View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Dundee Online Publications

University of Dundee

Citizen Science Projects (MOOC) 3.6

Woods, Mel; Coulson, Saskia; Ajates, Raquel; Amditis, Angelos ; Cobley, Andy; Domian, Dahlia

Publication date:
2020

Link to publication in Discovery Research Portal

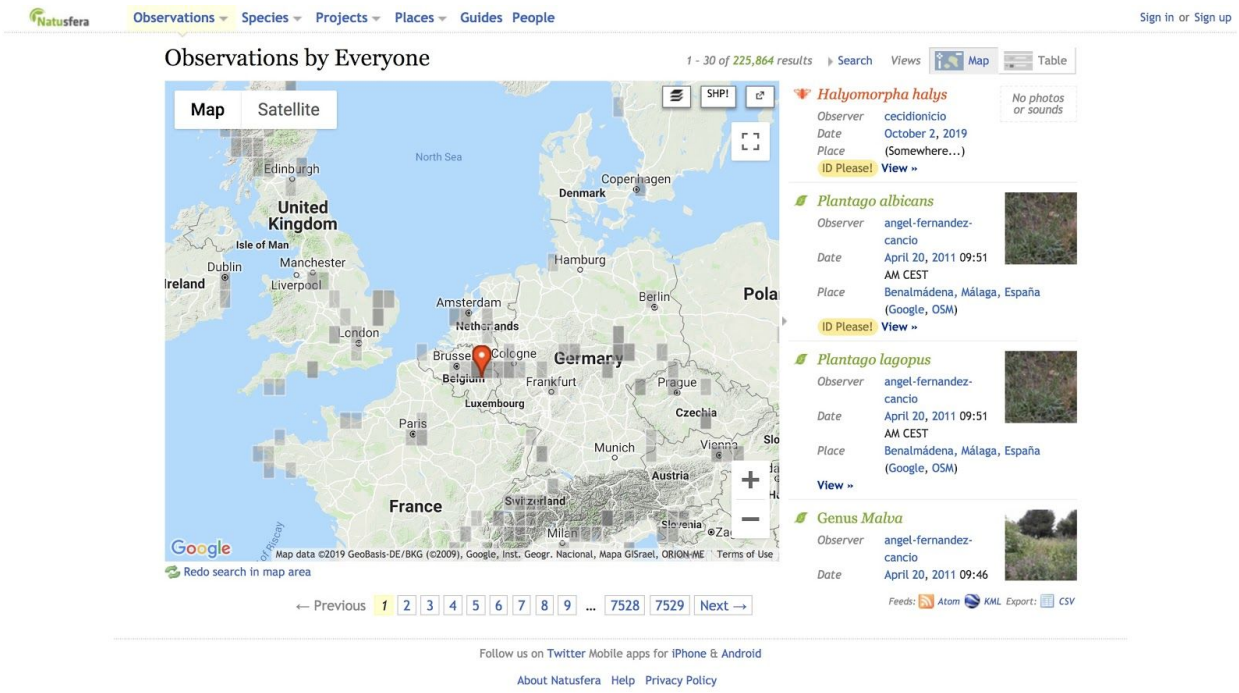
Citation for published version (APA):
Woods, M., Coulson, S., Ajates, R., Amditis, A., Cobley, A., Domian, D., Hager, G., Ferri, M., Fraisl, D., Fritz, S., Gold, M., Karitsioti, N., Masó, J., McCallum, I., Tomei, G., Monego, M., Moorthy, I., Prat, E., Tsertou, A., ... Wehn, U. (2020). Citizen Science Projects (MOOC) 3.6: Options for data analysis: tools & tips. WeObserve.

General rights
Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 20. Apr. 2021



Analysing citizen science data does not have to be different than analysing other types of data. In the previous step [(3.5 How to understand and interpret data)](https://www.futurelearn.com/courses/weobserve-the-earth/3/steps/870657) we discussed ways that help you gain a clear understanding of your dataset. Here we will provide some tips and suggested tools to help you analyse your data. Offering data visualisations and data analysis tools can help to maintain citizen motivation and interest in the project, while also providing powerful information also for researchers or policymakers.

First thing in data analysis is allowing the data to be downloaded and analyzed by the users themselves. Like in the previous step, you can make your data available in several formats: CSV, XLS, KML, or even in more structured ways like the [Sensor Observation Service (SOS)](https://www.ogc.org/standards/sos) or an [Application Programming Interface (API)](https://medium.com/@perrysetgo/what-exactly-is-an-api-69f36968a41f). These will help make data available, shareable and interoperable with other data sources.

[WeObserve](https://www.weobserve.eu/) is looking at ways to share citizen science data with a Community of Practice focused on Interoperability (the ability of computer systems or software to exchange and use information). We can see a example of this in the [Ground Truth 2.0](https://gt20.eu/) project, where a [map browser](http://www.ogc3.uab.cat/gt20/) was created which offers data from the six different observatories in the project from the same access point. In this case, protocols and rules to standardise the data from the [Open Geospatial Consortium (OGC)](https://www.ogc.org/) were used to enable interoperability.

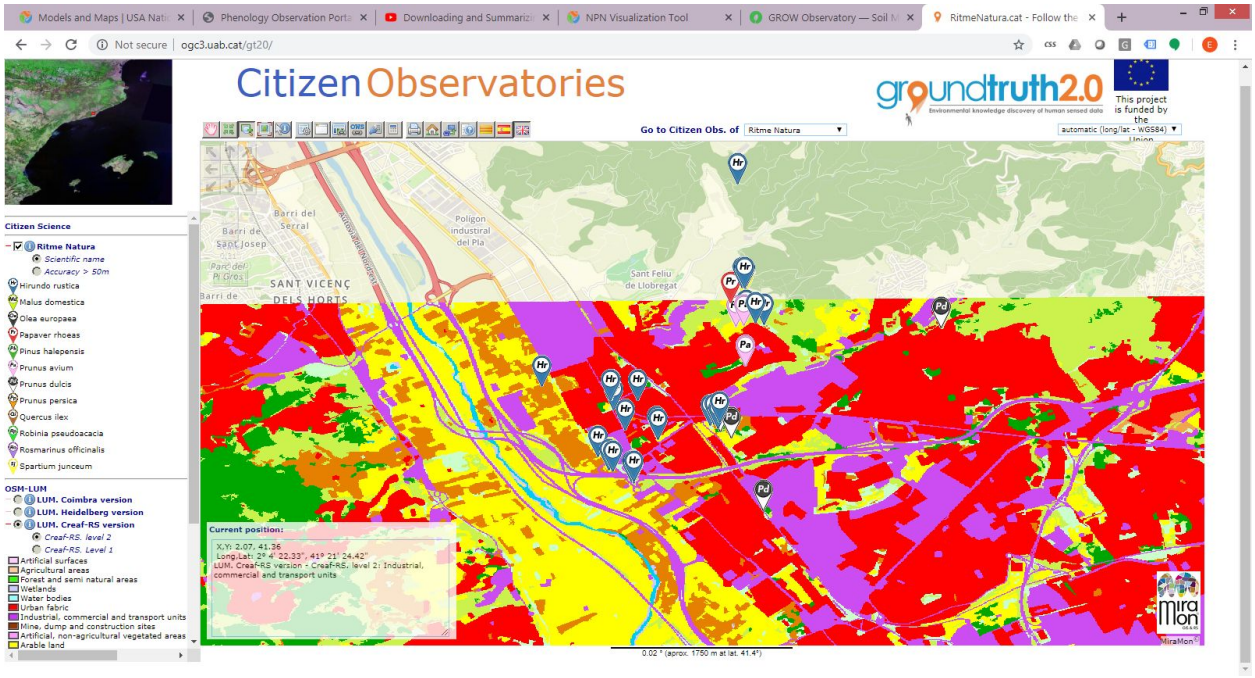
Sometimes it is not easy to get a clear picture of the information, and additional analysis is needed so researchers or policy makers can make concrete conclusions and present their findings. In these cases, it helps to ask experts or researchers in the field to make recommendations on the best way to process and present the data. Simple and easy to interpret graphs or maps are always a good option.

Still, using data analysis approaches on citizen science platforms can be challenging because it may require manual or assisted interpretation that doesn't happen automatically. This might mean that you need to check your data in order to ensure quality control and find inaccurate values or possible outliers. Such wrong values should be removed before analysing, like you would do with any other kind of data. This quality control will also help to avoid or detect a possible systematic deviation in the monitoring process.

In the last step, we discussed the visualisation of data and how sometimes maps are used to represent the location of the observations. There are many examples out there, for example, the [Natusfera](https://natusfera.gbif.es/) platform, or the [GroundTruth 2.0](http://www.ogc3.uab.cat/gt20/) map.

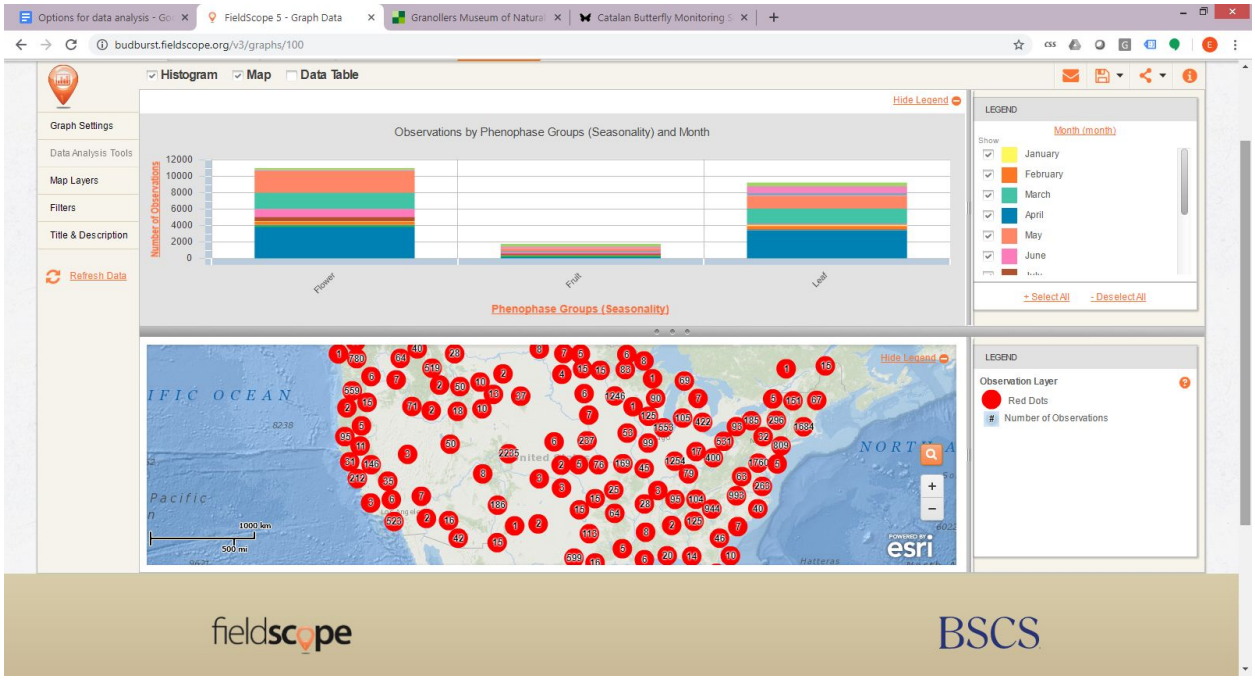
One of the first things you can do when starting data analysis is to use complementary information. For example, you could show the collected observations on a map or combine the data with additional measurements from other sources that can help to explain the distribution or trends of your data. In the image below, for example, you can see observations from the [RitmeNatura Citizen Observatory](http://ritmenatura.cat/). This shows the seasonal changes of vegetation species in combination with a land use map of the area, which helps in illustrating the kind of environment where those species live. This additional information can help users to find trends in data collected and give a broader information to the single observations.

In the example above, these same observations could be combined for example with other external data about extreme or average temperature if available to help interpret the seasonal changes of vegetation according to climatic conditions . This could be done by cross-referencing the location of the observations with a temperature map or by adding the corresponding temperature to each observation thus obtaining combined information for each data point.



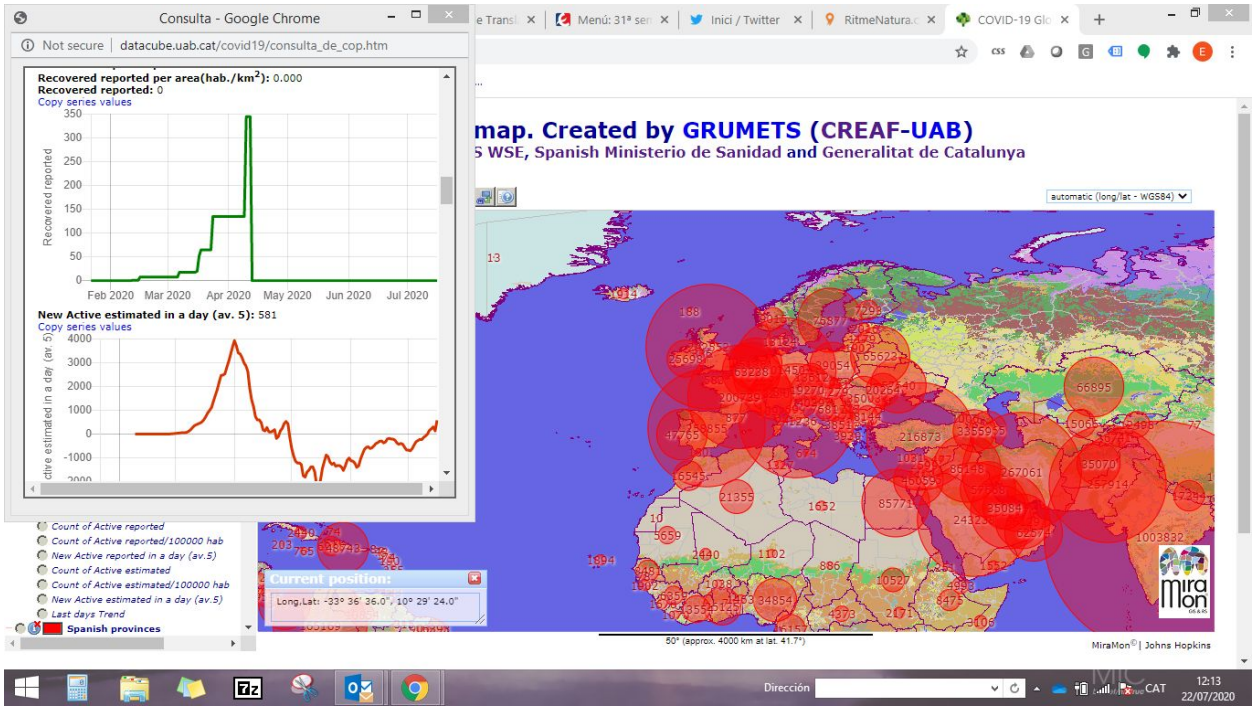
[Citizen science observations for seasonal changes in plant species shown in combination with land use information and without it. When combined with additional data, patterns of distribution can easily be detected.](https://ugc.futurelearn.com/uploads/assets/f4/02/f402a40e-d06d-4184-9222-28cb4d57cf11.png) ©Ground Truth 2.0 [visualisation and data quality tool](http://www.ogc3.uab.cat/gt20/)

Maps and graphs are easy to create and a helpful way to represent data. When you combine data with additional information, you can easily create visualisations like scatter plots, calendars, activity curves, charts and maps. Graphs can be pre-defined or created by users on the fly (if they can define the data points they want to combine). In the previous step, we mentioned some useful tools have been used to play with data. Here we offer you two more tools that you can use to create map and graph visualisations of your data. [Google Maps](https://www.google.es/maps/?hl=es) and [Google Charts](https://developers.google.com/chart) are cheap and easy tools to help observatories do analysis. [FieldScope](http://www.fieldscope.org/) is a powerful platform that provides mapping and analysis tools for citizen science.



[These are map and graph visualisations provided by FieldScope for the BudBurst project.](https://ugc.futurelearn.com/uploads/assets/01/d0/01d04090-7e1d-41c0-90f7-cfea66d1b326.png) ©[FieldScope](http://www.fieldscope.org/)

Other examples of data analysis is the COVID-19 datacube created with the [MiraMon Map Browser](https://github.com/joanma747/MiraMonMapBrowser) that offers a lot of analytical functions that could also be applied to [citizen science data](http://datacube.uab.cat/covid19/). As you can see in the image, data are viewed in a visual format and several graphs are provided when data are consolidated.



To summarise, we recommend using simple tools for data visualisation and options for evaluating data such as maps, graphs, statistical data, charts, data summaries, and, of course, the possibility to download data for further analysis.

##Share your thoughts!

Have a look at the [Field Scope Visualisations Page](http://www.fieldscope.org/visualizations/) to get a sense of the different options for the way data can be analysed.

+ Which do you find most informative, and why?
+ Would you use any of these options in your project?

Please share your thoughts below.