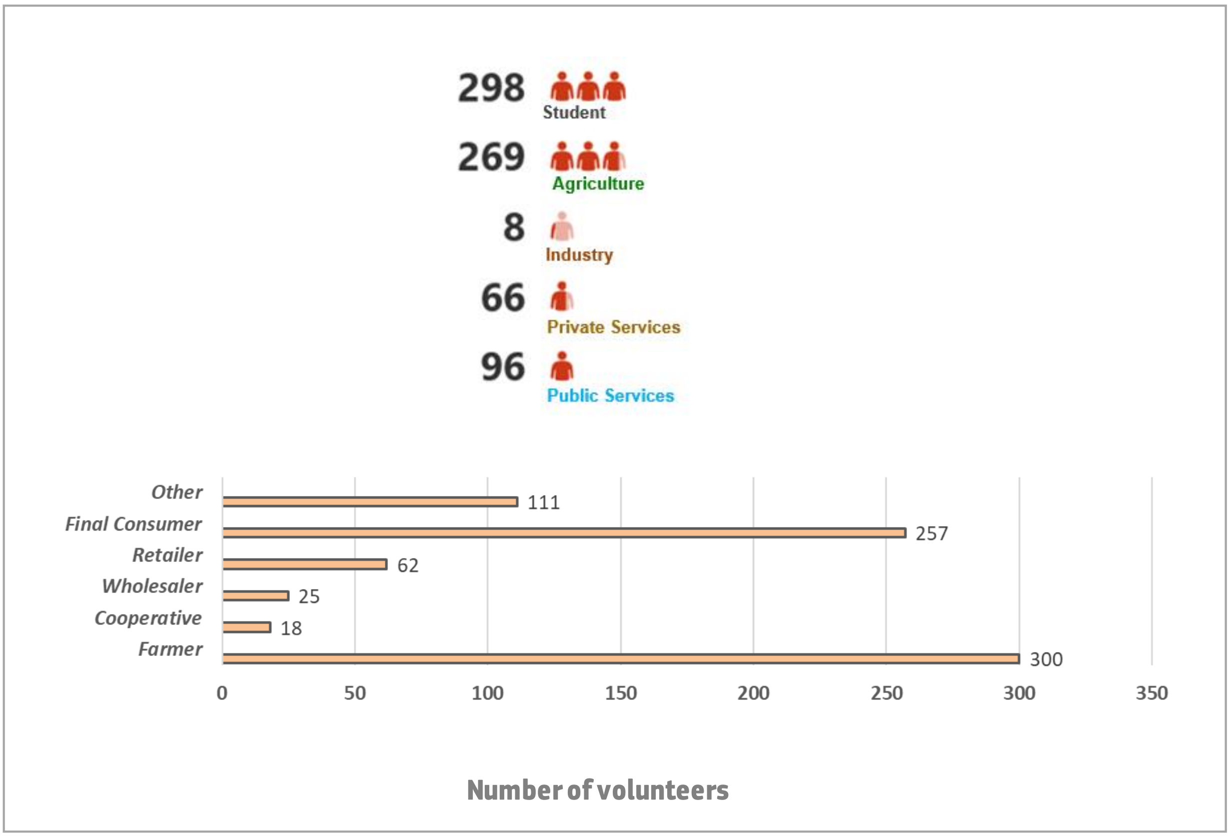
Figure 46. Age (A) and education level (B) of the enlisted volunteer crowd, based on profile



([36]) An end date for new registrations was fixed because the duration of the project was limited to eight months of data collection, and the workload of registering new volunteers was initially unknown. However, the experience in this project has brought enough information to automate this process so that it could be done continuously in real time.

([37]) This contained many duplicate and incomplete submissions, so they were thoroughly checked to ensure that unique IDs were assigned per phone number and volunteer entity.

Figure 47. Composition of the crowd by profession (upper) and market niche represented (lower)



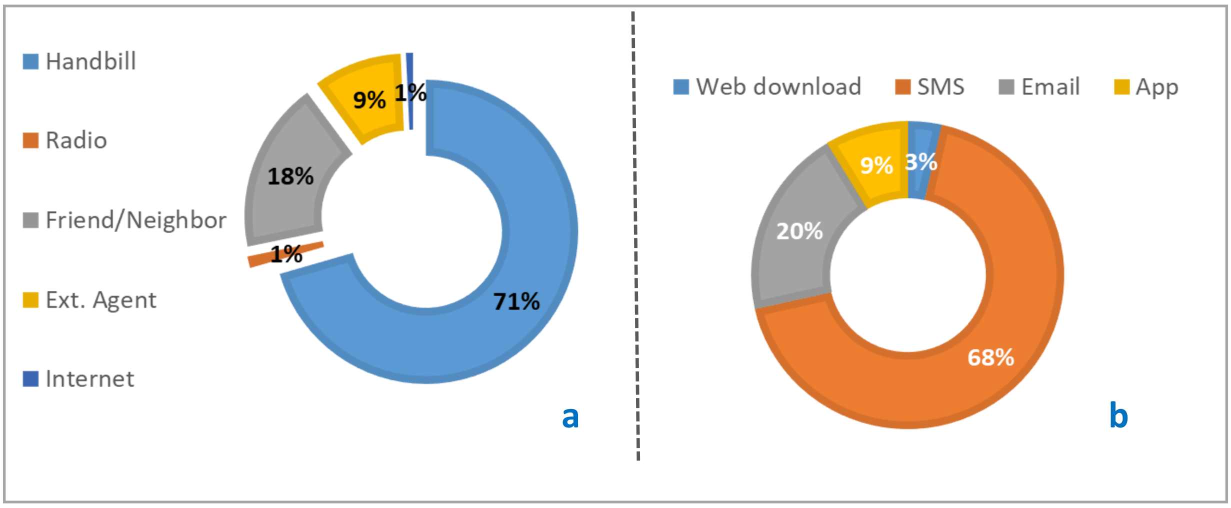
9.2 CROWD MOTIVATION

The initial data collected from the prospective volunteers was also useful in identifying the most promising opportunity for engaging and retaining the enlisted volunteers. This understanding of the preferences of the crowd provided an objective basis for decisions regarding rewards, communication and behavioural nudges. For instance, approximately 60% of the prospective volunteers indicated that they are participating in the initiative for professional reasons, while a lesser percentage (~40%) indicated that that they are motivated by monetary gain. Similarly, 8 out of 10 volunteers indicated that they are interested in receiving food price data and that they are willing to participate in the initiative without monetary rewards. Most of the volunteers (80%) indicated a preference for bank transfer to receive any reward.

9.3 CROWD ENGAGEMENT PREFERENCE

The outcome of enlisting, engaging, and retaining volunteers is likely influenced by the approach/mode of communication. The direct and indirect communication pathways which were adopted to disseminate information about the project and invite prospective volunteers had strongly contrasting outcomes. Seventy percent (70%) of the volunteers indicated that they heard about the initiative through flyers, while barely 1% (9 volunteers) heard through radio (Figure 48a). This contradicts the initial expectation that radio-disseminated adverts would enable coverage of a wider area, based on the assumption that prospective volunteers in the region listen to radio. Also, the production of the advert and cost of airtime dwarfs the cost of designing and producing flyers, yet the impact is negligible in terms of reaching the target audience. There is clear indication that SMS-based messaging is the preferred method for receiving information or updates, and this is likely due to the non-invasiveness of text messages and also due to the fact that SMSs are simpler, shorter and do not require Internet access. Besides SMS is a well-established and well-know technology. At least 6 out of 10 volunteers indicated that they prefer to receive food price data through SMS (Figure 48b).

Figure 48. Communication preference of volunteers of the crowd in Kano and Katsina States, as indicated by the initial source of information (a) and preferred means of receiving food price data (b).



9.4 VOLUNTEER DATA SUBMISSION TREND

At the end of the project period, the number of data records submitted exceeds the initial target of ~7 000 submissions ([38]). Overall, the average submission rate was 70 records/day, with maximum submissions attained shortly after the pilot phase; i.e. after more volunteers were added to the pool for the full roll-out phase (Figure 49). Disaggregating the rate of submission per major commodity type (Figure 50) shows that the daily data points for local commodities were comparable, while Indian rice was less reported. Overall, the most submissions based on "day of the week" were observed on Friday, however, the distribution showed that submission rates were not dramatically different across different days of the week (Figure 51). This suggests that the system set-up was sufficiently optimal to avoid potential bias in the crowd submission of data relative to the day of the week. The weekly temporal progression of data submission indicates that volunteer participation was stable ([39]) during the pilot phase, soared for ~8 weeks at the onset of the roll-out phase, and sustained a sinusoidal pattern afterwards, with upward trends likely triggered by the nudges or economic incentives, or a combination of both (Figure 52). The undulating pattern is likely due to a combination of factors, including: i. the SMS-based behavioural nudges, which may have induced the volunteers to improve their participation, and ii. the extended election period in Nigeria, which may have limited the motivation of volunteers to visit markets for survey of commodity prices.

([38]) The initial target of 7 000 submissions actually far exceeds the requirements of the EC-JRC, per the original technical specification, which stipulated a benchmark of ~1 000 data points (on a weekly submission basis).

([39]) Further check of the actual submission IDs may reveal further information about the volunteer consistency.

Figure 49. Daily frequency of food price data submission by volunteers of the crowd during the entire period (Sept. 2018 – June, 2019). The orange-highlighted region indicates the pilot phase, while the blue-highlighted region represents the full roll-out phase.

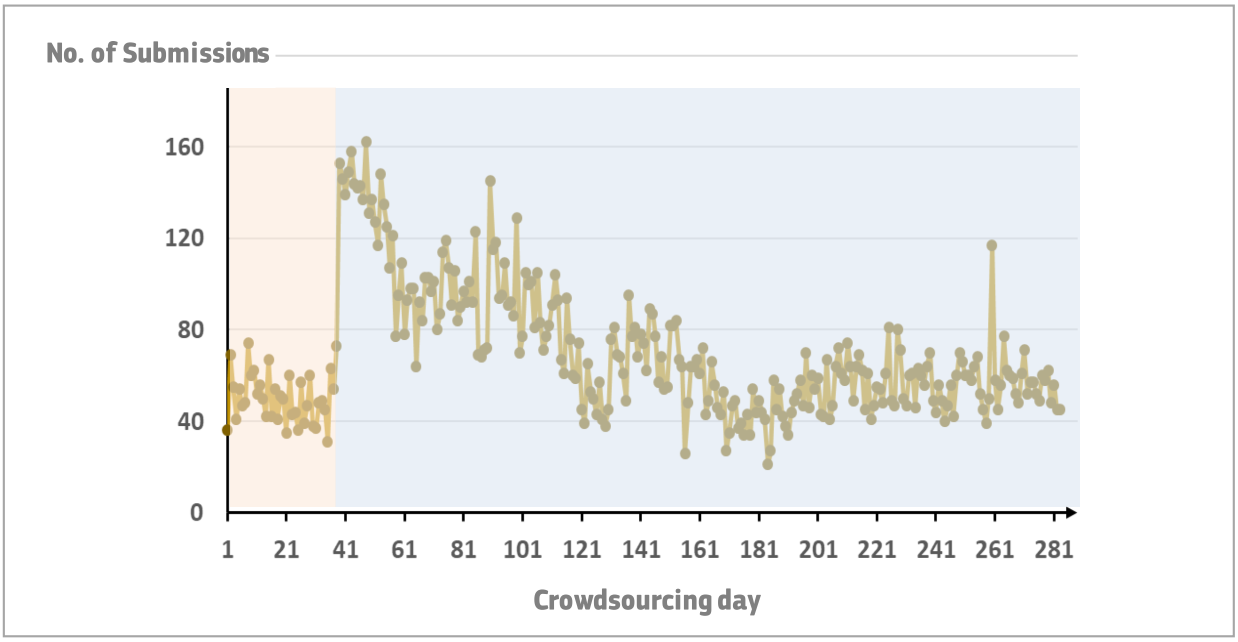


Figure 50. Daily frequency of food price data submission by the volunteer crowd during the entire period (Sept. 2018 – June, 2019), disaggregated by commodity

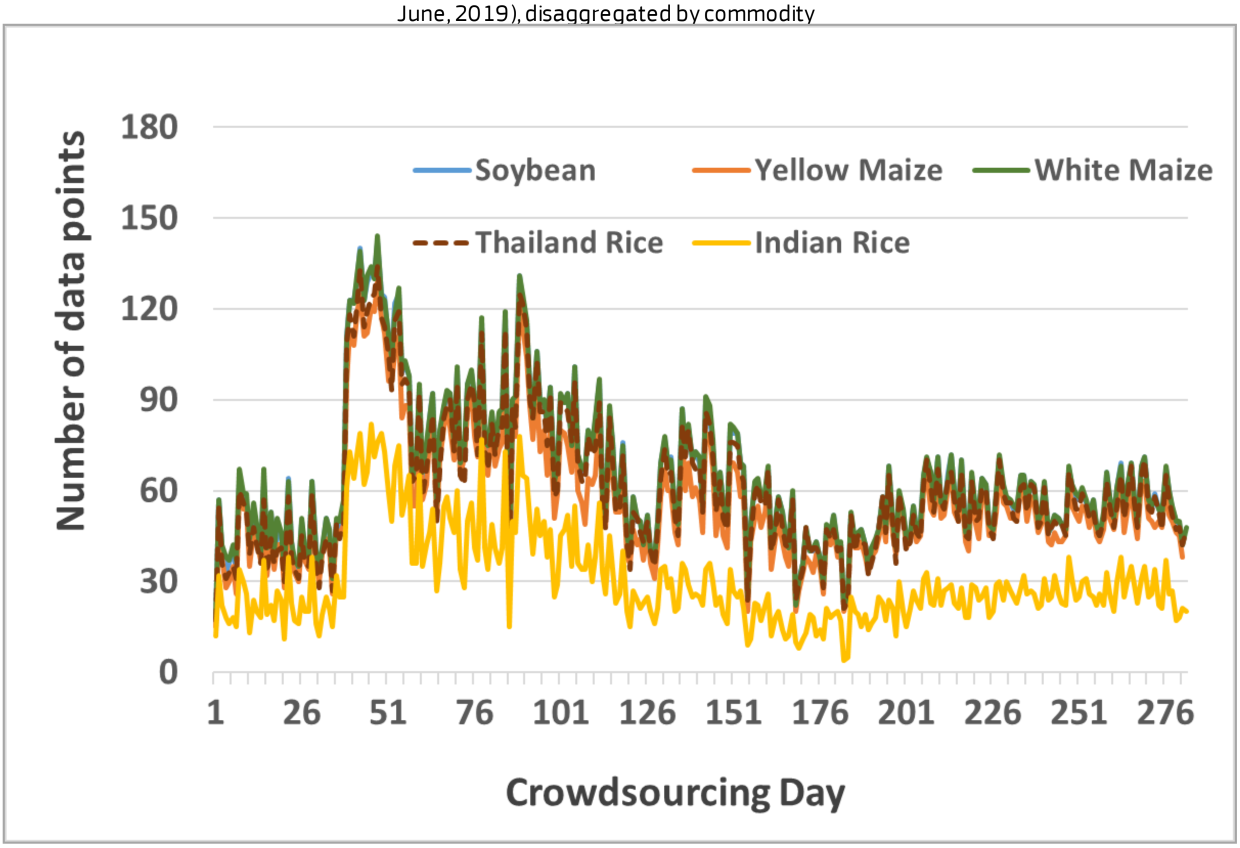


Figure 51. Frequency of food price data submission by volunteers for each day of the week across the entire period

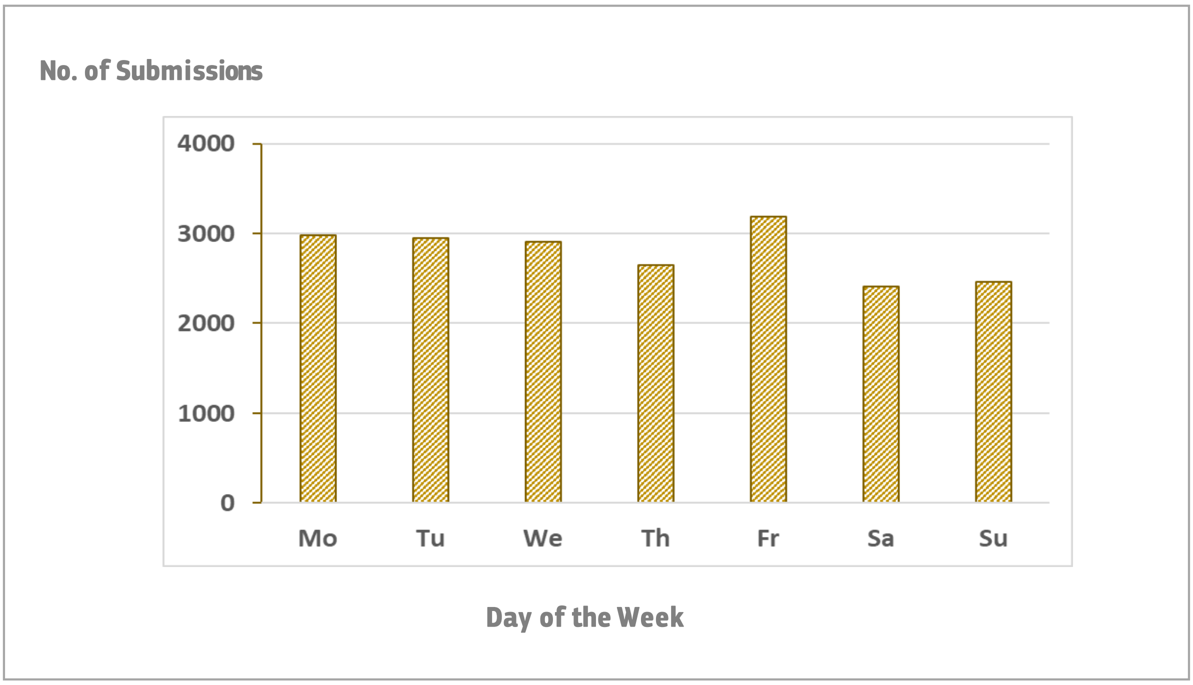


Figure 52. Weekly frequency of food price data submission by the volunteer crowd across the entire period

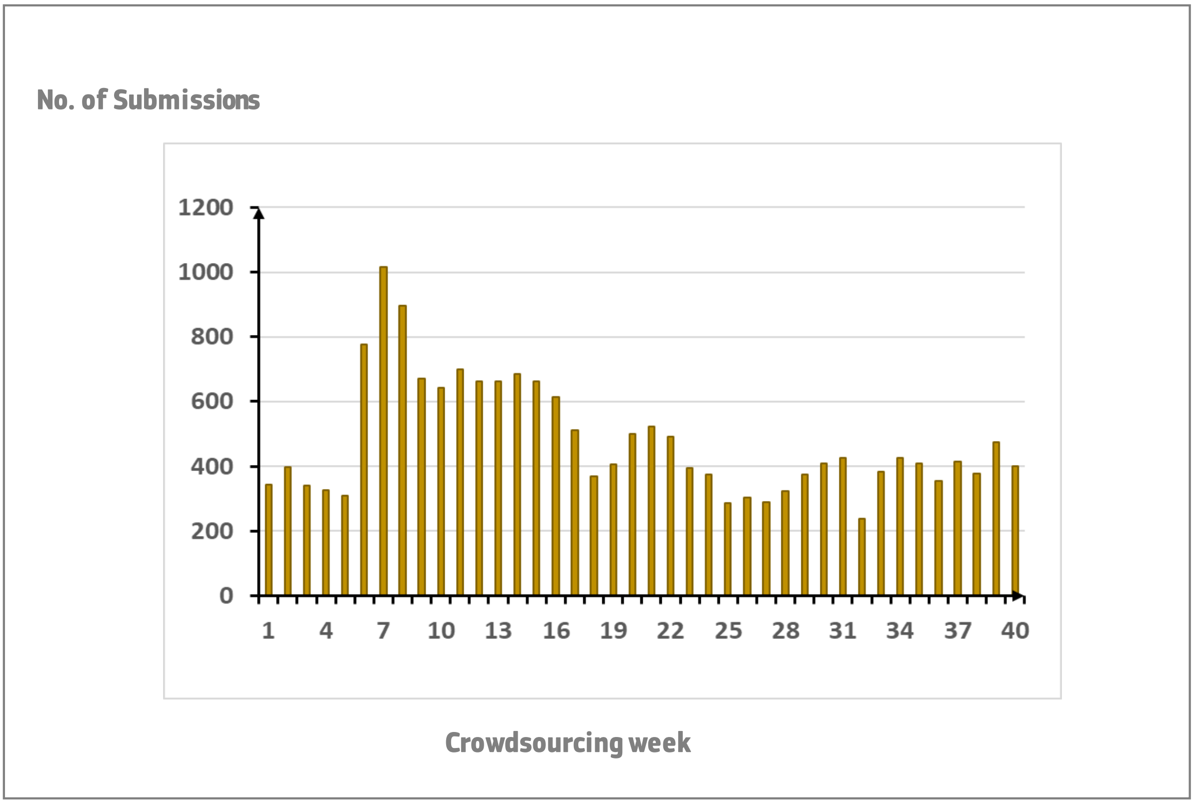
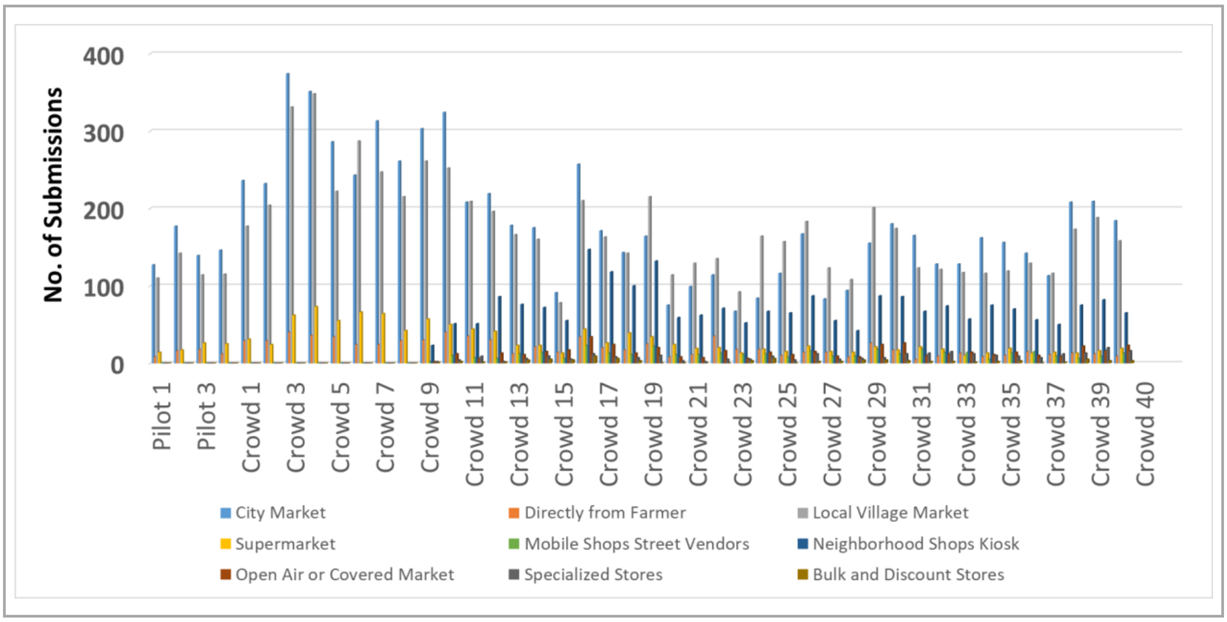


Figure 53. Weekly frequency of food price data submission disaggregated by market segments across the entire period, including the Pilot Phase (Pilot 1-4) and full roll-out weeks (Crowd 1-40).



9.5 COMMODITY FOOD PRICE TRENDS

The average weekly data showed that price of locally produced commodities (Soybeans, Beans, and Maize) slightly increased at the onset of the data collection period (mainly during the pilot phase), which was close to the period of harvest in the region. However, there was a subsequent decline in the prices of these locally produced commodities after the harvest period (from Week 3, i.e. around mid-October). Contrastingly, prices of imported Indian and Thailand rice were relatively stable, despite bi-weekly/monthly market swings. The consistent flow of data which reveals the nuances of commodity price data before and after harvest demonstrates the potential reliability of the crowdsourcing system (Figure 54). Maize showed the most visible price changes, with major declines observed shortly after the harvest period, prior to the sowing period. The initial major decline is likely associated with the supply of new grains, while later decline prior to the sowing period may be due to market-related dynamics, where traders typically store maize grain for 4-6 months after harvest in expectation of higher prices. Quite often, a large sell-off occurs prior to the sowing seasons that traders (and farmers) who had stored grains would have money to buy farm inputs for the next sowing season. Such sell-offs are suspected to have triggered a downward swing in market prices. The commodity prices reported within each market segment showed a hierarchical evolution of prices from the farm gate (i.e. directly from farmer) to consumer (i.e. urban supermarket), notwithstanding minor artefacts of outlying data points (Figure 55).

Figure 54. Weekly average of commodity prices submitted by the crowd during the FPCA project implementation phase

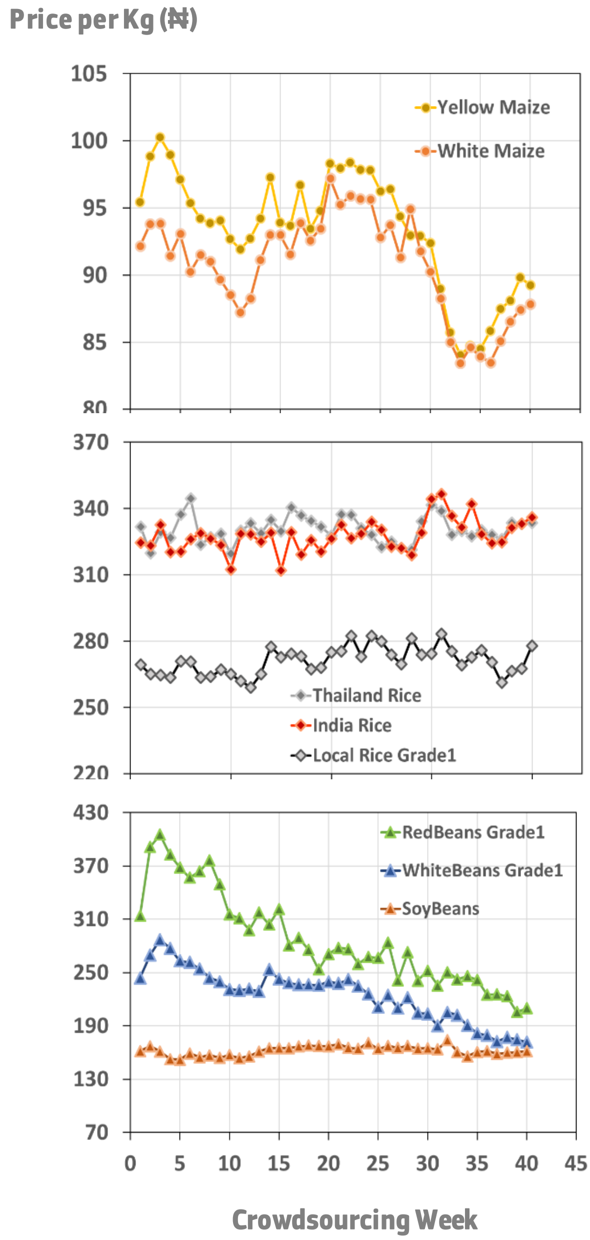
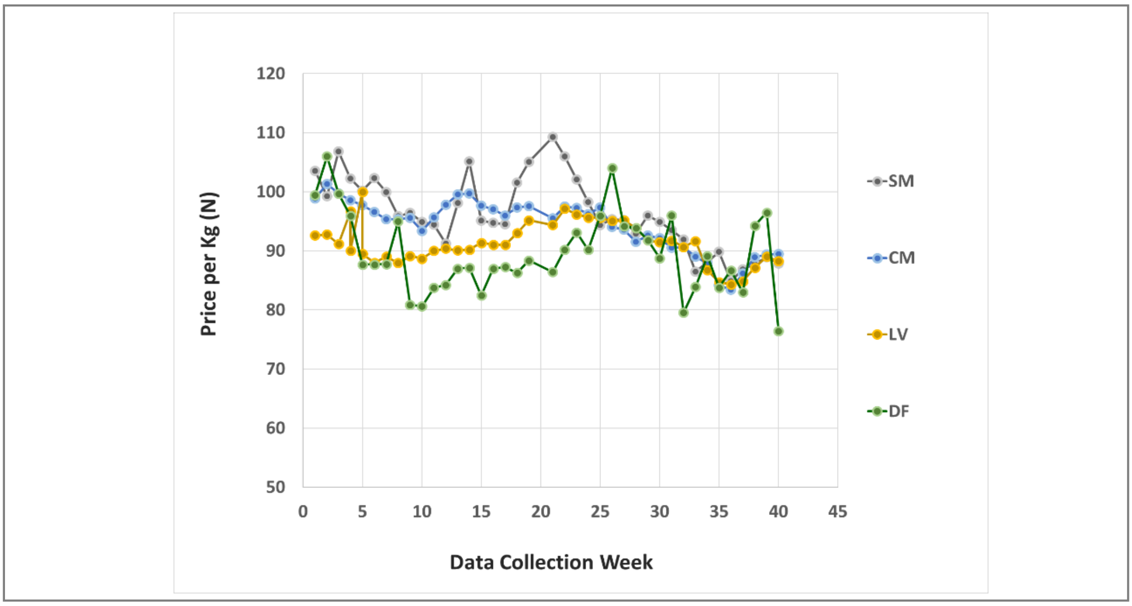


Figure 55. Average weekly crowdsourced price of Yellow Maize at various market segments



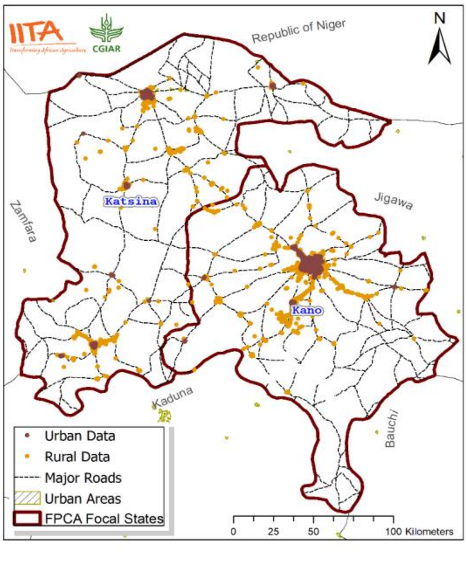
Note: DF: Directly from Farmer; LV: Local Village; CM: City Market; SM: Supermarket.

9.6 SPATIAL ANALYSES OF COMMODITY PRICES

9.6.1 Spatial coverage of commodity price data

In addition to the temporal richness of the crowdsourced data, we also achieved a measure of balance in the spatial coverage of rural vs. urban areas. As shown in Figure 56, major clusters of price submission occurred within the city limits, while the data submitted from rural areas is more dispersed and often along major roads, where most rural trades occur. Despite the difference in the rural-urban spatial dispersion pattern of submitted prices (depending on scale), the number of data records submitted in rural areas (outside city limits) was comparable to those submitted in urban areas (within city limits), with an almost equal split between the two.

Figure 56. Rural vs. Urban distribution of food price data submitted during the crowdsourcing period in Kano and Katsina states of Nigeria



9.6.2 Inter-temporal dynamics of price variability and spatial patterns

The last part of this section relates explicitly to spatial analysis and to the dynamics of crowdsourced prices, with the explicit aim of analysing the inter-temporal dynamics of price variability and spatial patterns. The method used is more descriptive than an explicit modelling strategy, since we illustrate the dynamics of price and its geo-spatial diffusion using static and dynamic maps.

The spatial analysis concentrated particularly on monthly data, and on one single product (namely: Rice local) for the price type retail.

For the sake of illustrating the potential of the proposed methodology, the following visual outputs were produced:

— Choropleth maps of the crowdsourced prices at LGA level

— Interpolated map of crowdsourced prices

Results are reported in Figures 57 to 62.

The choropleth maps are reported in Figure 57 and Figure 58. Our choropleth maps use differences in shading within predefined areas, the LGAs, to indicate the average values of prices for a particular commodity in those areas. Figure 57 and Figure 58 report on the retail price of local rice in Kano and Katsina States respectively. In both States, lighter shading in the months January to April (post-harvest) reflects lower prices. From then in May and June (planting) the shading darkens indicating higher prices. However, we note that crowdsourcing prices do not cover all areas (LGAs), so we will use a spatial interpolation technique presented below in Figure 59 to Figure 62.

We apply spatial interpolation to estimate values at unknown points using points with known values. For example, it is useful to make a map for the price of a certain commodity for a State, when we do not have enough evenly spread data points that cover the entire region. Figure 59 and Figure 62 present the interpolated maps of crowdsourced prices for the retail price of local rice in Kano and Katsina States respectively. The maps show a price evolution similar to that shown by the choropleth maps.

In addition, the interpolated maps allow us to visualize the dynamics of price diffusion in space. Figure 60 shows that the highest values for the price of local rice in Kano were concentrated in the Gwarko LGA in October 2018, but that in November this concentration moves to the LGA close to the capital city of Kano, Warama, then to Kiru in December and finally stays between Kiru, Warama and to Wuchi LGAs until June 2019. Similarly, Figure 61 shows that the highest prices for local rice in Katsina were concentrated in the capital city and in the LGAs of Bindawa in October 2018, and then moved to the LGA of Gkaranchi in November, to the capital city and the LGA of Mai'adua in December, and then after a period of lower prices high prices are concentrated in Dutsin-Ma from April to June 2019.

Figure 57. Choropleth map of crowdsourced retail prices for local rice in Kano (10/2018-6/2019)

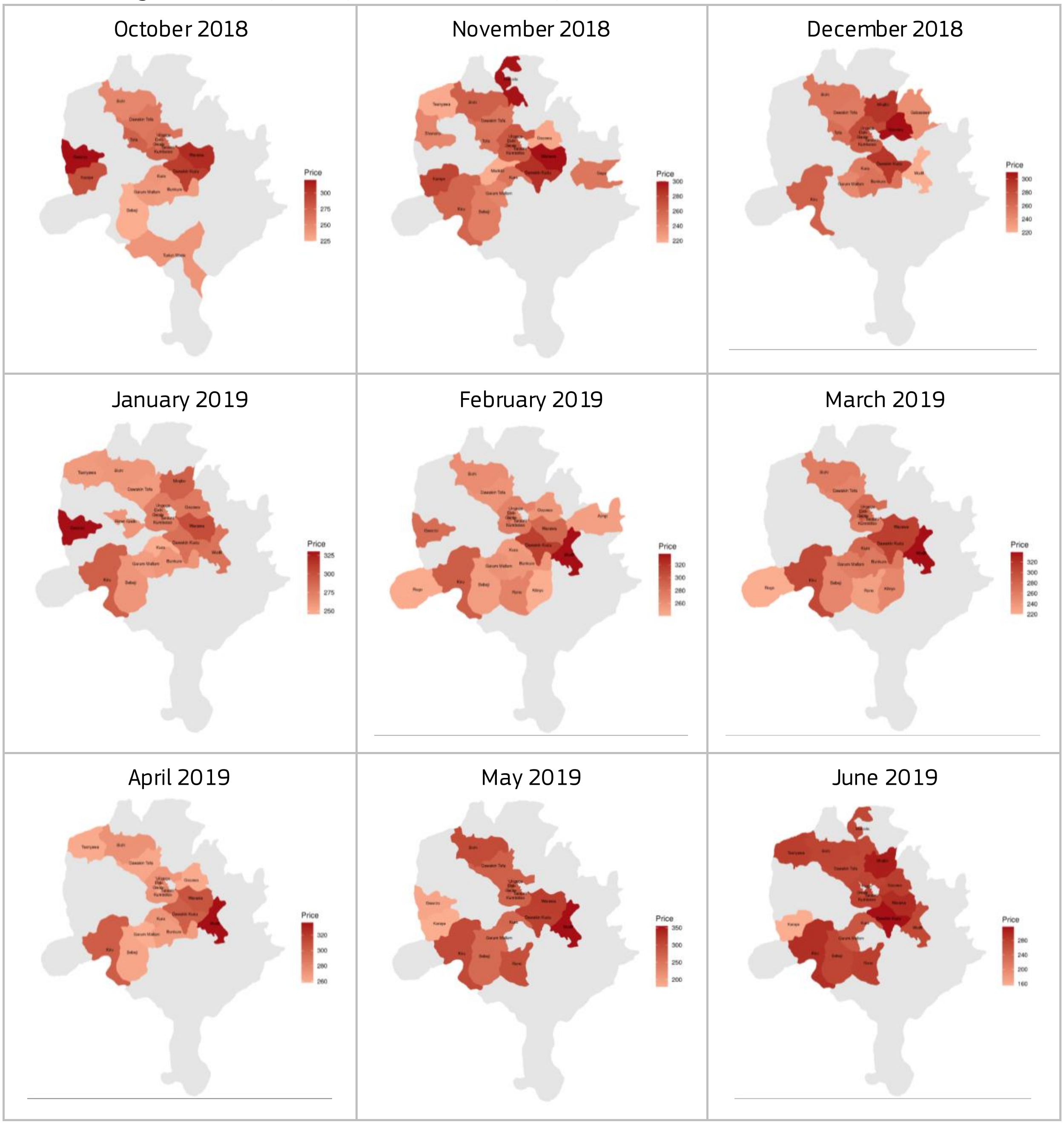


Figure 58. Choropleth map of crowdsourced retail prices for local rice in Katsina (10/2018-4/2019)

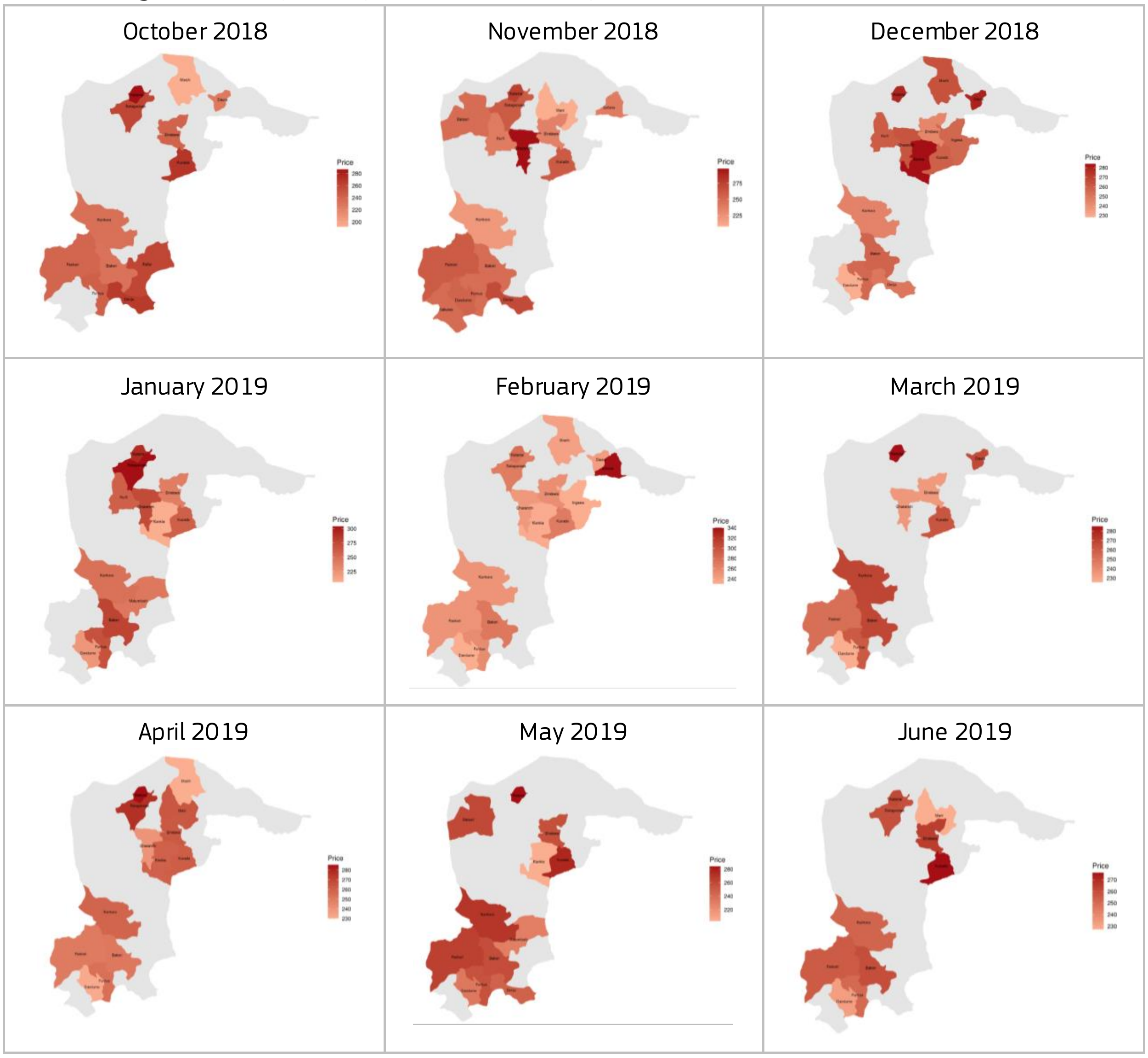


Figure 59. Interpolated map of crowdsourced prices for local rice in Kano (10/2018-12/2018)

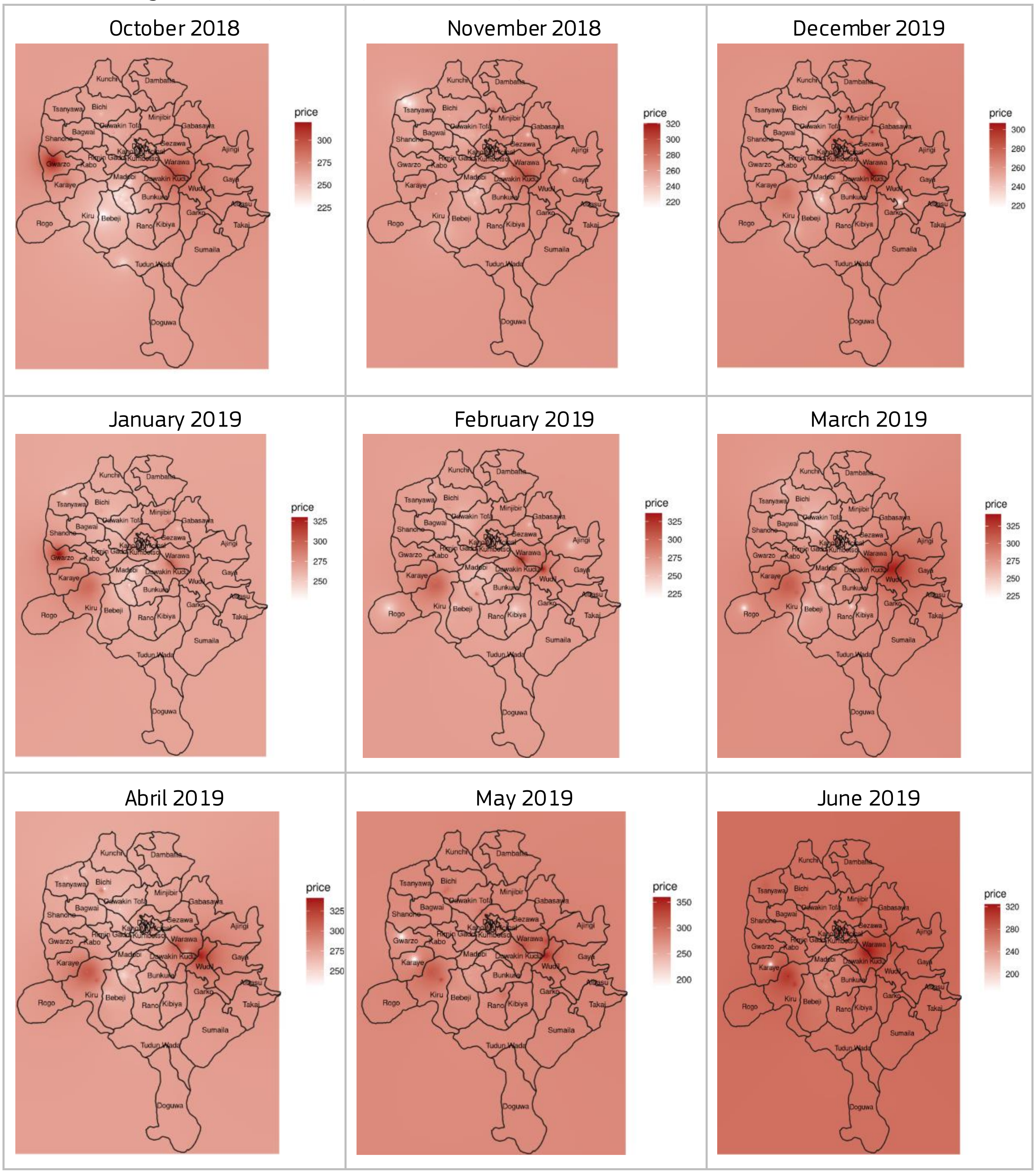


Figure 60. Diffusion of high values for local rice retail prices in Kano in the four months from Oct 2018 to Jun 2019

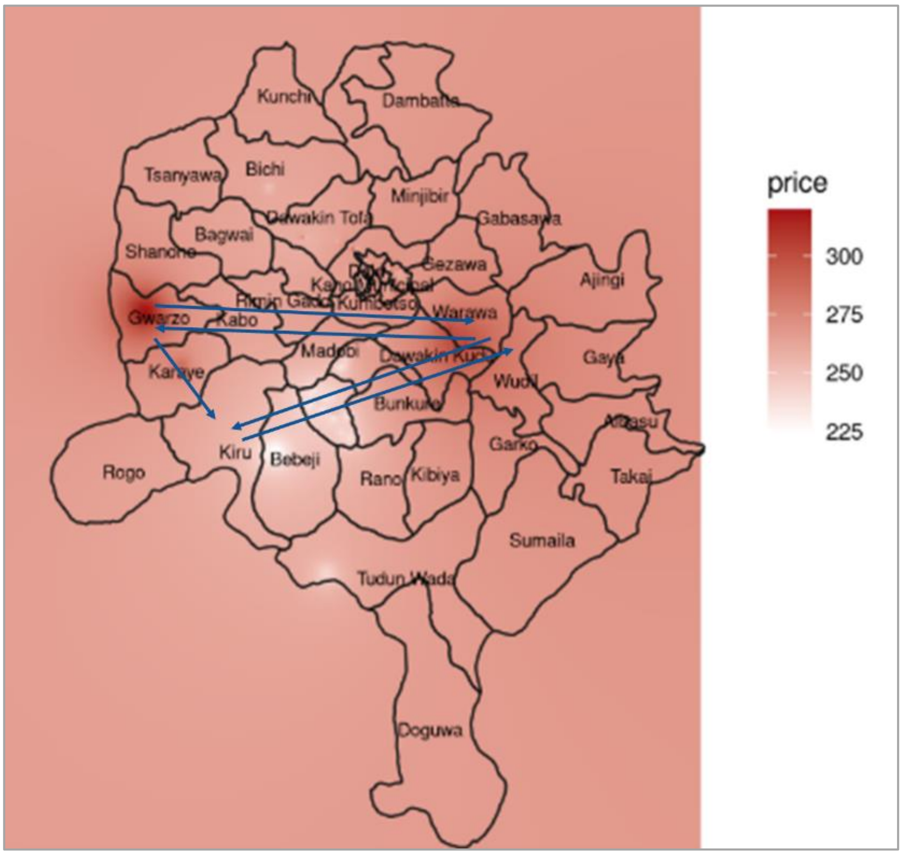


Figure 61. Diffusion of high values for local rice in Katsina in the four months from October 2018 to June 2018

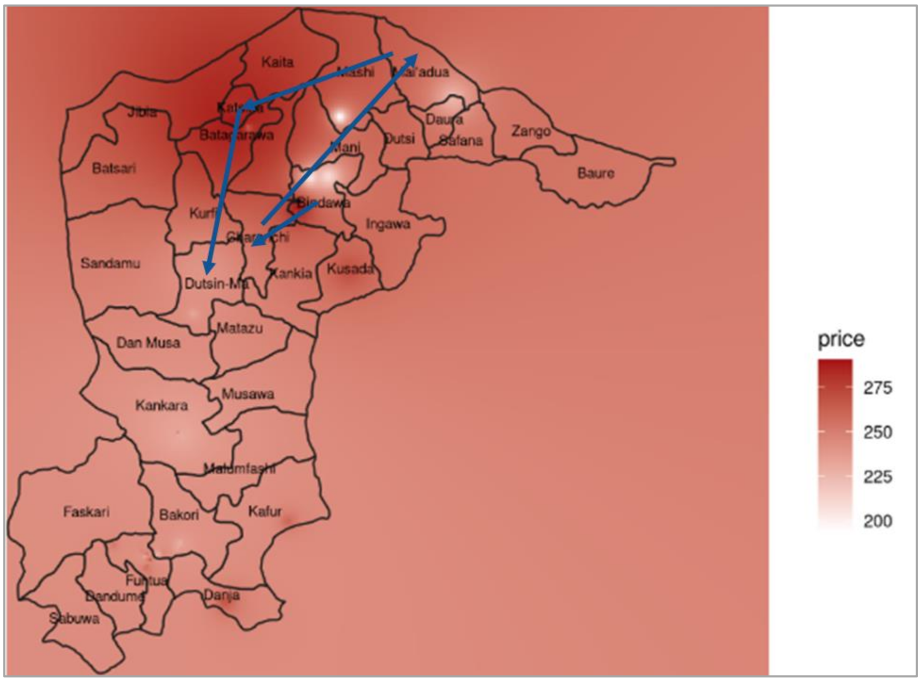
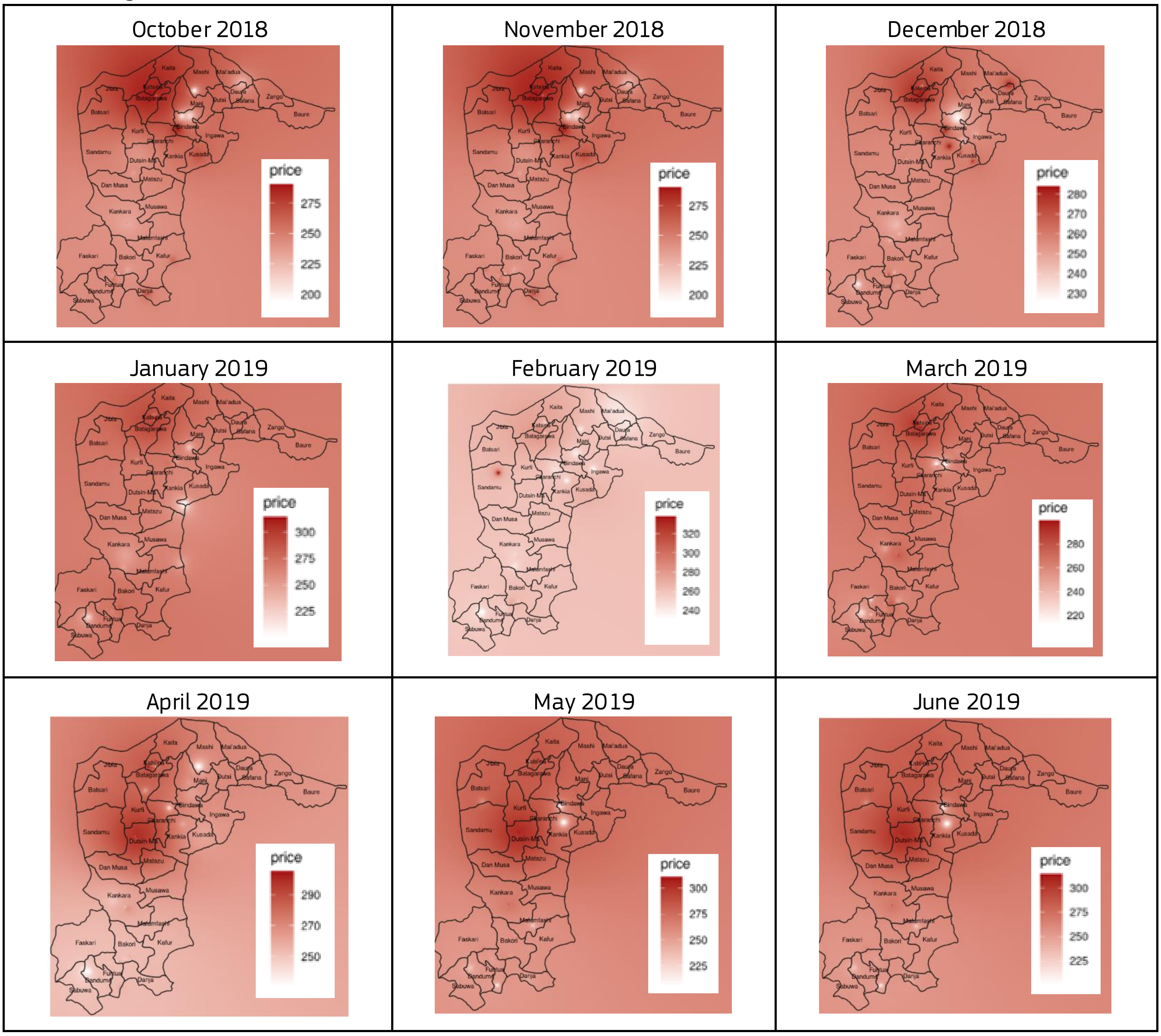


Figure 62. Interpolated map of crowdsourced retail prices for local rice in Katsina (9/2018-12/2018)



Figures 57 to 62 show the potential of a sophisticated spatial analysis of the crowdsourced data collection with the aim of monitoring the space-time dynamics of price changes and anticipating local price trends and possible food crises. Although this task was not undertaken in the current project, there is indeed an extensive literature in spatial econometrics that suggests the use of ad hoc models to account for the space-time diffusion of price changes (see e. g. Arbia, 2014).

9.6.3 Local spatial autocorrelation statistics

Figure 63 reports the Moran's local maps and the Local Indicator of Spatial Association (LISA) indices (Anselin, 1995). LISA indices estimate the spatial autocorrelation between areas/points and their neighboring areas/points and allow for the decomposition of global indicators of spatial autocorrelation (i.e. Moran's I), into the contribution of each observation. LISA indices may be interpreted as indicators of hot or cold spots or may serve to identify spatial outliers (Anselin, 1995).

A potential use of this spatial analysis is to examine spatial patterns of food market prices and their changes. This is an important application for policy and decision making, as the linkage between geography and food market prices and their changes could be used for planning food security or market interventions.

Relevant information on geographical aspects and dynamics of market prices can be captured by associating prices with their geographical location. The LISA analysis serves to produce a spatial layer showing spatial associations by associating a certain attribute data (e.g. food prices) with locational information. Our crowdsourced data are a collection of points described by their geographic coordinates (i.e. latitudes and longitudes) recorded automatically during the data submission, which correspond to a exact location in the surface of the earth. In our study, a high-high point association of the LISA statistics suggest that the point price is higher than the average of the whole study area and so are its neighbors (i.e. a hot spot). A low-low point association suggest tha the point price and its neighbours are lower than the average (i.e. a cold spot) A high-low point association indicates that the point price is higher than its neighbours and a low-high association indicates that the point price is lower than its neighbours.

In these maps (Figure 63), the red points refer to locations where we observe a high value for price surrounded by high values (marked as HH), the blue colour represent points where we observe a low value for price surrounded by low values (marked as LL), the orange points refer to observations where we observe a low value for price surrounded by high values (marked as LH). Finally, the light blue points represent points where we observe high values for price surrounded by low values (marked as HL).

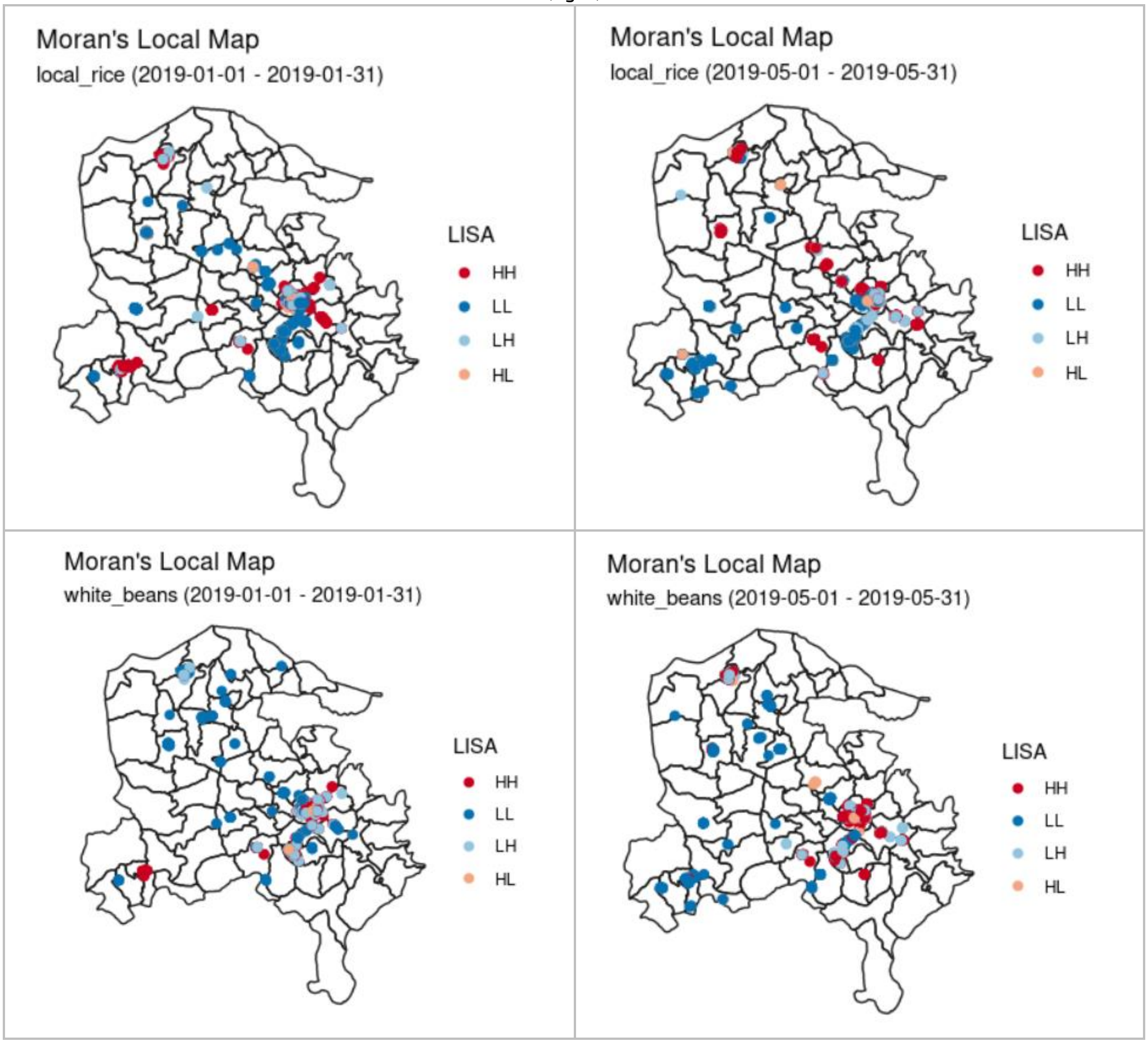
While HH and LL points reveal the presence of positive spatial correlations between prices, the LH and HL points highlight the presence of negative spatial correlation thus suggesting a formal way of identifying spatial "outliers" or anomalous values.

Box 14. Interpretation of the Local Indicators of Spatial Association (LISA) indices.

Inspecting the maps for two different months shows that prices of local rice are mostly clustered in either hot (mainly in the cities Kano, Katsina and Funtua) or cold spots. The presence of LH potential spatial outliers is limited to capital cities, and HL potential spatial outliers are quasi-inexistent.

Hot and cold spots are also observable in the case of white beans, but with a lesser presence of hot spots. The presence of LH potential spatial outliers is also limited to capital cities, and HL potential spatial outliers are also only a few.

Figure 63. Detection of hot and cold spots and spatial outliers through the calculation of Local Indicators of Spatial Association (LISA) for retail prices of local rice (above) and white beans (below) in January 2019 (left) and May 2019 (right)



10 Challenges and conclusions

Food market price information is not only an essential input for food security early warnings systems, but also a pre-requisite for the proper functioning of the market. Information plays a critical role in building inclusive markets that provide incentives for all market actors to participate. For this, it is crucial to explore cost-efficient data retrieval solutions that provide relevant and reliable data for consumers, producers governments and other organisations to make decisions, and that can serve as a public good by creating information feedback loops to market actors.

In this final section, we review the main characteristics of the FPCA approach, the performance under real-life implementation and best practices that can be considered when designing and rolling out a smartphone-based crowdsourcing initiative to collect and disseminate food prices. We focus on five main aspects including (i) the composition of the crowd, with attention to possible biases related to voluntary participation (such as self-selection bias); (ii) the trade-off between the convenience of using an existing IT platform and the need to cover the particularities of the crowdsourcing approach to be implemented; (iii) the system of incentives (both monetary and non-monetary) used to promote data contributions; (iv) the challenge of ensuring, monitoring and evaluating the quality of the final dataset in crowdsourcing as well as the challenge of processing a very large number of data; (v) the challenges of data dissemination and developing a web dashboard for data visualisation, including metrics to monitor price changes and trends. Finally, we draw some conclusions regarding the overall effort of obtaining crowdsourced food price data in the context of a developing country.

10.1 THE COMPOSITION OF THE CROWD

One of the main characteristics of crowdsourcing exercises is the importance of the crowd composition, both in quantitative and qualitative terms. The crowd is a self-selecting entity that is created by engaging individuals through outreach activities. This means that the crowd might suffer from self-selection bias, which may not be representative of all intended participants. In our case, the crowd was initially engaged via leaflets and radio commercials, then the word of mouth (e.g. through networks of family, friends, and neighbors) played an important role in the diffusion of the initiative.

In an ideal world, one would like to enrol only crowd individuals capable of minimising intentional and unintentional errors when submitting data. For this, volunteers of the crowd require moderate competence with regards to two different sets of skills to achieve meaningful results relative to the desired quality, volume, and consistency of data flow. The first requirement is familiarity with smartphone technology, and access to the internet. Non-possession of smartphones may reduce the chances of enlisting volunteers from rural areas, where smartphone and internet penetration is still low. Moreover, these areas might also be those in which a price monitoring system is less present. Second, the language in which the app is programmed might also create barriers to becoming a member of the crowd. This is particularly relevant for our case, as the app was only available in English, which might have limited the participation of prospective crowd participants who only speak the local language(s).

While it is important to recognise these inherent constraints, one cannot circumvent the first one if smartphone-based crowdsourcing is the objective. Moreover, the relevance of such self-selection bias may wane over time as literacy levels and access to modern communication infrastructure becomes equitable. The chances of creating a successful crowd may be relatively higher within urban and peri-urban areas (compared to rural areas), where there is a good chance of attracting prospective crowd members who can meet the basic requirements for delivery of quality data, with the different types of market well represented.

In our particular application, we observe that there is still an impact of poor (or lack-of) access to technology and connectivity. For example, in the south-eastern region of Kano state, volunteer sign-ups and data submission was very low, despite the distribution of flyers in LGAs within this region. This may be caused by the poor internet coverage in the region (Stryjak & Sivakumaran, 2019), which is predominantly rural, with low literacy levels and limited smartphone penetration.

As mitigating actions to avoid these biases, further efforts to engage volunteers in areas with low sign-up might be needed. In addition, if the price-monitoring system is to be set-up in an area with low connectivity, investments in network coverage might be needed. In addition, one should always compare the basic demographics of the crowd with secondary data related to the area of study, to make sure no specific group of consumers, farmers, wholesalers or retailers is being left out of the crowd. For instance, in order to avoid uneven participation of specific groups, targeted awareness-raising campaigns (e.g. through farmer organisations) are recommended. In delivering these, it is important to adequately identify the needs and potential contribution of the target group and select the most appropriate approach to reach them and influence their behaviour.

10.2 THE IT PLATFORM

A second aspect to consider when setting up a crowdsourcing initiative is that of the technical specifications and characteristics of the IT platform to achieve seamless data submission, visualisation and retrieval. During FPCA implementation, setting up and implementing the data collection process on a pre-configured open-source server and mobile-based tool (ONA and ODK respectively) allowed us to commence data collection immediately, and cover the expected food supply chain and products, without the need for significant investments in designing and developing a new data collection system. But it also imposed some limitations on our ability to enforce submission rules, or to reduce the complexity of the implementation workflow. For example, neither the ODK app nor the back-end server offered the option to fully restrict submission(s) based on phone ID or location. Thus, there were many instances where multiple registrations came from the same device. Also, subsequent assignment of IDs to registered volunteers could not be automated through the ODK platform, so this process was implemented manually. Both limitations resulted in an initial "mismatch" conundrum, where a single device is associated with multiple IDs, before we could detect the issue. This conundrum initially undermined the desired "perfect match" between device ID and VC ID as a control measure for data integrity, and could jeopardise the aspiration for automated relational data querying and analyses. It was possible to solve this *expost* manually, but involved a significant number of person-hours and the need to subjectively choose the ID per device that would be retained in the system.

Here, we would recommend thoroughly matching the characteristics of the available IT platforms to the needs of the crowdsourcing protocol, and adjusting the latter to the capacities of the former. In instances where existing platform does not fully match the needs, using an existing platform seems to be a cost-effective option for subsequent development and deployment of an integrated end-to-end crowdsourcing system, which can be deployed at regional or national scale. In any case, it is essential that privacy and data confidentiality requirements are met.

10.3 THE INCENTIVE

As participation in the crowd is voluntary, the system should include a set of incentives to maximise participation from the onset. In our case, we tested three types of incentives: monetary rewards, nudges based on social norms, and access to information. For the first case, the reward, most of the volunteers indicated a preference for rewards being monetary and implemented via direct transfer to their bank account. To implement this, options for partnering with major telecomms companies were explored. However, using a third-party mobile solution platform was sub-optimal, not only because of the short time frame of the project but also due to the lack of automatic linkage between the submission and reward systems. Moreover, the administrative procedures of the implementing partner (IITA) could not assure that the weekly financial rewards would be disbursed in a timely manner. Therefore, our reward system was based on monetary rewards, which were paid weekly through manual bank transfers. However, manual processing of transfers led to delays and errors in the reward implementation.

The aspiration to develop a fully integrated crowdsourcing system should be guided by prior understanding of potential issues that may be encountered. In particular, we recommend early contact with existing financial institutions who can facilitate the operational aspects of a "rewards" module within the envisioned system. As alternative if payments are made in the form of mobile money or credits, it can be automated through a script, on the condition that top-up is done though an internet platform.

Also, using behavioural tools in the form of nudges based on communicating social norms has proved to be immediately effective in terms of increasing the number of contributions, whereas giving access to the collectively produced dataset did not prove to significantly increase the number of voluntary contributions over the short time frame of the implementation. With regard the null effect of the latter, this can be only effective if the information disclosed is perceived as relevant by the crowd for their own use or for the community. However, several IT and internet connectivity issues limited the access of the crowd to the data which limited the crowd's ability to understand the usefulness of the data.

Interestingly, the effects of the economic incentive were additional to the effects of the nudge based on the communication on the social norm. We recommend including behavioural tools, such as different types of nudges that give feedback to the crowd, in the design of a crowdsourcing initiative, and exploring them with longer periods. For this we also recommend to better understand the type of information that is relevant to the crowd and the most adequate dissemination formats. In addition, improvements in the communication of the use of market information by the crowd could also contribute to understand usability and motivate their participation.

10.4 THE DATA QUALITY

One of the main concerns in crowdsourcing is the quality of the output (data) produced and how this is assured, including how time-costly it can be to work with a large volume of raw, unstructured, diverse and fast data that is collected as it comes from voluntary contributions without an a priori sample design. These are characteristics that crowdsourced data shares with other big data, giving raise to multiple opportunities and challenges, including those of accuracy, representativeness but also privacy, security and ethics.

The data management and quality assurance methodology we proposed and applied has proved to be an effective method for retrieving, cleaning, processing and aggregating the data in a reliable manner, if sufficient contributions are available. A series of algorithms have been developed and applied to extract the data from the web platform, transform it from unstructured to structured data, clean it based on spatial statistical methods and aggregate it (i.e. post-sample) in such a way that it resembles a formal sample design and allows for useful statistical inference. These steps where aimed at increasing reliability and usability.

To monitor the performance of the system, several indicators have been developed to assess different aspects of the quality of the final dataset. In the future, this quality could be benchmarked against a global data quality measure and partial measures for each quality dimension (e.g. timeliness, accuracy, consistency, accessibility).

For this, the quality measures proposed to monitor and evaluate the quality performance would have to be standardised in order to achieve comparable measures for each dimension. This in turn would allow the calculation of a global measure, eventually requiring the establishment of weights for each quality dimension. Since there can be trade-offs between the different quality dimensions (e.g. timeliness for accuracy) it could be left to data users to set the weights. We recommend the dissemination of a quality indicator(s) (e.g. quality label(s)) together with the crowdsourced dataset, which may help generate trust in the data.

10.5 THE WEB DASHBOARD

We developed a web dashboard ([40]) as a user-friendly dissemination tool for the daily crowdsourced and validated information on food prices. The web dashboard not only provided access to processed and validated data, but also information on different metrics, allowing users to monitor price changes and trends in time and space by digging deeper into the data in an interactive way. For greater visibility the FPCA web dashboard has been included in the JRC's Data portal of agro-economic Modelling (DataM) integrated with the JRC's Data Catalogue (https://data.jrc.ec.europa.eu/), which is also integrated with the EU Open Data portal (https://data.europa.eu/euodp/en/home).

We recommend including dashboards in open data portals that can make them more relevant for internet search engines, which can improve outreach to more people (e.g., people looking for "food prices Nigeria" on Internet. In addition, depending on the budget and country's reality, this type of initiatives could also consider using Search Engine Optimisation (SEO) or targeted marketing/advertisement on search engines or social networks. With improved outreach, consideration could be given to using the web dashboard to recruit new participants who can replace those who stopped contributing.

One of the main advantages of web dashboards is that they reduce information search costs and the risk of information overload. This helps to reduce information asymmetries and the capacity to use the information to support decision- and policy-making. The importance of dashboards increases as the crowdsourced dataset becomes larger and more diverse. Yet, for dashboards to be useful, it is important to conduct a careful design that includes the collection and integration of feedback from data users with regard to the type of information and dissemination format. One could also consider exploring what type of nudges could promote the use of the dashboard, for example, include information about the number of visits or user comments.

We recommend engaging with crowd participants and potential data users, and providing them with the opportunity to give feedback, which may play a crucial role in realising the potential benefits of dashboards and explore the use of behavioural tools, such as nudges to promote its use. Finally, given country's reality

([40]) https://datam.jrc.ec.europa.eu/datam/mashup/FP_NGA/

and in case of technological barriers (i.e. internet connectivity), alternative ways of disseminating crowdsourcing information could be considered (e.g. SMSs, email or a simple dashboard elaborated using basic HTML and images, which are compatible even with the most basic browsers and smartphones).

10.6 CONCLUSIONS

The FPCA project presented in this report is a novel prototype for crowdsourcing food price data which has been tested in two regions of northern Nigeria where contextual realities are limiting for traditional ground-level data collection, beyond the overarching cost issues related to extensive ground data collection by official statistical agencies ([41]). The JRC developed the FPCA with the support of experts on spatial statistics, which was then implemented in Nigeria by IITA. The methodology collects and processes citizen-generated data in order to make quality information available through a web dashboard in real time. The methodology is scalable by design and can be adopted within and beyond national boundaries.

This rich geo-referenced data, which can be accessed in real time, is a valuable asset for food security assessment, early warning of market pulses, and support for decisions related to commodities and trade. In addition, the spatial dimension of the data opens a door for spatial-temporal analysis. The proposed quality methodology and performance metrics also provide a strong basis for objective assessment of the developed crowdsourcing methodology, and can be applied to evaluate other existing or future approaches to data crowdsourcing.

Finally, it is important to engage citizens as data curators, not just as data consumers. For this, different types of incentives and, in particular, behavioural tools can be explored and incorporated in the design. This FPCA model of data crowdsourcing incorporates and balances the critical elements that can sustain such desirable engagement of citizens across diverse applications and national frontiers.

This study has served to identify the potentials and challenges of crowdsourcing data on food prices. Notably, the granularity and high-frequency of the collected data make real-time spatial-temporal analysis possible, the application of which may be essential for policy and market decisions and fast response by associating in real-time geographic location and food prices/price changes and possibly other relevant variables (e.g. people incomes). Further, the developed quality methodology reduces the trade-off between the timeliness of crowdsourcing and the accuracy of more traditional data collection methods. However, our application has also identified areas that could benefit from further research. These include a global metric for evaluating the quality of a crowdsourced dataset, and developing the feedback loop with data users, particularly market actors. Furthermore, the methodology could be expanded to other regions (including developed economies such as the EU), and to other data types and fields where data is sparse or costly to obtain through traditional methods. Moreover, the approach could be used for crowdsourcing data coming from sensors and not from the active participation of citizens (i.e. crowdsensing). Indeed, adequate methods and tools for quality in crowdsourcing may encourage the emergence of innovative applications to improve data availability to support better policy and decision-making. Finally, the challenge remains of how to integrate this type of data collection, which shares specific characteristics with other big data, as a complementary source of information in institutions/organisations that collect data using traditional methods. This integration could reduce the additional investment needed for the setup and could help to maximise complementarities.

([41]) For instance, issues of security threat (e.g., terrorism or war) impose severe limitations on the deployment of trained staff and constrain the need to collect robust data across space and time.

References

AGRA. (2017). *Africa Agriculture Status Report: The Business of Smallholder Agriculture in Sub-Saharan Africa (Issue 5)*. Nairobi, Kenya.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, *27*(2), 93–115.

APA. (2018). *Psychology – APA Dictionary of Psychology.* Retrieved from https://dictionary.apa.org/psychology

Arbia, G. (1993). The use of GIS in spatial statistical surveys. *International Statistical Review/Revue Internationale de Statistique*, 339–359.

Arbia, G. (2014). *A primer for spatial econometrics: with applications in R*. Springer.

Arbia, G., Solano-Hermosilla, G., Micale, F., Nardelli, V., & Genovese, G. (2018). *Post-sampling crowdsourced data to allow reliable statistical inference: the case of food price indices in Nigeria*. Retrieved from http://meetings3.sis-statistica.org/index.php/sis2018/49th/paper/viewFile/1090/64

Arbia, G., Solano-Hermosilla, G., Micale, F., Nardelli, V., & Genovese, G. (2020). From mobile app-based crowdsourcing to crowd-trusted food price estimates in Nigeria: a post-sampling strategy for optimal statistical inference. *In Preparation*.

Babagana, M., Ismail, M., Mohammed, B. G., Dilala, M. A., Hussaini, I., & Zangoma, I. M. (2018). Impacts of Boko Haram Insurgency on Agricultural Activities in Gujba Local Government Area, Yobe State, Nigeria. *International Journal of Contemporary Research and Review*, *9*(12), 20268–20282.

Beręsewicz, M., Lehtonen, R., Reis, F., Di Consiglio, L., & Karlberg, M. (2018). *An overview of methods for treating selectivity in big data sources*. Statistical Working Papers. Eurostat. Doi: https://doi.org./10.2785/312232.

Blaschke, S. M., Carroll, P. P., Chaves, D. R., Findley, M. G., Gleave, M. C., Morello, R. N., & Nielson, D. L. (2013). Extrinsic, Intrinsic, and Social Incentives for Crowdsourcing Development Information in Uganda: A Field Experiment. *Article Manuscript*.

Bobek, D. D., Roberts, R. W., & Sweeney, J. T. (2007). The social norms of tax compliance: Evidence from Australia, Singapore, and the United States. *Journal of Business Ethics*, *74*(1), 49–64.

Bott, M., & Young, G. (2012). The role of crowdsourcing for better governance in international development. *Praxis: The Fletcher Journal of Human Security*, *27*(1), 47–70.

Brabham, D. C. (2013). *Crowdsourcing*. Mit Press.

Brossard, D., Lewenstein, B., & Bonney, R. (2005). Scientific knowledge and attitude change: The impact of a citizen science project. *International Journal of Science Education*, *27*(9), 1099–1121.

Buchbinder, K. S. (2017). Crowdsourced Data Collection Provides On-the-Ground Insights. Retrieved from Digital@DAI website: https://dai-global-digital.com/crowdsourced-data-collection-provides-on-the-ground-insights.html

Chengalur-Smith, I. N., Ballou, D. P., & Pazer, H. L. (1999). The impact of data quality information on decision making: an exploratory analysis. *IEEE Transactions on Knowledge and Data Engineering*, *11*(6), 853–864.

Cullis, J., Jones, P., & Savoia, A. (2012). Social norms and tax compliance: Framing the decision to pay tax. *The Journal of Socio-Economics*, *41*(2), 159–168.

Damgaard, M. T., & Nielsen, H. S. (2018). Nudging in education. *Economics of Education Review*, *64*, 313–342.

Daniel, F., Kucherbaev, P., Cappiello, C., Benatallah, B., & Allahbakhsh, M. (2018). Quality control in crowdsourcing: A survey of quality attributes, assessment techniques, and assurance actions. *ACM Computing Surveys (CSUR)*, *51*(1), 7.

Davis, B., Di Giuseppe, S., & Zezza, A. (2017). Are African households (not) leaving agriculture? Patterns of households' income sources in rural Sub-Saharan Africa. *Food Policy*, *67*, 153–174.

Deichmann, U., Goyal, A., & Mishra, D. (2016). *Will digital technologies transform agriculture in developing countries?* The World Bank.

DGINS. (2013). *Scheveningen Memorandum - Big Data and Official Statistics*. Retrieved from https://ec.europa.eu/eurostat/documents/42577/43315/Scheveningen-memorandum-27-09-13

DGINS. (2018). *Bucharest Memorandum - on Official Statistics in a Datafied Society (Trusted Smart Statistics)*.

Donmez, A., Solano-Hermosilla, Gloria Bougay, Vladimir Subbaraman, B., M'barek, R., Santini, F., Adam, A., Bahemuka, Stephen Chinganya, O. J. M., Njosa, I. F. (2017). *Using web and mobile phone technologies to collect food market prices in Africa, Approaching real-time data and use of crowdsourcing, 2013 – 2016. JRC Working Papers JRC104311, Joint Research Centre (Seville site).*

EC. (2015). *Launching the EU International Cooperation and Development Results Framework. Commission Staff Working Document SWD (2015) 80 Final. Brussels: EC*.

EC. (2017). *Indicators to measure Social Protection Performance. Implications for EC Programming. Tools and Methods Series, concept paper n. 5. Brussels: EC*.

ESS. (2018). *ESS Handbook - Methodology for data validation v1.1 - Rev2018*. Retrieved from https://ec.europa.eu/eurostat/cros/content/ess-handbook-methodology-data-validation-v11-rev2018-0_en

ESS. (2019). *Quality Assurance Framework of the European Statistical System V2.0*. Retrieved from https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V1-2final.pdf/bbf5970c-1adf-46c8-afc3-58ce177a0646

Estellés-Arolas, Enrique, González-Ladrón-De-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, *38*(2), 189–200.

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd*, *96*(34), 226–231.

Eurostat. (2014). *ESS Handbook for Quality Reports*.

FAO. (2018). *Assessing the Quality of Agricultural Market Information Systems A self-assessment guide*. Retrieved from http://www.fao.org/3/i8935en/I8935EN.pdf

Flückiger, Y., & Seth, N. (2016). Sustainable Development Goals: SDG indicators need crowdsourcing. *Nature*, *531*(7595), 448.

Green, R., Cornelsen, L., Dangour, A. D., Turner, R., Shankar, B., Mazzocchi, M., & Smith, R. D. (2013). The effect of rising food prices on food consumption: systematic review with meta-regression. *Bmj*, *346*, f3703.

GSMA. (2018). *The mobile economy: West Africa 2018*. GSMA London.

GSMA. (2019). *The mobile economy 2019*. GSMA London.

Haile, M. G., Kalkuhl, M., & Usman, M. A. (2015). Market information and smallholder farmer price expectations. *African Journal of Agricultural and Resource Economics*, *10*(311-2016–5639), 297–311.

Hamadeh, N., Rissanen, M., & Yamanaka, M. (2013). Crowd-sourced price data collection through mobile phones. *Collaboration in Research and Methodology for Official Statistics. World Bank. Http://Ec. Europa. Eu/Eurostat/Cros/Sites/Crosportal/Files/NTTS2013fullPaper_82-v2. Pdf*.

Hansen, M. H., Hurwitz, W. N., & Madow, W. G. (1953). *Sample survey methods and theory. V. 1. Methods and applications. V. 2. Theory*. Wiley.

HDX. (2019). OCHA Nigeria - Administrative Boundaries (levels 0 - 3) and senatorial districts. Retrieved from https://data.humdata.org/organization/ocha-nigeria?ext_cod=1&q=&ext_page_size=25

Hernandez, M., Jamison, J., Korczyc, E., Mazar, N., & Sormani, R. (2017). *Applying behavioral insights to improve tax collection: experimental evidence from Poland*. © 2017 International Bank for Reconstruction and Development / The World Bank 1818 H Street NW, Washington DC 20433.

Hofmokl, J. (2010). The Internet commons: towards an eclectic theoretical framework. *International Journal of the Commons*, *4*(1), 226–250.

Hosseini, M., Phalp, K., Taylor, J., & Ali, R. (2019). On the configuration of crowdsourcing projects. In *Crowdsourcing: Concepts, Methodologies, Tools, and Applications* (pp. 33–52). IGI Global.

Howe, J. (2006). The rise of crowdsourcing. *Wired Magazine*, *14*(6), 1–4.

Ismail, E. F. A. (2010). *Markets and Price Monitoring: Training Manual*. Khartoum, Sudan.

Juran, J., & Godfrey, A. B. (1999). Quality handbook. *Republished McGraw-Hill*, 173–178.

Kalkuhl, M., von Braun, J., & Torero, M. (2016). *Volatile and Extreme Food Prices, Food Security, and Policy: An Overview BT - Food Price Volatility and Its Implications for Food Security and Policy* (M. Kalkuhl, J. von Braun, & M. Torero, Eds.). https://doi.org/10.1007/978-3-319-28201-5_1

Kaufmann, N., Schulze, T., & Veit, D. (2011). More than fun and money. Worker Motivation in Crowdsourcing-A Study on Mechanical Turk. *AMCIS*, *11*(2011), 1–11. Detroit, Michigan, USA.

Kormawa, P., & Ogundapo, A. T. (2004). *Local weights and measures in Nigeria: a handbook of conversion factors*. Ibadan: International Institute of Tropical Agriculture.

Kroese, F. M., Marchiori, D. R., & de Ridder, D. T. D. (2015). Nudging healthy food choices: a field experiment at the train station. *Journal of Public Health*, *38*(2), e133–e137.

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends® in Econometrics*, *4*(3), 165–224.

Linder, N., Lindahl, T., & Borgström, S. (2018). Using behavioural insights to promote food waste recycling in urban households—Evidence from a longitudinal field experiment. *Frontiers in Psychology*, *9*, 352.

Little, R. J. A. (1994). A class of pattern-mixture models for normal incomplete data. *Biometrika*, *81*(3), 471–483.

Lozano-Gracia, N., & Young, C. (2014). *Housing consumption and urbanization*. The World Bank.

Matheus, R., Janssen, M., & Maheshwari, D. (2018). Data science empowering the public: Data-driven dashboards for transparent and accountable decision-making in smart cities. *Government Information Quarterly*.

Mooney, P., & Minghini, M. (2017). *A review of OpenStreetMap data*.

NBS. (2010). National Manpower Stock and Employment Generation Survey: Household and Micro Enterprise. Retrieved from www.nigerianstat.gov.ng/download/7

OECD-OPSI. (2019). Unlocking the potential of crowdsourcing for public decision-making with artificial intelligence. Retrieved from Organization for Economic Cooperation and Development (OECD), Observatory of Public Sector Innovation (OPSI) website: https://oecd-opsi.org/innovations/unlocking-the-potential-of-crowdsourcing-for-public-decision-making-with-artificial-intelligence/

Ovaga, O. H. (2012). The Socio-Economic Implications of Book-Haram Activities in Northern Nigeria. *Review of Public Administration & Management*, *1*(2), 19–37.

Pedersen, J., Kocsis, D., Tripathi, A., Tarrell, A., Weerakoon, A., Tahmasbi, N., … De Vreede, G.-J. (2013). Conceptual foundations of crowdsourcing: A review of IS research. *2013 46th Hawaii International Conference on System Sciences*, 579–588. IEEE.

R Core Team. (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. Retrieved from https://www.r-project.org/

Seid, Y., & Fonteneau, F. (2017). *Leveraging crowdsourcing techniques and technologies to generate better agricultural information: three case studies, in Indonesia, Nigeria and the Philippines. FAO, Statistics Division*.

Short, C., Barreiro-Hurle, J., & Balié, J. (2014). Policy or markets? An analysis of price incentives and disincentives for rice and cotton in selected African countries. *Canadian Journal of Agricultural Economics/Revue Canadienne d'agroeconomie*, *62*(4), 441–469.

Solano-Hermosilla, G., Barreiro-Hurle, J., Adewopo, J., & Gorrín González, C. (2020). Increasing engagement in Price crowdsourcing initiatives: using nudges in Nigeria. *In Preparation*.

Soliman, W., & Tuunainen, V. K. (2015). Understanding continued use of crowdsourcing systems: An interpretive study. *Journal of Theoretical and Applied Electronic Commerce Research*, *10*(1), 1–18.

StataCorp. (2017). *Stata: Release 15. Statistical Software. College Station, TX: StataCorp LLC.*

Stryjak, J., & Sivakumaran, M. (2019). The Mobile Economy 2019. *GSMA Intelligence*.

Sunstein, C. R. (2013). Nudges. gov: Behavioral economics and regulation. *Forthcoming, Oxford Handbook of Behavioral Economics and the Law (Eyal Zamir and Doron Teichman Eds.)*.

Sunstein, C. R. (2014a). Nudging: a very short guide. *Journal of Consumer Policy*, *37*(4), 583–588.

Sunstein, C. R. (2014b). *Why nudge?: The politics of libertarian paternalism*. Yale University Press.

Surowiecki, J. (2005). The wisdom of crowds: why the many are smarter than the few. London. *Abacus: New Edition*, *39*.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

The Royal Society. (2012). *Science as an open enterprise*. Retrieved from https://royalsociety.org/-/media/policy/projects/sape/2012-06-20-saoe.pdf

UN. (2001). United Nations Group of Experts on Geographical Names. Retrieved from https://unstats.un.org/unsd/geoinfo/UNGEGN/links.html

UN. (2015). National Quality Assurance Frameworks. Retrieved from https://unstats.un.org/unsd/dnss/QualityNQAF/nqaf.aspx

Vetrò, A., Canova, L., Torchiano, M., Minotas, C. O., Iemma, R., & Morando, F. (2016). Open data quality measurement framework: Definition and application to Open Government Data. *Government Information Quarterly*, *33*(2), 325–337.

Villa, J. M. (2016). diff: Simplifying the estimation of difference-in-differences treatment effects. *The Stata Journal*, *16*(1), 52–71.

WFP. (2016). 2016 - Mobile Vulnerability Analysis & Mapping (mVAM). Retrieved from https://www.wfp.org/content/2016-mobile-vulnerability-analysis-mapping-mvam

Wiggins, A., & Crowston, K. (2011). From conservation to crowdsourcing: A typology of citizen science. *2011 44th Hawaii International Conference on System Sciences*, 1–10. IEEE.

Winner, L. (1984). Mythinformation in the high-tech era. *Bulletin of Science, Technology & Society*, *4*(6), 582–596.

World Bank. (2015). *Operational Guidelines and Procedures for Measuring the Real Size of the World Economy*.

Zeug, H., Zeug, G., Bielski, C., Solano-Hermosilla, G., & M'barek, R. (2017). *Innovative Food Price Collection in Developing Countries: Focus on Crowdsourcing in Africa*. https://doi.org/10.2788/53432

Ziegler, S., Crettaz, C., Hazan, M., Alexandrou, P., Filios, G., Nikoletseas, S., … Angelopoulos, C. M. (2017). *Combining Internet of Things and Crowdsourcing for Pervasive Research and End-user Centric Experimental Infrastructures (IoT Lab)*.

List of abbreviations and definitions

AI	Artificial Intelligence
AMIS	Agricultural and Market Information System
AfDB	African Development Bank
API	Application programming interface
APS	Agricultural Performance Survey
DID	Difference-in-Difference
EAs	Extension Agents
EC	European Commission
EC-JRC	European Commission Joint Research Centre
ESS	European Statistical System
FAO	Food and Agriculture Organisation of the United Nations
FEWNET	Famine Early Warning Systems Network
HDX	Humanitarian Data Exchange
ICP	International Comparison Programme
IITA	International Institute for Tropical Agriculture
IoT	Internet of Things
JRC	Joint Research Centre
KPI	Key Performance Indicator
LGA	Local Government Area
LI	Local Inspector
NAERLS	National Agricultural Extension, Research, and Liaison Services
NBS	National Bureau of Statistics
NYSC	National Youth Service Corp
OECD	Organisation for Economic Cooperation and Development
OPSI	Observatory of Public Sector Innovation
QAF	Quality Assessment Framework
SGR	Strategic Grains Reserve
VC	Volunteer crowd
VGI	Volunteered Geographic Information
WeNR	Wageningen Environment and Research
WFP	World Food Programme
WB	World Bank

List of boxes

List of figures

List of tables

Annexes

Annex 1. Smartphone app questionnaire/data submission form

Table 21.Data submission form

type	name	label	required	constraint	relevant	appearance	repeat	calculation	hint	media::image
start	start	Complete this form by Swiping forward and backward or Using the arrow navigation buttons								
today	today									
deviceid	deviceid									
calculate	timeStart							now()		
note	partner_logo	**Welcome. We are excited that you are a committed and valued volunteer for the '*Food Price Data Crowdsourcing*'. Ensure that you provide accurate and timely data. Only accurate submissions will be rewarded on a "First-submit, First-rewarded" basis. Note that there is a maximum limit on the number of submissions rewarded daily/weekly/monthly, so you should submit each record as soon you complete the survey.**								**Partnerlogos_alt.png**
text	**VC_ID**	**What is your Volunteer ID?**	**yes**						This is the ID that was sent to you in the confirmation text message.	
note	note_begin	**Disclaimer:** By accessing this form, you are agreeing to participate in a data collection task as a volunteer, with the likelihood of earning a reward (up to N8000 per month) for participation. You are undertaking this solely at your discretion and you are free to opt-out at any point during the duration of the initiative. By continuing to complete this survey, you accept that non of the collaborating institutions bear any liability for any risk or costs associated								

type	name	label	required	constraint	relevant	appearance	repeat	calculation	hint	media::image
		with your decision to provide daily food price data from your neighbouring market on a daily basis. You also agree that all data collected becomes public good and can be used to inform institutions and the public about market trends and threats relative to the target commodities.								
note	note_begin2	**Important Info! You should only submit one completed form per day because we will select first (quality-proven) submissions on a daily basis - all volunteers have a fair chance to be rewarded. Note that you may be TOTALLY Blacklisted if you send multiple submissions in a single day, because this will be in contradiction of the rule - No Spamming Allowed!**								
begin group	Section_A	**Let's get some information about the market where you are**								
geopoint	gps	Stand close to the selling point (market or store) and record the location (GPS coordinates) of the market location where you are planning to collect data.	yes						Make sure you **TURN-ON/ENABLE** location in your phone. Once this is done coordinates will be logged automatically	
select one from market_type	market_type	What type of market are you collecting this data from?	yes							
select one from yes_no	buying	Are you selling or buying the agricultural/food product of the reported price?	yes							
select one from buying_purpose	buying_purpose	What is your purpose of buying or selling?	yes							
decimal	market_distance	Aproximately, how far is this market place from your house (in Kilometers)?	yes							

type	name	label	required	constraint	relevant	appearance	repeat	calculation	hint	media::image
end group										
note	note_crops	**Now to the fun part! You will submit the price of 4 food commodities - Maize, Rice, Beans, Soybeans. Ask the seller(s) in the market politely for the price of the commodities. If one seller does not sell all the food commodities or varieties, move on to another seller.**								
begin group	Section_B1	**Information on Yellow Maize**								
note	yellow_maizepic	Ask and confirm if the seller has yellow maize (dried grain) in stock (i.e. for sale)								**YellowMaize.png**
select one from yes_no	maize_yellow	Are you providing the price for yellow maize?	yes							
select one from packaging_maize	packaging_Ymaize	Select the packaging unit for the yellow maize sold here	yes		${maize_yellow}='Yes'					
integer	price_Ymaize	What is the price per unit measure or package of this maize variety?	yes		${maize_yellow}='Yes'				*in Naira* (₦)	
end group										
begin group	Section_B2	**Proceed to provide price data on White Maize**								
note	white_maizepic	Ask and confirm if the seller has white maize (dried grain) in stock (i.e. for sale)								**WhiteMaize.png**
select one from yes_no	maize_white	Are you providing the price for white maize?	yes							
select one from packaging_maize	packaging_Wmaize	Select the packaging unit for the White maize sold here	yes		${maize_white}='Yes'					
integer	price_Wmaize	What is the price per unit measure or package of this white maize variety?	yes		${maize_white}='Yes'				*in Naira* (₦)	
end group										
begin group	Section_B3	**Proceed to provide price data on imported Thailand Rice**								
note	ThailandRice_Pic	Ask and confirm if the seller has								**ThailandRice.png**

type	name	label	required	constraint	relevant	appearance	repeat	calculation	hint	media::image
		Thailand Rice in stock (i.e. for sale)								
select one from yes_no	Thailand	Are you providing the price for Thailand Rice?	yes							
select one from packaging_rice	packaging_Thailandrice	Select the packaging unit for the Thailand rice sold here	yes		${Thailand}='Yes'					
integer	price_Thailand	What is the price per packaging unit of this maize variety?	yes		${Thailand}='Yes'				*in Naira* (₦)	
end group										
begin group	Section_B4	**Proceed to provide price data on Imported Indian Rice**								
note	IndianRice_Pic	Ask and confirm if the seller has Indian Rice in stock (i.e. for sale)								**IndianRice.png**
select one from yes_no	Indian	Are you providing the price for Indian Rice?	yes							
select one from packaging_rice	packaging_Indianrice	Select the packaging unit for the Indian rice sold here?	yes		${Indian}='Yes'					
integer	price_Indian	What is the price per packaging unit of this Inidan Rice?	yes		${Indian}='Yes'				*in Naira* (₦)	
end group										
begin group	Section_B5	**Proceed to provide price data on Local Rice**								
note	LocalRice_Pic	Ask and confirm if the seller has Local Rice in stock (i.e. for sale)								**LocalRice_Grade1.png**
select one from yes_no	local_rice	Are you providing the price for Local Rice?	yes							
integer	local_types	How many grades of local rice is sold here?	yes	. <= 4	${local_rice}='Yes'				*Grades are mostly differentiated by their prices*	
begin repeat	repeat_local_rice	Note that you will answer next 4 questions for each grade of local rice i.e. if you specify 2 grades, the next 4 questions will repeat 2 times.								
select one from local_type_list	local_type_list	Select type/grade of local rice	yes		${local_rice}='Yes'		${local_types}			
select one from packaging_local_rice	packaging_local_rice	Select the packaging unit for this local rice grade	yes		${local_rice}='Yes'		${local_types}			
integer	price_rice_local	What is the price per selected packaging unit of this rice grade?	yes		${local_rice}='Yes'		${local_types}		*in Naira* (₦)	
end repeat										
end group										

type	name	label	required	constraint	relevant	appearance	repeat	calculation	hint	media::image
begin group	Section_B6	**Proceed to provide price data on Red Beans**								
note	Red_BeansPic	Ask and confirm if the seller has Red/Brown Beans in stock (i.e. for sale)								**BrownBeans_Grade1.png**
select one from yes_no	red_beans	Are you providing the price for Red beans?	yes							
integer	redbeans_types	How many grades of Red beans is sold here?	yes		${red_beans}='Yes'					
begin repeat	repeat_redbeans	Note that you will answer next 4 questions for each grade of red beans i.e. if you specify 2 grades, the next 4 questions will repeat 2 times.								
select one from local_type_list	local_type_list	Select type/grade of Red beans	yes		${red_beans}='Yes'		${Redbeans_types}			
select one from packaging_local_rice	packaging_redbeans	Select the packaging unit for this Red beans grade	yes		${red_beans}='Yes'		${Redbeans_types}			
integer	price_red_beans	What is the price per selected packaging unit of this red beans grade?	yes		${red_beans}='Yes'		${Redbeans_types}		*in Naira* (₦)	
end repeat										
end group										
begin group	Section_B7	**Proceed to provide price data on White Beans**								
note	Whitebean_pic	Ask and confirm if the seller has White Beans in stock (i.e. for sale)								**WhiteBeans_Grade1.png**
select one from yes_no	White_beans	Are you providing the price for White beans?	yes							
integer	Whitebeans_types	How many grades of White beans is sold here?	yes		${White_beans}='Yes'					
begin repeat	repeat_whitebeans	Note that you will answer next 4 questions for each grade of white beans i.e. if you specify 2 grades, the next 4 questions will repeat 2 times.								
select one from local_type_list	local_type_list	Select type/grade of White beans	yes		${White_beans}='Yes'		${Whitebeans_types}			
select one from packaging_local_rice	packaging_whitebeans	Select the packaging unit for this White beans?	yes		${White_beans}='Yes'		${Whitebeans_types}			
integer	price_White_beans	What is the price per selected packaging unit of this red beans grade?	yes		${White_beans}='Yes'		${Whitebeans_types}		*in Naira* (₦)	
end repeat										
end group										
begin group	Section_B8	**Proceed to provide price data on Soybean**								
note	Soybean_pic	Ask and confirm if the seller has Soybeans in								**Soybeans.png**

type	name	label	required	constraint	relevant	appearance	repeat	calculation	hint	media::image
		stock (i.e. for sale)								
select one from yes_no	Soybean	Are you providing the price for Soybeans?	yes							
select one from packaging_soybean	packaging_soybean	Select the packaging unit for the Soybean sold here?	yes		${Soybean}='Yes'				*Mudu, Bag...*	
integer	price_soybean	What is the price for this smallest packaging unit of the soybean variety?	yes		${Soybean}='Yes'				*in Naira (₦)*	
end group										
text	seller_phone	For verification, please provide phone number of the seller		regex(., '[0]{1}[1-9]{1}[0-9]{9}')					*Although this is optional, it is better to provide the phone number of the seller if available. This will improve your chances of receiving reward.*	
note	end_note	Great! This is the end of the questionnaire. You may now proceed to the next stage of submission after confirming that all required questions have been answered. You will be notified by the end of the week if your submission meets the required target (quality and timing) for our daily reward. Keep submitting, one survey per day!								

Table 22. Choices in the data submission form

list_name	name	label
market_type	Supermarket	Supermarket
market_type	Neighborhood_shops_kiosk	Neighborhood shops kiosk
market_type	Open air_or_covered market	Open air or covered market
market_type	Mobile_shops_street vendors	Mobile shops street vendors
market_type	Bulk_and_discount stores	Bulk and discount stores
market_type	Specialised_stores	Specialised stores
market_type	Directly_from_Farmer	Directly from farmer
market_type	Local_village_market	Local village market
market_type	City_market	City market
buying	selling	Selling
buying	buying	Buying
yes_no	Yes	Yes
yes_no	No	No
packaging_maize	100kg	100kg bag
packaging_maize	50kg	50kg bag

packaging_maize	25kg	25kg bag
packaging_maize	10kg	10kg bag
packaging_maize	5kg	5kg bag
packaging_maize	1kg	1kg bag
packaging_maize	Mudu/Kwano/Tiyya	Mudu/Kwano/Tiyya
maize_variety1	White	White
maize_variety1	Yellow	Yellow
sorghum_variety1	Red	Red (Kaura)
sorghum_variety1	White	White
packaging_rice	100kg	100kg bag
packaging_rice	50kg	50kg bag
packaging_rice	25kg	25kg bag
packaging_rice	10kg	10kg bag
packaging_rice	5kg	5kg bag
packaging_rice	1kg	1Kg bag
packaging_rice	Mudu/Kwano/Tiyya	Mudu/Kwano/Tiyya
local_type_list	Grade_1	Grade 1
local_type_list	Grade_2	Grade 2
local_type_list	Grade_3	Grade 3
local_type_list	Grade_4	Grade 4
local_type_list	Grade_5	Grade 5
packaging_local_rice	100kg	100kg bag
packaging_local_rice	50kg	50kg bag
packaging_local_rice	25kg	25kg bag
packaging_local_rice	10kg	10kg bag
packaging_local_rice	5kg	5kg bag
packaging_local_rice	1kg	1Kg bag
packaging_local_rice	Mudu/Kwano/Tiyya	Mudu/Kwano/Tiyya
packaging_sorghum	100kg	100kg bag
packaging_sorghum	50kg	50kg bag
packaging_sorghum	25kg	25kg bag
packaging_sorghum	10kg	10kg bag
packaging_sorghum	5kg	5kg bag
packaging_sorghum	1kg	1Kg bag
packaging_sorghum	Mudu/Kwano/Tiyya	Mudu/Kwano/Tiyya
packaging_soybean	100kg	100kg bag
packaging_soybean	50kg	50kg bag
packaging_soybean	25kg	25kg bag
packaging_soybean	10kg	10kg bag
packaging_soybean	5kg	5kg bag
packaging_soybean	1kg	1Kg bag
packaging_soybean	Mudu/Kwano/Tiyya	Mudu/Kwano/Tiyya
buying_purpose	Consumption	Consumption

buying_purpose	Feeding Animals	Feeding Animals
buying_purpose	Reselling	Reselling
buying_purpose	Processing	Processing
buying_purpose	I'm just a price observer/reporter	I'm just a price observer/reporter

Annex 2. Smartphone app profile form

Table 23. Profile Data Form

type	name	label	required	constraint	media::image	relevant	appearance	calculation	Hint	body::accuracy Threshold
start	start	Start								
today	today									
deviceid	deviceid									
calculate	timeStart							now()		
phonenumber	phonenumber	submitters phone number								
begin group	**introduction_group**									
note	logo	**Collaborating partner logos**			**Partnerlogos_alt.png**					
note	note_begin	Welcome! Congrats on your prequalification as a potential volunteer for the food price data crowdsourcing task which is being led by EC-JRC, in collaboration with IITA and WUR. Your participation in this project will help us improve the access to agricultural and food market information. Thank you for taking the time to complete this survey. Move to the next page for few more notes.								
note	note_begin2	By completing this 'Profile Form' and successfully submitting, we will confirm your capability to send weekly data on food prices. This step is very critical for your final onboarding, so we encourage you to provide all requested information with clarity and accuracy. This form should take about 2 minutes to complete and send. Note that the entire exercise is set-up to be fair to all, yet competitive. Here's your chance to earn up to N8,000 recharge credit every month for the duration of the initiative. Now hurry up, complete the form, and submit without delay.								

type	name	label	required	constraint	media::image	relevant	appearance	calculation	Hint	body::accuracy Threshold
note	note_begin3	[PLEASE READ] Disclaimer: Personal data is collected through the mobile phone app only to the extent necessary for voluntary participation and to process related reward for participants. We will not disclose personal information to third parties except when exclusively necessary for the fulfilment of associated reward(s). The generated data will be made available to the EU and any relevant third parties without disclosing personal information or by anonymising the identity of participants either by total exclusion of personally identifiable information or use of aliases. IITA will only keep personal data for the time necessary to fulfil the purpose of data collection, quality monitoring, and further processing. You can contact Helen Peter (h.peter@cgiar.org) for any question or complain.								
select_one yes_no	Consent	Do you wish to continue to complete this form as an indication of your interest to participate, and consent to the use of submitted data by IITA and partners for research purposes?	yes							
end group	**introduction_group**									
geopoint	gps	Record the location (GPS coordinates) of your home/business/school	yes						Make sure you **TURN-ON/ENABLE** location in your phone. Once your GPS is switched on, stand outside (away from trees or buildings for 30secs), your coordinates will be automat	4

type	name	label	required	constraint	media::image	relevant	appearance	calculation	Hint	body::accuracyThreshold
									ically recorded. Proceed after seeing the recorded coordinates.	
begin group	**profile_group**			${Consent} = 'yes'						
note	note_profile	Few questions about you. We strongly recommend that you provide correct information that can be verified for future compensation purposes!								
text	VC_name	What is your name (Surname First)	yes							
integer	VC_ID	Input your VC_ID (if you have it, otherwise, skip)								
integer	Receive_Text	Receive Text Message for Performance (new volunteers should skip this!)							New volunteers should skip this!	
integer	Unique_Coding	Indicate if this is a unique or duplicate submission (new volunteers should skip this!)							New volunteers should skip this!	
integer	Reward_Class	Indicate the reward level for VC (Skip this!)							Volunteers should skip this!	

type	name	label	required	constraint	media::image	relevant	appearance	calculation	Hint	body::accuracyThreshold
text	VC_phonenum	what is your phone number	yes							
integer	VC_age	What is your age?	yes	.<=100						
select_one zm1iv99	VC_gender	What is your gender?	yes							
select_one ua2em60	HH_head	Are you the head of household?	yes							
select_one rn2ds93	VC_occupation	What is your sector of occupation?	yes							
select_one yj7os17	VC_Literacy	What is the highest education level that you have completed?	yes				field-list			
select_one hz7ia02	VC_Category	Are you a farmer, market trader, or final food consumer/buyer?	yes				field-list			
integer	HH_Pop	How many people live in your household?	yes							
integer	HH_Children	How many children (< 5 years old) live in your household?	yes						Enter '0' if there is none	
integer	HH_Pop_OnFarm	How many people in your household are working on your farm(s)?	yes			selected(${VC_Category}, 'farmer')				
integer	HH_Pop_OffFarm	How many people in your household are working outside or off the farm(s)?	yes			selected(${VC_Category}, 'farmer')				
note	note_media	These questions do not have any impact on your qualification for the task. We only want to understand your communication preference								
integer	Years_Smartphone	How many years have you been using a smartphone?	yes	. <= 20						
select_one mz5fy54	News_connect	Do you read/ listen to news (printed version, on internet, TV or radio)?	yes				field-list			
select_multiple mt04k45	News_source	From what source?	yes			${News_connect} = 'yes'	field-list			
select_one wj6fi42	Coop_member	Are you member of cooperative or any type of farmer association?	yes							
note	note_crowdsourcing	These questions do not have any impact on your qualification for the task. We only want to understand your preference and find out the best way to make the experience rewarding for you.								
select_one yes_no	VC_experience	Have you ever participated in a pure data crowdsourcing exercise before?	yes							

type	name	label	required	constraint	media::image	relevant	appearance	calculation	Hint	body::accuracy Threshold
select_one yes_no	VC_experience_phone	Did you use smartphone to collect and submit data, as part of a larger crowd?	yes			${VC_experience} = 'yes'				
select_multiple bs5yj80	VC_Motivation	What is your motivation behind your willingness to participate in this food price data crowdsourcing initiative?	yes							
select_one er7ev64	Preferred_reward	What is your preferred reward for full partiticipation in this crowdsourcing task?	yes				field-list			
select_one sm49r29	Freq_datasubmit	How often do you think you'll be able to submit data on food prices?	yes				field-list			
select_one ga3zc40	Datafeedback_interest	Are you interested in recieving information/aggregated data on market/prices based on the aggregated data from other volunteer crowd members?	yes							
select_one tp2rh06	Freq_DataFeedback	On which frequency are you interested in getting information/analysis on food prices based on the data provided by crowd?	yes			${Datafeedback_interest} = 'yes'	field-list			
select_multiple up6xw18	GeoScale_DataFeedback	At what geographical level would you like to receive the aggregate food prices?	yes			${Datafeedback_interest} = 'yes'	field-list			
select_one jh7ng91	DataFeedback_NoReward	If we are able to send you periodic data on food prices, would you be interested in providing food price data without any monetary/voucher reward?	yes			${Datafeedback_interest} = 'yes'	field-list			
select_one fm94f69	DataFeedback_conditions	Under what conditions would you be interested in receiving this data?	yes			${Datafeedback_interest} = 'yes'	field-list			
select_one kf23y09	DataFeedback_communication	In case we can provide the data/analysis you need, which is the most appropiate transmission/communication mode?	yes			${Datafeedback_interest} = 'yes'	field-list			
end group	**profile_group**									
note	consent_unapproved	Without your consent, we will not be able to enlist your participation in this project. If you erroneously selected "No", please use the back arrow to return to previous page and select "Yes", otherwise please proceed to the next pay to finalise and submit the form.		${Consent} = 'No'						
select one from	info_source	How did you first hear about this	yes							

type	name	label	required	constraint	media::image	relevant	appearance	calculation	Hint	body::accuracy Threshold
info_source		food price crowdsourcing initiative?								
note	note_end	Thank you for completing these quick questions. Please proceed to follow the submission instructions as highlighted in the SOP on the Project's website. If you are one of the first 100 submitters, we will send the FREE N500 voucher, and if you're among the first 1000 qualified candidates, we will contact you within the next one week with your VC_ID. Happy Submission								

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from the EU Bookshop at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

KJ-NA-30027-EN-N

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

Publications Office
of the European Union

doi:10.2760/366116

ISBN 978-92-76-14305-5