Hazard Analysis and Critical Control Point System: A Compliance Analysis of Food Safety in Portugal

Tatiana Sofia Branco de Sousa Pinto

09/**2019**

Hazard Analysis and Critical Control Point System: A Compliance Analysis of Food Safety in Portugal

Tatiana Sofia Branco de Sousa Pinto

Orientador(es)

Professora Doutora/Brígida Mónica Faria/Escola Superior de Saúde-Instituto Politécnico do Porto e Laboratório de Inteligência Artificial e Ciência e Computadores

Professor Doutor/Luís Paulo Reis/Faculdade de Engenharia da Universidade do Porto e Laboratório de Inteligência Artificial e Ciência de Computadores

Dissertação apresentada para cumprimento dos requisitos necessários à obtenção do grau de Mestre em **Bioestatística e Bioinformática Aplicadas à Saúde** pela Escola Superior de Saúde do Instituto Politécnico do Porto.

P.PORTO

**ESCOLA
SUPERIOR
DE SAÚDE**

Agradecimentos

À Professora Doutora Brígida Mónica Faria, minha orientadora de mestrado, pela permanente disponibilidade, interesse e constante apoio durante todo o processo de desenvolvimento do presente projeto de dissertação, assim como pelas oportunidades de divulgação científica que proporcionou. Sem a orientação e a disponibilidade e aconselhamento constantes da Professora nada disto seria possível de concluir.

Ao Professor Doutor Luís Paulo Reis, meu coorientador de mestrado, pela simpatia, preocupação e amabilidade em ter permitido ter feito parte do projeto IA.SAE.

A todos os elementos do LIACC, mais concretamente ao Doutor Tiago Santos, pela paciência em procurar comigo todos os erros de código e percalços que iam aparecendo durante esta dissertação.

À Soraia e ao Gustavo, meus companheiros e amigos de mestrado pela constante disponibilidade em me ajudar e apoiar nesta aventura de abertura do Mestrado em Bioestatística e Bioinformática Aplicadas à Saúde.

Ao Campelos e ao Filipe, os meus melhores amigos, por estarem sempre ao meu lado na vida académica e a partir de agora para o resto da vida. Sem eles nunca teria coragem de me aventurar num mestrado tão diferente de tudo o que aprendi durante a minha licenciatura, porque caso algo corresse mal, eles estariam prontos para me apoiar e encaminhar, como sempre fizeram.

Ao meu irmão que, apesar de nem sequer saber o nome do mestrado que frequento, está sempre na primeira fila a apoiar-me e a gabar-me.

Aos meus pais pelos valores e princípios que me transmitiram, e por me proporcionarem todas as condições para estar onde estou hoje.

Sempre defendi que os filhos são o reflexo dos pais que têm, e se hoje estou orgulhosa do meu percurso a vocês o devo.

Sou imensamente feliz por vossa causa e sei que para vocês não há maior agradecimento que este. Adoro-vos muito.

Resumo

Atualmente, a sociedade moderna está mais preocupada e alerta sobre perigos como a falta de segurança, a higiene e o controle de sistemas e serviços que podem afetar o consumidor. O sistema de Análise de Perigos e Ponto de Controlo Crítico (HACCP) é baseado numa metodologia preventiva, para evitar potenciais riscos que podem causar danos aos consumidores, eliminando ou reduzindo estes perigos. Este sistema tenta garantir que os alimentos não seguros não sejam disponibilizados ao consumidor.

A Autoridade de Segurança Alimentar e Económica (ASAE) é um órgão de polícia criminal especializado nas áreas de segurança alimentar e inspeção económica em Portugal. É responsável pela avaliação e comunicação de riscos na cadeia alimentar, bem como pelo cumprimento legal de atividades económicas nos setores alimentar e não alimentar. O sistema HACCP é reconhecido mundialmente e desenvolvido para evitar riscos potenciais ao consumidor sendo também seguido pela ASAE. A ASAE gera um grande e complexo volume de dados nas diversas inspeções que realiza, na receção de reclamações, na sua classificação, no seu registo e na monitorização até o final do processo.

Os dados analisados nesta dissertação referem-se a reclamações sobre entidades e pedidos de informações de 2014 a 2018.

Nos anos de 2014 a 2018, fazendo uma análise estatística com o Python, pode-se verificar que as reclamações relacionadas ao sistema HACCP são maiores do que em 2014, 2017 e 2018. A percentagem relativa à segurança alimentar é superior à segurança económica em três anos, entre 2014 e 2018.

Assim sendo, uma vez que o sistema HACCP é responsável por garantir a segurança alimentar, este estudo torna-se relevante devido à sua análise de dados, expressa com medidas estatísticas e visualização gráfica de reclamações relacionadas a este sistema.

Palavras-chave: Saúde Pública; Segurança Alimentar; Aplicações de Análise de Dados; HACCP

Abstract

Nowadays, modern society is more concerned and alert about the dangers, lack of security, hygiene, control of systems and services that can affect the consumer. The Hazard Analysis and Critical Control Point (HACCP) system is based on a preventive methodology to avoid potential hazards that may cause harm to consumers by eliminating or reducing these hazards. This system attempts to ensure that unsafe food is not made available to the consumer.

The Food and Economic Security Authority (ASAE) is a criminal police body specialized in the areas of food security and economic inspection in Portugal. It is responsible for risk assessment and communication in the food chain, as well as legal compliance with economic activities in the food and non–food sectors. The HACCP system is recognized worldwide and designed to avoid potential consumer risks and is also followed by ASAE. ASAE generates a large and complex volume of data in the various inspections it performs, in receiving complaints, classifying them, registering them and monitoring them until the end of the process.

The data analyzed in this dissertation refer to complaints about entities and information requests from 2014 to 2018.

In the years 2014 to 2018, making a statistical analysis with Python, it can be seen that complaints related to the HACCP system are higher than in 2014, 2017 and 2018. The percentage relative to food security is higher than the economic security in three between 2014 and 2018.

Therefore, since the HACCP system is responsible for ensuring food safety, this study becomes relevant due to its data analysis, expressed with statistical measures and graphical visualization of complaints related to this system.

Keywords: Public Health, Food safety, Data Analysis Applications, HACCP

Abbreviations and Acronyms

AI– Artificial intelligence

ASAE– Food and Economic Security Authority

CCPs– Critical Control Points

CDA– Confirmatory Data Analysis

CL– Critical Limit

DM– Data Mining

DMT– Data mining techniques

EC– European Community

EDA– Exploratory Data Analysis

FMEA – Failure, Mode and Effect Analysis

HACCP – Hazard Analysis and Critical Control Point

IA.SAE– Artificial Intelligence in Economic Security Authority

ID– Identification Number

KDD– Knowledge Discovery in Databases

LIACC– Laboratory of Artificial Intelligence and Computer Science

NASA – National Aeronautics and Space Administration

Pandas– Python Data Analysis Library

RAPEX– Rapid Alert System for Food and Feed

RASFF– Rapid Alert System for Food and Feed

TQM – Total Quality Management

Table of Contents

Figures

Tables

1. Introduction

This chapter presents the motivation to carry out this dissertation project in the context of the combination of food safety areas related to HACCP and Data Analysis, intrinsically related to the exploratory statistical analysis. Following are the objectives to be achieved, according to the organizational and academic component inherent to the project, and the organization of this document is presented.

1.1 Motivation

The Food and Economic Security Authority (ASAE) has followed with special attention the growth of technological innovation in Information Technologies and Systems, with the capacity to control and store the data produced in the daily tasks of all its stakeholders. ASAE aims improving public intervention and assistance to the population where this technological innovation may help.

ASAE has followed, with special attention, the growing of technological innovation and today has created the need to have the capacity to control and store the data produced in the daily tasks of all its stakeholders, to improve public health and automation of inspector's tasks. ASAE generates a large and complex volume of data in the various inspections it performs, in the reception of complaints, in its classification, in its registration and the monitoring until the end of the process. This factor enhances the use of specialized storage tools capable of transforming data into information that positively influences decision making on inspection days.

According to this expectation and with the predominant characteristic of the large volume of data, the techniques and tools for the analysis of these arise as a response, more concretely, through Data Mining. With the constant growth of databases appeared the need to organize all that information and improve the inspection system that already exists (3). Data Mining and Data Analysis techniques have achieved impressive results in several industries (5). Also, Bioinformatic and Biostatistics techniques spread and have more applications in several areas related to public services and economic sectors(6). Therefore, ASAE responsible for the Economic and Food Security area may take advantage of advances in these fields. To extract knowledge from data and provide the best solutions to ASAE objectives it was purposed the Artificial Intelligence in Economic Security Authority project (IA.SAE project) (7). This project involves two major organizations: ASAE and LIACC (Laboratory of Artificial Intelligence and Computer Science) and has the collaboration of the School of Health – Polytechnic of Porto. Through the data provided by ASAE, as well as through information gathered from all meetings with ASAE members, the project aims to obtain the maximum knowledge about the topics involved. Moreover, it has the objective to develop a less

dubious method of entry for who will file the complaint or even through the ASAE portal and develop a geo-referencing application to facilitate inspections in the field to the inspectors (7).

The easy and intuitive completion of a food-related complaint means there is no time-wasting for any further requests from the complainant for information or errors in filing the complaint. Thus, there is a faster response by inspectors to verify the food safety conditions of the establishment, thereby improving public health. To achieve formal declarations, the organization intends to use Data Analysis techniques, since it recognizes its capacity to respond to a project with the characteristics of IA.SAE.

The theme of this dissertation is inserted in the IA.SAE project focusing on the reported entities that are related to non-compliance with HACCP and from it try to understand the most common violations.

To understand from the many data provided by ASAE those that are related to non-compliance with the HACCP system and to analyze its recurrence is the motivational factor of this dissertation since the HACCP system is directly related to public health (8).

Due to the large volume of data that will be used in this work to analyze the compliance of the HACCP system, it is easier to conclude with visualization of data, using for example graphical representations (9).

The present chapter starts with this motivation combining the compliance of HACCP system in the food sector with Data Analysis techniques to extract more information and intrinsically related to the IA.SAE project (to be better explained in chapter 2). Next, it will be presented the main and specific dissertation objectives.

The environment involved in ensuring food and economic security generates a large and complex volume of data. This data can be generated by the interactions with complainants, in the designation of type of data entry, in the assessment of the type of complaint, in the investigation involved in the complaint until its archiving. This factor is a potentiator of the use of exploratory analysis tools in order to extract knowledge. According to this expectation and with the predominant characteristic of the large volume of data, the techniques and tools of the Database Knowledge Discovery Process arise as a response, more concretely, through Data Analysis.

Data Analysis techniques have achieved impressive results in several industries. As such the Food Safety sector needs to take benefits of advances in this exciting field.

Artificial Intelligence and Computer Science Laboratory (LIACC), characterized by its avant-garde vision and through an innovative project, intends to develop a support platform that facilitates the work of the ASAE inspectors when deciding the entities to inspect in a certain area.

LIACC intends, through the number of associated complaints, as well as through information acquired from the geographic location, to provide inspectors with the shortest and most efficient route to be able to analyze the maximum number of entities in the shortest time possible. By improving the efficiency of inspector routes, the number of inspected entities increases, increasing the guarantee of food safety for consumers.

Understanding how complaints related to food safety, more specifically with the non-compliance of the HACCP system received by the ASAE have significance in relation to all complaints received, is a factor that is the main motivation for the development of this study.

1.2 Dissertation Objectives

This dissertation main objective is the development of a Data Analysis procedure capable of analyzing data from ASAE database focusing on exploratory data analysis. To enable this global objective, the research performed comprises several other objectives:

- Characterize HACCP system and all the rules and processes involved in its application;
- Understand several possible ways that individuals in different countries have to report complains to competent authorities;
- Perform a data analysis using Descriptive statistics in order to better understand and characterize the available data from ASAE;
- Do an exploratory data analysis in all the files given by ASAE;
- Show the evolution of complains between 2014 and 2018;
- Analyze complaints related to the HACCP system;
- Relate the number of denunciations of the districts and islands of Portugal in relation to their population density.

The theme of this dissertation is inserted in the IA.SAE project focusing on the reported entities that are related to non-compliance with HACCP and from it try to understand the most common violations.

1.3 Organization of The Dissertation

This document is organized as follows:

- Introduction – this chapter intends by a brief introduction to the themes that make up the dissertation project, to describe some factors that motivated the development of the study.

- Background and Related Work– where will be put in context the concepts related to this dissertation

- An explanation to elucidate the definitions related to HACCP System, such as scrutinizing the process that have been implemented since a complaint is done until inspection by ASAE's agents;

- A description of what ASAE represents, the organizations with which it cooperates in the world, how ASAE works, and an explanation of the involvement of this food and economic security authority with this dissertation.

- An insight into Big Data, the types of data that exist, and the visual tools that enable knowledge extraction with this data type

- Methodology – this chapter explains the methods used during the practical application of the dissertation project, as well as the bibliographic research strategy for the literature review;

- Results– This chapter presents the results obtained from the study done in this dissertation.

- Analysis and Discussion of Results – This chapter presents an analysis and reflection the results

- Conclusions and Future Work – This chapter presents a conclusion about the theme of study of this dissertation and what can be done in the future using the developed work.

2. Background and Related Work

This chapter aims to elucidate, through bibliographic research, concepts and entities relevant to this dissertation. In this chapter the theoretical themes of this dissertation are explained and put into context.

A presentation, exemplification and clarification of the topics of interest are made, which in this case are the system under study which is the Hazard Analysis and Critical Control Point (HACCP); the Food and Economic Safety Authority (ASAE), the entity that made the data available and the given data type, Big Data.

2.1 HACCP System

Hazard Analysis and Critical Control Point (HACCP) system is based on a preventive methodology, in order to avoid potential hazards that can cause harm to consumers by eliminating or reducing problems, and to ensure that unsafe food is not made available to the consumer. The HACCP system is based on the application of technical and scientific principles in the production and handling of food aspects, from the beginning of their production to the final disposal to the consumer.

2.1.1 Historical and Conceptual Context

According to FAO (Food and Agriculture Organization) in 1998, the Hazard Analysis and Critical Control Point (HACCP) had its origins in the 50s / 60s (1). In the 1950s, there was a revolution in the quality of Japanese products; the improvement of the quality of these products was attributed to the application of Dr. W. E. Deming's theories of Quality Management (1). He and other researchers developed the concept of Total Quality Management (TQM), in which emphasis is given to the system's vision as in order to minimize manufacturing costs and improve product quality (2).

In 1959, the company Pillsbury joined with United States Army Research Laboratory to the National Aeronautics and Space Administration (NASA) program as a contractor and began working on cube-sized foods for space flight (3). The program was producing foods that would not crumble in zero gravity and provide adequate nutrition was only one of the problems; food safety was another (4). However, it became difficult to ensure that 100% of these foods would not be contaminated which could cause an aborted space mission (5).

Inspections of the ingredients, products and laboratory analysis did not ensure that all production was safe (5). Thus, there is a need to adopt a preventive approach to the quality control and food safety system (5). In pursuit of this new approach, Pillsbury found that the Natick Soldier Systems Center used an analysis system for medical supplies – "Failure, Mode and Effect Analysis"– FMEA

(6). The FMEA system consists of finding out what can go wrong at each stage of an operation, as well as the causes and effects resulting from these potential failures, before implementing effective control mechanisms (7). After evaluating the FMEA method, Pillsbury adopted this analysis model, inserted some modifications and the concept of HACCP was created (8).

In 1971, Pillsbury presented the concept of HACCP at a food protection conference, the American National Conference of Food Protection (9). Following this conference, the Food and Drug Administration (FDA) hired Pillsbury to provide professional training in HACCP Systems to its staff (10). In 1973, Pillsbury published the first HACCP document (11). In 1985, the National Academy of Science – NAS, United States, published a paper: An Evaluation of the Role of Microbiological Criteria for Foods and Food Ingredients, which concluded that a preventive system (HACCP) was essential for the control of microbiological hazards and that inspection of the final product was inadequate for the prevention of food–borne diseases (12).

In 2006, Regulation (EC) No 852/2004 on the hygiene of food kinds stipulates that all food operators should establish, implement and maintain a permanent process or processes based on the 7 principles of HACCP (13):

- Principle 1 – Conduct a Hazard Analysis

The application of this principle involves listing the steps in the process and identifying where significant hazards are likely to Occur. The HACCP team will focus on hazards that can be prevented, eliminated or controlled by the HACCP plan. A justification for including or excluding the hazard is reported and possible control measures are identified.

- Principle 2 – Identify the Critical Control Points

A critical control point (CCP) is a point step or procedure at which control can be applied and a food safety hazard can be prevented, eliminated or reduced to acceptable levels. The HACCP team will use a CCP decision tree to help identify the critical control points in the process. A critical control point may control more than one food safety hazard or in some cases more than one CCP is needed to control a single hazard. The number of CCP's needed depends on the processing steps and the control needed to assure food safety.

- Principle 3 – Establish Critical Limits

A critical limit (CL) is the maximum and/or minimum value to which a biological, chemical, or physical parameter must be controlled at a CCP to prevent, eliminate, or reduce to an acceptable level the occurrence of a food safety hazard. The critical limit is usually a measure such as time, temperature, water activity (Aw), pH, weight, or some other measure that is based on scientific literature and/or regulatory standards.

- Principle 4– Monitor CCP

The HACCP team will describe monitoring procedures for the measurement of the critical limit at each critical control point. Monitoring procedures should describe how the measurement will be taken, when the measurement is taken, who is responsible for the measurement and how frequently the measurement is taken during production.

- Principle 5 – Establish Corrective Action

Corrective actions are the procedures that are followed when a deviation in a critical limit occurs. The HACCP team will identify the steps that will be taken to prevent potentially hazardous food from entering the food chain and the steps that are needed to correct the process. This usually includes identification of the problems and the steps taken to assure that the problem will not occur again.

- Principle 6 – Verification

Those activities, other than monitoring, that determine the validity of the HACCP plan and that the system is operating according to the plan. The HACCP team may identify activities such as auditing of CCP's, record review, prior shipment review, instrument calibration and product testing as part of the verification activities.

- Principle 7 – Recordkeeping

A key component of the HACCP plan is recording information that can be used to prove that the food was produced safely. The records also need to include information about the HACCP plan. Record should include information on the HACCP Team, product description, flow diagrams, the hazard analysis, the CCP's identified, Critical Limits, Monitoring System, Corrective Actions, Recordkeeping Procedures, and Verification Procedures.

2.1.2 Prerequisites

In order to prevent, eliminate or only reduce the hazards that may contaminate the food during its production process and subsequent distribution, requirements which, once contemplated, have led to the effective application of the HACCP system should be taken into account (1).

The prerequisites control the hazards associated with the surrounding environment in the food production process, while the HACCP system controls the hazards associated with the production process (20).

The prerequisites are set out in Annex II of Regulation (EC) No 852/2004 of 29 April, and its application is not applicable by paragraph a) of paragraph 1 of Article 6 of Decree-Law No. 113/2006 of June 12 (21).

2.1.2.1 Structure and Equipment

The premises of establishments where food is produced or handled must be located in places where there are no industries or other agents likely to present contamination problems (22). The area should be cleaned, favoring pest control and the necessary infrastructures should be available – electricity, gas, sewage system, potable water supply, telephone network and road access (1). Premises must be designed in such a way that enables a continuous flow of operations and a level of hygiene under operational conditions appropriate to the production/marketing of safe food (23). Lighting systems, air treatment (filtration, temperature and humidity) appropriate to the operations carried out, as well as drainage systems on the floors, should be designed where necessary (8). Areas where incompatible activities are carried out (which may lead to cross-contamination) must be physically separated and in order to facilitate product flows, waste and people not presenting contamination hazards (2). Specific areas should be provided for receiving raw materials, storing raw materials, storing packaging materials, storing cleaning products, preparing cleaning solutions, washing equipment and utensils, producing, storing final product and shipping, facilities sanitary facilities, locker rooms, offices and social areas and temporary storage areas for waste (24). Information devices must be provided for restricted access to the work area and emergency exits (25).

The equipment, utensils and interfaces with the accessories have been influenced by the functions of the people for whom they were designed, safe for operations, protect food independently and do not constitute their own, a source of chemical, physical or microbiological contamination (17). The equipment shall be designed, constructed and installed in such a way as to promote proper maintenance, sanitation and inspection (26). The building materials must be resistant, easy to clean and non-toxic, and the equipment must be hygienically designed (27).

2.1.2.2 Hygiene Plan

One of the key measures to ensure food safety is undoubtedly an effective hygiene program for installations, equipment, utensils and surfaces that directly or indirectly contact food products (28). The hygiene plan aims to remove all dirt from surfaces, equipment and utensils (8). The hygiene plan must, in addition to this information, indicate the periodicity, the person responsible for the hygiene, the person responsible for checking the hygiene and their records (29). It should also elucidate the hygiene priorities, which surface/equipment should be sanitized first and which ones should be considered next (1).

Since most micro-organisms have the capacity to adapt, the disinfectant products to be used should be rotated in such a way that they do not become resistant to the products (30). Good hygiene requires efficient and regular cleaning of establishments, equipment and utensils (31). All equipment and utensils that contact food must be cleaned after each use and/or at the end of each working period and whenever warranted (32).

After the definition of the chemicals to be used in the cleaning and disinfection operations, a hygiene plan should be established, which should include, among others, items such as: what should be cleaned (zones, structure/equipment), what should cleaner/disinfectant (cleaner/disinfectant used), when it should be cleaned (hygienic periodicity), how it should be cleaned (equipment used for cleaning / disinfection and cleaning / disinfecting instructions) and who should clean (responsible for performing the operation) (23). The hygiene plan should be posted in a visible place (33).

The use of a cleaning record shows the application of the established Sanitation Plan, and this record should include, among others, items such as: date, products that were used (detergent / disinfectant that were used), how it was cleaned (cleaning accessories which were used) and who cleaned (person responsible for cleaning) (16).

2.1.2.3   Pest Control

At all stages of production, processing, storage and distribution, food must be protected from contamination that would render it unfit for human consumption, harmful to health or contaminated in such a way that it would not reasonably be expected to be consumed in that state (30).

Pests include outbreaks of animal species which, when present in a food establishment, may contact and contaminate food and may lead to the existence of unsafe products (34). The most common plagues are rodents (mice and rats), crawlers (ants and cockroaches flying insects (flies, mosquitoes and moth), birds (pigeons and sparrows) and less frequent dogs and cats (2). Pest infestations can occur in places that favor their proliferation and where food exists (27). They can enter the premises through their transport, along with the raw materials and materials, through the transport vehicles, equipment and / or manipulators (35).

Good hygiene practices should be employed in order to avoid creating a favorable environment for pest development (36).

2.1.2.4 Water supply

Food business operators shall provide an adequate supply of potable water as provided for in Chapter VII of Annex II to Regulation (EC) No 852/2004 and its amendments (21).

Decree-Law no. 306/2007, of August 27, amended by Decree-Law no. 92/2010, of July 26, establishes the water quality regime for human consumption, with the objective of protecting human health from harmful effects resulting from the possible contamination of this water and ensure the universal availability of clean, clean and desirably balanced water in its composition (37). If the water is distributed by a water supply management entity, the food company will only have to certify (and be able to show) that water meets these requirements (38). This can be achieved by requesting the management entity the results obtained in the analytical control of the water of the supply zone to which the company belongs (23). Management entities are obliged to provide these results to users of the water it supplies.

2.1.2.5 Waste Management

Companies should implement waste management systems appropriate to the size and type of activity, considering the steps of segregation, storage and routing of the different types of waste produced (8). The accumulation of waste should not be allowed in the areas of reception, storage, handling, dispatch or marketing of food products, as well as areas adjacent to these activities which could provide an environment conducive to product contamination (29).

The company should identify each of the types of waste generated and predict the respective quantities produced. With this information you will be able to select companies suitable for your treatment / referral and establish service contracts (21). Particular attention should be paid to the frequency of collection and the appropriate locations for temporary storage of waste (2). These should be stored in sealed bags, in appropriate containers with lids (2). Sanitation plans should include hygiene of these places and their containers at a periodicity appropriate to the absence of pests or odors (39).

Residue collection circuits should also be established in order to minimize or eliminate the likelihood of cross-contamination (30). This should take into account not only the physical space but also the time lag of the most critical operations in terms of exposure of food to this type of contamination (30). It should be noted that the facilities must be provided with suitable containers for the disposal of waste, lined with plastic bags and with a foot operated lid (24).

2.1.2.6 Materials in contact with food

Specific areas should be provided for receiving raw materials, storing raw materials, storing packaging materials, storing cleaning products, preparing cleaning solutions, washing equipment and utensils, producing, storing final product and shipping, facilities sanitary facilities, locker rooms, offices and social areas and temporary storage areas for waste (24).

Information devices must be provided for restricted access to the work area and emergency exits (25).

The equipment, utensils and interfaces with the accessories have been influenced by the functions of the people for whom they were designed, safe for operations, protect food independently and do not constitute their own, a source of chemical, physical or microbiological contamination (17).

The equipment shall be designed, constructed and installed in such a way as to promote proper maintenance, sanitation and inspection (26). The building materials must be resistant, hygienical and non-toxic, and the equipment must be hygienically designed (27).

2.1.2.7 Personal hygiene

The human being is one of the major sources of food contamination by microorganisms, since many of them live and develop in several places of the human organism, namely hair, nasal mucosa, mouth, throat, intestinal tract, hands, nails and skin in general (40).

Those who work with food should understand hygiene as a way to protect their health and the health of consumers (41). Thus, in each company and according to the activity to be developed, rules of hygiene, conduct and health should be established, with a view to minimizing the spread of diseases by this way (42). These rules should normally include procedures for hand hygiene, body hygiene, uniforms, glove procedures, behavior while performing their functions, procedures for measuring and reporting individual health status (43). Nor should we forget the definition and disclosure of standards for visitors (42).

2.2 ASAE – Food and Economic Security Authority

The ASAE is an acronym for Food and Economic Security Authority (3). ASAE is specialized in the areas of food security and economic supervision (3). The ASAE is a criminal police body, under the Ministry of Economy and responsible for the evaluation and communication of risks in the food chain, as well as for the discipline of the exercise of economic activities in the food and non-food sectors (44). ASAE acts as an oversight and market control body in the following areas of

intervention: food safety, tourism and business practices, security of products and facilities and intellectual property and industrial property (44).

2.2.1 Historical and Conceptual Context

The origins of the ASAE go back to the Administrative Inspection Police section created in 1893 in the Civil Police of Lisbon, by decree of King D. Carlos I (45). The Administrative Inspection Police had already many missions similar to those of the current ASAE, being responsible for the combat economic crimes and public health (46).

As part of the reform of the police services of 1918, the Administrative Police (also referred to as the "Administrative and Sanitary Police") was created, succeeding the Administrative Inspection Police of Lisbon, but with a national scope (45). In 1931, the General Inspectorate of Food Inspection Services was created to defend public health, with the main objective of combating food falsification (47). The dispersion of the Civic Police in several police forces leads to the empowerment of the Public Security Police (PSP) in 1935, which also inherits the powers of the former Administrative Police and General Inspectorate of Food Inspection Services (45).

During World War II the goal was to coordinate the rationing of supplies and to fight economic crime, so in 1943 the General Intendency of Supply was created within the Ministry of Economy and assumes the economic surveillance powers that had stayed with the PSP in 1935(45). In 1965, the General Intendency of Supply was restructured and transformed into the General Inspection of Economic Activities (47).

After April 25, 1974, the General Inspection of Economic Activities is restructured, becoming the General Directorate for Economic Control(48) and in 1984 was renamed "Directorate General for Economic Inspection" (49). In 1993, as part of its modernization process, the Directorate General for Economic Inspection underwent a deep restructuring, returning to the designation of "General Inspectorate of Economic Activities" (49).

Over time, the high efficiency, demonstrated by the General Intendency of Supply, causes it to transfer control functions to sectors previously assigned to other public departments (50). In 2005, the IGAE absorbed the Directorate-General for Food Quality Supervision and Control and the Portuguese Food Safety Agency, as well as the supervisory functions of the Regional Directorates of Agriculture, the Directorate-General for the Protection of Crops and the Directorate-General for Fisheries, becoming the Food and Economic Security Authority (51).

ASAE is a criminal police body whose acronym stands for "Food and Economic Security Authority" that is responsible for enforcing legislation in the food sector and in the non-food sector (52). With

the constant growing of databases appeared the need to organize all that information and improve the inspection system that already exists (3). The ASAE's mission is to supervise and prevent compliance with the regulatory legislation for the exercise of economic activities in the food and non-food sectors, as well as the evaluation and communication of risks in the food chain (44). ASAE is the national connection body with its counterparts at European and international level (53). ASAE is the one of the few organizations in the world that brings economic security together with food security (54), therefore it is responsible for guarantee the compliance of HACCP system (4).

2.2.2 Organizations that cooperate with ASAE

In a globalized world, international relations play an increasingly important role. The establishment of relations of international proximity with other organizations is increasingly (55). As already mentioned, ASAE is one of the few organizations in the world responsible for food and economic inspection (54). However, there are organizations in the world that cooperate with ASAE and have similar functions (53).

The ASAE establishes a relationship with Portuguese-speaking countries and territories(56), Africa (57), America (58), Asia (59), Europe (60) and the Middle East (61). Cooperation between ASAE and other similar organizations at international level aims to develop bilateral relations and monitor the implementation of existing protocols, as well as to promote the conclusion of new protocols in international relations (62).

Articulating ASAE representation internationally, especially within the framework of the European Union, is important to maintain a regular dialogue with organizations relevant to ASAE action on multilateral issues of mutual interest, promoting knowledge, exchanges of experience, mutual understanding and trust (62).

It is also important to promote, monitor and develop cooperation within the community of Portuguese speaking countries, consolidating partnerships already established (62). Table 1 shows ASAE and other organizations cooperation around the world.

Table 1– ASAE and other similar associated organizations

Region	Countries	Organization Acronym	Organization	Association Year	Mission
Europe	Germany	BfR(63)(64)	Federal Institute for Risk Assessment	2015	Identifying risks and protecting consumer health
		BVL(65)	Federal Office of Consumer	2015	Make risk communication more transparent and

				Protection and Food Safety		manage risks before they escalate into crises.
	Spain	AECOSAN (66)	Spanish Agency of Consumption, Food Security and Nutrition	2018	Establish and perform all functions related to the protection of consumer health, ensuring the achievement and maintenance of safety at all stages of the food chain.
	France	ANSES(67)	National Agency for Sanitary Security and Food, Environment and Labor	Cooperation protocol in negotiation	Assess the health risks to clarify public action
Portuguese-speaking countries and territories	Angola	SIC(68)	Criminal Investigation Service	2016	Criminal Investigation and Investigation of Economic Activities
		IGC(69)	General Inspection of Commerce	2015	Supervise, monitor and evaluate the implementation of plans and programs approved for trade, as well as to comply with the principles and rules of organization, operation and interests of the organs and services of the Ministry of Commerce
		MINHOTUR (70)	Ministry of Hospitality and Tourism	2011	Propose the formulation, conduct, control, evaluate and implement the Executive's policy in the field of hotel and tourism.
		INADEC(71)	National Institute of Consumer Protection	2010	Promote the policy of safeguarding Consumer Rights, as well as coordinate and implement measures to protect, inform and educate, and support consumer organizations.
	Cape Verde	ARFA(72)	Regulatory and Oversight Agency for Pharmaceutical and Food Products	2014	Regulatory body in the area of medicinal products for human use, cosmetics, chemical products, biocides, foodstuffs intended for human and veterinary use, foods with functional properties and novel foods, supplements and food additives.
		IGAE(73)	General Inspection of Economic Activities	2014	Promote preventive and repressive actions in the area of anti-economic and public health infractions.

	Guinea Bissau	IGCA	General Inspection of Commerce and Handicraft	2017	Inspection of trade and trade in hand–made products
	Macao	IACM – CSA	Macao Civic and Municipal Affairs Bureau – Center for Food Security	2016	Monitoring compliance with standards and supervising the quality of foodstuffs
		DSE(74)	Directorate of Economic Services of the Government of the Macao Special Administrative Region	2016	Study, elaboration and execution of economic policy, in the scope of economic activities and intellectual property.
		CC–RAEM	Consumer Council	2014	It pronounces itself on the policies of defense of the consumer
	Mozambique	INAE(75)	National Inspection of Economic Activities	2010	Ensure compliance with legality in the exercise of economic activities and create a good business environment in the country
	Sao Tome and Principe	DRCAE(76)	Direction of Regulation and Control of Economic Activities	2016	To promote the regulation, control and inspection of all food and economic activities carried out on or related to the national territory
	Timor	AIFAESA(77)	Authority for Inspection and Supervision of Economic, Health and Food Activities	2012	Ensure that food quality control activities are carried out, as well as discipline in the exercise of economic activities in the food and non–food sectors.
	Brazil	ANVISA(78)	National Health Surveillance Agency	2014	To promote the protection of the health of the population by means of sanitary control of the production and consumption of products and services subject to sanitary surveillance, including environments, processes, inputs and related technologies, as well as the control of ports, airports, borders and bonded sites.
		INMETRO(79)	National Institute of Metrology,	Cooperation protocol in negotiation	Promote confidence in Brazilian society in measurements and products, through

		Quality and Technology		metrology and conformity assessment, promoting the harmonization of consumer relations, innovation and competitiveness of the Country.
		DIPOV(80)	Department of Inspection of Products of Plant Origin	2018	Planning and execution of the inspection and inspection of the production, circulation and commercialization of products of plant origin, as well as wine and grape and wine derivatives.
		DIPOA(81)	Department of Inspection of Products of Animal Origin	Cooperation protocol in negotiation	Inspection of animal products under the Ministry of Agriculture
Africa	Algeria	DGCERF(82)	General Directorate of Economic Control and Repression of Frauds	2015	Regulates all economic acts, companies and consumers
	Morocco	ONSSA(83)	National Office of Sanitary Safety of Food Products	2015	Competences among others in the food range and in the control and registration of veterinary medicines; It also controls food additives, packaging materials, products and materials that come in contact with food and fertilizers and irrigation water products
	Tunisia	Ministry of Commerce of Tunisia – Directorate–General for Cooperation and Trade(84)		2015	It oversees small businesses, as well as small food industries, chemical industries, machinery and small–scale electronics industry; clothing and leather
America	Canada	CFIA(85)	Canadian Food Inspection Agency	2014	It has competence in the area of food safety and works in partnership with industry, consumers and federal, provincial and municipal organizations
	USA	FDA(86)	Food and Drug Authority	2015	Protection and promotion of public health through the control and supervision of food safety, tobacco products, dietary supplements, prescription and pharmaceutical drugs,

					vaccines, biopharmaceuticals, blood transfusions, medical devices, electromagnetic radiation, cosmetics and animal feed and veterinary products
		USDA(87)	U.S. Department of Agriculture	2018	Develop and implement policies related to agriculture and livestock, promote trade in agricultural goods, ensure food security, protect natural resources and support rural communities
Asia	China	IETP(88)	Zhuhai Public Technical Service Platform for Import and Export	2017	Supervision of food activities
		SAIC(89)	State Administration for Industry and Commerce	2016	Responsible for market supervision and regulation
		CFDA(90)	China Food and Drug Administration	2016	Competence in the area of regulation for food and drug safety
		AQSIQ(91)	General Administration of Quality Supervision Inspection and Quarantine	2014	Is responsible for national quality, metrology, health inspection quarantine entry and exit of animals and plants, import and export food safety, certification and accreditation, standardization, as well as administrative law enforcement
		SZFDA(92)	Food and Drug Administration of Shenzhen Municipality	2015	Administrative and technical supervision of medicines, medical devices, health food and cosmetics and their respective quality control
		CFS(93)	Centre for Food Safety of Hong Kong	2015	Ensures that food sold in Hong Kong is safe and suitable for consumption through tripartite collaboration between government, food trade and consumers
	South Korea	NIFDS(94)	National Institute of Food and Drug Safety	2017	It is responsible for the protection of public health, ensuring the safety of food,

					medicines and medical devices
	Japan	FSC(95)	Food Safety Commission	2015	Risk assessment that has competence in the scientific evaluation of food safety risks to human health
	India	FSSAI(96)	Food Safety and Standards Authority	2018	Provision of healthy and safe food for human consumption, with skills at all stages of the food chain, from manufacturing, food storage, distribution and sale
Middle East	Saudi Arabia	SFDA(97)	Saudi Food and Drugs Authority	Cooperation protocol in negotiation	Regulatory Authority for Food, Medicines and Medical Devices which is primarily concerned with the safety of food for human and animal consumption and is also responsible for biological and chemical substances as well as electronic products
	United Arab Emirates	MCCE(98)	Ministry of Climate Change and Environment	2014	Responsible for improving environmental protection and preserving and developing natural resources
	Turkey	GDFC(99)	General Directorate of Food and Control	2015	Responsible for the areas of food safety as well as quality control of agricultural products

2.2.3 ASAE work methods

The inspection proceeds proactively or reactively according to criteria previously established in the ASAE Inspection Plan (100). ASAE acts when (101): there are non-conformities detected in the food available for sale to the final consumer, when complaints are made directly to the ASAE through the website, letter / fax, telephone or face-to-face contact, when there are emergency situations (eg food poisoning) or food crises, receive information / news or alerts received at ASAE through the Rapid Alert System for Food and Feed (RASFF) (102) or through the Rapid Alert System for Food and Feed (RAPEX), when national or community restrictive measures are taken for non-food products presenting a risk and requiring rapid intervention and when requests are made for collaboration from other authorities. Figure 1 presents a general workflow of complaints from report to finalization.
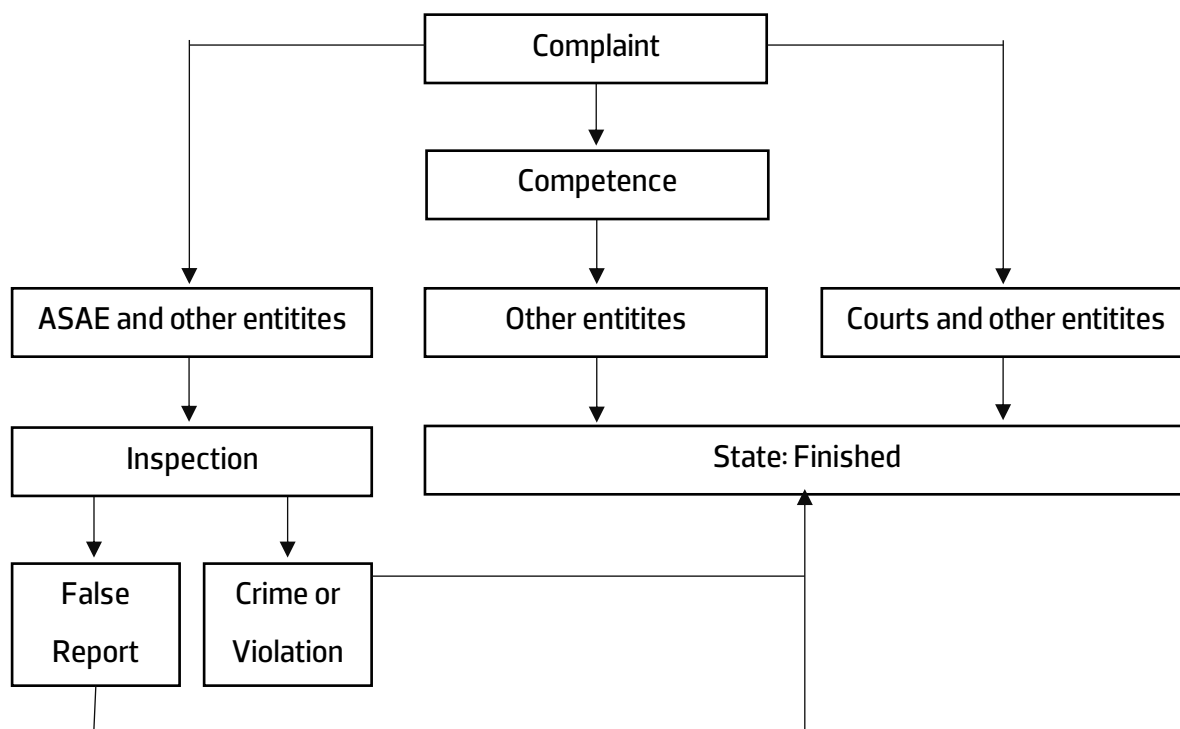
Figure 1 – From complaint to finalization

The ASAE receives the complaint through the site, letter, fax, telephone or face to face and it is ascertained who is responsible to respond to the complaint (101). If it is the responsibility of ASAE and also involve other entities, an inspection is made to evaluate if it is a false report, if it is filed. If there is a crime or violation measures are applied accordingly (103), the complaint is filed. If the initial assessment to investigate the entity responsible for responding to the complaint is sent to courts and other entities or only to other entities, the complaint is automatically filed.

2.2.4 IA.SAE project objectives

The easy and intuitive completion of a food-related complaint means there is no time-wasting for any further requests from the complainant for information or errors in filing the complaint. Thus, there is a faster response by inspectors to verify the food safety conditions of the establishment, thereby improving public health. To achieve formal declarations, the organization intends to use Data Analysis techniques, since it recognizes its capacity to respond to a project with the characteristics of IA.SAE.

ASAE recognized in the LIACC the opportunity to improve, through the IA.SAE project.

The main goal is to develop of models of risk analysis and selection of economic agents to be monitored, based on the latest techniques of Artificial Intelligence (AI) and Computational Learning and in the databases available and to be developed at ASAE.

For the risk analysis and inspection models to be created:

- Use all the information available in the ASAE databases to improve prevention in the areas of food security and economic surveillance;
- Use of techniques of fusion and extraction of information, learning machine, optimization, data and text mining, simulation of the behavior of diversified economic agents and intelligent visualization of the information.

Problems to be solved and modules to be developed:

- Intelligent and semi-automatic analysis of the complaints and denunciations received by electronic means and carry out their separation and prioritization;
- Analyze the notifications regarding the realization of sale and promotions s with semi-automatic and intelligent analysis. Intelligent inspection of sales/promotions with generation of georeferenced routes of inspection of undue/suspicious sales;
- Analyze and optimize (using machine learning) the global risk matrices based on consumption volume, default rate, product / service and other factors;
- Through the fusion of information related to each economic agent, its history and current information from multiple sources, and the simulation of the behavior of these economic agents, will be made an intelligent selection of economic agents to supervise and generate flexible georeferenced inspection plans;
- Create a configuration module and intelligent visualization of information;
- System will be fed with information / corrections of the operators and specialists and with results of each inspection so as to allow its continuous learning throughout its operation.

Measures should be included for data protection and safeguarding the ethical and legal aspects associated use of the data. With this effect, each element of the project signed a document to guarantee data protection.

The aim of this project is not to increase the number of criminal misconduct and prosecutions, but to achieve a real and clear awareness of the economic agents that the violation will be more easily identified, thus allowing an increase in national food and economic security.

2.3   Big Data

The exponential increase in the volume of data has led to an information and knowledge revolution (104). It is now a key aspect of research and strategy building to gather meaningful information and insights from existing data (105).

The term Big Data has been used worldwide to describe exponential data growth since the beginning of this century (106). The concept defines a set of technologies, processes and practices that allow organizations to take advantage of the analysis of the large volume of data for decision making (107). The biggest source of Big Data are people, and so the term is present in all areas that involve human activity (108). Since Big Data is associated with several areas and different sectors, the definition of the term varies according to the specificity of each area(109). The first definition describes Big Data as being a technology for processing a high volume of data, with high speed and of several natures (109).

Management challenges are related to security in collecting, processing and storing data, which translates into ethical and privacy issues controlled by policies and rules within an organization or other context that involves storage (110).

Big Data is a set of theoretical concepts that are the basis for the development of tools for collecting, storing and processing data, challenges arise in the collection, storage, analysis, sharing, privacy and data visualization (106). Data visualization is an indispensable technique in acquiring knowledge, which is submerged in the abundant volume given (113). The result of an analysis with charts and diagrams has been used for a long time but the complex data set makes the task harder to perceive (113). The development of a visualization technique for the heterogeneous data that exists today is useful to simplify and transform the information into a representation that is perceptible to the human being (113).

2.3.1   Data analysis concept

Data Analysis involves extracting, cleaning, transforming, modeling and visualization of data with an intention to uncover meaningful and useful information that can help in deriving conclusion and take decisions (114). Data mining is a technique that focus on modeling and knowledge discovery for predictive purposes and data analysis focus only in descriptive purposes (115).

In statistical applications, data analysis can be divided into descriptive statistics, exploratory data analysis (EDA) (114), and confirmatory data analysis (CDA) (116). EDA focuses on discovering new features in the data while CDA focuses on confirming or falsifying existing hypotheses (114).

The management and analysis of data have come to take significant dimensions within the area of information technology and so it is necessary to have some things in mind (113):

Data are the source of information and hence of knowledge and wisdom, so data must be presented in an appropriate format, leading to a favorable outcome for decision making. Preserve the quality of the data throughout the process of collection and construction of the final result. To produce the highest value, it is crucial that the data remain intact. The success of using data analysis practices does not only depend on how data is processed. The way the end result is built, presented and communicated to the user has a huge impact on data mining. The transformation of data into knowledge is the purpose of data visualization that is employed to achieve understanding in the heterogeneous dataset (117). An organization can extract value either in data visualization, in responding to specific business issues, or in finding solutions or new questions through exploratory need (117).

The data visualization fits in the area of health and food safety so it is seen as a practice or a discipline (113). Data visualization conveys a story or an event, its result being precise, attractive, timely and relevant (117). Practicing data visualization while improving decision making helps increase the power of analysis across the organization (113). The exploratory power of data visualization allows the user to base their decisions on visual means (113). Visual means is more effective than raw data. In general tools offer users visual aids to expose data and interaction features for quick and advantageous analysis.

2.3.2   Data Visualization

The Data/Information Visualization is a research area, related to Data Analysis, that uses visual representations with the help of information and communication technology (computers and software), seeking to aid and facilitate the process of cognition and obtaining information in a database (119).

Several researches related to information visualization have converged in the sense that, in order to make sense of the data, to make more rational and sustainable decisions, it is necessary to know and apply some rules (120). The word visual means to make something visual or visible, since the concept of visualization is defined as the transformation of abstract concepts into real or mentally visible images or, in computational context, visualizing is the conversion of numbers or categories into a graphic format that can be easily understood (117).

Typically, scientific contributions in the information visualization domain can be accurately mapped to specific parts of the pipeline (120). In the work they have developed they also leave the idea of

user interaction in the pipeline, since the visualization technique is not a static process (120). The pipeline begins by transforming the raw data into data formats suitable for viewing (120). This transformation aims at a data representation that is normalized in terms of content and structure so that the visualization can be separated from the input data (120). This is an important strategy that allows to adapt techniques to different hypothesis and data sets (120). It may involve trivial operations such as converting one data format to another, but in many cases, it is also necessary to identify and deal with incomplete, inaccurate, or erroneous data (120). Depending on the application, the result of this step is well-defined data for visualization (120). The second stage of the Card's visualization pipeline is the mapping of standardized but unprocessed data into the visual space (120). This mapping can be considered as the main transformation that forms the actual view (120). This is why different visualization techniques can be differentiated in this part of the pipeline (120). Visual space is described by a series of visual attributes that inherently represent the basic tools of visualization techniques (120).

The visualization process models are a reference for the user, presenting a set of tasks that should facilitate the creation of a visual representation (120). First the raw data is collected and transformed into data tables, through the data transformation process (120). Then the data tables must be modified to visual structures, through visual mapping methods, in order to achieve effective visualization (120). In the visualization process, three phases are identified (120):

1. Pre-processing: the modification of raw data in logical relations, more structured for human understanding;

2. Visual mapping: association between data and graphical representations;

3. Representation: the image of the visual representation is created.

The visualization process is represent in Figure 2.
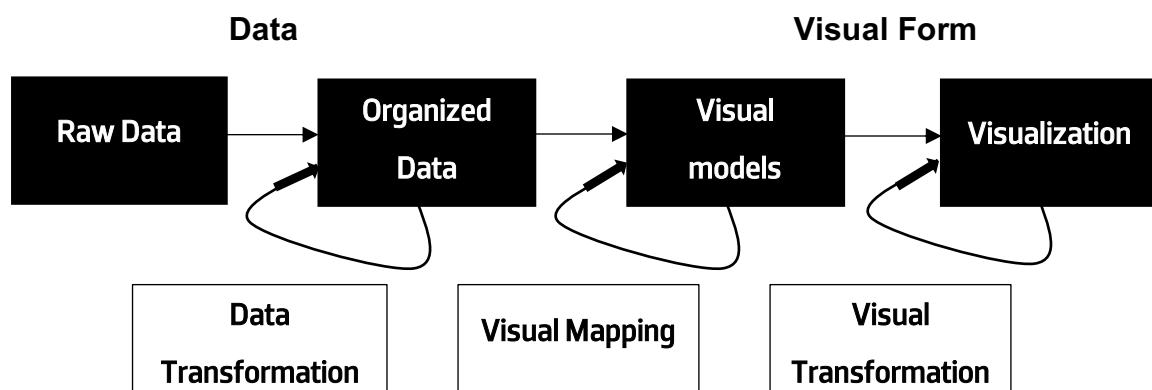
Figure 2- Data Visualization processing model (adapted from Card99)

The arrows indicate the flow of transformations that go from the untreated data to the visualization by the user and vice versa. Raw data can be in different formats and are transformed into data tables (120). Data tables are designed in visual structures, graphic representations defined by a spatial structure, symbols and graphic properties (120). Finally, these structures are visualized according to the interaction mechanisms chosen by the user and the result must be equal to an effective visualization (120).

Later in 2007, in the book Visualizing Data written by Benjamin Fry, the steps for visualizing data are seen as parts of a single process (121). The methods themselves are not new, but the steps have been isolated within individual fields (121):

1. Acquire: Obtain the data;
2. Parse: Provide some structure for the data's meaning and order it into categories;
3. Filter: Remove non useful data;
4. Mine: Apply statistical methods or data mining as a way to discover patterns or just to put the data in the mathematical context;
5. Represent: Choose a basic visual model (bar graph, list, or tree);
6. Refine: Improve the basic representation to make it clearer and more visually appealing;
7. Interact: Add methods for manipulating the data or controlling what features are visible.

2.3.3   Data Types

The graphical representations allow to identify phenomena and tendencies not visible with other techniques, thanks to the crossing of variables in the data (120). Visualizations are created through the crossing of quantitative and categorical data that justify the choice of the type of graphic representation to use (117).

To describe visualization methods taking into account the quality of the attributes and what they are capable of transmitting, it is useful to consider quantitative data and categorical information(117). There are three nominal, ordinal and interval measurement scales (122). The scale of the representation of categorical data types in a graphical representation, generally vertical and horizontal axes (122). Quantitative information alone does not only represent numbers, but also data that tells what numbers measure (122).

For this reason there are not several types of data but two variables in question – the numbers – quantitative data and – the phenomena that the numbers measure – categorical data (122). Quantitative data are only numbers, in percent, integer or decimal, are values that can represent or have a link with a category (123).

Grouping information into categories reduces large amounts of stored knowledge (123). Categories are the key to organizing spatial knowledge, that is, they are known for their relationship with the world (123). We can categorize relationships and events by assigning them sets of specifications (124). Properties are considered characteristics or attributes of a category (113).

The categorical scales, when defined in the graphs, assume three types (125):

- Nominal: Discrete and unordered elements, which even belong to the same category have no direct relationship to each other.

- Ordinal: Elements that have an intrinsic order, and just like the nominal scales do not relate to each other.

- Interval: Elements that have an intrinsic order and represent quantitative values. This scale begins as a quantitative scale that becomes a categorical scale that divides into intervals.

Information Visualization has evolved as the approach that transforms complex sets of data into useful information for the user through a fast and appealing way (126). The basic method of the user is to generate interactive visual representations of information that explore perceptual capabilities of the human visual system and the interactive features of the cognitive problem solution (126).

2.4   Conclusions

Regarding the HACCP system, during the development of the literature review, it was clear that its definition is global and objective.

It addresses public health and hygiene expectations from a preventive point of view and explains how organizations involved in the food industry should be organized from the "meadow to the plate".

ASAE is the authority responsible for food safety oversight in Portugal and HACCP is the system in place to ensure food safety. So ASAE is responsible for enforcing the HACCP system in Portugal. To study compliance with the HACCP system in Portugal requires access to databases with information on complaints received at ASAE.

As ASAE is the sole authority responsible for this topic, it is expected that a large amount of data is stored.

Throughout this dissertation it will be verified that the data available for the study are considered big data.

Big data analysis requires easy to interpret and intuitive information visualization methods that is intended with this dissertation.

ASAE cooperates with various organizations worldwide whose HACCP system is in place.

Since the HACCP system is globally implemented and big data is a growing area of study, the subject of this dissertation is a scientific contribution to the current state of food security in Portugal.

3. Methodology

This chapter presents the general research strategy that outlines how this study was performed. The general principles and process of knowledge discovery in databases are described and also several methods used are presented.

3.1 Principles of Knowledge Discovery in Databases

The data volume, generated and stored in the course of any activity, exceeds the capacity of human analysis and makes it impossible to extract knowledge from the same data, without resorting to a system that automates this process (112).

This context justifies the existence of the research area of Knowledge Discovery in Databases (KDD), generally defined as "the non-trivial process of identifying valid and potentially useful patterns, perceptible from the data" (112) , in which the principles associated to it are influenced by the areas of Artificial Intelligence, Automatic Learning, Pattern Recognition, Statistics, Database, Information Systems, among others (134).

The method used to search the patterns in the data is called Data Mining and is considered the core step of the process of Knowledge Discovery in Databases, which is developed in several phases (135). These patterns may or may not represent useful knowledge, being one of the phases that normally requires the participation of the user (112).

3.2 Process of Knowledge Discovery in Databases

Regarding the Knowledge Discovery in Databases, the mentioned patterns can be characterized by models, relations or structures in the data, that must be perceivable after a brief processing (112). The data represent a set of stored facts in the characterization of several patterns (104). The term process is associated to the execution of several iterative steps, started with the selection of data to be analyzed and ending with the interpretation of the results (134). The KDD process is interactive, since its development requires the participation of the user whenever decision making is necessary (136). The literature reviewed regarding the KDD process describes it through nine steps (112)(5):

- To develop learning in the domain of the application and perception of relevant knowledge about the domain, as well as the identification of the objectives to be reached in the process, in the perspective of the user and the environment in which it will develop;
- Selection of the data in which the discovery will focus, through the algorithms of Data Analysis;

- Pre-processing and treatment of the data set, through a set of strategies with respect to the appearance of incorrect or omitted data;

- Data transformation, which includes looking for appropriate settings to represent the data, in order to reduce the number of records and / or attributes under analysis;

- Selection of Data Mining techniques (classification, regression, clustering, among others) that best fit the defined objectives;

- Selection of Data Mining algorithms according to the strategy drawn to obtain knowledge about the data set, previously chosen and treated;

- Implementation of Data Mining algorithms in order to discover interesting and useful patterns. This step may be repeated until the desired results are useful in according to the objectives outlined;

- Interpretation of the discovered patterns, being possible the transformation of the same ones in user-perceived formats;

- Use of discovered knowledge, with the intention of using it in a different system, or only documenting the same, for later use by the interested parties.

The KDD process is represented in Figure 3.
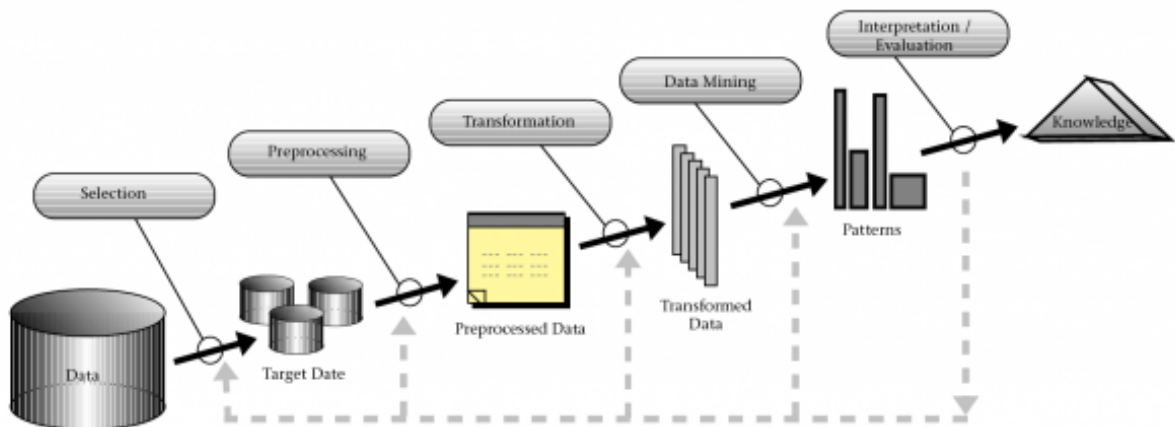
Figure 3– KDD Process (Usama Fayyad, 1996)

Data Mining phase represents 20% in the KDD process and is better supported by software than the other phases. All other phases, from the selection of data to the interpretation of the patterns found, are more a matter of "art" than a routine that can be automated (137).

3.3   Data Analysis vs Data Mining

The exponential increase in the volume of data has led to an information and knowledge revolution (104). It is now a key aspect of research and strategy building to gather meaningful information and insights from existing data (105). All this information is stored in a data warehouse, which is then used for Business Intelligence purpose.

Some definitions agree that Data Analysis and Data mining are two subsets of Business Intelligence (138). Data mining is a systematic and sequential process of identifying and discovering hidden patterns and information in a large dataset (135).

One of the best ways to analyse large datasets is through a high level programming language like Python.

3.4   Python – selected programming language

For this work it was selected Python because it is one of the most widely used programming languages in Data Science / Analysis and for a variety of reasons: it's easy to learn, powerful, complete, and has an endless list of modules and packages available for immediate use (127).

Python is a high-level programming language, which means it is a language with a relatively high level of abstraction, away from the machine code and closer to human language (128). Thus, high-level languages are not directly related to the computer's architecture (128). The programmer of a high-level language does not need to know processor characteristics because these features are excluded in the high-level language (128).

The choice of development tools is sometimes based on unique constraints or personal preferences. The main factors cited in the Python literature are (127):

- Software quality: For many, Python's focus on readability, consistency, and overall software quality sets it apart from other tools in the scripting world. The Python code is designed to be readable and therefore reusable and maintainable – much more than the traditional scripting languages. The uniformity of the Python code makes it easy to understand, even if you have not written it.

- Developer Productivity: Python increases developer productivity many times over compiled or static languages, such as C, C ++, and Java. The Python code is typically one-third to one-fifth the size of the C ++ or equivalent Java code. That means there's less to type, less to debug, and less to keep after the fact. Programs in Python also run immediately, without the time-consuming compilation and linking of some other tools, further increasing programmer speed.

- Program Portability: Most programs in Python are kept unchanged on all major computer platforms. In addition, Python offers several options for encoding graphical user interfaces, database access programs, Web-based systems, and more. Even the operating system interfaces, including program launch and directory processing, are as portable in Python as possible.

- Support libraries: Python comes with a large collection of pre-built and portable features, known as the standard library. This library supports a variety of programming tasks from the matching of text patterns to the network script. In addition, Python can be extended with both home libraries and a vast collection of software to support third-party applications. The Python third-party domain provides tools for building websites, numerical programming, and game development for example.

- Integration of component: Python scripts can communicate easily with other parts of an application, using a variety of integration mechanisms. These integrations allow Python to be used as a product customization and extension tool. Today, Python code can invoke C and C ++ libraries, it can integrate with Java components, it can communicate in structures like COM and .NET, and it can interact in networks with SOAP, XML-RPC and CORBA interfaces. It is not an autonomous tool.

- Pleasure: Because of the ease of use and the built-in Python toolkit, it can make scheduling more pleasure than task. While this may be an intangible benefit, its effect on productivity is an important asset. Of these factors, the first two (quality and productivity) are probably the most attractive benefits for most Python users.

- Software quality: By design, Python implements a deliberately simple and readable syntax and a highly coherent programming model. As a slogan at a recent Python conference attests, the end result is that Python seems to "fit its brain" – that is, language resources interact in a consistent and limited way, and follow naturally from a small set of concepts central banks. This makes language easier to learn, understand, and remember.

3.4.1   Pandas Library

Python has several libraries that make it easier and more efficient to perform specific tasks (129). The Pandas library stands for "Python Data Analysis Library" and is used for data analysis with Python and is one of the most popular and widely used tools for data collection and analysis. Pandas is an open source and free to use (under a Berkeley Software Distribution license) (130).

One of the advantages of Pandas is that it collects data (such as a CSV or TSV file or an SQL database) and creates a Python object with rows and columns called a data frame that is very similar to the table in a statistical software such as Excel (128).

It is easier to work compared to working with lists and / or dictionaries through loops or list comprehension (131). Pandas provides high-level data structures and own functions to make working with structured or tabular data fast, easy, and expressive (128). Pandas provides sophisticated indexing functionality to make it easier to reformat, cut data, execute aggregations, and select subsets of data (129).

Data manipulation, preparation and cleaning are important tasks in analyzing data that can be solved using Pandas (129).

To know which tools to use from the Pandas library, it is necessary to identify what type of data is presented to us by the imported file. According to the literature and as demonstrated in Figure 4, there are three types of data structures: series, data frame and panel (132).

DATA STRUCTURE	DIMENSIONALITY	SPREADSHEET ANALOG
Series	1D	Column
DataFrame	2D	Single Sheet
Panel	3D	Multiple Sheets

Figure 4– Pandas Data Structures (Matt Harrison, 2016)

The most commonly used data structures are the series and the data frames. A data frame is similar to a sheet with rows and columns, whereas the series are similar to a single column with data. A panel is a set of sheets, so in Pandas, a panel can have multiple data frames and each can give rise to multiple series (132).

The statistical studies for this dissertation were made using the high-level Python language in the Jupyter Notebook computing environment. The Jupyter Project runs several programming languages and the name of the project is a reference to the three main programming languages supported by Jupyter, Julia, Python and R (133).

3.5   Bibliographic Research Strategy

The bibliographic research strategy was based on the concepts of Food Control and Data Mining. The relationship with the concepts was constructed based on scientific studies, suggested by the advisor and the academic course of the author. This relationship proved to be useful for the definition

of key terms, which would influence the bibliographic research positively. The bibliographic research was developed in English and Portuguese, through the terms: Food Control, Food Control in the World, HACCP, ASAE, Data Analysis, Python, Biostatistics, Bioinformatics, Food Control Legislation, ASAE Reporting Methods. The results of the bibliographic research were achieved through crossings between the terms and using scientific content portals, Science Direct, SciELO, IEE Xplore, CiteSeerX, Wiley Online Library, Springer, Repositorium, Google Scholar. It is essential to note the importance of the protocol that the Polytechnic Institute of Porto and the Laboratory of Artificial Intelligence and Computer Science have with most of the mentioned institutions, facilitating access to certain documents. The results of the bibliographic research were not used in their entirety, since some of the terms used are transversal to several areas, resulting in a number of documents that might not meet what was intended. The need to carry out a reading of the summary of the different documents gathered allowed us to identify the documents with content that would justify a full reading. This process naturally fostered greater knowledge about relevant authors in the subject areas, leading to their individual research. The year and reputation of scientific articles were also taken into account.

3.6   Ethical Issues

The main ethical concerns in the course of this project are related to the confidentiality and protection of the data obtained, and do not appear to be any distribution or use of them, other than those originally envisaged with LIACC. All members involved in the project signed a confidentiality agreement aimed at protecting data from all those included in the ASAE database.

4. Results

This chapter presents the development of the practical component of the dissertation, where the main objective is to analyze the accomplishment of the HACCP system. Data Analysis techniques capable of exploring and extracting knowledge, as well to evidence standards in a set of data related to ASAE complaints were applied. The practical component was developed according to the KDD methodology, through its major five different phases: Selection of the data, Pre-processing and treatment of the data set, Data transformation, Data Analysis/Data Mining and Interpretation of the discovered patterns in order to extract knowledge. However, the mentioned phases are properly adapted, in order to reflect the activities developed in the present Dissertation project. It is also described the technical tools used, namely the Jupyter Notebook web environment whose language used is Pyhton, responsible for the visualization, validation and optimization of data. It is also responsible for the implementation of the techniques and Data Mining models necessary for the process of knowledge discovery.

4.1   Initial Data Format

The initial data were promptly provided by ASAE, after an extraction process executed on the database of the different excel files. Eleven files were extracted and made available in CSV or XSLX format, which aggregate information about the geographic distribution of the entities reported, the status of compliance with the complaints from 2014 to 2018 and the classification of each complaint in the same range of years. Documents that provide the designations of all the attributes that make up the tables were also available.

4.2   Datasets Description

The parameterizations mentioned allow to identify the different contents represented in the available tables. In the export2014 to export2018 tables, the information presented distinguishes the ASAE data entry between information requests and complaints between the years 2014 until 2018. The tables DENUNCIAS_2014 until DENUNCIAS_2018 contain information stored by ASAE that allow distinguishing the type of violation committed on each complaint. The Entities table aggregates information about all entities reported since the beginning of ASAE.

Table 2 presents the tables listed according to the content identified, as well as the number of attributes and records present individually.

Table 2– Datasets provided by ASAE

Dataset	Content	Atributes Number	Records Number
IR_Complaint2014	Number of complaints and information requests in 2014	11	22623
IR_Complaint2015	Number of complaints and information requests in 2015	11	25443
IR_Complaint2016	Number of complaints and information requests in 2016	11	24334
IR_Complaint2017	Number of complaints and information requests in 2017	11	25254
IR_Complaint2018	Number of complaints and information requests in 2018	11	23007
DENUNCIAS_2014	Description of the type of violation committed by complaint in 2014	40	20924
DENUNCIAS_2015	Description of the type of violation committed by complaint in 2015	40	22608
DENUNCIAS_2016	Description of the type of violation committed by complaint in 2016	40	20178
DENUNCIAS_2017	Description of the type of violation committed by complaint in 2017	40	21156
DENUNCIAS_2018	Description of the type of violation committed by complaint in 2018	40	19250
Entidades	All reported entities from the beginning of ASAE	16	2889714

4.3   Data Description

To achieve the desired objectives, for the practical component of this dissertation, it is essential to deepen the knowledge about the data that integrate the datasets. In this way, the parameters assigned by ASAE were again consulted, as well as visualization tasks were performed through the Pandas library on the Jupyter Notebook platform. This analysis allowed a data description report, which it is presented in Table 3. Attributes and identification of their format, as well as the sample of the registers that fill them are here presented.

In a first analysis it is possible to identify that the available datasets are based on the same subject, the complaints, and each one allows a different knowledge extraction of data. In order to increase the knowledge and understanding about the information contained in the different datasets a process was executed through the Jupyter Notebook web environment, which allowed the analysis of the different tables per year.

The datasets IR_Complaint 2014 until IR_Complaint 2018 have eleven attributes, as shown in Table 3 (for more information see annexes 1 until 6).

Table 3– Datasets IR_Complaint description

Attributes	Description	Variable Type
ID_CORRESPONDENCIA	Data entry number	Number
NID	Identification Number	Object
ESTADO_DOC	"Cumprido"– Solved "Pendente"– On Hold	Object
DT_REGISTO	Date of first and last registration	Date
DESC_CONTEUDO	Complaint description	Object
DESC_TP_DOC	How data was received	Object
DESC_ASSUNTO	"Denúncia"– Complaint "Pedido de informação"– Information request	Object
EMISSOR2	Complainant email or name	Object
RECLAMANTE	Complainant name	Object
ENT_VISADA	Reported entity	Object
EMAIL_CONTENT	Email content	Object

The datasets DENUNCIAS_2014 until DENUNCIAS_2018 have forty attributes, as shown in table 4 (for more information see annex 6 to 10) .

Table 4– Datasets DENUNCIAS description

Attributes	Variable Type	Description
NID	Object	Complaint identification number
Competência	Object	Organization responsible for handling the complaint
Estado	Object	Status of complaint
Estado Averiguação	Object	Status of inquiry
Data arquivo	Date Object	Filing date of the complaint
UO destino (1º Circuito)	Object	Operating unit (1st circuit)
Data de envio (1º Circuito)	Date	Date of dispatch (1st circuit)
Data de averig. (1ª FOF)	Date / Hour	Date and time of the investigation of the complaint in the FOF (File of Operator Supervised)
Local de registo	Object	Where the complaint was registered
Tipo	Object	How was the complaint received at ASAE
Tipo Remetente	Object	Type of sender that reported
Remetente	Object	Name of complainant
E–mail	Object	E-mail of complainant
Nome (e–mail)	Object	Name and/or e-mail of complainant

Remetida por Denunciante	Object	If the complaint was forwarded by the complainant
Tipo Denunciante	Object	Type of complainant
Denunciante	Object	Name of complainant
Motivo não preench. den.	Object	Reason for not completing the complaint
Entidade Visada	Object	Type of entity targeted
Motivo não preench. ent.	Object	Reason for not completing the entity field
Morada	Object	Entity's address
Cód. Postal	Object	Entity's Postal Code
Localidade	Object	Locality
Dist/Conc/Freg	Object	District / Council / Parish
Actividades	Object	Type of activity carried out in the reported entity
Área Geográfica/DR	Object	Region
Conteúdo	Object	Content of the complaint
Classificação Conteúdo	Object	Classification of content
Nº ofícios para Denunciante	Number	Number of offices for complainant
Nº Autoridades (Ofícios)	Number	Number of Authorities (Offices)
Ofícios para Denunciante	Object	Activity per complainant
Produtos	Object	Products available in the services described in the attributes "Activities"
Infracções	Object	Violations
Fiscalizações Relacionadas	Object	Related Inspections
Processos Relacionados	Object	Related processes
Suporte Digital	Object	"Digital Support:
Data de registo (criação)	Date/Hour	Registration Date (creation)
Utilizador (criação)	Object	User (creation)
Data de edição	Date/Hour	Date of issue
Utilizador (edição)	Object	User (editing)

The dataset Entites has sixteen attributes, as shown in table 5 (for more information see annex 11) .

Table 5–Dataset Entities description

Attributes	Variable Type	Description
RNPC	Object	RNPC–National Registry of Legal Persons
Relação	Object	Relation between reported entities
NIPC/NIF	Object	NIPC–Identification Number of Collective Person NIF–Tax Identification Number
Tipo	Object	Type of entity
Nome	Object	Entity Name
Código Postal	Object	Postal Code
Localidade	Object	Locality
Morada	Object	Adress
Dist/Conc/Freg	Object	District / Council / Parish

Natureza Jurídica RNPC	Object	RNPC legal nature
Código CAE Principal	Object	Primary CAE code CAE–Classification of Portuguese Economic Activities by Branch of Activity
Revisão CAE Principal	Object	Primary CAE review
Designação CAE Principal	Object	Primary CAE designation
Código CAE Secundário	Object	Secondary CAE Code
Revisão CAE Secundário	Object	Secondary CAE review
Designação CAE Secundário	Object	Secondary CAE designation

4.4   Datasets Cleaning

After the clear identification of the attributes needed to achieve the proposed categorization and forecasting tasks, it is necessary to carry out data cleaning operations with the aim of improving the data quality presented in the datasets IR_Complaint2014 until IR_Complaint2018, DENUNCIAS_2014 until DENUNCIAS_2018 and Entities. In this way, the treatments given to the attributes and records are presented in table 6, taking into account the problems and inconsistencies reflected in the data quality report. The treatment of missing data or data in non–agreement was done only for the columns of interest of each dataset because of the large amount of data to be analyze.

The IR_Complaint2014 until IR_Complaint2018 files were concatenated in one dataset designated op.

From the dataset IR_Complaint a distinction was made between information requests ("Pedidos de Informações") and complaints ("Denúncia") from 2014 until 2018 (for more information see annex 1 to annex 5).

It was also analyzed the solved information requests and complaints ("Cumprido") and the ones on hold to solved ("Pendente") .

In table 6 is shown the data quality and number of missing values.

Table 6–Cleaning Dataset IR_Complaint

Attribute	Omitted values (%)	Data Quality	Treatment carried out
ID_CORRESPONDENCIA	0	No identified problem	–
DESC_ASSUNTO	0	No identified problem	–
ESTADO_DOC	0	No identified problem	–

Attribute	Omitted values (%)	Data Quality	Treatment carried out
DT_REGISTO	0	Date with time: Format DD–MM–YYYY HH:MM:SS 2 date values separated by "/" Date followed with time for report entry	Replace "/" with "–" Split by "–" to separate year and time from day and month Split by " " to isolate the year from time

DENUNCIAS_2014 until DENUNCIAS_2018 files were concatenated in a dataset called Complaints. From this dataset was analyzed the violations committed by each complaint and were indentified the ones related with HACCP system from 2014 to 2018.

Table 7 represents the process of cleaning the dataset complaints.

Table 7– Cleaning Dataset Complaints

Attribute	Omitted values (%)	Data Quality	Treatment carried out
Estado	0	No identified problem	–
Estado averiguação	0	No identified problem	–
Competência	0,2	227 missing values	Assign the code "0"
Tipo	0	No identified problem	–
Tipo Remetente	0	No identified problem	–
Infracções	42	Each violation has an associated code to identify each type of violation. This code consists of letters and numbers separated by "." The number of each violation code allows to identify whether it is related to the HACCP system	Do the split for "." to study each part of the code that identifies the type of violation

The dataset "Entities" contains all the entities reported since the existence of ASAE.

From the dataset Entities the percentage of complaints by district and by islands in Portugal was studied to verify if the percentage of complaints is in accordance with the population density recorded in the last census conducted in 2011.

The dataset Entites exceed the number of lines allowed by excel.

In order to avoid losing data, the file was separated into three files in Sublime. Then, using Jupyter Notebook, these three files were added together in a single dataset called "Entities". Table 8 shows the cleaning process of this dataset.

Table 8–Cleaning Dataset Entities

Attribute	Omitted values (%)	Data Quality	Treatment carried out
Tipo	0,00003	4 missing value	There were four values that did not agree with this designation (45211, BRAGA/ BRAGA, R.Fialho De Almeida 1–4 E, Rua Da Porta Nova, 9) and were eliminated
Relação	0,0003	9 missing values	There were nine values that did not agree with this designation (46382,47712, 68100,620140,0000–000,1000–000, 4990–092,BRAGA,ÉVORA/ ESTREMOZ) were eliminated
Dist/Conc/Freg	1,5	42300 missing values	42300 missing values, assign "0" to these values Split by "/" to isolate districts

4.5 Construction and Data Transformation

The task of data construction and transformation aims to adapt the files information given by ASAE to the data analysis objectives, outlined for this work.

4.5.1 Dataset IR_Complaint

For studying the dataset IR_Complaint, was used Python Pandas library packages. In fact, by analyzing this dataset, it was possible to acquire information, such as: the purpose of knowing the type of contact with ASAE (Information Request or Complaint); the state in which this request for information or denunciation is (Completed or On Hold); and the year they were made; to detect and extract knowledge to compare the evolution, from 2014 until 2018.

After concatenating the datasets export2014, export2015, export2016, export2017 and export2018 in a single dataset called IR_Complaint the study was focused on the columns of interest.

From the column "DT_REGISTO" only concerns the year in which the complaint was made. The date is followed by time in column "DT_REGISTO" is in the format DD–MM–YYYY HH:MM:SS. To do a year-based study it is necessary to do two splits. The first split is to separate by "–" which separates the different elements of the date. This creates three new columns, one containing the day, another

that contains the month and other that is filled with the year and the time. The year and time are separated by one space. Only the column containing the year and time will be used. To separate the year from the complaint time a second split was made using one space " ". This creates two new columns. One with the year and another with the time. The column of interest is the one containing the year. The name given to the column of interest is the same as the original dataset "DT_REGISTO". The data from the new column created by the split was added to the original column "DT_REGISTO" of the dataset IR_Complaint. The column with the ID corresponding to each data entry was also used to be able to count the data of the columns of interest. The columns of interest for the study are "ID_CORRESPONDENCIA", "DT_REGISTO", "ESTADO_DOC" and "DESC_ASSUNTO". Table 9 represents data from columns of interest that were grouped together with the numpy package according to the columns of interest.

Table 9–Transformed Data of interest from Dataset IR_Complaint

DT_REGISTO	ESTADO_DOC	DESC_ASSUNTO	ID_CORRESPONDENCIA
2014	CUMPRIDO	Denúncia	15005
		PEDIDO DE INFORMAÇÕES	7231
	PENDENTE	Denúncia	310
		PEDIDO DE INFORMAÇÕES	77
2015	CUMPRIDO	Denúncia	14945
		PEDIDO DE INFORMAÇÕES	7953
	PENDENTE	Denúncia	2221
		PEDIDO DE INFORMAÇÕES	324
2016	CUMPRIDO	Denúncia	12894
		PEDIDO DE INFORMAÇÕES	8137
	PENDENTE	Denúncia	2873
		PEDIDO DE INFORMAÇÕES	428
2017	CUMPRIDO	Denúncia	12751
		PEDIDO DE INFORMAÇÕES	7854
	PENDENTE	Denúncia	4184
		PEDIDO DE INFORMAÇÕES	465
2018	CUMPRIDO	Denúncia	8481
		PEDIDO DE INFORMAÇÕES	6590
	PENDENTE	Denúncia	6692

		PEDIDO DE INFORMAÇÕES	1244

Table 9 allowed the construction of a bar chart to compare the number of complaints against the number of information requests in each year of study for a more intuitive visualization of the information (Figure 5).

Complaints vs Information Requests
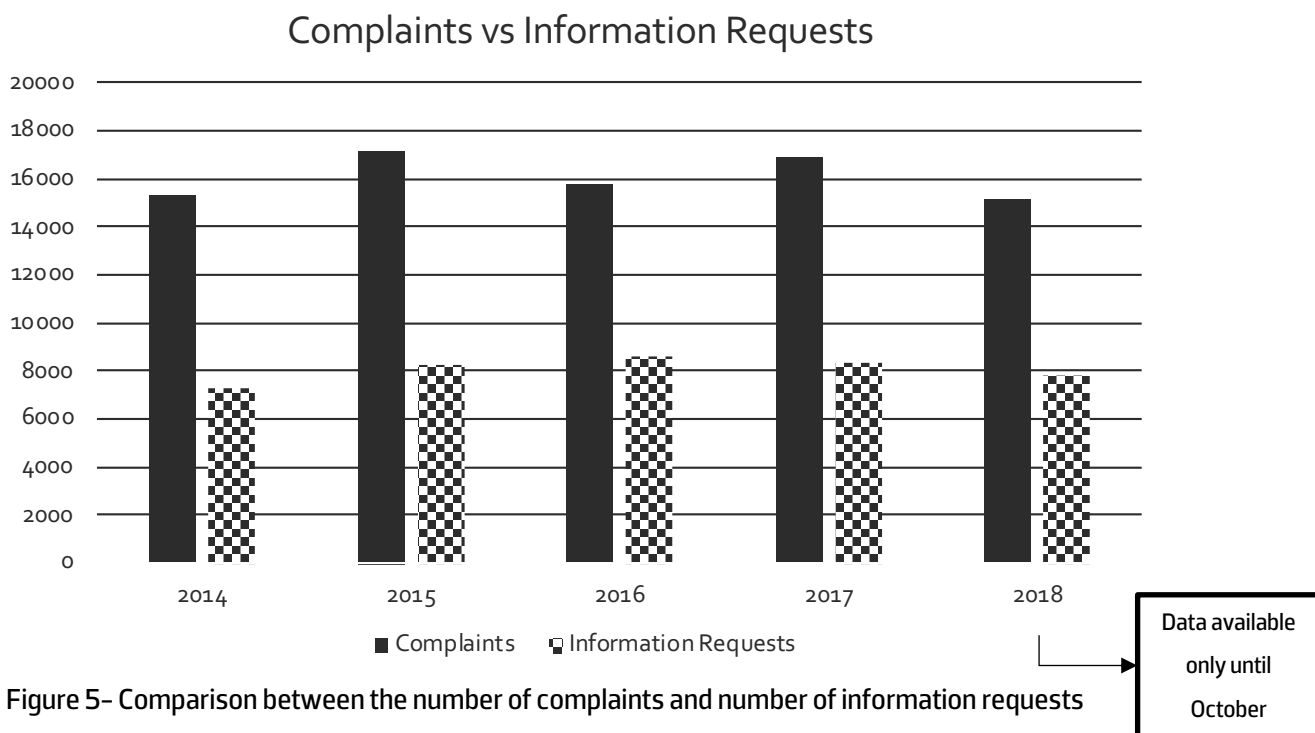
Figure 5– Comparison between the number of complaints and number of information requests

> Data available only until October

As can be in from Figure 5, regardless of the year, the number of complaints is always higher than the number of information requests. The year with the most complaints was 2015 with 17166 complaints, followed by 2017 with 16935 complaints. This is followed by the year of 2016 with 15769 complaints, 2014 with 15315 and lastly 2018 with 15163 complaints, which may be justified by the fact that the data provided for 2018 has information available until October of that year.

From table 9 it was also possible to analyze the number of information requests and complaints "Cumprido" (solved) or "Pendente" (on hold). A more schematic view is presented in Figure 6. Information request on hold is represented by "IR_Hold", information request solved is represented by "IR_Solved". If a complaint is solved is represented by "C_Solved" and if it is on hold it is represented by "C_Hold".

Complaints and Information Requests

2018	6692 / 7834 / 1244	15173
2017	4184 / 8319 / 465	16935
2016	2873 / 8565 / 428	15769
2015	2221 / 8277 / 324	17166
2014	310 / 7308 / 77	15315

■ Complaints   ▫ C_Hold   ▬ Information Request   ■ IR_Hold
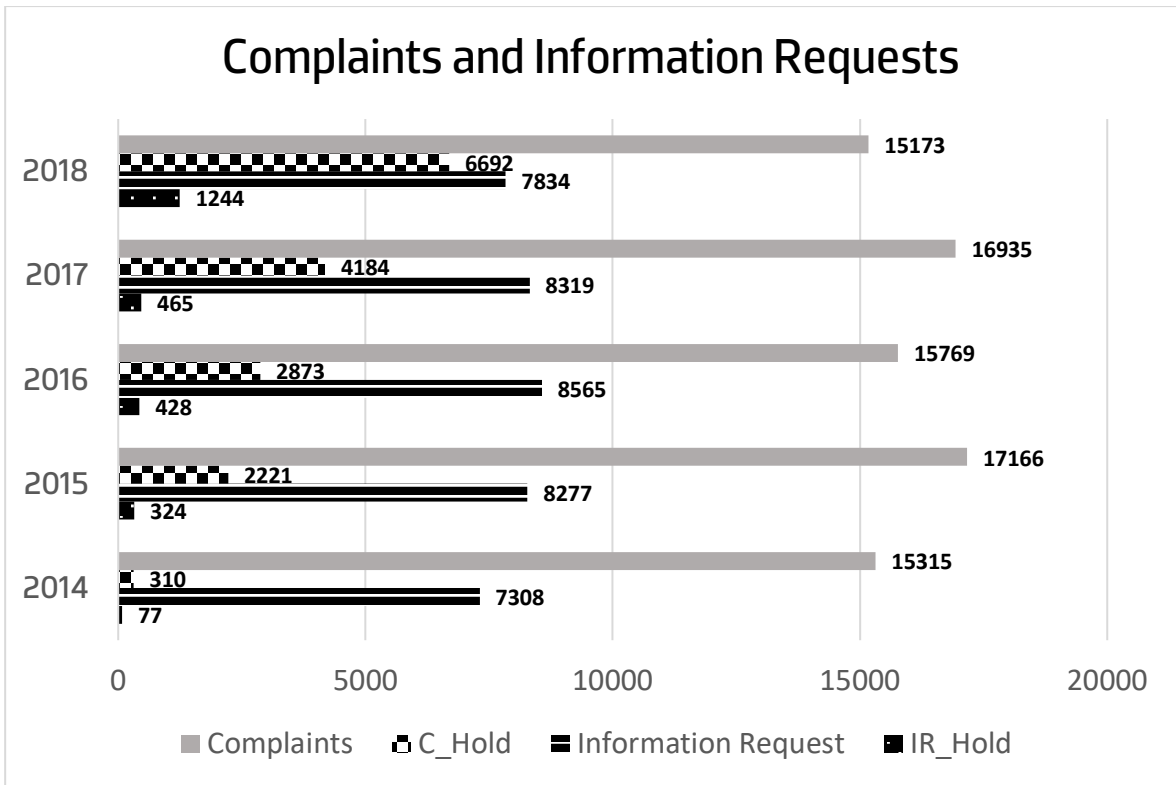
Figure 6– Comparison between complaints and information requests solved and on hold

According to Figure 6, there are always more complaints and information requests solved than on hold.

As the year of 2018 approaches, the number of complaints and information requests on hold increases because the closer to the current year, the less response time there has been.

4.5.2   Dataset entities

The file with the dataset of all entities reported since the beginning of ASAE, as mentioned, exceeded the limit of allowed lines for reading in excel with a total of 2889714 lines. The file was divided into three excel files (Entities, Entities2 and Entities3) and subsequently concatenated into a single dataset called "Entities".

This file, as it contains all reported entities registered in ASAE and includes the district of each one, allows the study of the distribution of complaints according to the district and the islands of Portugal. This distribution of complaints should be in accordance with the population distribution of Portugal. This means that it is expected that a district with a higher population density should have a higher percentage of complaints, than a district with a lower population density.

The file with all entities reported since the beginning of ASAE allows us to process all the complaints received by ASAE and to evaluate their distribution by district.

The column of interest in this dataset is the column "Dist/Conc/Freg" because it contains the District/Council/Parish. This column also contains the islands of Portugal. The district is isolated by split because it is in the "Dist/Conc/Freg" column, so the split is done by '/'. This split creates three new columns, a first one containing the district, a second one containing the council and the last one containing the parish. The first column of the split contains the districts and allows to count the percentage of complaints by district.

The total of rows in the "District" column after removing the missing values resulted in a total of 2847400 complaints per district. Once isolated the districts and after the analysis of the final dataset it was possible to analyze the percentage of entities reported by district and islands in relation to population density according with the last census made in 2011 as shown in Table 10.

Table 10– Distribution of complaints according to population density

Position	District	Population	Complaints (%)
1	Lisboa	2 884 984	25
2	Porto	2 397 191	15
3	Braga	924 351	7
4	Setúbal	880 765	9
5	Aveiro	814 456	7
6	Faro	569 714	5
7	Leiria	560 484	5
8	Coimbra	541 166	4
9	Santarém	454 947	4
10	Viseu	378 784	3
11	Madeira	272 706	2
12	Açores	249 459	1
13	Viana Do Castelo	252 952	2
14	Vila Real	214 490	2
15	Castelo Branco	196 989	2
16	Évora	174 490	2
17	Guarda	167 359	2
18	Beja	158 702	1
19	Bragança	140 385	1
20	Portalegre	120 585	1

It should be noted that in Portugal's islands belong to two archipelagos: Madeira and Açores.

Madeira archipelago has seven islands, two of which are inhabited: Madeira and Porto Santo. Açores archipelago has nine islands: Corvo, Flores, Faial, Graciosa, Pico, São Jorge, Terceira, Santa Maria and São Miguel. On the islands the ASAE only deals with complaints related to gambling and wine, so

despite the relative high population density of the islands of Açores, the percentage is not consistent with the rest.

Percentages of complaints agree with the population density of the respective district, except for Braga with no apparent reason and the autonomous region of the Açores for the aforementioned reason.

Through the study of dataset entities, it is possible to adapt the available information to a map of Portugal where it is possible to observe the percentage of complaints as shown in Figure 7.
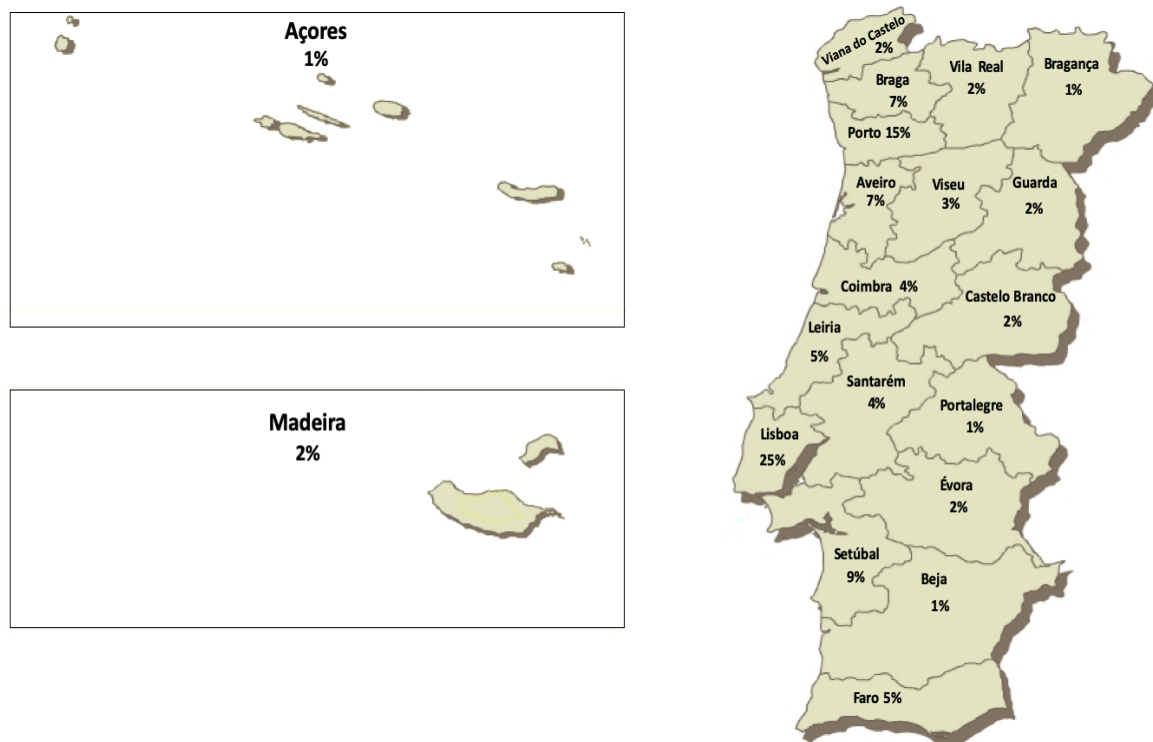
Figure 7– Percentage of complaints by district

Although the archipelago of Açores has a larger set of islands than archipelago of Madeira, the archipelago of Madeira has a higher population density than the archipelago of Açores.

In 2011, the last census was carried out, with 267785 habitants in the autonomous region of Madeira and 246746 habitants in the autonomous region of the Açores. This small variation justifies the 1% difference between the two sets of islands.

4.5.3   Dataset Complaints

The ASAE made available five files related to the complaints made between 2014 and 2018 called: DENUNCIAS_2014, DENUNCIAS_2015, DENUNCIAS_2016, DENUNCIAS_2017 and DENUNCIAS_2018.

These five files were concatenated into a dataset called Complaints. The columns of interest in this dataset are "Competência", "Estado averiguação", "Estado", "Infracções", "Dist/Conc/Freg" and "NID".

The first analysis of this dataset contains the following attributes that will be studied by district:

- Competence (Column–"Competência")

   The competent authority for responding to the complaint:

   – ASAE

   – ASAE e Tribunais (ASAE and Courts)

   – ASAE e outra Entidade (ASAE and other Entity)

   – Indeterminada (Undetermined)

   – Outra Entidade (Other Entity)

   – Outra Entidade e Tribunais (Other Entity and Courts)

   – Tribunais (Courts)

- State (Column–"Estado")

   The state of complaint can be solved "Cumprido" or on hold "Pendente".

- State of inquiry (Column–"Estado de averiguação")

   If a complaint is solved it can be filed "Arquivada" or filed and verified "Arquivada averiguada".

   If a complaint is on hold it can be in inquiry "Em averiguação" or to be inquired "Por averiguar".

- NID (Column–"NID")

   Is a number that identifies a complaint that is necessary to make a count when using numpy package.

To isolate the district from the column "Dist/Conc/Freg", a split was made with the same instructions of the split done in the dataset entities as well as the import of the column containing the districts. The column "Dist/Conc/Freg" in this dataset contains the islands of Portugal as in the dataset entities.

Through the group by function of the Pandas numpy package, the columns of interest are grouped, obtaining the information contained in Annex 13.

From Annex 13 it is possible to see that the most densely populated districts have the highest number of entries in any of the competent authorities. Also, there are more complaints with the status "Cumprido" than "Pendente." This means that there are more solved complaints than on hold. From the information obtained in Annex 13 it is possible to assess the competence of the establishments affected by the complaints. Figure 8 shows the percentage of complaints with the competent authority.
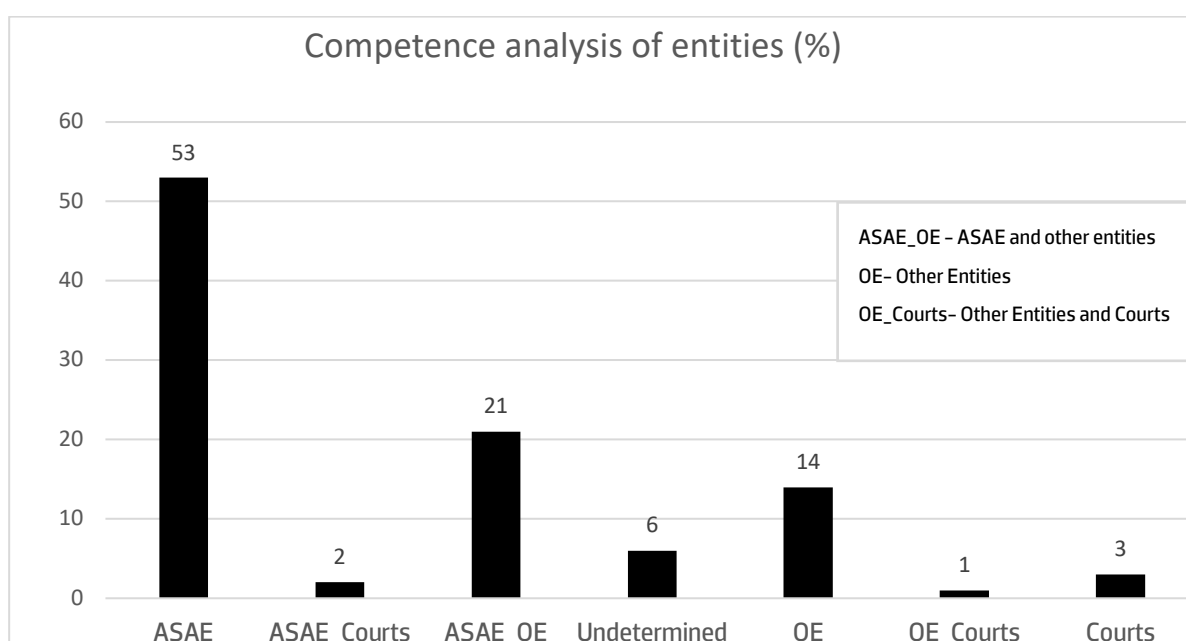
Figure 8– Competency assessment

As can be seen in Figure 8, 53% of the complaints received by ASAE are solely the responsibility of ASAE. The second highest value corresponds to the competence of ASAE and other entities with 21%. These data show that more than half of the complaints received by ASAE are entirely within its competence.

The second analysis of this dataset is based on the violations codes available in the column "Infracções" of the dataset. By continuing to analyze the dataset Complaints, it is possible to extract information regarding violations committed in each complaint.

In each complaint, more than one violation can be identified. The violations are represented by codes with brief descriptions in the column "Infracções". Each cell in this column corresponds to a complaint. And in that cell may be described more than one violation. This happens because each complaint may have more than one violation. The first step is, in the case of complaints with more than one violation, to separate the various violations.

Separating violations within the same complaint increased the number of rows in the dataset. This means that there were several complaints where more than one violation was committed, as can be seen in figure 9.

COMPARISON BETWEEN THE NUMBER OF ROWS IN THE DATASET

26083	28783	26628	29275	25882
20924	22608	20178	21156	19250
2014	2015	2016	2017	2018

— —Original Row   —○—Split Row
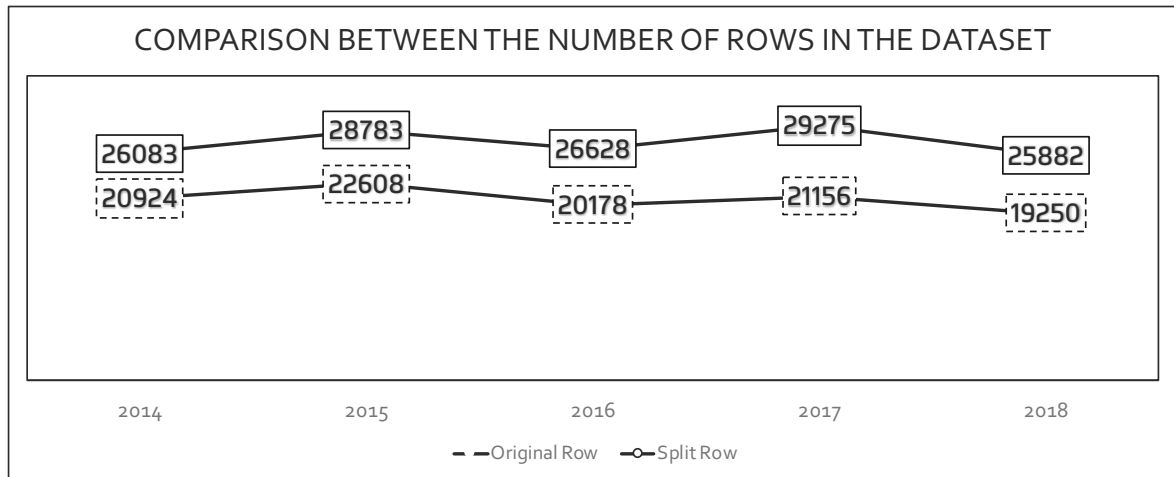
Figure 9– Comparison between the number of rows before and after split

As can be observed (Figure 9), in all datasets from 2014 until 2018 there are complaints with more than one violation committed. As the objective is the analysis of compliance with the HACCP system, it is important to analyze the violations committed related to this system in each year under study. With the different violations of each complaint separated, the percentage of violations can be analyzed by year.

In the "Infracções" column there are null and non-null values. Non-null values represent complaints of violations. Null values represent complaints without violations. Complaints that do not have the associated "Infracções" field have no violations committed.

In order to analyze the violations committed per year, it is necessary to use the column with the date called "Data de registo (criação)". In this column, there is information about the year and time the complaint was filled. Date and time are in the format "DD/MM/YY HH:MM". As with previous datasets, only the year is isolated and imported to the original dataset. Using the pandas numpy library, it is possible to study violations and non-violation complaints from 2014 until 2018 (figure 10).
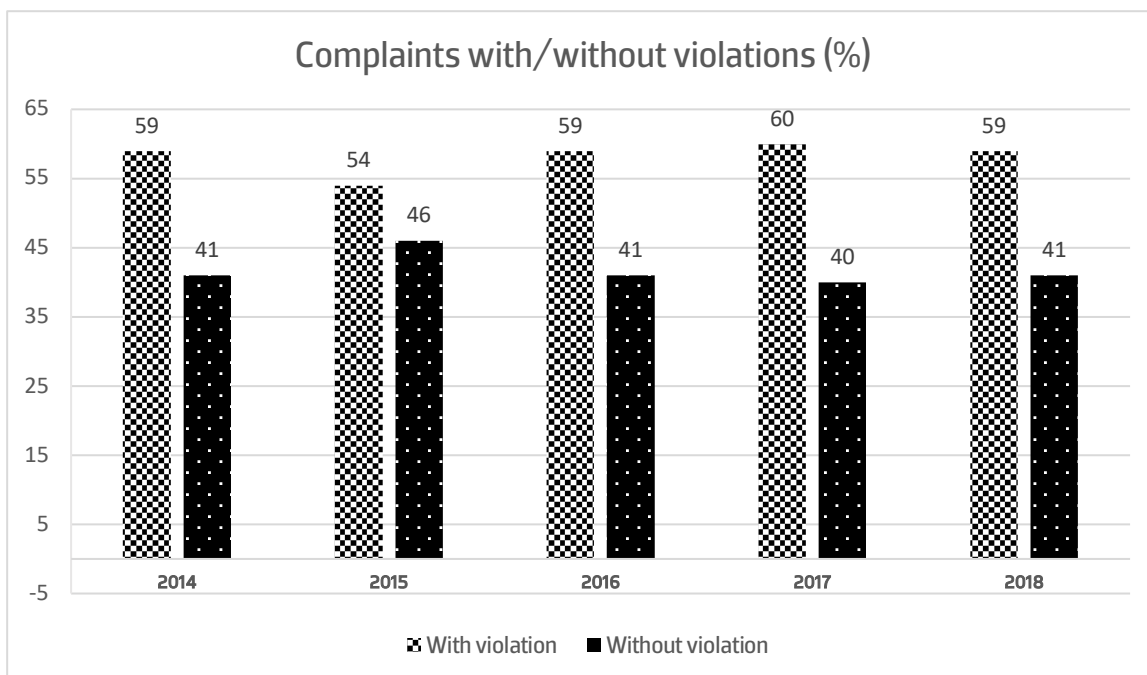
Figure 10– Complaint with and without violations from 2014 until 2018

From 2014 until 2018, the percentage of complaints with violations is always higher than the complaints without violations. In 2014, 2016 and 2018, 59% of complaints have violations and 41% of complaints have no violations. The year of 2015 was the year with less complaints with violations (54%) and with more complaints without violations than all the other years (46%). The year of 2017 was the year with more complaints with violations (60%) and with less complaints without violations than all the other years (40%).

Using only the complaints with violation it is possible to study the complaints related to HACCP system. Each violation code has numbers and letters. The first part of the violation code is a number. The violation code can start by the numbers 1, 2, 3 or 4. The number 1 refers to Segal (Segurança Alimentar) that means food safety. So, the violations that start with number one are the ones related to HACCP system.

If an violation code starts with the number 2 represents "FISEC-PPC" which means "Economic Inspection– Property Commercial Practices", if it starts with the number 3 represents "FISEC-SA" which means "Economic Inspection – Anonymous society" and if starts with number 4 represents "Diversos" which means "Others".

Knowing the meaning of each initial character of the violation code, it is possible to make a comparison, by year, of complaints related to the HACCP system (Segal) and those related to the remaining three categories. Figure 11 represents the complaint's categories in 2014.

Complaint's category in 2014 (%)
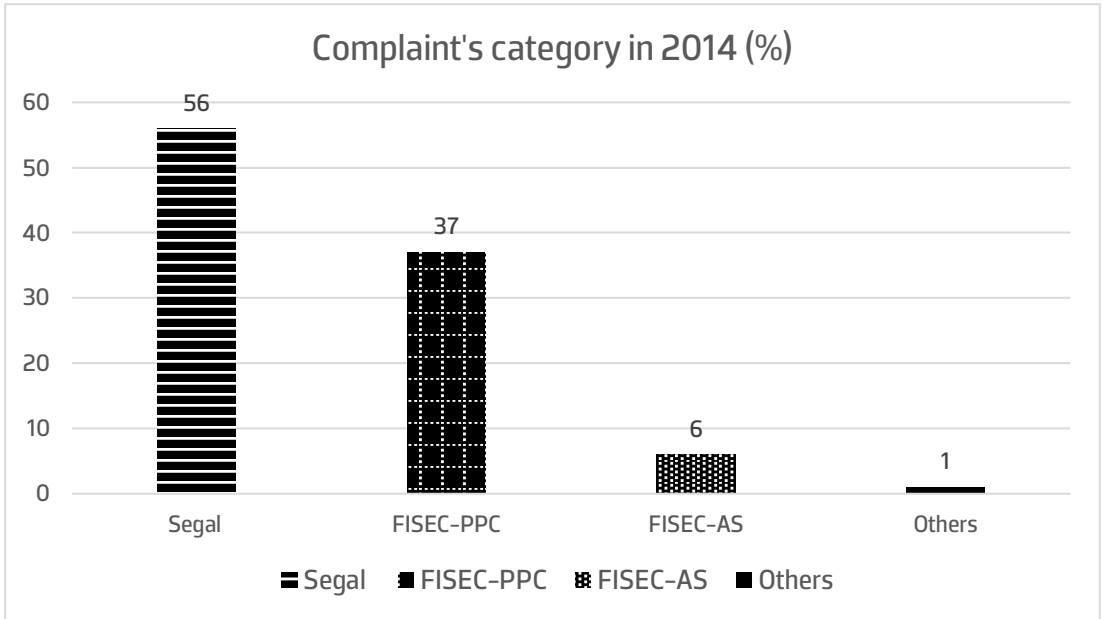
Figure 11– Complaint's category in 2014

In 2014, more than half of the complaints are related to HACCP system with a total of 56% (Figure 11). The complaints related to Economic Inspection– Property Commercial Practices represent 37% of the total complaints, the complaints related to Economic Inspection – Anonymous society represent 6% of the total complaints and the complaints related to other categories only represent 1% of the total complaints.

In 2015, the percentage of complaints related to HACCP system and to FISEC–PPC is equal (Figure 12).
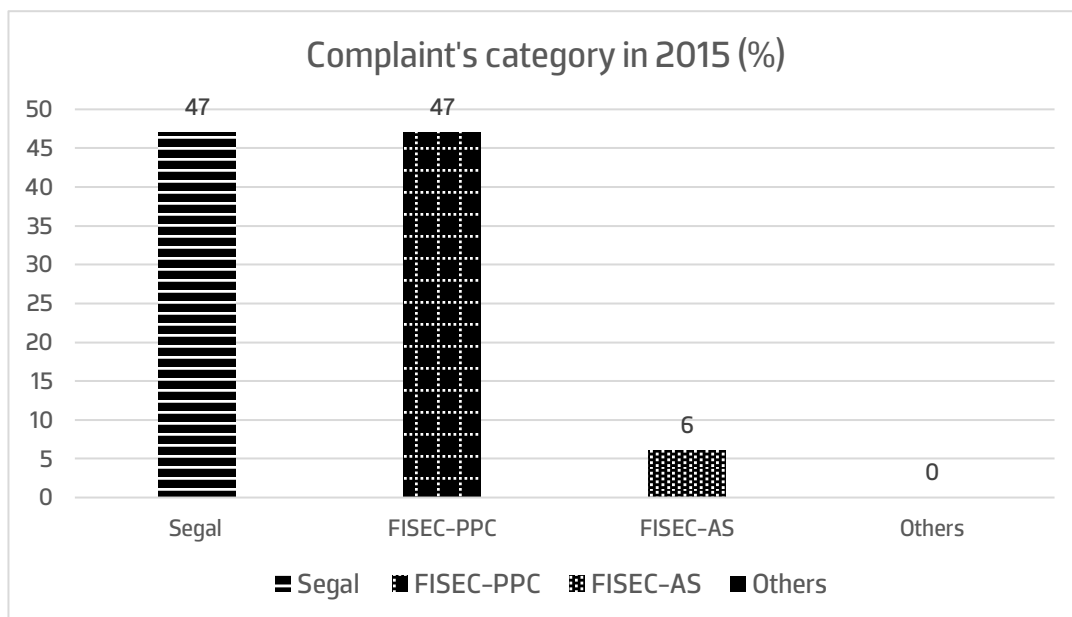
Figure 12- Complaint's category in 2015 (%)

As can be seen in Figure 12, in 2015, 47% of the complaints are related to HACCP system and with Economic Inspection- Property Commercial Practices, occupying these two categories a total of 94% of this year's total complaints.

The complaints related to Economic Inspection – Anonymous society represent 6% of the total complaints and there are none complaints related to other categories.

In 2016 the number of complaints related to FISEC–PPC exceeds the percentage of complaints related to HACCP system (Figure 13).
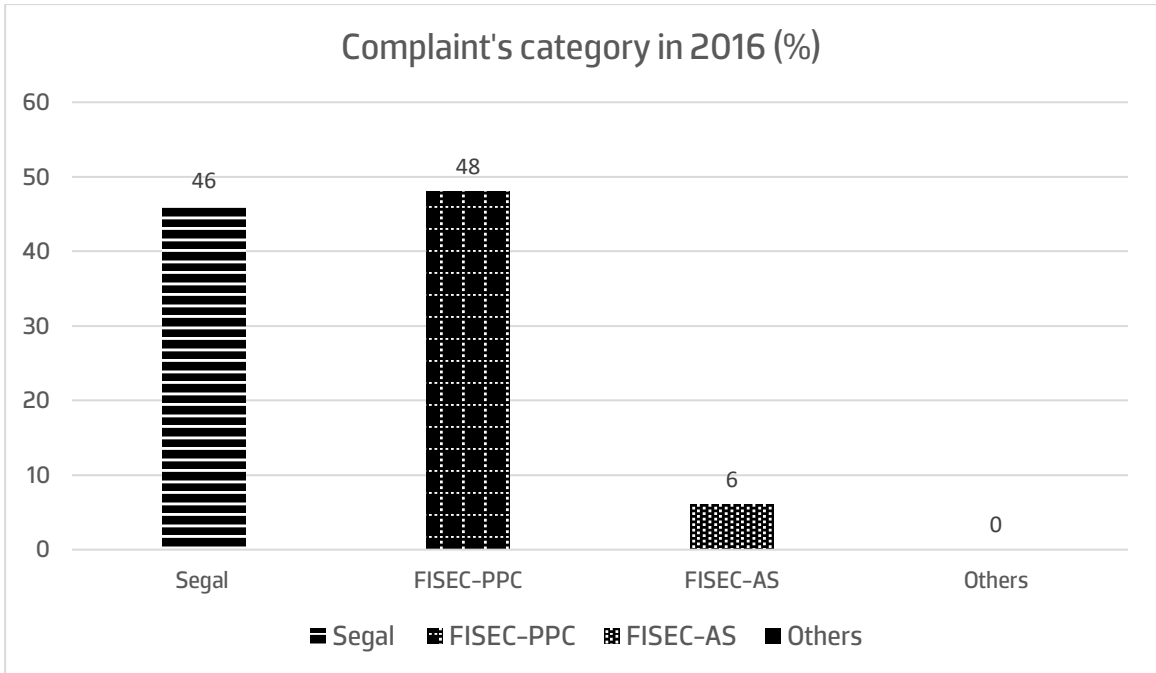
Figure 13- Complaint's category in 2016 (%)

Figure 13 presents data of 2016 and 46% of the complaints are related to HACCP system and 47% of the complaints are related to Economic Inspection- Property Commercial Practices, occupying these two categories a total of 94% of this year's total complaints again.

Since 2014, the percentage of complaints related to the HACCP system has been decreasing and the percentage of complaints related to the FISEC-PPC category has increased, surpassing in 2016 the Segal category by 1%.

The complaints related to the other two categories, FISEC-AS and Others, remain the same as the year of 2015. Complaints related to Economic Inspection – Anonymous society represent 6% of the total complaints and there are none complaints related to other categories.

In 2017 exactly half of the complaints are related to the HACCP system (Figure 14).
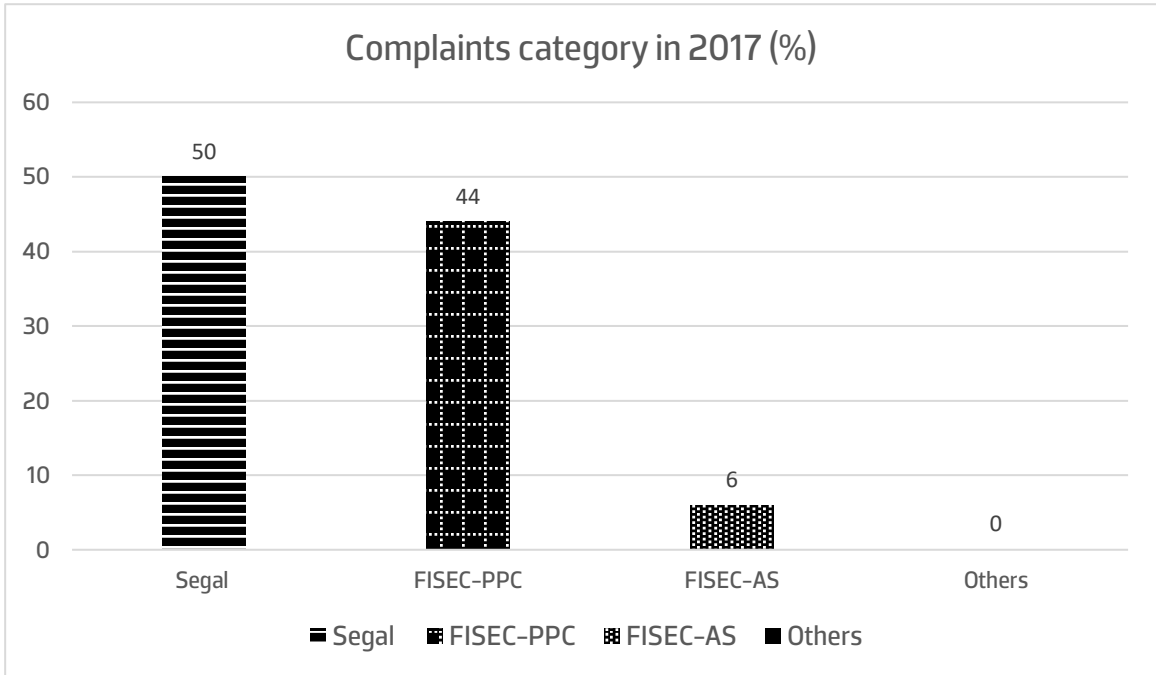
Figure 14– Complaint's category in 2017

As can be seen in Figure 14, in 2017, 50% of the complaints are related to HACCP system and 44% of the complaints are related to Economic Inspection– Property Commercial Practices, occupying these two categories a total of 94% of this year's total complaints again.

From 2016 to 2017 the percentage of complaints related to the HACCP system increased and the percentage of complaints related to the FISEC–PPC category decreased.

The complaints related to the other two categories, FISEC–AS and Others, remain the same since 2015.

In 2018, the percentage of complaints related to HACCP system exceeded 50% (Figure 15).
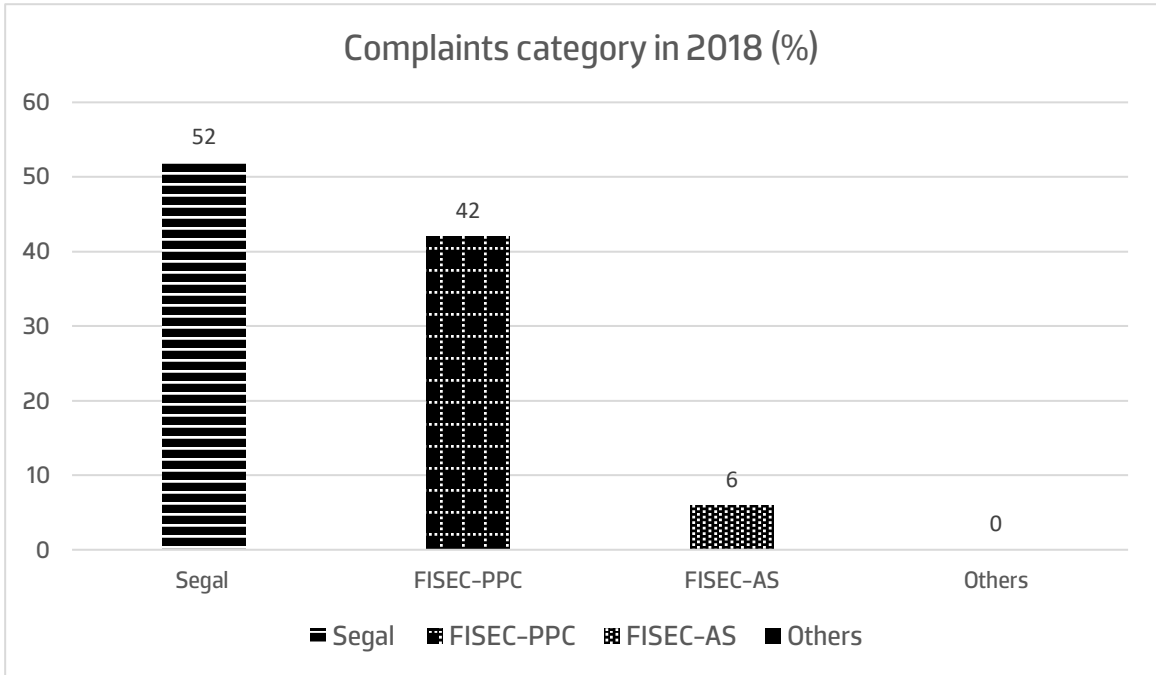
Figure 15– Complaints category in 2018

According to Figure 15, in 2018, the percentage of complaints related to HACCP system increased 2% compared to 2017 and the percentage of complaints related to FISEC–PPC decreased 2% compared with the same year. FISEC–AS continues with 6% of the total complaints and there are no complaints related with other categories.

With the complaints analyzed by the four categories from 2014 to 2018 and knowing the percentage of complaints related to the HACCP system (category Segal), it is possible to build the graph of Figure 16.
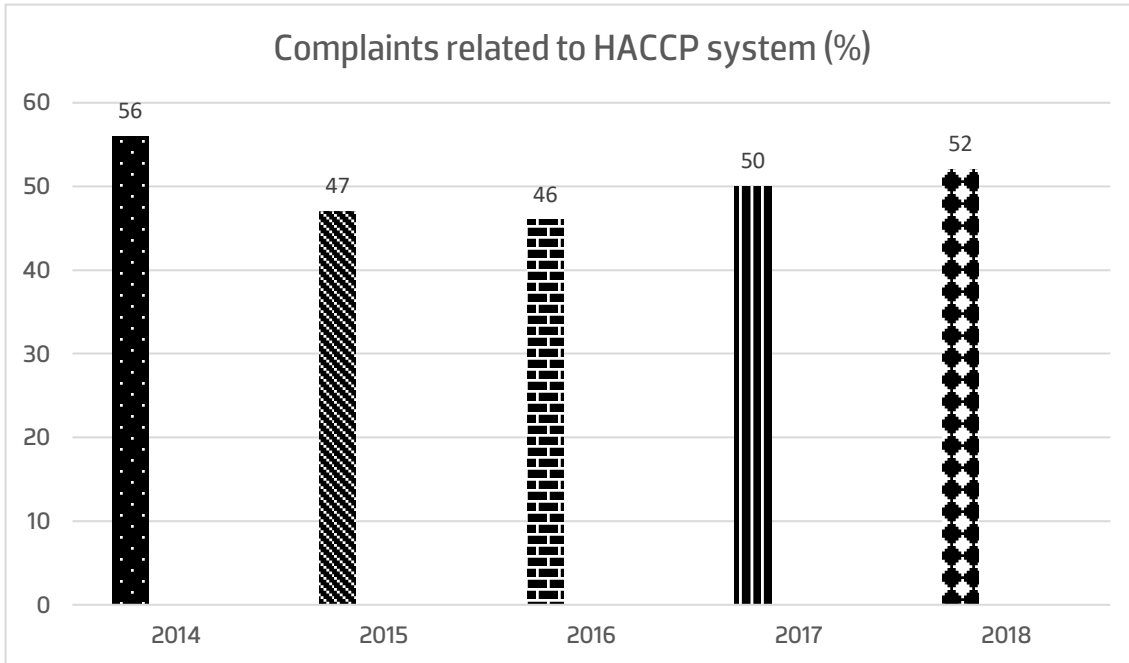
Figure 16– Complaints related to HACCP system from 2014 to 2018

As can be seen from the analysis of Figure 16, the year with the most complaints related to the HACCP system was 2014 with 56% of the total complaints. From 2014 to 2015 there was a 9% decrease in complaints related to the HACCP system, again decreasing 1% for 2016. From 2016 to 2017 there was a 4% increase in complaints related to the HACCP system, reaching 50%. From 2017 to 2018 the increase was 2%, this last year having 52% of the complaints received related to the HACCP system.

The complaints related to HACCP system have a violation code associated. According to the information provided by the ASAE, the first character of the code allows us to distinguish complaints related to the HACCP system. However, it is important to see more specifically what that violation is. For this, it was analyzed the second part of the code which, in this case, is a letter. For this, a new split is done to isolate the second part of the violation code. Then the two columns obtained from the original column "Infracções" are merged. This new column only contains the first two characters of the original violation code.

Thus, we can study which types of complaints related to the HACCP system were made between 2014 and 2018 according to the values in the Table 11.

Table 11– Complaints related to HACCP

Violation Code	Type of Violation	(%) Violations
1A	Clandestine slaughter	1,50

1B	Food kinds	39,0
1C	Food Hygiene and Safety	48,0
1D	Veterinary medicines and veterinary products	0,20
1E	Preservation / Species Handling	0,01
1F	Products not subject to specific regulations	0,00
1G	Labeling, presentation and advertising of foodstuffs (general rules)	6,00
1H	Materials and articles in contact with foodstuffs	0,03
1I	Feeding animals	0,02
1J	Veterinary checks	0,02
1K	Foodstuffs with toast	0,00
1L	Catering establishments and beverages	4,50
1M	Cultures	0,06
1N	Animal by-products and derived products not intended for human consumption	0,03
1O	Livestock activity	0,08
1P	Contracts for the purchase and sale of raw cow's milk	0,00

According to Table 11, the majority of the percentage of HACCP-related violations are directly related to food. The category "Food Kinds" e "Food Hygiene and Safety" are responsible for 87% of HACCP-related offenses (Figure 17).
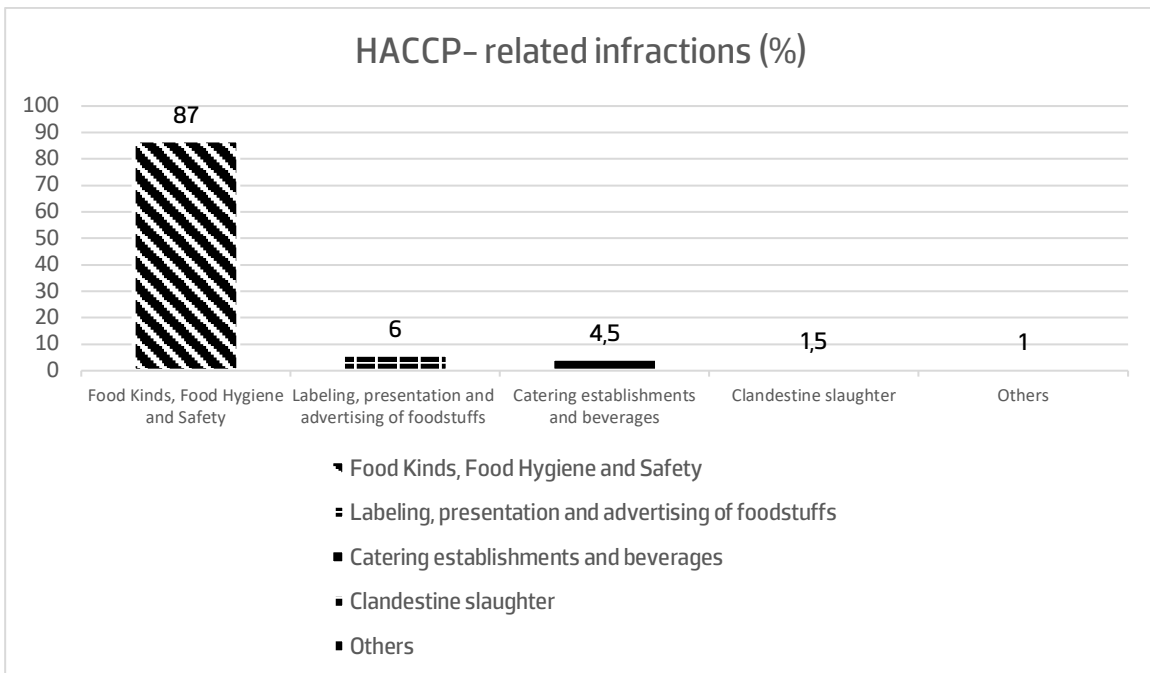
Figure 17- Type of violations related to HACCP system

As can be seen there are few specific categories that stand out statistically. The "Food Kinds" and "Food Hygiene and Safety" categories are responsible for 87% of all HACCP related violations. This is followed by the "Labeling, presentation and advertising of foodstuffs" category with only 6% of the violations, the "Catering establishments and beverages" category with a value of 4% and the "Clandestine slaughter" category with only 1.5% of the total violations related to the HACCP system. The remaining twelve categories related to the HACCP system only occupy 1% of total violations. It is also possible to further deepen the specificity of each complaint by analyzing the third part of the code for each offense. The third part of the code consists of 2 numbers. Another split was made to isolate the third part of the code. Then this third part was added to the column that had the first and second part of the code. There are some codes whose specification is extended up to six parts. However, the cases in which this happens are few. This new split allowed the violations to be distributed across 182 categories. If in the previous split where the study was done in 16 categories, the distribution had a high statistical discrepancy in the distribution of percentages, it would be expected that this new split would not contribute relevant knowledge to the study.

The study of the codes of violations was done up to the three parts that define the code and since there are already very small values of percentages of violations it is not justified to deepen the study as can be seen in annex 12.

5. Analysis and Discussion of Results

Dataset IR_Complaint results

In Figure 5 there is a comparison between the number of information requests and the number of complaints from 2014 to 2018. According to Figure 5 there is a large discrepancy between the number of complaints and the number of information requests. This is not surprising because the Portuguese population associate ASAE as an entity to report a situation and not so much to request information. From 2014 to 2018, the year of 2015 was the year with the most complaints, followed by the years of 2017, 2016, 2014 and 2018. Complaints and information requests in 2018 are only available until October. The number of complaints in 2018 is very close with 2014 and 2016, and the number of complaints by the end of the year is likely to be higher than 2014 and 2016.

In Figure 6 a comparison is made between the number of information requests solved and on hold and the number of complaints solved and on hold.

The number of information requests solved increases consecutively until 2016, decreasing in 2017 and further decreasing in 2018.

The total number of complaints and information requests in 2018 is lower than in the remaining years because there is only information until October this year.

Proportionally, there are more inquiries resolved than complaints. This is because for an information request to be solved, only an answer from an ASAE worker is necessary. To solve a complaint, an inspection of the reported establishment by an ASAE inspector is required. Therefore, it takes more time to resolve a complaint than a request for information.

As the year approaches the number of information requests on hold gradually increases until 2018, increasing significantly in the year of 2018. It is normal that in 2018 there are still on hold information requests because this data was provided by ASAE at the end of 2018. The number of complaints solved decreases and the number of complaints on hold as 2018 approaches as expected due to the type of process that must be taken for the complaint to be solved.

That been said it justifies a 29% increase in the number of complaints on hold between 2014 and 2018.

Dataset Entities results

Table 10 shows the percentage of complaints per district/archipelago in relation to the population density of each district/archipelago. Porto, Lisbon and Braga are responsible for almost half of Portugal's complaints with 47% of total complaints. All districts and archipelagos have percentages of complaints that are in accordance with their population density, except the Azores archipelago.

With its population density, this archipelago should have 2% of total complaints. This disagreement with the other data may be due to the fact that in the archipelagos ASAE only receives complaints related to wine and gambling and these areas are not much explored in this archipelago.

Dataset Complaints results

Figure 8 compares the responsible entities with the percentage of complaints they receive. Of the total complaints received, ASAE is involved in 76% of complaints. ASAE itself is responsible for 53% of complaints and 21% of complaints relate to ASAE and other entities (ASAE_OE).

It can also be noted that 14% of the total complaints are the responsibility of other entities (OE), 6% of the complaints could not be attributed (Undetermined), 3% of the complaints are the responsibility of the courts and 2% of the complaints. are the responsibility of ASAE and courts (ASAE_Courts). Only 1% of complaints are the responsibility of other entities and the courts (OE_Courts). Figure 9 shows that there are several complaints between 2014 and 2018 with more than one violation per complaint. In the range of years under study in each of the years it appears that after a split to separate the different violations of the same complaint the number of rows increases by at least 6000 rows.

Continuing the analysis of the information obtained in Annex 13 it is possible to obtain the information in Figure 10. From 2014 to 2018 there are more complaints with violations than complaints without violations. This result is to be expected because in order to have a complaint there is supposed to be legal non-compliance and usually non-compliance is related to violations. 2017 was the year where there were the most complaints with violations with 60% of that year's complaints with violations.

The years of 2014, 2016 and 2018 had 59% of total complaints of violations. The year 2015 was the year that the studied year interval had the least complaints of violations.

The codes that define the violations of complaints allow us to distinguish the category of violations. For example, if the violation code starting with "1" belongs to the Segal category. This category is an acronym that stands for food safety. Complaints in this category are those related to the HACCP system.

From Figure 11 to Figure 15 we can see the distribution of violations according to the categories to which they belong from 2014 to 2018, respectively.

Looking at Figures 11 to 15, it can be seen that in 2014 56% of complaints received were related to the HACCP system (Figure 11), from 2014 to 2015 there was a 9% decrease in complaints in this category (Figure 12) and From 2015 to 2016 there was a decrease of 1% (Figure 13). In contrast to

the previous years from 2016 to 2017 there was a 4% increase in HACCP-related complaints (Figure 14) and between 2017 and 2018 there continued to be an increase although this time it was 2% to 52% (Figure 15).

Figure 16 represents the fluctuation of values described in this last paragraph as it shows the percentages of complaints related only to the HACCP system from 2014 to 2018. Knowing the percentage of complaints related to the HACCP system from 2014 to 2018, it is important to find out what type of violations are committed in complaints received at ASAE. Table 11 contains part of the code of the violation committed with its designation and percentage. A total of 99% of complaints belong to only five of the sixteen types of violations. Food Hygiene and Safety has the highest percentage, accounting for 48% of the total complaints related to the HACCP system. As this concept of "Food Hygiene and Safety" is very comprehensive it is natural that it is the type of violation with the highest percentage.

Another equally broad genre is "Food Kinds", which accounts for 39% of violations. Following are violations related to labeling, presentation and advertising of foodstuffs (general rules) with 6% of the total complaints, catering establishments and beverages with 4.5% of the complaints and last of these five types of violations responsible for 99% of the complaints. Complaints is clandestine slaughter with 1.50% of total complaints.

As noted at the end of Chapter 4, the code analysis of each violation may be extended to more characters of the same, however the discrepancy of values does not justify it as found in Annex 12.

6. Conclusions and Future Work

This dissertation applies several phases of the process of knowledge discovery in databases, which allowed the identification of Data Analysis techniques, capable of extracting knowledge from the datasets provided by ASAE. The process was preceded by a review of the literature in food safety, HACCP system and exploratory statistical analysis. This literature review reveals a huge concern by governs and scientific community all around the world related with security and safety.

There are not many studies that look into the most common complaints of the HACCP system, but there are several studies that demonstrate the importance of this system for public health.

The World Health Organization has recognized the importance of the HACCP system for prevention of foodborne diseases for over 40 years. HACCP system became more important in 1993 when the Codex Guidelines for the Application of the Hazard Analysis Critical Control Point (HACCP) system were adopted by the Codex Alimentarius Commission.

The knowledge discovery process was applied in the practical component of this dissertation, by adapting the KDD methodology to the objectives of identifying complaints related to the HACCP system, according to the data extracted from the ASAE database. This extraction resulted in several datasets that were somewhat interrelated, but which provided different information.

The study of the datasets made possible to analyze the complaints related to the HACCP system and its distribution by district, by the competent unit for the treatment of the complaint, by the state of the complaint and by the state of investigation of the complaint.

From the dataset IR_Complaint as expected the number of complaints was higher than the number of requests for information in the studied data ranges. Evaluating the status of the complaint or request information it is verified that as the current year approaches, the number of complaints and information requests on hold is increasing because sufficient response time is required to match either the complaint that was the request information.

The results of the IR_Complaint dataset are in line with what you would expect because taking into account that ASAE is the food and economic security authority is more likely to receive complaints than requests for information. It also makes sense that with the approaching date of the current year there are fewer requests for information and denunciations fulfilled and more outstanding precisely for the response time needed for the state to be fulfilled.

The dataset entities allowed to study the distribution of the percentage of complaints in Portugal, including the archipelagos since the beginning of ASAE.

As it was possible to verify the percentage of complaints is higher in the districts with more population density such as Lisboa, Porto and Braga.

The archipelago of Madeira and the Açores have a low percentage of complaints although their population density is not so low because the ASAE only deals with complaints of gambling and wine in those regions.

Within these complaints related to the HACCP system, it was possible to verify that most of them are related to food and hygiene, food safety and food kinds, and there is also an expression of complaints related to the labelling, presentation and advertising of food products, catering establishments and with clandestine slaughter.

In an attempt for a more in-depth analysis of the type of violations committed in the complaints a scrutiny of the type of violations was made and it was verified that the majority of them were with the foodstuffs in general, as expected, and with the length failure of general and specific hygiene requirements related to food safety and hygiene. Moreover, all percentages were much lower than these two, which clearly shows which type of violation is most committed in Portugal.

The results obtained in the study of this dissertation are in agreement with what would be expected in the proposal of the theme for this dissertation.

As ASAE is the food and economic security authority unit, it is normal for 50% of the complaints to be related to the food sector. It also makes sense that ASAE receives more complaints than requests for information, not only because not everyone knows that ASAE has a platform that allows a question to be sent and subsequently clarified, as well as the information available online clarify your doubts without having to contact ASAE. For the complaint to be made, it is necessary to have ASAE contact.

The results obtained for the analysis of the complaints according to the population density are also in agreement with what would be expected. With regard to the analysis of the types of offenses committed, it was not surprising that the highest percentage of offences committed are related to food and economic security, because on a day-to-day basis, most of the time a particular establishment and ASAE are related to the lack of hygiene conditions and/or to the potential hazards that these conditions could pose to food safety.

The study of datasets made it possible to obtain a great deal of information about the state of the HACCP system in Portugal.

The ASAE is an authority and as such is more associated by the population with the complaint than the request for information. Therefore, between 2014 and 2018 ASAE always received more complaints than requests for information.

Complaints are submitted to ASAE by people. The more people there are in a given place, the more supply of establishments there will be and consecutively more complaints. It is therefore possible to

establish a relationship between the population density of a district and the archipelago and the percentage of total occupancy complaints.

As already mentioned, ASAE is better known for food security than economic security. If ASAE is best known for food safety, it is no wonder that most complaints are related to the HACCP system. HACCP system failures vary from country to country since according to the latest US Food and Drog Administration data the five major causes of HACCP system failure are sanitation monitoring; pest control; manufacturing, processing, packing and holding controls; sanitary operations and plant maintenance and personal (by this order). In Portugal, 99% of HACCP system failures are related to food hygiene and safety; food kinds; labeling, presentation and advertising of foodstuffs, catering establishments and beverages and lastly clandestine slaughter, respectively.

It can be concluded that the results revealed models capable of exploring, extracting and evidencing the most committed offences in Portugal.

It should also be noted that this dissertation project fills a gap observed in the scientific community, which aims at the lack of studies that clarify the different types of violations that can be committed and which are in fact the most committed violations in Portugal, which seems to be relevant because it is a matter related to food safety and ultimately, public health.

Several limitations occur, mainly because of the new data protection law and by the high number of data to analyze and all associated problems. Another aspect is related with the denunciations under study, which were filled out by the complainant himself through a form available on the ASAE website, there are always inconsistencies in filling out that varying according to the degree of understanding of each person. The forwarding of complaints within the ASAE itself as well as the attribution of violation codes is also manual, which can be translated into writing errors. These factors will lead to missing data or meaningless data for the study in question which also conditions the extraction of knowledge from the databases available.

It is, therefore, necessary to monitor the evolution of the IA.SAE project in the future, which is already based on the exploratory analysis of databases, focusing on food safety, and in particular the HACCP system.

References

1. The food safety hazard guidebook. Choice Rev Online. 2013;

2. Holah J. Hygiene in Food Processing and Manufacturing. In: Food Safety Management: A Practical Guide for the Food Industry. 2014.

3. ASAE. Missão, Visão e Valores [Internet]. 2018 [cited 2018 Dec 5]. Available from: https://www.asae.gov.pt/asae20/missao-visao-e-valores.aspx

4. Mil-Homens S. ASAE- HACCP [Internet]. 2007 [cited 2019 Jan 19]. Available from: http://www.asae.gov.pt/

5. Chawla N V. Data Mining for Imbalanced Datasets: An Overview. In: Data Mining and Knowledge Discovery Handbook. 2010.

6. Wang B, Li R, Perrizo W. Big Data Analytics in Bioinformatics and Healthcare. Big Data Analytics in Bioinformatics and Healthcare. 2015.

7. IA.SAE - Inteligência Artificial na Segurança Alimentar e Económica". Portugal: financiado pelo FCT/MCTES através de fundos nacionais (PIDDAC), inserindo-se no Programa Iniciativa Nacional Competências Digitais e.2030, inscrito no Plano Nacional de Reformas; INCoDe.2030, 2018.

8. Varzakas T, Tzia C. Handbook of food processing: Food safety, quality, and manufacturing processes. Handbook of Food Processing: Food Safety, Quality, and Manufacturing Processes. 2015.

9. Friendly M. A Brief History of Data Visualization. In: Handbook of Data Visualization.

10. Ramos EJ. História e Evolução do HACCP. 2017;

11. Kanji GK. Total quality management: The second industrial revolution. Total Qual Manag. 1990;

12. Skovgaard N. Scientific Criteria to Ensure Safe Food. Int J Food Microbiol. 2004;

13. Bauman HE. Introduction to HACCP. In: HACCP. 2012.

14. YU¨KSEL I, EDITORS EEAAD. HANDBOOK OF SEAFOOD QUALITY AND SAFETY MAINTENANCE AND APPLICATIONS. 2016. 14–15 p.

15. Schneider H, Stamatis DH. Failure Mode and Effect Analysis: FMEA from Theory to Execution. Technometrics. 2006;

16. Wallace CA, Mortimore SE. HACCP. In: Handbook of Hygiene Control in the Food Industry: Second Edition. 2016.

17. Figueiredo VF de, Costa Neto PL de O. Implantação do HACCP na indústria de alimentos. Gestão & Produção. 2005;

18. Skura BJ. An Evaluation of the Role of Microbiological Criteria for Foods and Food Ingredients. Vol. 19, Canadian Institute of Food Science and Technology Journal. 2013. 143 p.

19. Europeu P, Europeu C. Regulamento (CE) n°765/2008 do Parlamento Europeu e do Conselho de 9 de Julho de 2008. J Of da União Eur. 2008;30–47.

20. Stankovic I. Codex Alimentarius. In: Encyclopedia of Food and Health. 2015.

21. Regulamento (CE) n.º 852. Regulamento (UE) n° 852/2004 do parlamento europeu e do conselho. J Of da União Eur. 2004;139:54.

22. Lin YY, Chang JY, Lin CT. Identification and prediction of dynamic systems using an interactively recurrent self-evolving fuzzy neural network. IEEE Trans Neural Networks Learn Syst. 2013;

23. Codex Alimentarius Comission. General Principles of Food Hygiene. Codex Alimentarius Comission. 2003.

24. Melero B, Diez AM, Rovira J. Basic Sanitation. In: Handbook of Fermented Meat and Poultry: Second Edition. 2014.

25. Djekic I, Smigic N, Kalogianni EP, Rocha A, Zamioudi L, Pacheco R. Food hygiene practices in different food establishments. Food Control. 2014;

26. Barach JT. FSMA and Food Safety Systems: Understanding and Implementing the Rules. Wiley. 2017.

27. Safefood. Cleaning and Disinfection in Food Processing Operations. Safefood 360, Inc. 2012;

28. Forsythe SJ, Hayes PR. Food Hygiene, Microbiology and HACCP. Food Hygiene, Microbiology and HACCP. 2013.

29. Nines M. An industry perspective. EM Air Waste Manag Assoc Mag Environ Manag. 2010;

30. Moazami N. Biological Control. In: Comprehensive Biotechnology, Second Edition. 2011.

31. Sperber WH. HACCP and transparency. Food Control. 2005;

32. Amorim J, Novais MR. Guia para Controlo da Segurança Alimentar em Restaurantes Europeus. Inst Nac Saúde Dr Ricardo Jorge. 2015;

33. Doménech E, Escriche I, Martorell S. Assessing the effectiveness of critical control points to guarantee food safety. Food Control. 2008;

34. Abrol DP, Shankar U. Integrated Pest Management. In: Breeding Oilseed Crops for Sustainable Production: Opportunities and Constraints. 2015.

35. Lokunarangodage C, Wickramasinghe I, K. K. D. S R. Impact of HACCP Based Food Safety Management Systems in Improving Food Safety of Sri Lankan Tea Industry. J Tea Sci Res. 2016;

36. Ryan JM. In-transit container sanitation standards: packaging and control of packaging. In:

Guide to Food Safety and Quality During Transportation. 2017.

37. Jur C, Europeu P. I – princípios e as regras necessárias para simplificar o livre acesso e exercício das actividades de serviços. 2010;

38. Damikouka I, Katsiri A, Tzia C. Application of HACCP principles in drinking water treatment. Desalination. 2007;

39. Bell CH. Pest Management. In: Food Safety Management: A Practical Guide for the Food Industry. 2014.

40. Doyle MP, Buchanan RL. Food Microbiology: Fundamentals and Frontiers. Food Microbiology. 2013.

41. Taylor TM, Davidson PM. Chemical Preservatives and Natural Antimicrobial Compounds. In: Food Microbiology: Fundamentals and Frontiers, Third Edition. 2014.

42. Kafetzopoulos DP, Psomas EL, Kafetzopoulos PD. Measuring the effectiveness of the HACCP Food Safety Management System. Food Control. 2013;

43. Kafetzopoulos DP, Gotzamani KD. Critical factors, food quality management and organizational performance. Food Control. 2014;

44. DL194/. Decreto–Lei nº 194/2012 ASAE. Diário da República 1ª série. 2012;

45. Pública P de S. A História [Internet]. Available from: http://www.psp.pt/Pages/apsp/historia.aspx?menu=1&submenu=4

46. Decreto–Lei n.º 32945. Imprensa Nac Lisboa.

47. ASAE AS dos F da. ASAE – Autoridade de Segurança Alimentar e Económica [Internet]. Available from: http://asf-asae.pt/asf-asae/historia/

48. Decreto–Lei n.º 329-D/74 de 7 de outubro. Diário da República.

49. Decreto–Lei n.º 23/84. Diário da República. 1984;

50. Decreto–Lei n.º 14/93. Diário da República. 1993.

51. Diário da República. Decreto–Lei n.º 237/2005. 2005;7486–97.

52. ASAE. A direção [Internet]. 2018 [cited 2018 Dec 2]. Available from: https://www.asae.gov.pt/asae20/conselho-cientifico.aspx

53. ASAE. ASAE-Cooperação Internacional Bilateral. https://www.asae.gov.pt/cooperacao/internacional/bilateral/medio-oriente.aspx. 2018.

54. ASAE. Como Actua a ASAE [Internet]. 2018 [cited 2018 Nov 20]. Available from: https://www.asae.gov.pt/inspecao-fiscalizacao/como-atua-a-asae.aspx

55. Melo RV de. ASAE como entidade dinamizadora das boas práticas europeias para os PALOP [Internet]. Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa; 2018. Available from: http://hdl.handle.net/10362/57718

56.   Alencastro ER de. Análise Comparativa da Gestão das políticas públicas no âmbito sanitário alimentar: o caso de Portugal e do Brasil [Internet]. nstituto de Investigação Interdisciplinar da Universidade de Coimbra; 2018. Available from: http://hdl.handle.net/10316/80695

57.   ASAE. ASAE e Argélia [Internet]. 2015 [cited 2019 Jan 19]. Available from: https://www.asae.gov.pt/cooperacao-nacional-e-internacional/cooperacao-internacional-/assinatura-do-protocolo-com-a-argelia.aspx

58.   ASAE. ASAE- América [Internet]. 2018 [cited 2019 Jan 19]. Available from: https://www.asae.gov.pt/cooperacao/internacional/bilateral/america.aspx

59.   ASAEnews nº 107. Delegação da ASAE na China durante o mês de abril. 2017.

60.   ASAE. ASAE- Europa. 2018.

61.   ASAE. ASAE- Médio Oriente. 2018.

62.   Emprego M da E e do. Diário da República n.º 163/2012. Diário da República. 2012;Série I de.

63.   ARFA, BfR e ASAE lançam almanaque e compromissam cooperação [Internet]. [cited 2019 Apr 22]. Available from: http://www.arfa.cv/index.php/noticias/1028-arfa-bfr-e-asae-lancam-almanaque-e-compromissam-cooperacao

64.   Homepage – BfR [Internet]. [cited 2019 Apr 22]. Available from: https://www.bfr.bund.de/en/home.html

65.   BVL – Federal Office of Consumer Protection and Food Safety [Internet]. [cited 2019 Apr 22]. Available from: https://www.bvl.bund.de/EN/Home/homepage_node.html

66.   Aecosan – Agencia Española de Consumo, Seguridad Alimentaria y Nutrición [Internet]. [cited 2019 Apr 22]. Available from: http://www.aecosan.msssi.gob.es/AECOSAN/web/home/aecosan_inicio.htm

67.   anses. Agência francesa de segurança alimentar, ambiental e do trabalho.

68.   SIC – Serviços de Investigação Criminal – Mais um site WordPress [Internet]. [cited 2019 Apr 22]. Available from: http://policiaeconomica.gv.ao/

69.   IGC - Sobre Nós [Internet]. [cited 2019 Apr 30]. Available from: http://igc.co.ao/perfil.html

70.   Minhotur Divulga Plano Director Nacional do Turismo - WELCOME TO ANGOLA [Internet]. [cited 2019 Apr 30]. Available from: http://www.welcometoangola.co.ao/_plano_director_do_turismo

71.   Portal do cidadão | Ver Prestador de Serviço [Internet]. [cited 2019 Apr 30]. Available from: http://www.cidadao.gov.ao/VerPrestadorServico.aspx?id=314

72.   Home [Internet]. [cited 2019 Apr 22]. Available from: http://www.arfa.cv/

73.   Home [Internet]. [cited 2019 Apr 22]. Available from: http://www.igae.cv/

74.   Missão e Atribuições – DSE [Internet]. [cited 2019 Apr 22]. Available from:

https://www.economia.gov.mo/pt/web/public/pg_dse_md

75. Governo regula fiscalização da INAE / Notícias / Imprensa / Início - Portal do Governo de Moçambique [Internet]. [cited 2019 Apr 30]. Available from: http://www.portaldogoverno.gov.mz/index.php/por/Imprensa/Noticias/Governo-regula-fiscalizacao-da-INAE

76. ASAE assina protocolo e presta formação profissional a congénere são-tomense [Internet]. [cited 2019 Apr 30]. Available from: https://www.dnoticias.pt/hemeroteca/580381-asae-assina-protocolo-e-presta-formacao-profissional-a-congenere-sao-tomense-KHDN580381

77. início - TATOLI Agência Noticiosa de Timor-Leste [Internet]. [cited 2019 Apr 30]. Available from: http://www.tatoli.tl/pt/

78. Página Inicial da Anvisa - Anvisa [Internet]. [cited 2019 Apr 22]. Available from: http://portal.anvisa.gov.br/

79. Inmetro | Portal do Inmetro [Internet]. [cited 2019 Apr 22]. Available from: http://www4.inmetro.gov.br/

80. Dipov — Ministério da Agricultura, Pecuária e Abastecimento [Internet]. [cited 2019 Apr 22]. Available from: http://www.agricultura.gov.br/assuntos/inspecao/produtos-vegetal

81. Dipoa — Ministério da Agricultura, Pecuária e Abastecimento [Internet]. [cited 2019 Apr 22]. Available from: http://www.agricultura.gov.br/assuntos/inspecao/produtos-animal

82. DGCCRF - Direction Générale de la Concurrence, de la Consommation et de la Répression des Fraudes | Le portail des ministères économiques et financiers [Internet]. [cited 2019 Apr 22]. Available from: https://www.economie.gouv.fr/dgccrf

83. ONSSA - Accueil [Internet]. [cited 2019 Apr 22]. Available from: http://www.onssa.gov.ma/fr/index.php

84. Ministère du Commerce de la République Tunisienne [Internet]. [cited 2019 Apr 30]. Available from: http://www.commerce.gov.tn/Fr/accueil_46_50

85. Welcome to the Canadian Food Inspection Agency - Canadian Food Inspection Agency [Internet]. [cited 2019 Apr 22]. Available from: http://www.inspection.gc.ca/eng/1297964599443/1297965645317

86. U S Food and Drug Administration Home Page [Internet]. [cited 2019 Apr 22]. Available from: https://www.fda.gov/

87. USDA [Internet]. [cited 2019 Apr 22]. Available from: https://www.usda.gov/

88. Secretariado Permanente testemunha celebração de protocolo entre IETP e ASAE | Fórum para a Cooperação Económica e Comercial entre a China e os Países de Língua Portuguesa

(Macau) [Internet]. [cited 2019 Apr 30]. Available from: http://www.forumchinaplp.org.mo/permanent-secretariat-witnesses-signing-of-a-cooperative-protocol-between-ietp-and-asae/?lang=pt

89. [Internet]. [cited 2019 Apr 22]. Available from: http://www.saic.gov.cn/

90. CFDA China Agent [Internet]. [cited 2019 Apr 30]. Available from: https://www.qservegroup.com/en/services/expertise-fields/regulatory-affairs/legal-representation/cfda-china-agent

91. General Administration of Quality Supervision,Inspection and Quarantine of the PRC (AQSIQ) [Internet]. [cited 2019 Apr 22]. Available from: http://english.aqsiq.gov.cn/

92. 3rd Shenzhen Food Safety Risk Communication Forum 2018 | CODEXALIMENTARIUS [Internet]. [cited 2019 Apr 30]. Available from: http://www.fao.org/fao-who-codexalimentarius/news-and-events/news-details/en/c/1176381/

93. Centre for Food Safety – Home [Internet]. [cited 2019 Apr 22]. Available from: https://www.cfs.gov.hk/english/index.html

94. [Internet]. [cited 2019 Apr 22]. Available from: http://www.nifds.go.kr/en/index.do

95. ENGLISH | [Internet]. [cited 2019 Apr 22]. Available from: http://www.fsc.go.jp/english/

96. Home [Internet]. [cited 2019 Apr 22]. Available from: https://fssai.gov.in/home

97. Saudi Food and Drug Authority [Internet]. [cited 2019 Apr 30]. Available from: https://www.sfda.gov.sa/en/Pages/default.aspx

98. UAE Ministry of Climate Change and Environment [Internet]. [cited 2019 Apr 22]. Available from: https://www.moccae.gov.ae/

99. General Directorate Of Food And Control [Internet]. [cited 2019 Apr 22]. Available from: https://www.tarimorman.gov.tr/GKGM/Sayfalar/EN/AnaSayfa.aspx

100. Parlamento Europeu. Regulamento (CE) nº882/2004, relativo aos controlos oficiais realizados para assegurar a verificação do cumprimento da legislação e aos alimentos para animais e aos géneros alimentícios e das normas relativas à saúde e ao bem-estar dos animais, do Parlame. 2001;1–8.

101. Como Atua a ASAE [Internet]. [cited 2019 Apr 30]. Available from: https://www.asae.gov.pt/inspecao-fiscalizacao/como-atua-a-asae.aspx

102. Artigo 36º [Internet]. 2002 [cited 2019 Apr 30]. Available from: https://www.mendeley.com/reference-management/web-importer/#id_2

103. Procedimentos Operacionais [Internet]. [cited 2019 May 2]. Available from: https://www.asae.gov.pt/inspecao-fiscalizacao/procedimentos-operacionais.aspx

104. Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics.

Bioinformatics. 2007;23(19):2507–17.

105. Maneth S, Poulovassilis A. Data science. Computer Journal. 2017.

106. Lele A. Big data. In: Smart Innovation, Systems and Technologies. 2019.

107. Sowmya R, Suneetha KR. Data Mining with Big Data. In: Proceedings of 2017 11th International Conference on Intelligent Systems and Control, ISCO 2017. 2017.

108. boyd danah, Crawford K. CRITICAL QUESTIONS FOR BIG DATA. Information, Commun Soc. 2012;

109. Gartner Inc. What Is Big Data? – Gartner IT Glossary – Big Data. Gartner IT Glossary. 2013.

110. Marx V. The big challenges of big data. Nature. 2013;

111. Saporito P. The 5 V's of Big Data. Best's Rev. 2013;

112. Fayyad U, Piatetsky-Shapiro G, Smyth P. From Data Mining to Knowledge Discovery in Databases. AI Mag. 1996;

113. Keim D, Qu H, Ma KL. Big-data visualization. IEEE Computer Graphics and Applications. 2013.

114. Abzalov M. Exploratory data analysis. In: Modern Approaches in Solid Earth Sciences. 2016.

115. Blog | Spotless [Internet]. [cited 2019 May 7]. Available from: https://web.archive.org/web/20171018181046/https://spotlessdata.com/blog/explori ng-data-analysis

116. Exploratory and Confirmatory Analysis: What's the Difference? I Sisense [Internet]. [cited 2019 May 7]. Available from: https://www.sisense.com/blog/exploratory-confirmatory-analysis-whats-difference/

117. Fisher EC, Marean CW. Data visualization. Encycl Earth Sci Ser. 2017;

118. Kotu V, Deshpande B. Predictive Analytics and Data Mining: Concepts and Practice with RapidMiner. Predictive Analytics and Data Mining: Concepts and Practice with RapidMiner. 2014.

119. Chen M, Laramee RS, Ebert D, Hagen H, van Liere R, Ma KL, et al. Data, Information, and Knowledge in Visualization. IEEE Comput Graph Appl. 2009;

120. Card SK, Mackinlay JD, Shneiderman B. Information visualization: using vision to think. In: Readings in information visualization: using vision to think. 1999.

121. Fry B. Visualizing Data: Exploring and Explaining Data With the Processing Environment. American Scientist. 2007.

122. Ware C. Information Visualization: Perception for Design: Second Edition. Information Visualization: Perception for Design: Second Edition. 2004.

123. Jick TD. Mixing qualitative and quatitative methods. Adm Sci Q. 1979;

124. IRMA IRMA. Big Data: Concepts, Methodologies, Tools, and Applications. 2016. 2200 p.

125. Kühne R. Categorical Data Analysis. In: The International Encyclopedia of Communication Research Methods. 2017.

126. Handbook of Data Visualization. Handbook of Data Visualization. 2007.

127. Lutz M. Learning Python, 5th Edition. CEUR Workshop Proceedings. 2015.

128. McKinney W. Python for Data Analysis Data Wrangling with Pandas, NumPy, and IPython. Climate Change 2013 – The Physical Science Basis. 2018.

129. McKinney W. Python for Data Analysis second edititon [Internet]. 2017. 1–470 p. Available from: papers3://publication/uuid/82323D3A-5062-440A-981F-9DB5F38BBC63

130. DataCamp. Python for Data Science. Cheatsheet. 2001;

131. McKinney W. Data Structures for Statistical Computing in Python. In: Proceedings of the 9th Python in Science Conference. 2010.

132. Harrison M, Prentiss M. Learning the Pandas Library: Python Tools for Data Munging, Analysis, and Visual. 2016;

133. Jupyter Community PJ. Project Jupyter. Project Jupyter. 2018.

134. Maimon O, Rokach L. Introduction to Knowledge Discovery and Data Mining. In: Data Mining and Knowledge Discovery Handbook. 2010.

135. Witten IH, Frank E, Hall MA. Data Mining Practical Machine Learning Tools And Techniques. Data Mining. 2005.

136. Silva MP dos S. Mineração de Dados – Conceitos , Aplicações e Experimentos com Weka. Dados. 2009;

137. Jothi N, Rashid NA, Husain W. Data Mining in Healthcare – A Review. In: Procedia Computer Science. 2015.

138. Provost F, Fawcett T. Data Science and its Relationship to Big Data and Data-Driven Decision Making. Big Data. 2013;

Annexes

Annex 1 – Data description of dataset IR_Complaint2014

Attributes	Description	Variable Type	Export 14
ID_CORRESPONDENCIA	Data entry number	Number	[3336477:3696004]
NID	NID– Identification Number E–Entity 8–Number of entry 14–Year SC–Central Services	Object	[E/8/14/SC:E/143685/14/SC]
ESTADO_DOC	"Cumprido"– Solved "Pendente"– On Hold	Object	Cumprido–22236; Pendente–387
DT_REGISTO	Date of first and last registration	Date	[02/01/2014:31/12/2014]
DESC_CONTEUDO	Complaint description	Object	–
DESC_TP_DOC	How the data was received	Object	E–MAIL
DESC_ASSUNTO	"Denúncia"– Complaints "Pedido de informações"– Information requests	Object	Denúncia–15315; Pedido de Informações–7308
EMISSOR2	Complainant email or name	Object	–
RECLAMANTE	Complainant name	Object	–
ENT_VISADA	Denounced entity	Object	–
EMAIL_CONTENT	Email content	Object	–

Annex 2– Data description of dataset IR_Complaint2015

Attributes	Description	Variable Type	Export 15
ID_CORRESPONDENCIA	Data entry number	Number	[3696005:3942747]
NID	NID– Identification Number E–Entity 1–Number of entry 15–Year SC–Central Services	Object	[E/1/15/SC:E/106250/15/SC]
ESTADO_DOC	"Cumprido"– Solved "Pendente"– On Hold	Object	Cumprido–22898;Pendente–2545
DT_REGISTO	Date of first and last registration	Date	[02/01/2015:31/12/2015]
DESC_CONTEUDO	Complaint description	Object	–
DESC_TP_DOC	How the data was received	Object	E–MAIL
DESC_ASSUNTO	"Denúncia"– Complaints "Pedido de informações"– Information requests	Object	Denúncia–17166; Pedido de Informações–8277
EMISSOR2	Complainant email or name	Object	–
RECLAMANTE	Complainant name	Object	–
ENT_VISADA	Denounced entity	Object	–
EMAIL_CONTENT	Email content	Object	–

Annex 3– Data description of dataset IR_Complaint 16

Attributes	Description	Variable Type	Export 16
ID_CORRESPONDENCIA	Data entry number	Number	[3942748:4180996]
NID	NID– Identification Number E–Entity 1–Number of entry 16–Year SC–Central Services	Object	[E/1/16/SC:E/93157/16/SC]
ESTADO_DOC	"Cumprido"– Solved "Pendente"– On Hold	Object	Cumprido–21033;Pendente–3301
DT_REGISTO	Date of first and last registration	Date	[01-01-2016:31-12-2016]
DESC_CONTEUDO	Complaint description	Object	–
DESC_TP_DOC	How the data was received	Object	E–MAIL
DESC_ASSUNTO	"Denúncia"– Complaints "Pedido de informações"– Information requests	Object	Denúncia–15769; Pedido de Informações–8565
EMISSOR2	Complainant email or name	Object	–
RECLAMANTE	Complainant name	Object	–
ENT_VISADA	Denounced entity	Object	–
EMAIL_CONTENT	Email content	Object	–

Annex 4–Data description of dataset IR_Complaint2017

Attributes	Description	Variable Type	Export 17
ID_CORRESPONDENCIA	Data entry number	Number	[4180997:4363535]
NID	NID– Identification Number E–Entity 1–Number of entry 17–Year SC–Central Services	Object	[E/1/17/SC:E/59541/17/SC]
ESTADO_DOC	"Cumprido"– Solved "Pendente"– On Hold	Object	Cumprido–20605;Pendente–4649
DT_REGISTO	Date of first and last registration	Date	[01-01-2017:31-12-2017]
DESC_CONTEUDO	Complaint description	Object	–
DESC_TP_DOC	How the data was received	Object	E–MAIL
DESC_ASSUNTO	"Denúncia"– Complaints "Pedido de informações"– Information requests	Object	Denúncia–16935; Pedido de Informações–8319
EMISSOR2	Complainant email or name	Object	–
RECLAMANTE	Complainant name	Object	–
ENT_VISADA	Denounced entity	Object	–
EMAIL_CONTENT	Email content	Object	–

Annex 5–Data description of dataset IR_Complaint2018

Attributes	Description	Variable Type	Export 18
ID_CORRESPONDENCIA	Data entry number	Number	[4363536:4645634]
NID	NID– Identification Number E–Entity 1–Number of entry 18–Year SC–Central Services	Object	[E/1/18/SC:E/147130/18/SC]
ESTADO_DOC	"Cumprido"– Solved "Pendente"– On Hold	Object	Cumprido–15071;Pendente–7936
DT_REGISTO	Date of first and last registration	Date	[01-01-2017:31-12-2017]
DESC_CONTEUDO	Complaint description	Object	–
DESC_TP_DOC	How the data was received	Object	E–MAIL
DESC_ASSUNTO	"Denúncia"– Complaints "Pedido de informações"– Information requests	Object	Denúncia–15173; Pedido de Informações–7834
EMISSOR2	Complainant email or name	Object	–
RECLAMANTE	Complainant name	Object	–
ENT_VISADA	Denounced entity	Object	–
EMAIL_CONTENT	Email content	Object	–

Annex 6–Data description of dataset DENUNCIAS_2014

Attributes	Variable Type	Description	DENUNCIAS_2014
NID	Object	Complaint identification number: E–Entity 15–Number of entry 14–Year SC–Central Services	[E/15/14/SC:E/143643/14/SC]
Competência	Object	Organization responsible for handling the complaint: "ASAE"–ASAE "ASAE e Tribunais"– ASAE and Courts "ASAE e outra Entidade"– "ASAE and other Entity" "Indeterminada"–Undetermined "Outra Entidade"– Another Entity "Outra Entidade e Tribunais"–Other Entity and Courts "Tribunais"– Courts	ASAE–9113 ASAE e Tribunais–281 ASAE e outra Entidade–4567 Indeterminada–1536 Outra Entidade–4076 Outra Entidade e Tribunais–132 Tribunais–1219
Estado	Object	Status of complaint: Solved – "Cumprido" On Hold – "Pendente"	Cumprido–20443 Pendente–489
Estado Averiguação	Object	Status of inquiry: Filed – "Arquivada" Filed In Found – "Arquivada Averiguada" Inquiry– "Em averiguação" To inquire – "Por averiguar"	Arquivada–15498 Arquivada Averiguada–4945 Em averiguação–65 Por averiguar–416
Data arquivo	Date Object	Filing date of the complaint	[02/01/14:19/02/18)
UO destino (1º Circuito)	Object	GPEF– Office of Strategic Planning and Training NIP–Nucleus of Procedural Instruction UNIIC–National Information and Criminal Investigation Unit NIF I –Nucleus of Inspection and Inspection I NIF II–Nucleus of Inspection and Inspection II NBE– Nucleus of Specialized Brigades NIF IV–Nucleus of Inspection and Inspection IV NIF V– Inspection and Inspection Center V – Coimbra NIF VII –Nucleus of Inspection and Inspection VII NIF VIII –Nucleus of Inspection and Inspection VIII NIF IX –Nucleus of Inspection and Inspection IX DGC –Division of Management of Administrative Offenses SECIG–Secretariat of the Inspector General	/SC/IG/DAJC–19 /SC/IG/UNIIC–945 /SC/IG/UNIIC/DAPI–119 /SC/IG/UNO–2 /SC/IG/UNO/DEPO–115
UO destino (1º Circuito)	Object	GRI–Office of International Relations DCAAI –Inspection Activity Coordination and Evaluation Division DIP –Division of Public Information SC –Central Services URN –Northern Regional Unit URC –Regional Center Unit UO XI–EVR –Evora UO XII–FAR –Faro IG –General Inspector SGAAL –Sub–Inspector General Administration and Logistics Area SGAT –Sub–Inspector General Technical Area	/URC/UO V–CBR Norte/NIF V–4 /URC/UO VI–CTB–338 /URN/DIR–3150
UO destino (1º Circuito)	Object	URS –Regional Southern Unit SECSIG –Secretariat of the General Sub–Inspectors (Administrative and Technical Area) DAJC –Department of Legal Affairs and Administrative Offenses DRA –Division of Food Risks DAL –Department of Administration and Logistics DGRH –Human Resources Management Division RH –Human Resources DAG –Division of Management Support DSPD –Digital Support and Expertise Division UNO –National Operations Unit DAPI –Division of Analysis and Information Research UCII –Central Research and Intervention Unit	/SC/IG/UNO/DIP–176 /SC/SECIG–3 /URC/DIR–1392 /URC/UO IV–CBR/NIF IV–1

UO destino (1º Circuito)	Object	NIF III –Nucleus of Inspection and Inspection III UO V–CBR Norte –Coimbra / North UO IV–CBR –Coimbra UO VI–CTB –Castelo Branco NIF VI –Nucleus of Inspection and Inspection VI UO VIII–LSB Oeste –Lisbon / West UO VII–LSB –Lisbon UO X–STR –Santarem NIF X –Nucleus of Inspection and Inspection X NIF XI –Nucleus of Inspection and Inspection XI NIF XII –Nucleus of Inspection and Inspection XII DAS –Support and Security Division	/URN/UO III–MIR–277 /URS–2 /URS/DIR–4413 /URS/NATA–1 /URS/UO IX–LSB Sul–2
UO destino (1º Circuito)	Object	GCAAI –Office of Coordination and Evaluation of the Insight Activity NIC I –Nucleus of Criminal Investigation I NIC II– Nucleus of Criminal Investigation II NIF V –Tondela SC –Counter–Order Section NIF V.TInspection and Inspection Center V – Tondela SIC –Criminal Investigation Section LBPV –Laboratory of Beverages and Viticultural Products UO II–BRC –Barcelos NIIP –Nucleus of Investigation and Procedural Instruction NATA –Nucleus of Technical and Administrative Support	/URS/UO IX–LSB Sul–2 /URS/UO VII–LSB–30 /URS/UO VIII–LSB Oeste–2 /URS/UO X–STR–468 /URS/UO XI–EVR–406
UO destino (1º Circuito)	Object	UO III–MIR– Mirandela DGRP –Asset Management Division UO IX–LSB Sul –Lisbon / South EXP –Office hour TES –Treasury SCO –Accounting and Budget Service SAP –Provisioning Service SPMA –Seized Heritage and Material Services SGV –Car Management Service DIR –Office of the Directorate of the Southern Regional Unit	/URS/UO XII–FAR–826 /URS/UO XII–FAR/NIF XII–1
UO destino (1º Circuito)	Object	"DEPO –Division of Studies and Operational Planning DCO –Operational Control Division DRAL –Department of Food Risks and Laboratories LFQ– Physical Chemistry Laboratory LM –Microbiology Laboratory UO I–PRT –Porto"	/URN/UO I–PRT/NIF I–2 /URN/UO II–BRC /NIF II–2
Data de envio (1º Circuito)	Date	Date of dispatch (1st circuit)	[02/01/14:02/01/15]
Data de averig. (1ª FOF)	Date / Hour	Date and time of the investigation of the complaint in the FOF (File of Operator Supervised)	[27/03/14 16:27 ; 21/03/16 15:30]
Local de registo	Object	Where the complaint was registered. Same definition of object attribute acronyms	/SC/IG/UNO/DIP–20892 /SC/SGAAL/DAL/DAG/EXP–1 /URS/NATA–31
Tipo	Object	How was the complaint received at ASAE: "AUTO"– judicial order "CARTA"– letter "E–MAIL" "FAX" "OFÍCIO"– offical communication "OUTRO"– other "PRESENCIAL" – presential "TELEFONEMA"– phone call	AUTO–1 CARTA–3670 E–MAIL–15315 FAX–58 OFFICIAL COMMUNICATION–1 OTHER–3 PRESENCIAL–3 TELEFONEMA–1873
Tipo Remetente	Object	Type of sender that reported: "Agente Económico" –Economic Agent "Agente Económico (anónimo)" –Economic Agent (anonymous) "Associação Representativa" –Repressentative Association "Câmara Municipal" –Town Council "Organismo Oficial" –Official Body "Outro" –Other "Particular" –Private "Tribunal" –Court	Agente Económico–777 Agente Económico (anónimo)–9081 Associação Representativa–123 Câmara Municipal–186 Organismo Oficial–1039 Outro–98 Particular–9616 Tribunal–4

Remetente	Object	Name of complainant	–
E-mail	Object	E-mail of complainant	–
Nome (e-mail)	Object	Name and/or e-mail of camplainant	–
Remetida por Denunciante	Object	If the complaint was forwarded by the complainant: "SIM" –Yes "NÃO" –No	SIM–20197 NÃO–727
Tipo Denunciante	Object	Type of complainant: "Agente Económico" –Economic Agent "Agente Económico (anónimo)" –Economic Agent (anonymous) "Associação Representativa" –Repressentative Association "Câmara Municipal" –Town Council "Organismo Oficial" –Official Body "Outro" –Other "Particular" –Private "Tribunal" –Court	Agente Económico – 790 Agente Económico (anónimo) – 9118 Associação Representativa – 96 Câmara Municipal – 154 Organismo Oficial – 535 Outro – 77 Particular – 10152 Tribunal – 2
Denunciante	Object	Name of complainant	–
Motivo não preench. den.	Object	Reason for not completing the complaint: "Informação imperceptível"-Imperceptible information "Não fornecido"–Not provided	Informação imperceptível–4 Não fornecido–9914
Entidade Visada	Object	Type of entity targeted	–
Motivo não preench. ent.	Object	Reason for not completing the entity field "Informação imperceptível"-Imperceptible information "Não fornecido"–Not provided	Informação imperceptível–3 Não fornecido–1727
Morada	Object	Entity's address	–
Cód. Postal	Object	Entity's Postal Code	–
Localidade	Object	Locality	–
Dist/Conc/Freg	Object	District / Council / Parish	–
Actividades	Object	"Boite" (III.1.3.2) –Boite "Cafetaria" (III.1.2.5) – Coffee Shop "Café" (III.1.2.3) – Cafe "Casa de Pasto" (III.1.1.3) – House of Grass "Cervejaria" (III.1.2.2) – Brewery "Clube noturno" (III.1.3.1) – Nightclub "Dancing" (III.1.3.6) – Dancing "Fast-food" (III.1.1.8) – Fast-food "Gelataria" (III.1.2.7) – Ice cream parlor "Marisqueira" (III.1.1.2) – Seafood restaurant "Pastelaria" (III.1.2.4) – Pastry "Pizzaria" (III.1.1.4) – Pizzeria "Snack-bar" (III.1.1.5) – Snack bar	"Boite" (III.1.3.2) –1 "Cafetaria" (III.1.2.5) –60 "Café" (III.1.2.3) –2043 "Casa de Pasto" (III.1.1.3) –18 "Cervejaria" (III.1.2.2) –27 "Clube noturno" (III.1.3.1) –1 "Dancing" (III.1.3.6) –9 "Fast-food" (III.1.1.8) –5 "Gelataria" (III.1.2.7) –27 "Marisqueira" (III.1.1.2) –7 "Pastelaria" (III.1.2.4) –670 "Pizzaria" (III.1.1.4) –42 "Snack-bar" (III.1.1.5) –102
		"Taberna" (III.1.2.6) – Tavern "Take-away" (III.1.1.7) – Take away "Atelier de Arranjos de Costura (IX.9)\n+ Lavandarias/Engomadorias" (IX.39) – Atelier of Sewing Arrangements (IX.9) \ n + Laundry / Ironing (IX.39) "Cabeleireiros (IX.12)\n+ Tatuagens/Piercing's" (IX.52) – Hairdressers (IX.12) \ n + Tattoos / Piercing's (IX.52)	"Taberna" (III.1.2.6) –76 "Take-away" (III.1.1.7) –15 Atelier de Arranjos de Costura (IX.9)\n+ Lavandarias/Engomadorias (IX.39) –1 Cabeleireiros (IX.12)\n+ Tatuagens/Piercing's (IX.52) –1

Actividades	Object	"Sem actividade identificada (Z)\n+ "Pastelaria""(III.1.2.4) – No identified activity (Z) \ n + "Pastry" (III.1.2.4) "Sem actividade identificada (Z)\n+ Clínicas Médicas" (IX.16) – No identified activity (Z) \ n + Medical Clinics (IX.16)"Sem actividade identificada (Z)\n+ Comercialização de Aparelhos e Equipamentos para Comunicações" (VIII.9) – No identified activity (Z) \ n + Marketing of Communications Devices and Equipment "Sem actividade identificada (Z)\n+ Mercados Municipais" (V.19) – No identified activity (Z) \ n + Municipal Markets "Sem actividade identificada (Z)\n+ Oficinas de Automóveis e Motociclos" (IX.42) – No identified activity (Z) \ n + Automobile and Motorcycle Workshops "Tabaco/Tabacarias" (VIII.22) – Tobacco / Tobacco Shop	Sem actividade identificada (Z)\n+ "Pastelaria" (III.1.2.4) –1 Sem actividade identificada (Z)\n+ Clínicas –1 Médicas (IX.16) –1 Sem actividade identificada (Z)\n+ Comercialização de Aparelhos e Equipamentos para Comunicações (VIII.9) –1 Sem actividade identificada (Z)\n+ Mercados Municipais (V.19) –1 Sem actividade identificada (Z)\n+ Oficinas de Automóveis e Motociclos (IX.42) –1 Tabaco/Tabacarias (VIII.22) –38
Actividades	Object	"Administração de Condomínios" (IX.1) – Administration of Condos "Agentes de Gás" (VIII.1) – Gas Agents "Agências Funerárias" (IX.5) – Funeral Agencies "Agências de Publicidade e Marketing" (IX.2) – Advertising and Marketing Agencies "Agências de Venda de Bilhetes" (IX.3) – Ticket Agencies "Sapateiros" (IX.49) – Shoemakers "Seguradoras" (IX.50) – Insurance companies "Sem actividade identificada" (Z) – No identified activity "Serralharias" (IX.51) – Locksmiths "Ópticas" (VIII.42) – Optics "Tratamento de resíduos" (II.26) – Waste treatment	Administração de Condomínios (IX.1) –55 Agentes de Gás (VIII.1) –27 Agências Funerárias (IX.5) –68 Agências de Publicidade e Marketing (IX.2) –15 Agências de Venda de Bilhetes (IX.3) –1 Sapateiros (IX.49) –7 Seguradoras (IX.50) –30 Sem actividade identificada (Z) –3037 Serralharias (IX.51) –37 Ópticas (VIII.42) –25 Tratamento de resíduos (II.26) –5
Actividades	Object	"Sex–shop e outras atividades de serviços pessoais diversos" (VIII.52) – Sex shop and other miscellaneous personal service activities "Smartshops"(X.18) – Smartshops "Stand de Automóveis Novos" (VIII.53) – New Cars Stand "Stand de Automóveis Usados" (VIII.54) – Used Cars Stand "Stand de Máquinas e Motores (inclui equipamentos de construção civil)" (VIII.57) – Machinery and Engine Stand (includes construction equipment) "Sucatas" (VIII.58) – Scrap "Suiniculturas" (I.3.2) – Pig farms "Supermercados/Minimercados/Mercearias" (V.20) – Supermarkets / Minimarkets / Groceries	Sex–shop e outras atividades de serviços pessoais diversos (VIII.52) –22 Smartshops (X.18) –1 Stand de Automóveis Novos (VIII.53) –18 Stand de Automóveis Usados (VIII.54) –276 Stand de Máquinas e Motores (inclui equipamentos de construção civil) (VIII.57) –2 Sucatas (VIII.58) –37 Suiniculturas (I.3.2) –12 Supermercados/Minimercados/Mercearias (V.20) –1097
Actividades	Object	"Supermercados/Minimercados/Mercearias \n+ "Café"" (III.1.2.3) – Supermarkets / Minimarkets / Groceries \ n + "Cafe" "Supermercados/Minimercados/Mercearias\n+ "Snack-bar"" (III.1.1.5) – "Acondicionamento e conservação de hortofrutícolas frescos (inclui produtos minimamente processados – centros normalização)" (II.11.1) – Supermarkets / Minimarkets / Groceries \ n + "Snack bar" "Talhos" (V.1) – Butchers "Tatuagens/Piercing's" (IX.52) – Tattoos / Piercing's "Telecomunicações" (IX.54) – Telecommunications "Transportador" (IX.67) – Carrier "Transportes Públicos" (IX.56) – Public Transport	Supermercados/Minimercados/Mercearias (V.20)\n+ "Café" (III.1.2.3) –4 Supermercados/Minimercados/Mercearias (V.20)\n+ "Snack-bar" (III.1.1.5) –1 Acondicionamento e conservação de hortofrutícolas frescos (inclui produtos minimamente processados – centros normalização) (II.11.1) –7 Talhos (V.1) –158 Tatuagens/Piercing's (IX.52) –26 Telecomunicações (IX.54) –533 Transportador (IX.67) –42 Transportes Públicos (IX.56) –36
Actividades	Object	"Turismo de Natureza" (IX.22) – Tourism of Nature "Táxis" (IX.53) – Taxis "Venda Automática Produtos Alimentares" (V.22) – Sell Automatic Food Products "Venda ambulante" (V.16) – Street vending "Venda de artigos de desporto" (VIII.64) – Sale of sporting goods "Vendas à distância (por Catálogo e Internet)" (VII.1) – Distance selling (by Catalog and Internet) "Veículos em fim de vida (parques de sucata)" (X.15) – End–of–life vehicles (scrap yards) "Vidreiras/Cristais" (VIII.8) – Glassware / Crystals "Várias espécies" (I.3.7) – Several species "Óleos novos e usados" (X.10) – New and Used Oils	Turismo de Natureza (IX.22) –2 Táxis (IX.53) –103 Venda Automática Produtos Alimentares (V.22) –20 Venda ambulante (V.16) –149 Venda de artigos de desporto (VIII.64) –8 Vendas à distância (por Catálogo e Internet) (VII.1) –964 Veículos em fim de vida (parques de sucata) (X.15) –7 Vidreiras/Cristais (VIII.8) –3 Várias espécies (I.3.7) –9 Óleos novos e usados (X.10) –2

Área Geográfica/DR	Object	Same attribute definitions of "UO destino (1º Circuito)" except: "Sem identificação localidade"– No location identified	/URS/UO VII–LSB–23 /URS/UO VIII–LSB Oeste–1 /URS/UO X–STR–492 /URS/UO XI–EVR–468 /URS/UO XII–FAR–862 Sem identificação localidade–7293
Conteúdo	Object	Content of the complaint	–
Classificação Conteúdo	Object	"Contraordenação\n+ Indefinido" – Infringement \ n + Undefined "Crime\n+ Contraordenação" –Crime\n+ Infringement "Conflito de Consumo" – Consumer Conflict "Contraordenação" –Infringement "Crime"–Crime "Indefinido" –Undefined	Contraordenação\n+ Indefinido–1 Crime\n+ Contraordenação–248 Conflito de Consumo –936 Contraordenação–9914 Crime–954 Indefinido–4394
Nº ofícios para Denunciante	Number	Number of offices for complainant	0–7492 1–13129 2–299 3–4
Nº Autoridades (Ofícios)	Number	Number of Authorities (Offices)	0–13245 1–5738 2–1545 3–317 4–67 5–4 6–8
Ofícios para Denunciante	Object	Activity per complainant	–
Produtos	Object	roducts available in the services described in the attributes "Activitie	Products available in the services described in the attributes "Activities"
Infracções	Object	Infractions	Due to the large number of infractions, this attribute will be analyzed and described in more detail later in this dissertation
Fiscalizações Relacionadas	Object	Related Inspections	–
Processos Relacionados	Object	Related processes	–
Suporte Digital	Object	Digital Support: "SIM"–Yes	SIM–5651
Data de registo (criação)	Date/Hour	Registration Date (creation)	[2014-01-02 08:44:00;2014-12-31 18:44:00]
Utilizador (criação)	Object	User (creation)	–
Data de edição	Date/Hour	Date of issue	[2014-01-02 09:49:00;2018-02-19 10:57:00]
Utilizador (edição)	Object	User (editing)	–

Annex 7–Data description of dataset DENUNCIAS_2015

Attributes	Variable Type	Description	DENUNCIAS_2015
NID	Object	Complaint identification number: E–Entity 1–Number of entry 15–Year SC–Central Services	[E/1/15/SC:E/99999/15/SC]
Competência	Object	Organization responsible for handling the complaint: "ASAE"–ASAE "ASAE e Tribunais"– ASAE and Courts "ASAE e outra Entidade"– "ASAE and other Entity" "Indeterminada"–Undetermined "Outra Entidade"– Another Entity "Outra Entidade e Tribunais"–Other Entity and Courts "Tribunais"– Courts	ASAE–8785 ASAE e Tribunais–686 ASAE e outra Entidade–4638 Indeterminada–2497 Outra Entidade–4409 Outra Entidade e Tribunais–152 Tribunais–1441
Estado	Object	Status of complaint: Solved – "Cumprido" On Hold – "Pendente"	Cumprido–19193 Pendente–3415
Estado Averiguação	Object	Status of inquiry: Filed – "Arquivada" Filed In Found – "Arquivada Averiguada" Inquiry– "Em averiguação" To inquire – "Por averiguar"	Arquivada–14158 Arquivada Averiguada–5035 Em averiguação–362 Por averiguar–3053
Data arquivo	Date Object	Filing date of the complaint	[02/01/15:22/11/18)
UO destino (1º Circuito)	Object	GPEF– Office of Strategic Planning and Training NIP–Nucleus of Procedural Instruction UNIIC–National Information and Criminal Investigation Unit NIF I –Nucleus of Inspection and Inspection I NIF II–Nucleus of Inspection and Inspection II NBE– Nucleus of Specialized Brigades NIF IV–Nucleus of Inspection and Inspection IV NIF V– Inspection and Inspection Center V – Coimbra NIF VII –Nucleus of Inspection and Inspection VII NIF VIII –Nucleus of Inspection and Inspection VIII NIF IX –Nucleus of Inspection and Inspection IX DGC –Division of Management of Administrative Offenses SECIG–Secretariat of the Inspector General	/SC/IG/DAJC–11 /SC/IG/UNIIC–832 /SC/IG/UNIIC/DAPI–1 /SC/IG/UNIIC/UCII–1 /SC/IG/UNO–2 /SC/IG/UNO/DEPO–113
UO destino (1º Circuito)	Object	GRI–Office of International Relations DCAAI –Inspection Activity Coordination and Evaluation Division DIP –Division of Public Information SC –Central Services URN –Northern Regional Unit URC –Regional Center Unit UO XI–EVR –Evora UO XII–FAR –Faro IG –General Inspector SGAAL –Sub–Inspector General Administration and Logistics Area SGAT –Sub–Inspector General Technical Area	/URC/UO VI–CTB/NIF VI–1 /URN/DIR–253 /URN/UO I–PRT–1890 /URN/UO I–PRT/NBE–1 /URN/UO I–PRT/NIF I–2
UO destino (1º Circuito)	Object	URS –Regional Southern Unit SECSIG –Secretariat of the General Sub–Inspectors (Administrative and Technical Area) DAJC –Department of Legal Affairs and Administrative Offenses DRA –Division of Food Risks DAL –Department of Administration and Logistics DGRH –Human Resources Management Division RH –Human Resources DAG –Division of Management Support DSPD –Digital Support and Expertise Division UNO –National Operations Unit DAPI –Division of Analysis and Information Research UCII –Central Research and Intervention Unit	/URS/UO IX–LSB Sul–927 /URS/UO VII–LSB–2893 /URS/UO VIII–LSB Oeste–427 /URS/UO X–STR–424

UO destino (1º Circuito)	Object	NIF III –Nucleus of Inspection and Inspection III UO V–CBR Norte –Coimbra / North UO IV–CBR –Coimbra UO VI–CTB –Castelo Branco NIF VI –Nucleus of Inspection and Inspection VI UO VIII–LSB Oeste –Lisbon / West UO VII–LSB –Lisbon UO X–STR –Santarem NIF X –Nucleus of Inspection and Inspection X NIF XI –Nucleus of Inspection and Inspection XI NIF XII –Nucleus of Inspection and Inspection XII DAS –Support and Security Division	/SC/IG/UNO/DIP–728 /SC/SECSIG–1 /SC/SGAAL/GCAAI–1 /URC/DIR–99 /URC/UO IV–CBR–676
UO destino (1º Circuito)	Object	GCAAI –Office of Coordination and Evaluation of the Insight Activity NIC I –Nucleus of Criminal Investigation I NIC II– Nucleus of Criminal Investigation II NIF V –Tondela SC –Counter–Order Section NIF V.TInspection and Inspection Center V – Tondela SIC –Criminal Investigation Section LBPV –Laboratory of Beverages and Viticultural Products UO II–BRC –Barcelos NIIP –Nucleus of Investigation and Procedural Instruction NATA –Nucleus of Technical and Administrative Support	/URN/UO II–BRC–922 /URN/UO II–BRC /NIF II–1 /URN/UO III–MIR–311 /URS/DIR–318
UO destino (1º Circuito)	Object	UO III–MIR– Mirandela DGRP –Asset Management Division UO IX–LSB Sul –Lisbon / South EXP –Office hour TES –Treasury SCO –Accounting and Budget Service SAP –Provisioning Service SPMA –Seized Heritage and Material Services SGV –Car Management Service DIR –Office of the Directorate of the Southern Regional Unit	/URC/UO V–CBR Norte–499 /URC/UO V–CBR Norte/NIF V–1 /URC/UO VI–CTB–378
UO destino (1º Circuito)	Object	"DEPO –Division of Studies and Operational Planning DCO –Operational Control Division DRAL –Department of Food Risks and Laboratories LFQ– Physical Chemistry Laboratory LM –Microbiology Laboratory UO I–PRT –Porto"	/URS/UO XI–EVR–423 /URS/UO XII–FAR–1027 /URS/UO XII–FAR/NATA–1
Data de envio (1º Circuito)	Date	Date of dispatch (1st circuit)	[02/01/2015:27/01/17]
Data de averig. (1ª FOF)	Date / Hour	Date and time of the investigation of the complaint in the FOF (File of Operator Supervised)	[23/01/2015 15:30; 04/03/2016 11:20]
Local de registo	Object	Where the complaint was registered. Same definition of object attribute acronyms	/SC/IG/UNO/DIP–22605 /URS/NATA–3
Tipo	Object	How was the complaint received at ASAE: "AUTO"– judicial order "CARTA"– letter "CIRCULAR"–Written document "E–MAIL" "FAX" "OFÍCIO"– offical communication "OUTRO"– other "PRESENCIAL" – presential "TELEFONEMA"– phone call	AUTO–1 CARTA–3257 CIRCULAR–1 E–MAIL–17166 FAX–33 ÓFÍCIO–3 PRESENCIAL–167 TELEFONEMA–1980
Tipo Remetente	Object	Type of sender that reported: "Agente Económico" –Economic Agent "Agente Económico (anónimo)" –Economic Agent (anonymous) "Associação Representativa" –Representative Association "Câmara Municipal" –Town Council "Organismo Oficial" –Official Body "Outro" –Other "Particular" –Private "Particular (anónimo)"– Private (anonymous) "Tribunal" –Court	Agente Económico–783 Agente Económico (anónimo)–10245 Associação Representativa–140 Câmara Municipal–245 Organismo Oficial–1048 Outro–84 Particular–10049 Particular (anónimo)– 3 Tribunal–1

Remetente	Object	Name of complainant	–
E-mail	Object	E-mail of complainant	–
Nome (e-mail)	Object	Name and/or e-mail of camplainant	–
Remetida por Denunciante	Object	If the complaint was forwarded by the complainant: "SIM" –Yes "NÃO" –No	SIM–21857 NÃO–751
Tipo Denunciante	Object	Type of complainant: "Agente Económico" –Economic Agent "Agente Económico (anónimo)" –Economic Agent (anonymous) "Associação Representativa" –Repressentative Association "Câmara Municipal" –Town Council "Organismo Oficial" –Official Body "Outro" –Other "Particular" –Private "Particular (anónimo)"– Private (anonymous) "Tribunal" –Court	Agente Económico – 833 Agente Económico (anónimo) – 10225 Associação Representativa – 99 Câmara Municipal – 212 Organismo Oficial – 553 Outro – 40 Particular – 10642 Particular (anónimo)–3 Tribunal – 1
Denunciante	Object	Name of complainant	–
Motivo não preench. den.	Object	Reason for not completing the complaint: "Informação imperceptível"–Imperceptible information "Não fornecido"–Not provided	Informação imperceptível–3 Não fornecido–10222
Entidade Visada	Object	Type of entity targeted	–
Motivo não preench. ent.	Object	Reason for not completing the entity field "Informação imperceptível"–Imperceptible information "Não fornecido"–Not provided	Não fornecido–2652
Morada	Object	Entity's address	–
Cód. Postal	Object	Entity's Postal Code	–
Localidade	Object	Locality	–
Dist/Conc/Freg	Object	District / Council / Parish	–
Actividades	Object	"Boite" (III.1.3.2) –Boite "Cafetaria" (III.1.2.5) – Coffee Shop "Café" (III.1.2.3) – Cafe "Casa de Pasto" (III.1.1.3) – House of Grass "Cervejaria" (III.1.2.2) – Brewery "Clube noturno" (III.1.3.1) – Nightclub "Dancing" (III.1.3.6) – Dancing "Fast-food" (III.1.1.8) – Fast-food "Gelataria" (III.1.2.7) – Ice cream parlor "Marisqueira" (III.1.1.2) – Seafood restaurant "Pastelaria" (III.1.2.4) – Pastry "Pizzaria" (III.1.1.4) – Pizzeria "Snack-bar" (III.1.1.5) – Snack bar	"Boite" (III.1.3.2) –1 "Cafetaria" (III.1.2.5) –71 "Café" (III.1.2.3) –2110 "Casa de Pasto" (III.1.1.3) –11 "Cervejaria" (III.1.2.2) –30 "Clube noturno" (III.1.3.1) –2 "Dancing" (III.1.3.6) –30 "Fast-food" (III.1.1.8) –22 "Gelataria" (III.1.2.7) –15 "Marisqueira" (III.1.1.2) –7 "Pastelaria" (III.1.2.4) –705 "Pizzaria" (III.1.1.4) –29 "Snack-bar" (III.1.1.5) –138
Actividades	Object	"Pub" (III.1.2.8)–Pub "Taberna" (III.1.2.6) – Tavern "Take-away" (III.1.1.7) – Take away "Atelier de Arranjos de Costura (IX.9)\n+ Lavandarias/Engomadorias" (IX.39) – Atelier of Sewing Arrangements (IX.9) \ n + Laundry / Ironing (IX.39) "Cabeleireiros (IX.12)\n+ Tatuagens/Piercing's" (IX.52) – Hairdressers (IX.12) \ n + Tattoos / Piercing's (IX.52) "Venda a retalho de medicamentos veterinários (IX.81)"–Retail sale of veterinary medicinal products (IX.81)	"Pub" (III.1.2.8)–1 "Taberna" (III.1.2.6) –49 "Take-away" (III.1.1.7) –8 Atelier de Arranjos de Costura (IX.9)\n+ Lavandarias/Engomadorias (IX.39) –1 Cabeleireiros (IX.12)\n+ Tatuagens/Piercing's (IX.52) –1 Venda a retalho de medicamentos veterinários (IX.81)–10

Actividades	Object	Venda Automática Produtos Alimentares (V.22) – Sell Automatic Food Products (V.22) Tabaco/Tabacarias (VIII.22)– Tobacco / Tobacco Stores (VIII.22) Aldeamento Turístico (IX.6)–Tourist Village (IX.6) Parques de Campismo e Caravanismo (III.1.3.7)\n+ Discoteca (III.1.3.7)–Campsites and Caravanning (IX.43) \ n + Discotheque (III.1.3.7) Hipermercados (V.21)\n+ Informática (IX.36)–Hypermarkets (V.21) \ n + Information Technology (IX.36)	Venda Automática Produtos Alimentares (V.22)–25 Tabaco/Tabacarias (VIII.22)–15 Aldeamento Turístico (IX.6)–6 Pronto–a–Vestir (VIII.47)\n+ Instituto de Beleza (IX.37)–1 Parques de Campismo e Caravanismo (IX.43)\n+ Discoteca (III.1.3.7)–1 Hipermercados (V.21)\n+ Informática (IX.36)–1
Actividades	Object	"Administração de Condomínios" (IX.1) – Administration of Condos "Agentes de Gás" (VIII.1) – Gas Agents "Agências Funerárias" (IX.5) – Funeral Agencies "Agências de Publicidade e Marketing" (IX.2) – Advertising and Marketing Agencies "Agências de Venda de Bilhetes" (IX.3) – Ticket Agencies "Sapateiros" (IX.49) – Shoemakers "Seguradoras" (IX.50) – Insurance companies "Sem actividade identificada" (Z) – No identified activity "Serralharias" (IX.51) – Locksmiths "Ópticas" (VIII.42) – Optics "Tratamento de resíduos" (II.26) – Waste treatment	Administração de Condomínios (IX.1) –71 Agentes de Gás (VIII.1) –43 Agências Funerárias (IX.5) –49 Agências de Publicidade e Marketing (IX.2) –6 Agências de Venda de Bilhetes (IX.3) –2 Sapateiros (IX.49) –18 Seguradoras (IX.50) –52 Sem actividade identificada (Z) –2743 Serralharias (IX.51) –57 Ópticas (VIII.42) –40 Tratamento de resíduos (II.26) –2
Actividades	Object	"Sex–shop e outras atividades de serviços pessoais diversos" (VIII.52) – Sex shop and other miscellaneous personal service activities "Smartshops"(X.18) – Smartshops "Stand de Automóveis Novos" (VIII.53) – New Cars Stand "Stand de Automóveis Usados" (VIII.54) – Used Cars Stand "Stand de Máquinas e Motores (inclui equipamentos de construção civil)" (VIII.57) – Machinery and Engine Stand (includes construction equipment) "Sucatas" (VIII.58) – Scrap "Suiniculturas" (I.3.2) – Pig farms "Supermercados/Minimercados/Mercearias" (V.20) – Supermarkets / Minimarkets / Groceries	Sex–shop e outras atividades de serviços pessoais diversos (VIII.52) –12 Smartshops (X.18) –1 Stand de Automóveis Novos (VIII.53) –48 Stand de Automóveis Usados (VIII.54) –234 Stand de Máquinas e Motores (inclui equipamentos de construção civil) (VIII.57) –3 Sucatas (VIII.58) –12 Suiniculturas (I.3.2) –15 Supermercados/Minimercados/Mercearias (V.20) –1234
Actividades	Object	"Supermercados/Minimercados/Mercearias \n+ "Café"" (III.1.2.3) – Supermarkets / Minimarkets / Groceries \ n + "Cafe" "Supermercados/Minimercados/Mercearias\n+ "Snack–bar"" (III.1.1.5) – "Acondicionamento e conservação de hortofrutícolas frescos (inclui produtos minimamente processados – centros normalização)" (II.11.1) – Supermarkets / Minimarkets / Groceries \ n + "Snack bar" "Talhos" (V.1) – Butchers "Tatuagens/Piercing's" (IX.52) – Tattoos / Piercing's "Telecomunicações" (IX.54) – Telecommunications "Transportador" (IX.67) – Carrier "Transportes Públicos" (IX.56) – Public Transport	Supermercados/Minimercados/Mercearias (V.20)\n+ "Café" (III.1.2.3) –3 Acondicionamento e conservação de hortofrutícolas frescos (inclui produtos minimamente processados – centros normalização) (II.11.1) –3 Talhos (V.1) –169 Tatuagens/Piercing's (IX.52) –28 Telecomunicações (IX.54) –606 Transportador (IX.67) –45 Transportes Públicos (IX.56) –64
Actividades	Object	"Turismo de Natureza" (IX.22) – Tourism of Nature "Táxis" (IX.53) – Taxis "Venda Automática Produtos Alimentares" (V.22) – Sell Automatic Food Products "Venda ambulante" (V.16) – Street vending "Venda de artigos de desporto" (VIII.64) – Sale of sporting goods "Vendas à distância (por Catálogo e Internet)" (VII.1) – Distance selling (by Catalog and Internet) "Veículos em fim de vida (parques de sucata)" (X.15) – End–of–life vehicles (scrap yards) "Vidreiras/Cristais" (VIII.8) – Glassware / Crystals "Várias espécies" (I.3.7) – Several species "Óleos novos e usados" (X.10) – New and Used Oils	Táxis (IX.53) –120 Venda Automática Produtos Alimentares (V.22) –25 Venda ambulante (V.16) –170 Venda de artigos de desporto (VIII.64) –18 Vendas à distância (por Catálogo e Internet) (VII.1) –963 Veículos em fim de vida (parques de sucata) (X.15) –20 Vidreiras/Cristais (VIII.8) –1 Várias espécies (I.3.7) –27 Óleos novos e usados (X.10) –5

Área Geográfica/DR	Object	Same attribute definitions of "UO destino (1º Circuito)"	/SC/IG/UNIIC-3 /URC-9 /URC/UO IV-CBR-914 /URC/UO V-CBR Norte-642 /URC/UO VI-CTB-427 /URN-26 /URN/UO I-PRT-2580 /URN/UO II-BRC-1267
Área Geográfica/DR	Object	Same attribute definitions of "UO destino (1º Circuito)"	/URN/UO III-MIR-369 /URS-30 /URS/UO IX-LSB Sul-1132 /URS/UO VII-LSB-3987 /URS/UO VIII-LSB Oeste-516 /URS/UO X-STR-518 /URS/UO XI-EVR-517 /URS/UO XII-FAR-1121 Sem identificação localidade-8550
Conteúdo	Object	Content of the complaint	
Classificação Conteúdo	Object	"Contraordenação\n+ Indefinido" – Infringement \ n + Undefined "Crime\n+ Contraordenação" –Crime\n+ Infringement "Conflito de Consumo" – Consumer Conflict "Contraordenação" –Infringement "Crime"–Crime "Indefinido" –Undefined	Contraordenação\n+ Indefinido-1 Crime\n+ Contraordenação–273 Conflito de Consumo -1188 Contraordenação-11010 Crime-881 Indefinido-6311 Contraordenação\n+ Conflito de Consumo-85 Crime\n+ Conflito de Consumo-1
Nº ofícios para Denunciante	Number	Number of offices for complainant	0–8869 1–13236 2–498 3–5
Nº Autoridades (Ofícios)	Number	Number of Authorities (Offices)	0–14612 1–5936 2–1666 3–298 4–82 5–4 6–10
Ofícios para Denunciante	Object	Activity per complainant	-
Produtos	Object	Products available in the services described in the attributes "Activities"	Products available in the services described in the attributes "Activities"
Infracções	Object	Infractions	Due to the large number of infractions, this attribute will be analyzed and described in more detail later in this dissertation
Fiscalizações Relacionadas	Object	Related Inspections	-
Processos Relacionados	Object	Related processes	-
Suporte Digital	Object	Digital Support: "SIM"-Yes	SIM-5495
Data de registo (criação)	Date/Hour	Registration Date (creation)	[02/01/15 07:43;31/12/15 20:47]
Utilizador (criação)	Object	User (creation)	-
Data de edição	Date/Hour	Date of issue	[02/01/15 09:51:00;22/11/18 08:14]
Utilizador (edição)	Object	User (editing)	-

Annex 8–Data description of dataset DENUNCIAS_2016

Attributes	Variable Type	Description	DENUNCIAS_2016
NID	Object	Complaint identification number: E–Entity 1–Number of entry 16–Year SC–Central Services	[E/1/16/SC;E/996/16/SC]
Competência	Object	Organization responsible for handling the complaint: "ASAE"–ASAE "ASAE e Tribunais"– ASAE and Courts "ASAE e outra Entidade"– "ASAE and other Entity" "Indeterminada"–Undetermined "Outra Entidade"– Another Entity "Outra Entidade e Tribunais"–Other Entity and Courts "Tribunais"– Courts	ASAE–8618 ASAE e Tribunais–422 ASAE e outra Entidade–3777 Indeterminada–2360 Outra Entidade–3727 Outra Entidade e Tribunais–131 Tribunais–1143
Estado	Object	Status of complaint: Solved – "Cumprido" On Hold – "Pendente"	Cumprido–16256 Pendente–3922
Estado Averiguação	Object	Status of inquiry: Filed – "Arquivada" Filed In Found – "Arquivada Averiguada" Inquiry– "Em averiguação" To inquire – "Por averiguar"	Arquivada–11510 Arquivada Averiguada–4746 Em averiguação–502 Por averiguar–3420
Data arquivo	Date Object	Filing date of the complaint	04/01/2016;22/11/18)
UO destino (1º Circuito)	Object	GPEF– Office of Strategic Planning and Training NIP–Nucleus of Procedural Instruction UNIIC–National Information and Criminal Investigation Unit NIF I –Nucleus of Inspection and Inspection I NIF II–Nucleus of Inspection and Inspection II NBE– Nucleus of Specialized Brigades NIF IV–Nucleus of Inspection and Inspection IV NIF V– Inspection and Inspection Center V – Coimbra NIF VII –Nucleus of Inspection and Inspection VII NIF VIII –Nucleus of Inspection and Inspection VIII NIF IX –Nucleus of Inspection and Inspection IX DGC –Division of Management of Administrative Offenses SECIG–Secretariat of the Inspector General	/SC/IG/DAJC–20 /SC/IG/UNIIC–826 /SC/IG/UNIIC/DAPI–1 /SC/IG/UNIIC/UCII–1 /SC/IG/UNO–2
UO destino (1º Circuito)	Object	GRI–Office of International Relations DCAAI –Inspection Activity Coordination and Evaluation Division DIP –Division of Public Information SC –Central Services URN –Northern Regional Unit URC –Regional Center Unit UO XI–EVR –Evora UO XII–FAR –Faro IG –General Inspector SGAAL –Sub-Inspector General Administration and Logistics Area SGAT –Sub-Inspector General Technical Area	/URC/UO IV–CBR–762 /URC/UO V–CBR Norte–532 /URC/UO V–CBR Norte/NIF V–1 /URC/UO VI–CTB–318
UO destino (1º Circuito)	Object	URS –Regional Southern Unit SECSIG –Secretariat of the General Sub-Inspectors (Administrative and Technical Area) DAJC –Department of Legal Affairs and Administrative Offenses DRA –Division of Food Risks DAL –Department of Administration and Logistics DGRH –Human Resources Management Division RH –Human Resources DAG –Division of Management Support DSPD –Digital Support and Expertise Division UNO –National Operations Unit DAPI –Division of Analysis and Information Research UCII –Central Research and Intervention Unit	/URN/UO III–MIR–306 /URS/NATA–1 /URS/UO IX–LSB Sul–926 /URS/UO VII–LSB–3075 /URS/UO VIII–LSB Oeste–402

UO destino (1º Circuito)	Object	NIF III –Nucleus of Inspection and Inspection III UO V–CBR Norte –Coimbra / North UO IV–CBR –Coimbra UO VI–CTB –Castelo Branco NIF VI –Nucleus of Inspection and Inspection VI UO VIII–LSB Oeste –Lisbon / West UO VII–LSB –Lisbon UO X–STR –Santarem NIF X –Nucleus of Inspection and Inspection X NIF XI –Nucleus of Inspection and Inspection XI NIF XII –Nucleus of Inspection and Inspection XII DAS –Support and Security Division	/SC/IG/UNO/DEPO–153 /SC/IG/UNO/DIP–467 /SC/SECIG–5
UO destino (1º Circuito)	Object	GCAAI –Office of Coordination and Evaluation of the Insight Activity NIC I –Nucleus of Criminal Investigation I NIC II– Nucleus of Criminal Investigation II NIF V –Tondela SC –Counter–Order Section NIF V.TInspection and Inspection Center V – Tondela SIC –Criminal Investigation Section LBPV –Laboratory of Beverages and Viticultural Products UO II–BRC –Barcelos NIIP –Nucleus of Investigation and Procedural Instruction NATA –Nucleus of Technical and Administrative Support	/URS/UO X–STR–403 /URS/UO XI–EVR–522 /URS/UO XII–FAR–904 /URS/UO XII–FAR/NATA–1
UO destino (1º Circuito)	Object	UO III–MIR– Mirandela DGRP –Asset Management Division UO IX–LSB Sul –Lisbon / South EXP –Office hour TES –Treasury SCO –Accounting and Budget Service SAP –Provisioning Service SPMA –Seized Heritage and Material Services SGV –Car Management Service DIR –Office of the Directorate of the Southern Regional Unit	/URN/UO I–PRT–2069 /URN/UO I–PRT/NIF I–4 /URN/UO II–BRC–904 /URN/UO II–BRC /NIF II–1
UO destino (1º Circuito)	Object	"DEPO –Division of Studies and Operational Planning DCO –Operational Control Division DRAL –Department of Food Risks and Laboratories LFQ– Physical Chemistry Laboratory LM –Microbiology Laboratory UO I–PRT –Porto"	/SC/SECSIG–1 /SC/SGAT/DRAL–1 /SC/SGAT/DRAL/DRA–1
Data de envio (1º Circuito)	Date	Date of dispatch (1st circuit)	[04/01/16:20/03/18]
Data de averig. (1ª FOF)	Date / Hour	Date and time of the investigation of the complaint in the FOF (File of Operator Supervised)	[17/04/15 10:30 ; 21/11/18 19:10]
Local de registo	Object	Where the complaint was registered. Same definition of object attribute acronyms	/SC/IG/UNO/DIP–20177 /SC/SGAAL/DAL/DAG/EXP–1
Tipo	Object	How was the complaint received at ASAE: "ATESTADO MÉDICO"–Medical Certificate "AUTO"– judicial order "CARTA"– letter "E–MAIL" "FAX" "OFÍCIO"– offical communication "PRESENCIAL" – presential "TELEFONEMA"– phone call	ATESTADO MÉDICO–1 AUTO–19 CARTA–2818 E–MAIL–15773 FAX–44 OFFICIAL COMMUNICATION–6 PRESENCIAL–277 TELEFONEMA–1240
Tipo Remetente	Object	Type of sender that reported: "Agente Económico" –Economic Agent "Agente Económico (anónimo)" –Economic Agent (anonymous) "Associação Representativa" –Repressentative Association "Câmara Municipal" –Town Council "Organismo Oficial" –Official Body "Organismo Oficial (anónimo)" –Official Body (anonymous) "Outro" –Other "Outro (anónimo)" –Other (anonymous) "Particular" –Private "Particular (anónimo)" –Private (anonymous) "Tribunal" –Court	Agente Económico–687 Agente Económico (anónimo)–9528 Associação Representativa–126 Câmara Municipal–330 Organismo Oficial–937 Organismo Oficial (anónimo)–1 Outro–93 Outro (anónimo)–3 Particular–8463 Particular (anónimo)–4 Tribunal–6

Remetente	Object	Name of complainant	–
E-mail	Object	E-mail of complainant	–
Nome (e-mail)	Object	Name and/or e-mail of camplainant	–
Remetida por Denunciante	Object	If the complaint was forwarded by the complainant: "SIM" –Yes "NÃO" –No	SIM–18698 NÃO–1480
Tipo Denunciante	Object	Type of complainant: "Agente Económico" –Economic Agent "Agente Económico (anónimo)" –Economic Agent (anonymous) "Associação Representativa" –Repressentative Association "Câmara Municipal" –Town Council "Organismo Oficial" –Official Body "Organismo Oficial (anónimo)" –Official Body (anonymous) "Outro" –Other "Outro (anónimo)" –Other (anonymous) "Particular" –Private "Particular (anónimo)" –Private (anonymous) "Tribunal" –Court	Agente Económico–736 Agente Económico (anónimo)–8680 Associação Representativa–92 Câmara Municipal–300 Organismo Oficial–548 Organismo Oficial (anónimo)–1 Outro–58 Outro (anónimo)–2 Particular–9749 Particular (anónimo)–6 Tribunal–6
Denunciante	Object	Name of complainant	–
Motivo não preench. den.	Object	Reason for not completing the complaint: "Informação imperceptível"–Imperceptible information "Não fornecido"–Not provided	Informação imperceptível–10 Não fornecido–8670
Entidade Visada	Object	Type of entity targeted	–
Motivo não preench. ent.	Object	Reason for not completing the entity field "Informação imperceptível"–Imperceptible information "Não fornecido"–Not provided	Informação imperceptível–45 Não fornecido–3028
Morada	Object	Entity's address	–
Cód. Postal	Object	Entity's Postal Code	–
Localidade	Object	Locality	–
Dist/Conc/Freg	Object	District / Council / Parish	–
Actividades	Object	"Cafetaria" (III.1.2.5) – Coffee Shop "Café" (III.1.2.3) – Cafe "Casa de Pasto" (III.1.1.3) – House of Grass "Cervejaria" (III.1.2.2) – Brewery "Clube noturno" (III.1.3.1) – Nightclub "Dancing" (III.1.3.6) – Dancing "Fast-food" (III.1.1.8) – Fast-food "Gelataria" (III.1.2.7) – Ice cream parlor "Marisqueira" (III.1.1.2) – Seafood restaurant "Pastelaria" (III.1.2.4) – Pastry "Pizzaria" (III.1.1.4) – Pizzeria "Pub (III.1.2.8)"–Pub "Snack-bar" (III.1.1.5) – Snack bar	"Cafetaria" (III.1.2.5) –119 "Café" (III.1.2.3) –1692 "Casa de Pasto" (III.1.1.3) –2 "Cervejaria" (III.1.2.2) –26 "Clube noturno" (III.1.3.1) –4 "Dancing" (III.1.3.6) –17 "Fast-food" (III.1.1.8) –38 "Gelataria" (III.1.2.7) –15 "Marisqueira" (III.1.1.2) –6 "Pastelaria" (III.1.2.4) –653 "Pizzaria" (III.1.1.4) –41 "Pub (III.1.2.8)"–1 "Snack-bar" (III.1.1.5) –102
		"Taberna" (III.1.2.6) – Tavern "Take-away" (III.1.1.7) – Take away Barbearias (IX.11)\n+ Cabeleireiros (IX.12)–Barbershops (IX.11) \ n + Hairdressers (IX.12) Indústria de Calçado (VIII.24)\n+ Indústria de Vestuário (VIII.25)–Footwear Industry (VIII.24) \ n + Apparel Industry (VIII.25) Papelarias (VIII.44)\n+ "Cafetaria" (III.1.2.5)–Stationery (VIII.44) \ n + "Cafeteria" (III.1.2.5) Supermercados/Minimercados/Mercearias (V.20)\n+ "Cafetaria" (III.1.2.5)–Supermarkets / Minimarkets / Groceries (V.20) \ n + "Cafeteria" (III.1.2.5)	"Taberna" (III.1.2.6) –30 "Take-away" (III.1.1.7) –7 Barbearias (IX.11)\n+ Cabeleireiros (IX.12)–1 Indústria de Calçado (VIII.24)\n+ Indústria de Vestuário (VIII.25)–3 Papelarias (VIII.44)\n+ "Cafetaria" (III.1.2.5)–1 Supermercados/Minimercados/Mercearias (V.20)\n+ "Cafetaria" (III.1.2.5) –2

Actividades	Object	"Administração de Condomínios" (IX.1) – Administration of Condos "Agentes de Gás" (VIII.1) – Gas Agents "Agências Funerárias" (IX.5) – Funeral Agencies "Agências de Publicidade e Marketing" (IX.2) – Advertising and Marketing Agencies "Agências de Venda de Bilhetes" (IX.3) – Ticket Agencies "Sem actividade identificada" (Z) – No identified activity "Serralharias" (IX.51) – Locksmiths "Ópticas" (VIII.42) – Optics "Tratamento de resíduos" (II.26) – Waste treatment	Administração de Condomínios (IX.1) –52 Agentes de Gás (VIII.1) –24 Agências Funerárias (IX.5) –37 Agências de Publicidade e Marketing (IX.2) –14 Agências de Venda de Bilhetes (IX.3) –3 Agências de Viagens (IX.4)–156 Sem actividade identificada (Z) –3037 Serralharias (IX.51) –26 Ópticas (VIII.42) –33 Tratamento de resíduos (II.26) –5
Actividades	Object	"Sex–shop e outras atividades de serviços pessoais diversos" (VIII.52) – Sex shop and other miscellaneous personal service activities "Stand de Automóveis Novos" (VIII.53) – New Cars Stand "Stand de Automóveis Usados" (VIII.54) – Used Cars Stand "Stand de Máquinas e Motores (inclui equipamentos de construção civil)" (VIII.57) – Machinery and Engine Stand (includes construction equipment) "Sucatas" (VIII.58) – Scrap "Suiniculturas" (I.3.2) – Pig farms "Supermercados/Minimercados/Mercearias" (V.20) – Supermarkets / Minimarkets / Groceries	Sex–shop e outras atividades de serviços pessoais diversos (VIII.52) –28 Stand de Automóveis Novos (VIII.53) –51 Stand de Automóveis Usados (VIII.54) –143 Stand de Máquinas e Motores (inclui equipamentos de construção civil) (VIII.57) –3 Sucatas (VIII.58) –6 Suiniculturas (I.3.2) –11 Supermercados/Minimercados/Mercearias (V.20) –1126
Actividades	Object	"Acondicionamento e conservação de hortofrutícolas frescos (inclui produtos minimamente processados – centros normalização)" (II.11.1) – Supermarkets / Minimarkets / Groceries \ n + "Snack bar" "Talhos" (V.1) – Butchers "Tatuagens/Piercing's" (IX.52) – Tattoos / Piercing's "Telecomunicações" (IX.54) – Telecommunications "Transportador" (IX.67) – Carrier "Transportes Públicos" (IX.56) – Public Transport	Acondicionamento e conservação de hortofrutícolas frescos (inclui produtos minimamente processados – centros normalização) (II.11.1) –3 Talhos (V.1) –117 Tatuagens/Piercing's (IX.52) –10 Telecomunicações (IX.54) –334 Transportador (IX.67) –46 Transportes Públicos (IX.56) –76
Actividades	Object	"Turismo de Natureza" (IX.22) – Tourism of Nature "Táxis" (IX.53) – Taxis "Venda Automática Produtos Alimentares" (V.22) – Sell Automatic Food Products "Venda ambulante" (V.16) – Street vending "Venda de artigos de desporto" (VIII.64) – Sale of sporting goods "Vendas à distância (por Catálogo e Internet)" (VII.1) – Distance selling (by Catalog and Internet) "Veículos em fim de vida (parques de sucata)" (X.15) – End–of–life vehicles (scrap yards) "Vidreiras/Cristais" (VIII.8) – Glassware / Crystals "Várias espécies" (I.3.7) – Several species "Óleos novos e usados" (X.10) – New and Used Oils	Turismo de Natureza (IX.22) –10 Táxis (IX.53) –104 Venda Automática Produtos Alimentares (V.22) –18 Venda ambulante (V.16) –117 Venda de artigos de desporto (VIII.64) –13 Vendas à distância (por Catálogo e Internet) (VII.1) –676 Veículos em fim de vida (parques de sucata) (X.15) –18 Vidreiras/Cristais (VIII.8) –3 Várias espécies (I.3.7) –4 Óleos novos e usados (X.10) –2
		Gas Agents (VIII.1) –Gas Agents (VIII.1) Temporary employment agencies (IX.77) –Temporary employment agencies (IX.77) Tourist Village (IX.6) –Tourist Village (IX.6) Local accommodation (IX.75) –Local accommodation (IX.75) No identified activity (Z) –No identified activity (Z) Boats Stand (VIII.55) –Boats Stand (VIII.55) Timesharing (IX.55) –Timesharing (IX.55) Retail trade in veterinary medicines (IX.81) –Retail trade services of veterinary medicines (IX.81) Supermarkets / Minimarkets / Groceries (V.20) \ n + Commercialization of Communications Equipment and Devices (VIII.9) "Tabaco/Tabacarias" (VIII.22) – Tobacco / Tobacco Shop	Agentes de Gás (VIII.1)–24 Agências de trabalho temporário (IX.77)–1 Aldeamento Turístico (IX.6)–2 Alojamento local (IX.75)–427 Sem actividade identificada (Z)–2917 Stand de Barcos (VIII.55)–2 Timesharing (IX.55)–2 Venda a retalho de medicamentos veterinários (IX.81)–6 Supermercados/Minimercados/Mercearias (V.20)\n+ Comercialização de Aparelhos e Equipamentos para Comunicações (VIII.9)–1 "Tabaco/Tabacarias" (VIII.22)–22

Área Geográfica/DR	Object	Same attribute definitions of "UO destino (1º Circuito)"	/SC–1 /SC/IG/UNIIC–51 /SC/IG/UNIIC/UCII–1 /SC/IG/UNO–1 /SC/IG/UNO/DEPO–11 /URC/UO IV–CBR–880 /URC/UO V–CBR Norte–621 /URC/UO VI–CTB–354 /URN–4 /URN/UO I–PRT–2608 /URN/UO II–BRC–1109 /URN/UO III–MIR–355 /URS–6 /URS/UO IX–LSB Sul–1092
Área Geográfica/DR	Object	Same attribute definitions of "UO destino (1º Circuito)" except: "Sem identificação localidade"– No +D39C39:C40	/URS/UO VII–LSB–3932 /URS/UO VIII–LSB Oeste–468 /URS/UO X–STR–468 /URS/UO XI–EVR–586 /URS/UO XI–EVR/NATA–1 /URS/UO XII–FAR–1006 Sem identificação localidade–6623
Conteúdo	Object	Content of the complaint	–
Classificação Conteúdo	Object	"Crime\n+ Contraordenação"–Crime \ n + Counter–Infringement "Conflito de Consumo"–Consumer Conflict "Contraordenação"–Infringement "Crime"–Crime "Indefinido"–Undefined	Crime\n+ Contraordenação–40 Conflito de Consumo–1190 Contraordenação–11261 Crime–1068 Indefinido–6108
Nº ofícios para Denunciante	Number	Number of offices for complainant	0–6207 1–13391 2–558 3–20 4–1 5–1
Nº Autoridades (Ofícios)	Number	Number of Authorities (Offices)	0–13084 1–5559 2–1302 3–184 4–48 6–1
Ofícios para Denunciante	Object	Activity per complainant	–
Produtos	Object	roducts available in the services described in the attributes "Activitie	Products available in the services described in the attributes "Activities"
Infracções	Object	Infractions	Due to the large number of infractions, this attribute will be analyzed and described in more detail later in this dissertation
Fiscalizações Relacionadas	Object	Related Inspections	–
Processos Relacionados	Object	Related processes	–
Suporte Digital	Object	Digital Support: "SIM"–Yes	SIM–4442
Data de registo (criação)	Date/Hour	Registration Date (creation)	[01/01/2016 09:54:00;2014–12–31 18:44:00]
Utilizador (criação)	Object	User (creation)	–
Data de edição	Date/Hour	Date of issue	[04/01/2016 11:15:00;22/11/2018 15:56:00]
Utilizador (edição)	Object	User (editing)	–

Annex 9–Data description of dataset DENUNCIAS_2017

Attributes	Variable Type	Description	DENUNCIAS_2017
NID	Object	Complaint identification number: E–Entity 1–Number of entry 17–Year SC–Central Services	[E/1/17/SC;E/9999/17/SC]
Competência	Object	Organization responsible for handling the complaint: "ASAE"–ASAE "ASAE e Tribunais"– ASAE and Courts "ASAE e outra Entidade"– "ASAE and other Entity" "Indeterminada"–Undetermined "Outra Entidade"– Another Entity "Outra Entidade e Tribunais"–Other Entity and Courts "Tribunais"– Courts	ASAE–10918 ASAE e Tribunais–147 ASAE e outra Entidade–3061 Indeterminada–2084 Outra Entidade–3882 Outra Entidade e Tribunais–163 Tribunais–901
Estado	Object	Status of complaint: Solved – "Cumprido" On Hold – "Pendente"	Cumprido–15559 Pendente–5597
Estado Averiguação	Object	Status of inquiry: Filed – "Arquivada" Filed In Found – "Arquivada Averiguada" Inquiry– "Em averiguação" To inquire – "Por averiguar"	Arquivada–11161 Arquivada Averiguada–4398 Em averiguação–683 Por averiguar–4914
Data arquivo	Date Object	Filing date of the complaint	[2017–01–09:2018–11–22]
UO destino (1º Circuito)	Object	GPEF– Office of Strategic Planning and Training NIP–Nucleus of Procedural Instruction UNIIC–National Information and Criminal Investigation Unit NIF I –Nucleus of Inspection and Inspection I NIF II–Nucleus of Inspection and Inspection II NBE– Nucleus of Specialized Brigades NIF IV–Nucleus of Inspection and Inspection IV NIF V– Inspection and Inspection Center V – Coimbra NIF VII –Nucleus of Inspection and Inspection VII NIF VIII –Nucleus of Inspection and Inspection VIII NIF IX –Nucleus of Inspection and Inspection IX DGC –Division of Management of Administrative Offenses SECIG–Secretariat of the Inspector General	/SC/IG/DAJC–16 /SC/IG/UNIIC–970 /SC/IG/UNO–2
UO destino (1º Circuito)	Object	GRI–Office of International Relations DCAAI –Inspection Activity Coordination and Evaluation Division DIP –Division of Public Information SC –Central Services URN –Northern Regional Unit URC –Regional Center Unit UO XI–EVR –Evora UO XII–FAR –Faro IG –General Inspector SGAAL –Sub-Inspector General Administration and Logistics Area SGAT –Sub-Inspector General Technical Area	/URC/UO VI–CTB–285 /URN/DIR–6 /URN/UO I–PRT–2332 /URN/UO I–PRT/NIF I–3
UO destino (1º Circuito)	Object	URS –Regional Southern Unit SECSIG –Secretariat of the General Sub-Inspectors (Administrative and Technical Area) DAJC –Department of Legal Affairs and Administrative Offenses DRA –Division of Food Risks DAL –Department of Administration and Logistics DGRH –Human Resources Management Division RH –Human Resources DAG –Division of Management Support DSPD –Digital Support and Expertise Division UNO –National Operations Unit DAPI –Division of Analysis and Information Research UCII –Central Research and Intervention Unit	/URN/UO II–BRC–1051 /URN/UO III–MIR–331 /URS/DIR–4 /URS/NATA–4 /URS/NIIP–1

UO destino (1º Circuito)	Object	NIF III –Nucleus of Inspection and Inspection III UO V–CBR Norte –Coimbra / North UO IV–CBR –Coimbra UO VI–CTB –Castelo Branco NIF VI –Nucleus of Inspection and Inspection VI UO VIII–LSB Oeste –Lisbon / West UO VII–LSB –Lisbon UO X–STR –Santarem NIF X –Nucleus of Inspection and Inspection X NIF XI –Nucleus of Inspection and Inspection XI NIF XII –Nucleus of Inspection and Inspection XII DAS –Support and Security Division	/SC/IG/UNO/DEPO-125 /SC/IG/UNO/DIP-1 /SC/SECIG-7
UO destino (1º Circuito)	Object	GCAAI –Office of Coordination and Evaluation of the Insight Activity NIC I –Nucleus of Criminal Investigation I NIC II– Nucleus of Criminal Investigation II NIF V –Tondela SC –Counter–Order Section NIF V.TInspection and Inspection Center V – Tondela SIC –Criminal Investigation Section LBPV –Laboratory of Beverages and Viticultural Products UO II–BRC –Barcelos NIIP –Nucleus of Investigation and Procedural Instruction NATA –Nucleus of Technical and Administrative Support	/URS/UO IX-LSB Sul-1495 /URS/UO VII-LSB-2119 /URS/UO VII-LSB/NIF VII-1
UO destino (1º Circuito)	Object	UO III–MIR– Mirandela DGRP –Asset Management Division UO IX–LSB Sul –Lisbon / South EXP –Office hour TES –Treasury SCO –Accounting and Budget Service SAP –Provisioning Service SPMA –Seized Heritage and Material Services SGV –Car Management Service DIR –Office of the Directorate of the Southern Regional Unit	/URS/UO VIII-LSB Oeste-1380 /URS/UO X-STR-537 /URS/UO X-STR/NATA-1
UO destino (1º Circuito)	Object	"DEPO –Division of Studies and Operational Planning DCO –Operational Control Division DRAL –Department of Food Risks and Laboratories LFQ– Physical Chemistry Laboratory LM –Microbiology Laboratory UO I–PRT –Porto"	/URS/UO XI-EVR-468 /URS/UO XI-EVR/NATA-1 /URS/UO XII-FAR-1074
Data de envio (1º Circuito)	Date	Date of dispatch (1st circuit)	[2017-01-04:2018-11-22]
Data de averig. (1ª FOF)	Date / Hour	Date and time of the investigation of the complaint in the FOF (File of Operator Supervised)	[2016-01-29 23:45:00 ; 2018-11-21 11:50:00]
Local de registo	Object	Where the complaint was registered. Same definition of object attribute acronyms	/SC/IG/UNO/DIP-21156
Tipo	Object	How was the complaint received at ASAE: "AUTO"– judicial order "AVISO"–Notice "CARTA"– letter "CIRCULAR"–Informative Circular "COMUNICAÇÃO DE SERVIÇO"–Service Communication "E-MAIL"–E-mail "FAX"-Fax	AUTO-91 AVISO-4 CARTA-2554 CIRCULAR-1 COMUNICAÇÃO DE SERVIÇO-1 E-MAIL-16927 FAX-2
Tipo	Object	"INFORMAÇÃO"–Information "OFÍCIO"–Offical Communication "OUTRO"–Ohter "PARECER"–Specialized Manifestation "PRESENCIAL" – Presential "PROPOSTA"–Proposal "REQUISIÇÃO"–Request "TELEFONEMA"– phone call	INFORMAÇÃO-1 OFÍCIO-134 OUTRO-2 PARECER-4 PRESENCIAL-396 PROPOSTA-1 REQUISIÇÃO-7 TELEFONEMA-1031

Tipo Remetente	Object	Type of sender that reported: "Agente Económico" –Economic Agent "Agente Económico (anónimo)" –Economic Agent (anonymous) "Associação Representativa" –Repressentative Association "Câmara Municipal" –Town Council "Organismo Oficial" –Official Body "Organismo Oficial (anónimo)" –Official Body (anonymous) "Organização Internacional"–Internacional Organization "Outro" –Other "Outro (anónimo)" –Other (anonymous) "Particular" –Private "Particular (anónimo)" –Private (anonymous) "Tribunal" –Court	Agente Económico–867 Agente Económico (anónimo)–10912 Associação Representativa–151 Câmara Municipal–355 Organismo Oficial–831 Organismo Oficial (anónimo)–2 Organização Internacional–1 Outro–164 Outro (anónimo)–3 Particular–7856 Particular (anónimo)–6 Tribunal–11
Remetente	Object	Name of complainant	–
E-mail	Object	E-mail of complainant	–
Nome (e-mail)	Object	Name and/or e-mail of camplainant	–
Remetida por Denunciante	Object	If the complaint was forwarded by the complainant: "SIM" –Yes "NÃO" –No	SIM–18957 NÃO–2199
Tipo Denunciante	Object	Type of complainant: "Agente Económico" –Economic Agent "Agente Económico (anónimo)" –Economic Agent (anonymous) "Associação Representativa" –Repressentative Association "Câmara Municipal" –Town Council "Organismo Oficial" –Official Body "Organização Internacional"–International Organization "Outro" –Other "Particular" –Private "Particular (anónimo)" –Private (anonymous) "Tribunal" –Court	Agente Económico–912 Agente Económico (anónimo)–9919 Associação Representativa–86 Câmara Municipal–237 Organismo Oficial–276 Organização Internacional–1 Outro–138 Particular–9578 Particular (anónimo)–7 Tribunal–2
Denunciante	Object	Name of complainant	–
Motivo não preench. den.	Object	Reason for not completing the complaint: "Informação imperceptível"–Imperceptible information "Não fornecido"–Not provided	Informação imperceptível–44 Não fornecido–9875
Entidade Visada	Object	Type of entity targeted	–
Motivo não preench. ent.	Object	Reason for not completing the entity field "Informação imperceptível"–Imperceptible information "Não fornecido"–Not provided	Informação imperceptível–99 Não fornecido–4275
Morada	Object	Entity's address	–
Cód. Postal	Object	Entity's Postal Code	–
Localidade	Object	Locality	–
Dist/Conc/Freg	Object	District / Council / Parish	–
Actividades	Object	Cabaret" (III.1.3.5)–Cabaret "Cafetaria" (III.1.2.5) – Coffee Shop "Café" (III.1.2.3) – Cafe "Casa de Pasto" (III.1.1.3) – House of Grass "Cervejaria" (III.1.2.2) – Brewery "Clube noturno" (III.1.3.1) – Nightclub "Dancing" (III.1.3.6) – Dancing "Fast-food" (III.1.1.8) – Fast-food "Gelataria" (III.1.2.7) – Ice cream parlor "Marisqueira" (III.1.1.2) – Seafood restaurant "Pastelaria" (III.1.2.4) – Pastry "Pizzaria" (III.1.1.4) – Pizzeria "Snack-bar" (III.1.1.5) - Snack bar	"Cabaret" (III.1.3.5)–1 "Cafetaria" (III.1.2.5) –77 "Café" (III.1.2.3) –1732 "Casa de Pasto" (III.1.1.3) –7 "Cervejaria" (III.1.2.2) –16 "Clube noturno" (III.1.3.1) –24 "Dancing" (III.1.3.6) –8 "Fast-food" (III.1.1.8) –44 "Gelataria" (III.1.2.7) –25 "Marisqueira" (III.1.1.2) –6 "Pastelaria" (III.1.2.4) –558 "Pizzaria" (III.1.1.4) –59 "Snack-bar" (III.1.1.5) –113

Actividades	Object	"Take-away" (III.1.1.7)–Take-away "Café" (III.1.2.3)\n+ "Pastelaria" (III.1.2.4)–Coffee" (III.1.2.3) \ n + "Pastry" (III.1.2.4) "Pizzaria" (III.1.1.4)\n+ "Gelataria" (III.1.2.7)–Pizzeria" (III.1.1.4) \ n + "Gelataria" (III.1.2.7) "Snack-bar" (III.1.1.5)\n+ "Cafetaria" (III.1.2.5)–Snack bar" (III.1.1.5) \ n + "Cafeteria" (III.1.2.5) "Snack-bar" (III.1.1.5)\n+ "Café" (III.1.2.3)–"Snack-bar" (III.1.1.5) \ n + "Cafe" (III.1.2.3) "Snack-bar" (III.1.1.5)\n+ "Café" (III.1.2.3)\n+ "Pastelaria" (III.1.2.4)–Snack-bar" (III.1.1.5) \ n + "Coffee" (III.1.2.3) \ n + "Pastry" (III.1.2.4) "Snack-bar" (III.1.1.5)\n+ Bar (III.1.2.1)\n+ "Café" (III.1.2.3)–Snack-bar" (III.1.1.5) \ n + Bar (III.1.2.1) \ n + "Coffee" (III.1.2.3)	"Take-away" (III.1.1.7)–13 "Café" (III.1.2.3)\n+ "Pastelaria" (III.1.2.4)–4 "Pizzaria" (III.1.1.4)\n+ "Gelataria" (III.1.2.7)–1 "Snack-bar" (III.1.1.5)\n+ "Cafetaria" (III.1.2.5)–1 "Snack-bar" (III.1.1.5)\n+ "Café" (III.1.2.3)–3 "Snack-bar" (III.1.1.5)\n+ "Café" (III.1.2.3)\n+ "Pastelaria" (III.1.2.4)–1 + "Snack-bar" (III.1.1.5)\n+ Bar (III.1.2.1)\n+ "Café" (III.1.2.3)–1
Actividades	Object	"Agências de Viagens (IX.4)\n+ ""Café"" (III.1.2.3)–Travel Agencies (IX.4) \ n +" "Coffee "(III.1.2.3) Agências de Viagens (IX.4)\n+ Empresas de Animação Turística (IX.23)–Travel Agencies (IX.4) \ n + Tourism Animation Companies (IX.23) Alojamento local (IX.75)\n+ Restaurante (III.1.1.1)–Local accommodation (IX.75) \ n + Restaurant (III.1.1.1) Alojamento local (IX.75)\n+ Restauração em lares (III.2.5)–Local accommodation (IX.75) \ n + Restoration in homes (III.2.5) Apartamentos Turísticos (IX.7)\n+ Alojamento local (IX.75) –Tourist Apartments (IX.7) \ n + Local accommodation (IX.75) Associação Recreativa ou Desportiva (IX.72)\n+ Restaurante (III.1.1.1)–Recreational or Sports Association (IX.72) \ n + Restaurant (III.1.1.1)	Agências de Viagens (IX.4)\n+ "Café" (III.1.2.3)–1 Agências de Viagens (IX.4)\n+ Empresas de Animação Turística (IX.23)–2 Alojamento local (IX.75)\n+ Restaurante (III.1.1.1)–6 Alojamento local (IX.75)\n+ Restauração em lares (III.2.5)–1 Apartamentos Turísticos (IX.7)\n+ Alojamento local (IX.75) –1 Associação Recreativa ou Desportiva (IX.72)\n+ Restaurante (III.1.1.1)–1
Actividades	Object	Atelier de Arranjos de Costura (IX.9)\n+ Lavandarias/Engomadorias (IX.39)–Atelier of Sewing Arrangements (IX.9) \ n + Laundry / Ironing (IX.39) Bar (III.1.2.1)\n+ "Café" (III.1.2.3)–Bar (III.1.2.1) \ n + "Coffee" (III.1.2.3) Sapatarias (VIII.51)–Shoe stores (VIII.51) Sapateiros (IX.49)–Shoemakers (IX.49) Seguradoras (IX.50)–Insurance companies (IX.50) Sem actividade identificada (Z)–No identified activity (Z) Sex-shop e outras atividades de serviços pessoais diversos (VIII.52)–Sex-shop and other miscellaneous personal service activities (VIII.52)	Atelier de Arranjos de Costura (IX.9)\n+ Lavandarias/Engomadorias (IX.39)–1 Bar (III.1.2.1)\n+ "Café" (III.1.2.3)–4 Sapatarias (VIII.51)–56 Sapateiros (IX.49)–5 Seguradoras (IX.50)–44 Sem actividade identificada (Z)–2938 Serralharias (IX.51)–32 Sex-shop e outras atividades de serviços pessoais diversos (VIII.52)–6
Actividades	Object	Stand de Automóveis Novos (VIII.53)–New Cars Stand (VIII.53) Stand de Automóveis Usados (VIII.54)–Used Cars Stand (VIII.54) Sucatas (VIII.58)–Scrap (VIII.58) Suiniculturas (I.3.2)– Pig farms (I.3.2) Supermercados/Minimercados/Mercearias (V.20)–Supermarkets / Minimarkets / Groceries (V.20) Tabaco/Tabacarias (VIII.22)–Tobacco / Tobacco (VIII.22) Talhos (V.1)–Butchers (V.1) Tatuagens/Piercing's (IX.52)–Tattoos / Piercing's (IX.52) Telecomunicações (IX.54)–Telecommunications (IX.54) Timesharing (IX.55)–Timesharing (IX.55) Transportador (IX.67)–Carrier (IX.67)	Stand de Automóveis Novos (VIII.53)–51 Stand de Automóveis Usados (VIII.54)–155 Sucatas (VIII.58)–21 Suiniculturas (I.3.2)–4 Supermercados/Minimercados/Mercearias (V.20)–910 Tabaco/Tabacarias (VIII.22)–38 Talhos (V.1)–153 Tatuagens/Piercing's (IX.52)–28 Telecomunicações (IX.54)–301 Timesharing (IX.55)–1 Transportador (IX.67)–68
Actividades	Object	Transportes Públicos (IX.56)–Public Transport (IX.56) Tratamento de resíduos (II.26)–Treatment of waste (II.26) Turismo de Natureza (IX.22)–Tourism of Nature (IX.22) Táxis (IX.53)–Taxis (IX.53) Venda Automática Produtos Alimentares (V.22)–Selling Food Products (V.22) Venda a retalho de medicamentos veterinários (IX.81)–Retail trade services of veterinary medicines (IX.81) Venda ambulante (V.16)–Street vending (V.16) Venda de artigos de desporto (VIII.64)–Sale of sporting goods (VIII.64) Vendas à distância (por Catálogo e Internet) (VII.1)–Distance selling (by Catalog and Internet) (VII.1)	Transportes Públicos (IX.56)–126 Tratamento de resíduos (II.26)–3 Turismo de Natureza (IX.22)–11 Táxis (IX.53)–56 Venda Automática Produtos Alimentares (V.22)–23 Venda a retalho de medicamentos veterinários (IX.81)–4 Venda ambulante (V.16)–130 Venda de artigos de desporto (VIII.64)–12 Vendas à distância (por Catálogo e Internet) (VII.1)–814
Actividades	Object	Veículos em fim de vida (parques de sucata) (X.15)–End-of-life vehicles (scrap yards) (X.15) Vidreiras/Cristais (VIII.8)–Glassware / Crystals (VIII.8) Várias espécies (I.3.7)–Several species (I.3.7) Ópticas (VIII.42)–Optics (VIII.42)	Veículos em fim de vida (parques de sucata) (X.15)–1 Vidreiras/Cristais (VIII.8)–3 Várias espécies (I.3.7)–5 Ópticas (VIII.42) 24

Área Geográfica/DR	Object	Same attribute definitions of "UO destino (1º Circuito)"	/SC–1 /SC/IG/DAJC–1 /SC/IG/UNIIC–179 /SC/IG/UNO/DEPO–19 /URC–7 /URC/DIR–1 /URC/UO IV–CBR–1075 /URC/UO V–CBR Norte–884 /URC/UO VI–CTB–400 /URN–17 /URN/DIR–1 /URN/UO I–PRT–3579 /URN/UO II–BRC –1581
Área Geográfica/DR	Object	Same attribute definitions of "UO destino (1º Circuito)" except: "Sem identificação localidade"– No location identification	/URN/UO III–MIR–453 /URS–21 /URS/DIR–1 /URS/UO IX–LSB Sul–2050 /URS/UO IX–LSB Sul/NIF IX–1 /URS/UO VII–LSB–3297 /URS/UO VIII–LSB Oeste–2099 /URS/UO X–STR–774 /URS/UO XI–EVR–653 /URS/UO XI–EVR/NIF XI–1 /URS/UO XII–FAR–1490 /URS/UO XII–FAR/NIF XII–1 Sem identificação localidade –2570
Conteúdo	Object	Content of the complaint	–
Classificação Conteúdo	Object	Conflito de Consumo\n+ Indefinido–Consumer Conflict \ n + Undefined Contraordenação\n+ Conflito de Consumo–Infringement\n+ Consumer Conflict Contraordenação\n+ Indefinido–Infringement\ n + Undefined Crime\n+ Conflito de Consumo–Crime \ n + Consumer Conflict Crime\n+ Contraordenação–Crime \ n + Infringement Conflito de Consumo–Conflict of Consumption Contraordenação–Infringement Crime–Crime Indefinido–Undefined	Conflito de Consumo\n+ Indefinido–2 Contraordenação\n+ Conflito de Consumo–44 Contraordenação\n+ Indefinido–1 Crime\n+ Conflito de Consumo–2 Crime\n+ Contraordenação–264 Conflito de Consumo–885 Contraordenação–12355 Crime–823 Indefinido–5455
Nº ofícios para Denunciante	Number	Number of offices for complainant	0–7167 1–13391 2–551 3–37 4–9 5–1
Nº Autoridades (Ofícios)	Number	Number of Authorities (Offices)	0–14620 1–5255 2–1048 3–193 4–34 5–5 6–1
Ofícios para Denunciante	Object	Activity per complainant	–
Produtos	Object	Products available in the services described in the attributes "Activities"	Products available in the services described in the attributes "Activities"
Infracções	Object	Infractions	Due to the large number of infractions, this attribute will be analyzed and described in more detail later in this dissertation
Fiscalizações Relacionadas	Object	Related Inspections	–
Processos Relacionados	Object	Related processes	–
Suporte Digital	Object	Digital Support: "SIM"–Yes	SIM–4318
Data de registo (criação)	Date/Hour	Registration Date (creation)	[2017–01–01 08:02:00;2017–12–30 23:48:00]
Utilizador (criação)	Object	User (creation)	–
Data de edição	Date/Hour	Date of issue	[2017–01–10 14:36:00;2018–11–22 15:59:00]
Utilizador (edição)	Object	User (editing)	–

Annex 10–Data description of dataset DENUNCIAS_2018

Attributes	Variable Type	Description	DENUNCIAS_2018
NID	Object	Complaint identification number: E–Entity 22–Number of entry 18–Year SC–Central Services	[E/22/18/SC;E/99992/18/SC]
Competência	Object	Organization responsible for handling the complaint: "ASAE"–ASAE "ASAE e Tribunais"– ASAE and Courts "ASAE e outra Entidade"– "ASAE and other Entity" "Indeterminada"–Undetermined "Outra Entidade"– Another Entity "Outra Entidade e Tribunais"–Other Entity and Courts "Tribunais"– Courts	ASAE–9449 ASAE e Tribunais–60 ASAE e outra Entidade–2866 Indeterminada–2482 Outra Entidade–3500 Outra Entidade e Tribunais–72 Tribunais–594
Estado	Object	Status of complaint: Solved – "Cumprido" On Hold – "Pendente"	Cumprido–10386 Pendente–8864
Estado Averiguação	Object	Status of inquiry: Filed – "Arquivada" Filed In Found – "Arquivada Averiguada" Inquiry– "Em averiguação" To inquire – "Por averiguar"	Arquivada–8810 Arquivada Averiguada–1576 Em averiguação–632 Por averiguar–8232
Data arquivo	Date Object	Filing date of the complaint	[2018–01–03:2018–11–22]
UO destino (1º Circuito)	Object	GPEF– Office of Strategic Planning and Training NIP–Nucleus of Procedural Instruction UNIIC–National Information and Criminal Investigation Unit NIF I –Nucleus of Inspection and Inspection I NIF II–Nucleus of Inspection and Inspection II NBE– Nucleus of Specialized Brigades NIF IV–Nucleus of Inspection and Inspection IV NIF V– Inspection and Inspection Center V – Coimbra NIF VII –Nucleus of Inspection and Inspection VII NIF VIII –Nucleus of Inspection and Inspection VIII NIF IX –Nucleus of Inspection and Inspection IX DGC –Division of Management of Administrative Offenses SECIG–Secretariat of the Inspector General	/SC/IG/DAJC–9 /SC/IG/UNIIC–1219 /SC/IG/UNIIC/DAPI–1 /SC/IG/UNIIC/UCII/NIC II–1
UO destino (1º Circuito)	Object	GRI–Office of International Relations DCAAI –Inspection Activity Coordination and Evaluation Division DIP –Division of Public Information SC –Central Services URN –Northern Regional Unit URC –Regional Center Unit UO XI–EVR –Evora UO XII–FAR –Faro IG –General Inspector SGAAL –Sub–Inspector General Administration and Logistics Area SGAT –Sub–Inspector General Technical Area	/SC/IG/UNO/DCO–2 /SC/IG/UNO/DEPO–153 /SC/IG/UNO/DIP–2 /SC/SECIG–1
UO destino (1º Circuito)	Object	URS –Regional Southern Unit SECSIG –Secretariat of the General Sub–Inspectors (Administrative and Technical Area) DAJC –Department of Legal Affairs and Administrative Offenses DRA –Division of Food Risks DAL –Department of Administration and Logistics DGRH –Human Resources Management Division RH –Human Resources DAG –Division of Management Support DSPD –Digital Support and Expertise Division UNO –National Operations Unit DAPI –Division of Analysis and Information Research UCII –Central Research and Intervention Unit	/SC/SGAAL/DAL–1 /URC/DIR–94 /URC/NATA–1 /URC/UO IV–CBR–650

UO destino (1º Circuito)	Object	NIF III –Nucleus of Inspection and Inspection III UO V–CBR Norte –Coimbra / North UO IV–CBR –Coimbra UO VI–CTB –Castelo Branco NIF VI –Nucleus of Inspection and Inspection VI UO VIII–LSB Oeste –Lisbon / West UO VII–LSB –Lisbon UO X–STR –Santarem NIF X –Nucleus of Inspection and Inspection X NIF XI –Nucleus of Inspection and Inspection XI NIF XII –Nucleus of Inspection and Inspection XII DAS –Support and Security Division	/URC/UO IV–CBR/NBE–1 /URC/UO V–CBR Norte–511 /URC/UO VI–CTB–249 /URN/DIR–103
UO destino (1º Circuito)	Object	GCAAI –Office of Coordination and Evaluation of the Insight Activity NIC I –Nucleus of Criminal Investigation I NIC II– Nucleus of Criminal Investigation II NIF V –Tondela SC –Counter–Order Section NIF V.TInspection and Inspection Center V – Tondela SIC –Criminal Investigation Section LBPV –Laboratory of Beverages and Viticultural Products UO II–BRC –Barcelos NIIP –Nucleus of Investigation and Procedural Instruction NATA –Nucleus of Technical and Administrative Support	/URN/NATA–1 /URN/UO I–PRT–1798 /URN/UO I–PRT/NIF I–1 /URN/UO II–BRC–830
UO destino (1º Circuito)	Object	UO III–MIR– Mirandela DGRP –Asset Management Division UO IX–LSB Sul –Lisbon / South EXP –Office hour TES –Treasury SCO –Accounting and Budget Service SAP –Provisioning Service SPMA –Seized Heritage and Material Services SGV –Car Management Service DIR –Office of the Directorate of the Southern Regional Unit	/URN/UO III–MIR–250 /URS/DIR–119 /URS/UO IX–LSB Sul–1310 /URS/UO VII–LSB–1223
UO destino (1º Circuito)	Object	"DEPO –Division of Studies and Operational Planning DCO –Operational Control Division DRAL –Department of Food Risks and Laboratories LFQ– Physical Chemistry Laboratory LM –Microbiology Laboratory UO I–PRT –Porto"	/URS/UO VIII–LSB Oeste–1405 /URS/UO X–STR–440 /URS/UO XI–EVR–376 /URS/UO XII–FAR–946
Data de envio (1º Circuito)	Date	Date of dispatch (1st circuit)	[2018-01-02:2018-11-22]
Data de averig. (1ª FOF)	Date / Hour	Date and time of the investigation of the complaint in the FOF (File of Operator Supervised)	[2017-03-18 03:00:00 ; 2018-11-21 11:50:00]
Local de registo	Object	Where the complaint was registered. Same definition of object attribute acronyms	/SC/IG/UNO/DIP–19250
Tipo	Object	How was the complaint received at ASAE: "AUTO"– judicial order "AVISO"–Notice "CARTA"– letter "CIRCULAR"–Informative Circular "E–MAIL"–E–mail "FAX"–Fax	AUTO–150 AVISO–4 CARTA–1561 CIRCULAR–1 E–MAIL–15266 FAX–5
Tipo	Object	"OFÍCIO"–Offical Communication "OUTRO"–Other "PRESENCIAL" – Presential "RECLAMAÇÃO"–Claim "REQUERIMENTO"–Application "PROPOSTA"–Proposal "REQUISIÇÃO"–Request "TELEFONEMA"– phone call	OFÍCIO–498 OUTRO–19 PRESENCIAL–396 "RECLAMAÇÃO"–1 "REQUERIMENTO"– 1 PROPOSTA–1 REQUISIÇÃO–7 TELEFONEMA–1031

Tipo Remetente	Object	Type of sender that reported: "Agente Económico" –Economic Agent "Agente Económico (anónimo)" –Economic Agent (anonymous) "Associação Representativa" –Representative Association "Câmara Municipal" –Town Council "Organismo Oficial" –Official Body "Organismo Oficial (anónimo)" –Official Body (anonymous) "Outro" –Other "Particular" –Private "Particular (anónimo)" –Private (anonymous) "Tribunal" –Court	Agente Económico-631 Agente Económico (anónimo)-11986 Associação Representativa-80 Câmara Municipal-243 Organismo Oficial-871 Organismo Oficial (anónimo)-2 Outro-116 Particular-5244 Particular (anónimo)-65 Tribunal-5
Remetente	Object	Name of complainant	-
E-mail	Object	E-mail of complainant	-
Nome (e-mail)	Object	Name and/or e-mail of camplainant	-
Remetida por Denunciante	Object	If the complaint was forwarded by the complainant: "SIM" –Yes "NÃO" –No	SIM-16037 NÃO-3213
Tipo Denunciante	Object	Type of complainant: "Agente Económico" –Economic Agent "Agente Económico (anónimo)" –Economic Agent (anonymous) "Associação Representativa" –Representative Association "Câmara Municipal" –Town Council "Instituição Da Ue"–Institution of the European Union "Outro" –Other "Particular" –Private "Particular (anónimo)" –Private (anonymous) "Tribunal" –Court	Agente Económico-724 Agente Económico (anónimo)-9783 Associação Representativa-43 Câmara Municipal-165 Instituição Da Ue-2 Organismo Oficial-353 Outro-98 Particular-8009 Particular (anónimo)-72 Tribunal-1
Denunciante	Object	Name of complainant	-
Motivo não preench. den.	Object	Reason for not completing the complaint: "Informação imperceptível"-Imperceptible information "Não fornecido"-Not provided	Informação imperceptível-45 Não fornecido-9737
Entidade Visada	Object	Type of entity targeted	-
Motivo não preench. ent.	Object	Reason for not completing the entity field "Informação imperceptível"-Imperceptible information "Não fornecido"-Not provided	Informação imperceptível-321 Não fornecido-6114
Morada	Object	Entity's address	-
Cód. Postal	Object	Entity's Postal Code	-
Localidade	Object	Locality	-
Dist/Conc/Freg	Object	District / Council / Parish	-
Actividades	Object	"Cafetaria" (III.1.2.5) - Coffee Shop "Café" (III.1.2.3) – Cafe "Casa de Pasto" (III.1.1.3) – House of Grass "Cervejaria" (III.1.2.2) - Brewery "Clube noturno" (III.1.3.1) – Nightclub "Dancing" (III.1.3.6) – Dancing "Fast-food" (III.1.1.8) - Fast-food "Gelataria" (III.1.2.7) - Ice cream parlor "Marisqueira" (III.1.1.2) – Seafood restaurant "Pastelaria" (III.1.2.4) – Pastry "Pizzaria" (III.1.1.4) - Pizzeria "Pub" (III.1.2.8)– Pub "Snack-bar" (III.1.1.5) – Snack bar	"Cafetaria" (III.1.2.5) –278 "Café" (III.1.2.3) -1114 "Casa de Pasto" (III.1.1.3) –2 "Cervejaria" (III.1.2.2) –17 "Clube noturno" (III.1.3.1) -16 "Dancing" (III.1.3.6) -6 "Fast-food" (III.1.1.8) -49 "Gelataria" (III.1.2.7) -30 "Marisqueira" (III.1.1.2) –2 "Pastelaria" (III.1.2.4) –483 "Pizzaria" (III.1.1.4) -77 "Pub" (III.1.2.8)-1 "Snack-bar" (III.1.1.5) -104

Actividades	Object	"Taberna" (III.1.2.6)-Tavern "Take-away" (III.1.1.7)-Take-away "Café" (III.1.2.3)\n+ "Pastelaria" (III.1.2.4)-Coffee" (III.1.2.3) \ n + "Pastry" (III.1.2.4) "Snack-bar" (III.1.1.5)\n+ "Café" (III.1.2.3)-"Snack-bar" (III.1.1.5) \ n + "Cafe" (III.1.2.3) Associação Recreativa ou Desportiva (IX.72)\n+ Espaço de Eventos (III.1.8)-Recreational or Recreational Association (IX.72) \ n + Area of Events (III.1.8) Bricolage e Jardim (III.17)\n+ Produtos da construção (tijolos, mosaicos, louças sanitárias, ferro, etc) (X.13)-DIY and Garden (VIII.17) \ n + Building products (bricks, mosaics, sanitary ware, iron, etc.) (X.13)	"Taberna" (III.1.2.6)-31 "Take-away" (III.1.1.7)-9 "Café" (III.1.2.3)\n+ "Pastelaria" (III.1.2.4)-1 "Snack-bar" (III.1.1.5)\n+ "Café" (III.1.2.3)-2 Associação Recreativa ou Desportiva (IX.72)\n+ Espaço de Eventos (III.1.8)-1 Bricolage e Jardim (VIII.17)\n+ Produtos da construção (tijolos, mosaicos, louças sanitárias, ferro, etc) (X.13)-2
Actividades	Object	Agentes de Gás (VIII.1)\n+ Produção de energia (VIII.69)-Gas Agents (VIII.1) \ n + Energy Production (VIII.69) Agências de Viagens (IX.4)\n+ Empresas de Animação Turística (IX.23)-Travel Agencies (IX.4) \ n + Tourism Animation Companies (IX.23) Agências de Viagens (IX.4)\n+ Táxis (IX.53)-Travel Agencies (IX.4) \ n + Taxis (IX.53) Animais domésticos e Animais de Estimação (VIII.2)\n+ Organização de eventos (IX.70)-Pets and Pets (VIII.2) \ n + Organization of events (IX.70) Artigos e Utilidades para o Lar (VIII.5)\n+ Mobiliário e Decoração (VIII.41)-Articles and Utilities for the Home (VIII.5) \ n + Furniture and Decoration (VIII.41)	Agentes de Gás (VIII.1)\n+ Produção de energia (VIII.69)-2 Agências de Viagens (IX.4)\n+ Empresas de Animação Turística (IX.23)-1 Agências de Viagens (IX.4)\n+ Táxis (IX.53)-2 Animais domésticos e Animais de Estimação (VIII.2)\n+ Organização de eventos (IX.70)-1 Artigos e Utilidades para o Lar (VIII.5)\n+ Mobiliário e Decoração (VIII.41)-1
Actividades	Object	Artigos e Utilidades para o Lar (VIII.5)\n+ Pilhas e acumuladores, Comercialização de ... (X.945)-Household Articles and Utilities (VIII.5) \ n + Batteries and accumulators, Marketing ... (X.945) Aviculturas (I.3.3)\n+ Cuniculturas (I.3.5)-Aviculturas (I.3.3) \ n + Cuniculars (I.3.5) Bancos (IX.10)\n+ Seguradoras (IX.50)-Banks (IX.10) \ n + Insurers (IX.50) Barbearias (IX.11)\n+ "Café" (III.1.2.3)-Barbershops (IX.11) \ n + "Coffee" (III.1.2.3) Sapateiros (IX.49)-Shoemakers (IX.49) Seguradoras (IX.50)-Insurance companies (IX.50) Sem actividade identificada (Z)-No identified activity (Z) Sex-shop e outras atividades de serviços pessoais diversos (VIII.52)-Sex-shop and other miscellaneous personal service activities (VIII.52)	Artigos e Utilidades para o Lar (VIII.5)\n+ Pilhas e acumuladores, Comercialização de ... (X.945)-2 Aviculturas (I.3.3)\n+ Cuniculturas (I.3.5)-1 Bancos (IX.10)\n+ Seguradoras (IX.50)-1 Sapateiros (IX.49)-5 Seguradoras (IX.50)-47 Barbearias (IX.11)\n+ "Café" (III.1.2.3)-1 Sem actividade identificada (Z)-2896 Serralharias (IX.51)-16 Sex-shop e outras atividades de serviços pessoais diversos (VIII.52)-12
Actividades	Object	Stand de Automóveis Novos (VIII.53)-New Cars Stand (VIII.53) Stand de Automóveis Usados (VIII.54)-Used Cars Stand (VIII.54) Sucatas (VIII.58)-Scrap (VIII.58) Suiniculturas (I.3.2)- Pig farms (I.3.2) Supermercados/Minimercados/Mercearias (V.20)-Supermarkets / Minimarkets / Groceries (V.20) Tabaco/Tabacarias (VIII.22)-Tobacco / Tobacco (VIII.22) Talhos (V.1)-Butchers (V.1) Tatuagens/Piercing's (IX.52)-Tattoos / Piercing's (IX.52) Telecomunicações (IX.54)-Telecommunications (IX.54) Transportador (IX.67)-Carrier (IX.67) Transporte de espécies protegidas (IX.85)-Transport of protected species (IX.85)	Stand de Automóveis Novos (VIII.53)-37 Stand de Automóveis Usados (VIII.54)-158 Sucatas (VIII.58)-21 Suiniculturas (I.3.2)-11 Supermercados/Minimercados/Mercearias (V.20)-663 Tabaco/Tabacarias (VIII.22)-44 Talhos (V.1)-102 Tatuagens/Piercing's (IX.52)-25 Telecomunicações (IX.54)-142 Transportador (IX.67)-58 Transporte de espécies protegidas (IX.85)-1
Actividades	Object	Transportes Públicos (IX.56)-Public Transport (IX.56) Tratamento de resíduos (II.26)-Treatment of waste (II.26) Venda Automática Produtos Alimentares (V.22)-Selling Food Products (V.22) Venda a retalho de medicamentos veterinários (IX.81)-Retail trade services of veterinary medicines (IX.81) Venda ambulante (V.16)-Street vending (V.16) Venda de artigos de desporto (VIII.64)-Sale of sporting goods (VIII.64) Vendas à distância (por Catálogo e Internet) (VII.1)-Distance selling (by Catalog and Internet) (VII.1)	Transportes Públicos (IX.56)-78 Tratamento de resíduos (II.26)-16 Táxis (IX.53)-51 Venda Automática Produtos Alimentares (V.22)-33 Venda a retalho de medicamentos veterinários (IX.81)-11 Venda ambulante (V.16)-117 Venda de artigos de desporto (VIII.64)-22 Vendas à distância (por Catálogo e Internet) (VII.1)-753
Actividades	Object	Veículos em fim de vida (parques de sucata) (X.15)-End-of-life vehicles (scrap yards) (X.15) Vidreiras/Cristais (VIII.8)-Glassware / Crystals (VIII.8) Várias espécies (I.3.7)-Several species (I.3.7) Ópticas (VIII.42)-Optics (VIII.42) Cabeleireiros (IX.12)\n+ "Café" (III.1.2.3)-Hairdressers (IX.12) \ n + "Coffee" (III.1.2.3) Cabeleireiros (IX.12)\n+ Instituto de Beleza (IX.37)-Hairdressers (IX.12) \ n + Beauty Institute (IX.37)	Veículos em fim de vida (parques de sucata) (X.15)-2 Vidreiras/Cristais (VIII.8)-1 Várias espécies (I.3.7)-9 Ópticas (VIII.42)-16 Cabeleireiros (IX.12)\n+ "Café" (III.1.2.3)-1 Cabeleireiros (IX.12)\n+ Instituto de Beleza (IX.37)-7

Área Geográfica/DR	Object	Same attribute definitions of "UO destino (1º Circuito)"	/SC/IG/DAJC-1 /SC/IG/UNIIC-230 /SC/IG/UNIIC/DAPI-1 /SC/IG/UNO/DEPO-8 /URC-7 /URC/DIR-18 /URC/UO IV-CBR-1007 /URC/UO V-CBR Norte-792 /URC/UO VI-CTB-373 /URN-9 /URN/DIR-22 /URN/UO I-PRT-3152 /URN/UO II-BRC-1306
Área Geográfica/DR	Object	Same attribute definitions of "UO destino (1º Circuito)" except: "Sem identificação localidade"– No location identification	/URN/UO III-MIR-405 /URS-8 /URS/DIR-46 /URS/UO IX-LSB Sul-1936 /URS/UO VII-LSB-1865 /URS/UO VIII-LSB Oeste-2225 /URS/UO X-STR-659 /URS/UO XI-EVR-582 /URS/UO XII-FAR-1362 Sem identificação localidade-3236
Conteúdo	Object	Content of the complaint	–
Classificação Conteúdo	Object	Conflito de Consumo\n+ Indefinido–Consumer Conflict \ n + Undefined Contraordenação\n+ Conflito de Consumo–Infringement\n+ Consumer Conflict Contraordenação\n+ Indefinido–Infringement\ n + Undefined Crime\n+ Contraordenação–Crime \ n + Infringement Conflito de Consumo–Conflict of Consumption Contraordenação–Infringement Crime–Crime Indefinido–Undefined	Conflito de Consumo\n+ Indefinido–4 Contraordenação\n+ Conflito de Consumo–21 Contraordenação\n+ Indefinido–5 Crime\n+ Contraordenação–215 Conflito de Consumo–585 Contraordenação–10654 Crime–1126 Indefinido–5951
Nº ofícios para Denunciante	Number	Number of offices for complainant	0–6602 1–12155 2–456 3–30 4–3 5–2 6–1 8–1
Nº Autoridades (Ofícios)	Number	Number of Authorities (Offices)	0-13303 1-4802 2-979 3-134 4-26 5-3 6-3
Ofícios para Denunciante	Object	Activity per complainant	–
Produtos	Object	Products available in the services described in the attributes "Activities"	Products available in the services described in the attributes "Activities"
Infracções	Object	Infractions	Due to the large number of infractions, this attribute will be analyzed and described in more detail later in this dissertation
Fiscalizações Relacionadas	Object	Related Inspections	–
Processos Relacionados	Object	Related processes	–
Suporte Digital	Object	Digital Support: "SIM"–Yes	SIM–4110
Data de registo (criação)	Date/Hour	Registration Date (creation)	[2018-01-01 10:27:00;2018-11-22 16:39:00]
Utilizador (criação)	Object	User (creation)	–
Data de edição	Date/Hour	Date of issue	[2018-01-02 11:31:00;2018-11-22 16:43:00]
Utilizador (edição)	Object	User (editing)	–

Annex 11–Data description of dataset Entities

Attributes	Variable Type	Description	Df
RNPC	Object	RNPC–National Registry of Legal Persons	–
Relação	Object	Relation between denounced entities	Estabelecimento–18934 Sede–2857278
NIPC/NIF	Object	NIPC–Identification Number of Collective Person NIF–Tax Identification Number	–
Tipo	Object	"Agente Económico"–Economic Agent "Associação Representativa"–Representative Association "Câmara Municipal"–Town hall "Organismo Oficial"–Official Body "Outro"–Other "Particular"–Private	Agente Económico–2876165 Associação Representativa–26 Câmara Municipal–1 Organismo Oficial–9 Outro–3 Particular–8
Nome	Object	Entity Name	–
Código Postal	Object	Postal Code	–
Localidade	Object	Locality	–
Morada	Object	Adress	–
Dist/Conc/Freg	Object	District / Council / Parish	This attribute will be analyzed later in this dissertation.
Natureza Jurídica RNPC	Object	RNPC legal nature	–
Código CAE Principal	Object	Primary CAE code CAE–Classification of Portuguese Economic Activities by Branch of Activity	–
Revisão CAE Principal	Object	Primary CAE review	–
Designação CAE Principal	Object	Primary CAE designation	–
Código CAE Secundário	Object	Secondary CAE Code	–
Revisão CAE Secundário	Object	Secondary CAE review	–

Annex 12– HACCP–related violations

Infraction Code	Designation	(%) Infractions
1A1	Clandestine slaughter	1,5
1B00	General	30,5
1B01	Sugar	0
1B02	Honey	0,08
1B03	Food additives and processing aids	0,008
1B04	Special food	0,008
1B05	Alcoholic beverages	2
1B06	Non–Alcoholic Beverages	1,32
1B07	Animals, meat and meat products	1,03
1B08	Cereals, legumes and derivatives	0,05
1B09	Stimulants, derivatives and substitutes	0,005
1B10	Fruits, vegetables and derivatives	0,15
1B11	Fats and their raw materials	0,36
1B12	Milk and dairy products	0,24
1B13	Eggs and egg products	0,08
1B14	Bread and related products	1,3
1B15	Fish and dairy products	0,35
1B16	Salt	0,01
1B17	Vinegar	0
1B18	Genetically modified organisms (GMO)	0,002
1B19	Food suplements	0,32
1B20	Organic Products	0,03
1C1	Emergency slaughter outside the slaughterhouse in circumstances other than those permitted in Annex III Regulation (EC) No 853/2004 or without observing the conditions imposed	0
1C2	Slaughter of animals in food establishment	0,004
1C3	Slaughter on the holding of poultry in circumstances other than those permitted in Annex III to Regulation (EC) No 853/2004 or without observing the conditions imposed	0,04
1C4	Packaging, packaging, labeling, storage or transport of fishery products not complying with the conditions laid down in Annex III Regulation (EC) No 853/2004	0,04
1C5	Storage and transport of meat by food business operators without compliance with conditions laid down in Annex III Regulation (EC) No 853/2004	0,009
1C6	Hunting of wild animals with a view to their placing on the market for human consumption by persons not having the training required by Annex III to Regulation	0
1C7	Placing on the market of products from imports and products intended for export which do not comply with Articles 10 and 11 of Regulation (EC) No 852/2004	0,002
1C8	Placing on the market of farmed game meat and wild game that has not undergone the operations imposed by Annex III to Regulation (EC) No 853/2004	0,03
1C9	Placing on the market of live bivalve molluscs, echinoderms, tunicates and live marine gastropods without complying with the conditions laid down	0,04
1C10	Placing on the market of fishery products containing toxins harmful to human health	0
1C11	Placing on the market of fishery products other than live bivalve molluscs, echinoderms, tunicates and marine gastropods, unless the required conditions are met	0,003
1C12	Placing on the market of products of animal origin manufactured in the Community by unregistered or non–approved establishments or which do not comply with the	0,007
1C13	Placing on the market of products of animal origin without a health or identification mark, the application of health marks or product identification	0,37
1C14	Continuity of establishment work to which the authorization is withdrawn, or, in the case of conditional authorization, the authorization is not extended or granted	0
1C15	Failure to comply with the labeling rules laid down in Annex III to Regulation (EC) No 853/2004	0
1C16	Failure by food business operators responsible for slaughterers of the obligations imposed by Section III of Annex II to Regulation	0,002
1C17	Existence of process or process based on HACCP principles that does not meet the requirements of Article 5 of Regulation (EC) No 852/2004	0,15
1C18	Operation of purification and shipping centers which do not comply with the requirements set out in Annex III to Regulation (EC) No 853/2004	0,002
1C19	Operation of slaughterhouses and their cutting plants which do not comply with the requirements laid down in Annex III to Regulation (EC)	0,008
1C20	Operation of establishments producing minced meat, meat preparations, mechanically separated meat and meat products which do not comply with the requirements	0,004
1C21	Preventing or creating obstacles to official controls	0
1C22	Importation of products of animal origin from third countries or establishments not included in the list of third countries or establishments	0
1C23	Failure to comply with the rules laid down in Annex III to Regulation (EC) No 853/2004 for the production and placing on the market of milk products	0,02
1C24	Failure to comply with the rules laid down in Annex III to Regulation (EC) No 853/2004 for the production, collection and placing on the market of raw milk	0,01
1C25	Non–compliance with the rules laid down in Annex III to Regulation (EC) No 853/2004 for the manufacture, handling, storage, labeling and marking of identification	0

1C26	Failure to comply with the rules for the tuning of live bivalve molluscs set out in Annex III to Regulation (EC) No 853/2004	0
1C27	Non-compliance with the rules for the handling of live bivalve molluscs set out in Annex III to Regulation (EC) No 853/2004	0
1C28	Non-compliance with the requirements for the production of live bivalve molluscs in Annex III to Regulation (EC) No 853/2004 and Regulation (EC) No 854/2004	0,002
1C29	Failure to comply with the hygiene requirements laid down in Annex III to Regulation (EC) No 853/2004 for the production of minced meat, meat preparations	0,01
1C30	Failure to comply with the specific requirements laid down for pectinidae in Annex III to Regulation (EC) No 853/2004	0,002
1C31	Failure to comply with general and specific hygiene requirements (Articles 3 and 4 of Regulation (EC) No 852/2004)	39
1C32	Failure of ships engaged in harvesting fishery products from their natural environment or in their handling or processing after harvest	0,007
1C33	Failure by establishments, including vessels, to handle fishery products, including frozen, mechanically separated and processed products	0,004
1C34	Non-compliance by collagen-producing establishments with the requirements set out in Annex III to Regulation (EC) No 853/2004	0
1C35	Non-compliance by establishments manufacturing gelatine with the requirements laid down in Annex III to Regulation (EC) No 853/2004	0
1C36	Breach by establishments which collect or process raw materials for the production of rendered fats and greaves	0,002
1C37	Breach by establishments treating stomachs, bladders and intestines of the requirements laid down in Annex III to Regulation (EC) No 853/2004	0,001
1C38	No process or procedure based on HACCP principles (Article 5 of Regulation (EC) No 852/2004)	8,8
1C39	Non-compliance with the special guarantees for the transits provided for in Article 6 of Regulation (EC) No 853/2004	0
1C40	No monitoring of consignments of products of animal origin by certificates or other documents required under Article 7 of Regulation (EC) No 853/2004	0,004
1C41	Failure to update the documents describing the process or processes based on HACCP principles	0,05
1C42	The non-affixing of an identification mark on products of animal origin or non-compliance with the requirements (Annex II to Regulation (EC) No 853/2004)	0,004
1C43	Non-preservation of documents describing the process or processes based on HACCP principles or other documents or records	0,02
1C44	Non-cooperation with the competent authorities in disregard of Article 4 (4) of Regulation (EC) No 853/2004	0,002
1C45	Preparation of frogs' thighs and snails for human consumption without complying with the requirements set out in Annex III to Regulation (EC) No 853/2004	0,002
1C46	Transport of live animals to slaughterhouses without complying with the requirements set out in Annex III to Regulation (EC) No 853/2004	0,004
1C47	Exchange or return of products ordered	0
1C48	Use of rockets in the restoration	0,04
1C49	Use of vessels in harvesting fishery products from their natural environment or in their handling or processing after harvest which do not meet the requirements	0
1C50	Use of an unauthorized substance to remove any surface contamination of products of animal origin in breach of the provisions of Annex III to Regulation (EC) No 853/2004	0
1C51	Use in establishments producing minced meat, meat preparations, mechanically separated meat and meat products of raw materials	0
1D01	Veterinary Medicines	0,2
1D02	Veterinary Products	0,012
1D03	Medical care centers	0,03
1.E01	Animals	0,007
1.E02	Endangered wild fauna and flora	0,002
1.E03	Introduction of non-indigenous species of flora and fauna	0,005
1F1	Marketing of uncontrolled products	0
1F2	Issuance of certificates that do not correspond to the actual state of the products	0
1F3	Non-compliance with requirements (in marketing)	0
1F4	Non-compliance with implementing rules	0
1G00	General	6
1G01	Nutrition labeling	0,04
1G02	Labeling of foodstuffs	0,2
1H00	General	0,03
1H01	Use of certain epoxy derivatives in mob	0
1I00	General	0,1
1I01	Additives in feedingstuffs	0,009
1I02	Protein products in animal feed	0
1I03	Undesirable substances in feedingstuffs	0,01
1I04	Animal feedingstuffs	0,07
1I05	Raw materials for animal feed	0,002
1I06	Intra-Community trade in animal feed	0

1I07	Medicated feed for animals	0,004
1J01	Veterinary checks applicable to intra-Community trade in products of animal origin	0,01
1J02	Financing rates for health inspections and controls of live animals and fresh meat	0,002
1J03	Veterinary checks on animal and animal products from third countries	0,002
1J04	Animal health control Bovine spongiform encephalopathy	0
1J05	Measures against certain animal diseases	0
1K1	Marketing of foodstuffs with direct blending of gifts	0
1K2	Lack of packaging requirements	0
1K3	Lack of gift requirements	0
1K4	Lack of requirements in the indirect blending of gifts with foodstuffs	0
1K5	Lack, insufficiency, inaccuracy of labeling	0
1L01	Security systems in r & b establishments with dance spaces	5
1M01	Seeds of agricultural and horticultural species	0,03
1M02	Fruit and vegetable materials	0,01
1M03	Vegetative propagating material of the vine	0,01
1M04	Seeds and potatoes	0,07
1M05	Preservation mixtures intended for use in the preservation of the natural environment	0
1N15	Failure to comply with the obligations laid down in Article 4 (1) and (2) of Regulation (EC) No 1069/2009 of the European Parliament and of the Council	0
1N16	Failure to comply with the general animal health restrictions provided for in Article 6 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council	0
1N17	Classification of animal by-products in breach of Articles 8 to 10 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council	0
1N18	Classification of derivatives in breach of Article 7 (2) of Regulation (EC) No 1069/2009 of the European Parliament and of the Council	0
1N19	Disposal or use of animal by-products or derived products in breach of Articles 12 to 14 of Regulation (EC) No 1069/2009	0
1N20	Use of animal by-products and derived products for purposes prohibited by Article 11 of the Regulation, as well as by Article 5 and Annex II of Regulation (EC) No 1069/2009 of the European Parliament and of the Council	0,006
1N21	Use of animal by-products and derived products for the purposes set out in Articles 16 to 18 of Regulation (EC) No 1069/2009 of the European Parliament	0
1N22	Disposal of animal by-products referred to in Article 19 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council without the authorization of DGAV	0
1N23	Use of alternative methods of using or disposing of animal by-products and derived products not authorized in accordance with Article 20 of Regulation	0
1N24	Collection or transport of animal by-products or derived products in breach of the rules laid down in Article 21 of Regulation (EC) No 1069/2009 of the European Parliament	0,006
1N25	Failure to comply with the rules on the traceability of animal by-products and derived products as defined in Article 22 of Regulation (EC) No 1069/2009 of the European Parliament	0
1N26	Activities covered by Article 23 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council without the registration provided for in Article 3 of Regulation (EC) No 1069/2009, of European Parliament and of the Council	0,006
1N27	Without prejudice to Article 3 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council, to carry out activities covered by Article 24 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council and the Council	0
1N28	Significant alteration or termination of the activities provided for in paragraphs 1 and 2 of article 3 of this Decree Law, without the	0
1N29	Failure to comply with the general hygiene rules provided for in Article 25 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council	0
1N30	Manipulation of animal by-products by food businesses in breach of the conditions laid down in Article 26 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council	0,006
1N31	Absence or non-application of internal controls provided for in Article 28 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council;	0
1N32	Inexistence, lack of application or lack of review of a permanent written procedure or procedures based on the principles of risk analysis	0
1N33	Processing and placing on the market of animal by-products and derived products for the feeding of farmed animals	0
1N34	Placing on the market and use of organic fertilizers and soil improvers not complying with the requirements of Article 32 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council	0
1N35	Placing on the market of petfood in breach of the requirements of Article 35 of Regulation (EC) No 1069/2009 of the European Parliament	0
1N36	Placing on the market of other derived products with the exception of the products referred to in Articles 31 to 33 and 35 of the Regulation, in disregard of the requirements	0
1N37	Importation and transit of animal by-products and derived products in breach of the rules laid down in Article 41 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council	0
1N38	Exports of animal by-products and derived products in breach of the rules laid down in Article 43 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council	0
1N39	Issuance of animal by-products and derived products to other Member States in breach of the rules laid down in Article 48 of Regulation (EC) No 1069/2009 of the European Parliament and of the Council	0
1N40	Elimination by incineration or recovery by co-incineration and the use as fuel for combustion of animal by-products and derived products	0
1N41	Landfill of certain Category 1 and 3 material in breach of the rules laid down in Article 7 and Chapter III of Annex VI to Regulation (EC) No 1069/2009 of the European Parliament and of the Council	0
1N42	Failure to comply with the requirements applicable to processing units and other establishments provided for in Article 8 and Chapter I of Annex IV to Regulation	0
1N43	Failure to comply with the hygiene and processing requirements applicable to processing units and other establishments provided for in Article 9.	0

1N44	Processing of animal by-products and derived products into biogas and compost in disregard of the requirements laid down in Article 10 and Annex V of Regulation	0
1N45	Failure to comply with the special rules on samples for research and diagnosis provided for in Article 11 and Annex VI of Regulation (EU) No 142/2011	0
1N46	Failure to comply with the special rules on commercial samples and exhibition articles provided for in Article 12 and Annex VI of Regulation (EU) No 142/2011	0
1N47	Failure to comply with the special rules on animal nutrition provided for in Article 13 and Annex VI of Regulation (EU) No 142/2011	0
1N48	Failure to comply with the special rules for feeding certain species inside and outside feeding grounds and foreseen zoos	0
1N49	Collection and disposal of animal by-products covered by Article 19 (1) (a) to (c), (e) and (f) of the Regulation, in disregard of specific rules	0
1N50	Failure to comply with the requirements for commercial documents and health certificates, identification, collection and transportation of animal by-products and	0
1N51	Failure to comply with the requirements relating to the layout of facilities and handling of by-products in establishments or facilities handling by-products	0
1N52	Failure to comply with the requirements for certain establishments and facilities approved for the manufacture of petfood	0
1N53	Failure to comply with the requirements for certain approved establishments and facilities for the storage and handling of animal by-products	0
1N54	Failure to comply with the requirements for certain establishments and facilities approved for the storage of derived products	0
1N55	Failure to comply with the requirements for the storage on the holding of animal by-products intended for subsequent disposal	0
1N56	Non-compliance with requirements for certain establishments and registered facilities handling animal by-products and derived products	0
1N57	Non-compliance with requirements for certain registered operators transporting animal by-products and derived products	0
1N58	Failure to comply with requirements concerning the handling of intermediate products transported to an establishment or facility	0
1N59	Failure to comply with the rules on the manufacture of derived products which are intended to be ingested by or applied to humans or animals	0
1N60	Failure to comply with the rules on import or transit through the national territory and export of animal by-products and derived products	0
1N61	Failure to comply with the rules concerning the placing on the market, including the import and export of Category 1 material	0
1N62	Failure to comply with the rules on import and transit of samples for research and diagnosis	0
1N63	Failure to comply with the rules on the import and transit of commercial samples and exhibition articles	0
1N64	Failure to pay to slaughterhouses, by the animal presenters, the fee referred to in paragraph 1 of article 8 of this decree law	0
1N65	Failure to pay to DGAV, by slaughter establishments, the amount paid and charged to the submitter of animals for slaughter	0
1N66	Failure to comply with the mention that must appear in the invoices, under the terms of article 10 of this decree law	0
1P1	Installation or the exercise of a livestock activity of Class 1 subject to the prior authorization scheme without the owner having followed the procedures	0,007
1P2	Failure to comply with the legal and regulatory terms and conditions for the exercise of the livestock activity set forth in the license	0,02
1P3	Installation or carrying on of a livestock activity of Class 2 subject to a prior declaration, without having been followed by the holder the procedures provided for	0,004
1P4	Installation or exercise of livestock farming of Class 3, without prior registration	0,02
1P5	Changes in livestock activities, without the	0,002
1P6	Non-compliance with the review conditions	0
1P7	Failure to notify the change of the holder of the livestock activity provided for in article 47	0
1P8	Absence of notification of suspension or cessation or resumption of the exercise of livestock	0
1P9	Failure to comply with the particular conditions for the exercise of livestock	0,03
1P10	Failure to comply with the obligations of filing livestock activities	0
1P11	Non-compliance with Article 66 for the transitional period for farms already licensed or authorized under previous schemes	0
1P12	Failure to comply with the provisions of Articles 67 and 73 in relation to the exceptional regularization arrangements for holdings already in existence on the date of publication of Article	0
1P13	Failure to comply with the standards set out in the directives referred to in Article 4	0
1Q1	Non-conclusion of a written contract for the purchase and sale of milk	0
1Q2	Delivery or receipt of milk without prior written agreement	0
1Q3	Lack of any of the mandatory elements of the contract of sale of milk	0
1Q4	Lack of indication in the contract of the combination of the factors of calculation of the price, when this is variable	0
1Q5	Conclusion of a contract for the purchase and sale of milk in breach of the terms and conditions of the elements of the contract	0
1Q6	Failure to comply with reporting obligations on milk purchasers	0

Annex 13 – Analysis of Competence, State and State of inquiry by District

District	Competence	State	State of inquiry	NID
Aveiro	ASAE	CUMPRIDO	Arquivada	2591
			Arquivada averiguada	2219
		PENDENTE	Em averiguação	222
			Por averiguar	725
	ASAE e Tribunais	CUMPRIDO	Arquivada	50
			Arquivada averiguada	19
		PENDENTE	Em averiguação	3
			Por averiguar	20
	ASAE e outra Entidade	CUMPRIDO	Arquivada	1490
			Arquivada averiguada	1035
		PENDENTE	Em averiguação	59
			Por averiguar	225
	Indeterminada	CUMPRIDO	Arquivada	531
			Arquivada averiguada	4
	Outra Entidade	CUMPRIDO	Arquivada	1743
			Arquivada averiguada	12
		PENDENTE	Por averiguar	7
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	44
	Tribunais	CUMPRIDO	Arquivada	235
			Arquivada averiguada	2
Beja	ASAE	CUMPRIDO	Arquivada	194
			Arquivada averiguada	454
		PENDENTE	Em averiguação	49
			Por averiguar	328
	ASAE e Tribunais	CUMPRIDO	Arquivada	1
			Arquivada averiguada	12
		PENDENTE	Por averiguar	10
	ASAE e outra Entidade	CUMPRIDO	Arquivada	86
			Arquivada averiguada	249
		PENDENTE	Em averiguação	19
			Por averiguar	162
	Indeterminada	CUMPRIDO	Arquivada	69
	Outra Entidade	CUMPRIDO	Arquivada	183
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	7
	Tribunais	CUMPRIDO	Arquivada	44

			Arquivada	2200
Braga	ASAE	CUMPRIDO	Arquivada	2200
			Arquivada averiguada	1999
		PENDENTE	Em averiguação	248
			Por averiguar	1096
	ASAE e Tribunais	CUMPRIDO	Arquivada	30
			Arquivada averiguada	32
		PENDENTE	Em averiguação	2
			Por averiguar	4
	ASAE e outra Entidade	CUMPRIDO	Arquivada	1338
			Arquivada averiguada	839
		PENDENTE	Em averiguação	143
			Por averiguar	460
	Indeterminada	CUMPRIDO	Arquivada	549
			Arquivada averiguada	4
		PENDENTE	Por averiguar	8
	Outra Entidade	CUMPRIDO	Arquivada	1444
			Arquivada averiguada	1
		PENDENTE	Por averiguar	6
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	48
	Tribunais	CUMPRIDO	Arquivada	273
Bragança	ASAE	CUMPRIDO	Arquivada	217
			Arquivada averiguada	328
		PENDENTE	Em averiguação	52
			Por averiguar	283
	ASAE e Tribunais	PENDENTE	Por averiguar	1
	ASAE e outra Entidade	CUMPRIDO	Arquivada	128
			Arquivada averiguada	191
		PENDENTE	Em averiguação	7
			Por averiguar	190
	Indeterminada	CUMPRIDO	Arquivada	106
		PENDENTE	Por averiguar	2
	Outra Entidade	CUMPRIDO	Arquivada	202
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	2
	Tribunais	CUMPRIDO	Arquivada	18
Castelo-Branco	ASAE	CUMPRIDO	Arquivada	587
			Arquivada averiguada	635
		PENDENTE	Em averiguação	131
			Por averiguar	133
	ASAE e Tribunais	CUMPRIDO	Arquivada	7
			Arquivada averiguada	13

		PENDENTE	Em averiguação	2
			Por averiguar	2
	ASAE e outra Entidade	CUMPRIDO	Arquivada	382
			Arquivada averiguada	336
		PENDENTE	Em averiguação	49
			Por averiguar	43
	Indeterminada	CUMPRIDO	Arquivada	118
			Arquivada averiguada	2
	Outra Entidade	CUMPRIDO	Arquivada	303
		PENDENTE	Por averiguar	1
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	8
	Tribunais	CUMPRIDO	Arquivada	87
Coimbra	ASAE	CUMPRIDO	Arquivada	1127
			Arquivada averiguada	2327
		PENDENTE	Em averiguação	272
			Por averiguar	333
	ASAE e Tribunais	CUMPRIDO	Arquivada	26
			Arquivada averiguada	30
	ASAE e outra Entidade	CUMPRIDO	Arquivada	374
			Arquivada averiguada	920
		PENDENTE	Em averiguação	85
			Por averiguar	145
	Indeterminada	CUMPRIDO	Arquivada	309
		PENDENTE	Por averiguar	1
	Outra Entidade	CUMPRIDO	Arquivada	892
			Arquivada averiguada	1
		PENDENTE	Por averiguar	2
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	10
	Tribunais	CUMPRIDO	Arquivada	177
			Arquivada averiguada	1
		PENDENTE	Por averiguar	1
Faro	ASAE	CUMPRIDO	Arquivada	974
			Arquivada averiguada	3608
		PENDENTE	Em averiguação	151
			Por averiguar	2166
	ASAE e Tribunais	CUMPRIDO	Arquivada	26
			Arquivada averiguada	70
		PENDENTE	Em averiguação	2
			Por averiguar	23
	ASAE e outra Entidade	CUMPRIDO	Arquivada	462

Região	Entidade	Estado	Resultado	Valor
			Arquivada averiguada	1218
		PENDENTE	Em averiguação	38
			Por averiguar	812
	Indeterminada	CUMPRIDO	Arquivada	725
			Arquivada averiguada	17
		PENDENTE	Por averiguar	5
	Outra Entidade	CUMPRIDO	Arquivada	1589
			Arquivada averiguada	10
		PENDENTE	Em averiguação	2
			Por averiguar	1
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	50
	Tribunais	CUMPRIDO	Arquivada	317
			Arquivada averiguada	5
Guarda	ASAE	CUMPRIDO	Arquivada	339
			Arquivada averiguada	387
		PENDENTE	Em averiguação	115
			Por averiguar	125
	ASAE e Tribunais	CUMPRIDO	Arquivada	10
			Arquivada averiguada	1
	ASAE e outra Entidade	CUMPRIDO	Arquivada	211
			Arquivada averiguada	283
		PENDENTE	Em averiguação	39
			Por averiguar	53
	Indeterminada	CUMPRIDO	Arquivada	83
	Outra Entidade	CUMPRIDO	Arquivada	272
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	2
	Tribunais	CUMPRIDO	Arquivada	19
Ilha Graciosa	Indeterminada	CUMPRIDO	Arquivada	1
	Outra Entidade	CUMPRIDO	Arquivada	1
Ilha Terceira	ASAE	CUMPRIDO	Arquivada	4
	Outra Entidade	CUMPRIDO	Arquivada	26
Ilha da Madeira	ASAE	CUMPRIDO	Arquivada	5
			Arquivada averiguada	5
		PENDENTE	Em averiguação	2
			Por averiguar	1

			Arquivada averiguada	1
	ASAE e outra Entidade	CUMPRIDO	Arquivada averiguada	1
	Indeterminada	CUMPRIDO	Arquivada	13
	Outra Entidade	CUMPRIDO	Arquivada	313
	Tribunais	CUMPRIDO	Arquivada	23
Ilha de Porto Santo	Outra Entidade	CUMPRIDO	Arquivada	14
Ilha de Santa Maria	Outra Entidade	CUMPRIDO	Arquivada	2
Ilha de São Jorge	Outra Entidade	CUMPRIDO	Arquivada	1
Ilha de São Miguel	ASAE	CUMPRIDO	Arquivada	4
			Arquivada averiguada	3
		PENDENTE	Por averiguar	2
	ASAE e outra Entidade	CUMPRIDO	Arquivada averiguada	3
	Indeterminada	CUMPRIDO	Arquivada	10
	Outra Entidade	CUMPRIDO	Arquivada	115
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	4
	Tribunais	CUMPRIDO	Arquivada	3
Ilha do Faial	Outra Entidade	CUMPRIDO	Arquivada	27
Ilha do Pico	ASAE	CUMPRIDO	Arquivada	1
	Outra Entidade	CUMPRIDO	Arquivada	14
Leiria	ASAE	CUMPRIDO	Arquivada	772
			Arquivada averiguada	1973
		PENDENTE	Em averiguação	210
			Por averiguar	614
	ASAE e Tribunais	CUMPRIDO	Arquivada	24
			Arquivada averiguada	38
		PENDENTE	Em averiguação	2
			Por averiguar	10
	ASAE e outra Entidade	CUMPRIDO	Arquivada	485
			Arquivada averiguada	1074
		PENDENTE	Em averiguação	80
			Por averiguar	247
	Indeterminada	CUMPRIDO	Arquivada	367
			Arquivada averiguada	3
		PENDENTE	Por averiguar	5
	Outra Entidade	CUMPRIDO	Arquivada	937
			Arquivada averiguada	3
		PENDENTE	Em averiguação	3
			Por averiguar	2
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	26

	Tribunais	CUMPRIDO	Arquivada	204
Lisboa	ASAE	CUMPRIDO	Arquivada	7935
			Arquivada averiguada	8115
		PENDENTE	Em averiguação	412
			Por averiguar	13618
	ASAE e Tribunais	CUMPRIDO	Arquivada	681
			Arquivada averiguada	471
		PENDENTE	Em averiguação	24
			Por averiguar	368
	ASAE e outra Entidade	CUMPRIDO	Arquivada	2494
			Arquivada averiguada	2560
		PENDENTE	Em averiguação	142
			Por averiguar	4715
	Indeterminada	CUMPRIDO	Arquivada	4020
			Arquivada averiguada	30
		PENDENTE	Em averiguação	2
			Por averiguar	20
	Outra Entidade	CUMPRIDO	Arquivada	8834
			Arquivada averiguada	23
		PENDENTE	Por averiguar	14
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	418
			Arquivada averiguada	1
	Tribunais	CUMPRIDO	Arquivada	4044
			Arquivada averiguada	8
		PENDENTE	Por averiguar	2
Portalegre	ASAE	CUMPRIDO	Arquivada	221
			Arquivada averiguada	514
		PENDENTE	Em averiguação	57
			Por averiguar	143
	ASAE e Tribunais	CUMPRIDO	Arquivada	9
			Arquivada averiguada	13
	ASAE e outra Entidade	CUMPRIDO	Arquivada	162
			Arquivada averiguada	290
		PENDENTE	Em averiguação	27
			Por averiguar	73
	Indeterminada	CUMPRIDO	Arquivada	75
	Outra Entidade	CUMPRIDO	Arquivada	192
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	7
	Tribunais	CUMPRIDO	Arquivada	24

Porto	ASAE	CUMPRIDO	Arquivada	9081
			Arquivada averiguada	2858
		PENDENTE	Em averiguação	655
			Por averiguar	2815
	ASAE e Tribunais	CUMPRIDO	Arquivada	210
			Arquivada averiguada	49
		PENDENTE	Em averiguação	4
			Por averiguar	73
	ASAE e outra Entidade	CUMPRIDO	Arquivada	4706
			Arquivada averiguada	1157
		PENDENTE	Em averiguação	311
			Por averiguar	1106
	Indeterminada	CUMPRIDO	Arquivada	1754
			Arquivada averiguada	10
		PENDENTE	Por averiguar	2
	Outra Entidade	CUMPRIDO	Arquivada	4548
			Arquivada averiguada	20
		PENDENTE	Por averiguar	8
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	114
	Tribunais	CUMPRIDO	Arquivada	1166
			Arquivada averiguada	4
Santarém	ASAE	CUMPRIDO	Arquivada	825
			Arquivada averiguada	1563
		PENDENTE	Em averiguação	169
			Por averiguar	896
	ASAE e Tribunais	CUMPRIDO	Arquivada	4
			Arquivada averiguada	22
		PENDENTE	Por averiguar	13
	ASAE e outra Entidade	CUMPRIDO	Arquivada	562
			Arquivada averiguada	818
		PENDENTE	Em averiguação	70
			Por averiguar	433
	Indeterminada	CUMPRIDO	Arquivada	292
			Arquivada averiguada	5
		PENDENTE	Por averiguar	2
	Outra Entidade	CUMPRIDO	Arquivada	813
			Arquivada averiguada	4
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	27

		PENDENTE	Por averiguar	1
	Tribunais	CUMPRIDO	Arquivada	90
Setúbal	ASAE	CUMPRIDO	Arquivada	2024
			Arquivada averiguada	2844
		PENDENTE	Em averiguação	294
			Por averiguar	3492
	ASAE e Tribunais	CUMPRIDO	Arquivada	21
			Arquivada averiguada	33
		PENDENTE	Em averiguação	4
			Por averiguar	38
	ASAE e outra Entidade	CUMPRIDO	Arquivada	1154
			Arquivada averiguada	1165
		PENDENTE	Em averiguação	118
			Por averiguar	1462
	Indeterminada	CUMPRIDO	Arquivada	925
			Arquivada averiguada	7
		PENDENTE	Por averiguar	4
	Outra Entidade	CUMPRIDO	Arquivada	2090
			Arquivada averiguada	7
		PENDENTE	Por averiguar	1
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	50
	Tribunais	CUMPRIDO	Arquivada	459
Viana do Castelo	ASAE	CUMPRIDO	Arquivada	766
			Arquivada averiguada	728
		PENDENTE	Em averiguação	72
			Por averiguar	307
	ASAE e Tribunais	CUMPRIDO	Arquivada	7
			Arquivada averiguada	8
	ASAE e outra Entidade	CUMPRIDO	Arquivada	306
			Arquivada averiguada	282
		PENDENTE	Em averiguação	27
			Por averiguar	149
	Indeterminada	CUMPRIDO	Arquivada	188
			Arquivada averiguada	2
		PENDENTE	Por averiguar	5
	Outra Entidade	CUMPRIDO	Arquivada	442
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	25
	Tribunais	CUMPRIDO	Arquivada	76
Vila Real	ASAE	CUMPRIDO	Arquivada	333

			Arquivada averiguada	439
		PENDENTE	Em averiguação	32
			Por averiguar	324
	ASAE e Tribunais	CUMPRIDO	Arquivada averiguada	6
		PENDENTE	Em averiguação	1
			Por averiguar	11
	ASAE e outra Entidade	CUMPRIDO	Arquivada	214
			Arquivada averiguada	196
		PENDENTE	Em averiguação	43
			Por averiguar	146
	Indeterminada	CUMPRIDO	Arquivada	124
		PENDENTE	Por averiguar	2
	Outra Entidade	CUMPRIDO	Arquivada	322
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	3
	Tribunais	CUMPRIDO	Arquivada	37
Viseu	ASAE	CUMPRIDO	Arquivada	775
			Arquivada averiguada	1377
		PENDENTE	Em averiguação	138
			Por averiguar	352
	ASAE e Tribunais	CUMPRIDO	Arquivada	32
			Arquivada averiguada	17
		PENDENTE	Por averiguar	1
	ASAE e outra Entidade	CUMPRIDO	Arquivada	358
			Arquivada averiguada	676
		PENDENTE	Em averiguação	55
			Por averiguar	134
	Indeterminada	CUMPRIDO	Arquivada	225
			Arquivada averiguada	5
	Outra Entidade	CUMPRIDO	Arquivada	576
			Arquivada averiguada	6
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	11
	Tribunais	CUMPRIDO	Arquivada	127
			Arquivada averiguada	1
Évora	ASAE	CUMPRIDO	Arquivada	352
			Arquivada averiguada	1000
		PENDENTE	Em averiguação	72
			Por averiguar	268
	ASAE e Tribunais	CUMPRIDO	Arquivada	12

			Arquivada averiguada	8
		PENDENTE	Por averiguar	7
	ASAE e outra Entidade	CUMPRIDO	Arquivada	164
			Arquivada averiguada	313
		PENDENTE	Em averiguação	34
			Por averiguar	97
	Indeterminada	CUMPRIDO	Arquivada	112
			Arquivada averiguada	5
	Outra Entidade	CUMPRIDO	Arquivada	321
			Arquivada averiguada	2
		PENDENTE	Por averiguar	1
	Outra Entidade e Tribunais	CUMPRIDO	Arquivada	13
	Tribunais	CUMPRIDO	Arquivada	51
			Arquivada averiguada	4

COMPLIANCE STUDY OF HAZARD ANALYSIS AND CRITICAL CONTROL POINT SYSTEM

Tatiana Sousa Pinto[1], Brigida Monica Faria[1,2], Luis Paulo Reis[2,3], Henrique Lopes Cardoso[2,3], Tiago Santos[2,3]

[1]*Escola Superior de Saúde – Instituto Politécnico do Porto (ESS-P.Porto), Porto, Portugal*
[2]*Laboratory of Artificial Intelligence and Computer Science (LIACC), University of Porto, Porto, Portugal*
[3]*Faculty of Engineering, University of Porto (FEUP), Porto, Portugal*

ABSTRACT

Hazard Analysis and Critical Control Point (HACCP) system is based on a preventive methodology to avoid potential hazards that can cause harm and to ensure that unsafe food is not made available to consumers. This system is recognized by the Economic and Food Safety Authority that is criminal police responsible for food safety and economic inspection in Portugal. Every day, Economic and Food Safety Authority generates a large and complex volume of data in the various inspections it performs, by the reception of complaints, in its classification, registration and in monitoring until the end of the process analysis. This study focus on the reported entities that are related to non-compliance with HACCP and from it try to understand the most common infractions. Results show values between 30% and 37% related to non-compliance of the HACCP system. As main conclusions, from 2014 to 2018, the number of these infractions maintained the same level and it will be important to understand if the relationship between these problems are related to legislation understanding or application.

KEYWORDS

Public Health, Food safety, Data analysis applications, HACCP.

1. INTRODUCTION

Nowadays, modern society is more concerned and alert about the dangers, lack of security, hygiene, control of systems and services that can affect the consumer. Hazard Analysis and Critical Control Point (HACCP) system is based on a preventive methodology, in order to avoid potential hazards that can cause harm to consumers by eliminating or reducing problems, and to ensure that unsafe food is not made available to the consumer (Wallace and Mortimore, 2016). HACCP system is based on the application of technical and scientific principles in the production and handling of food aspects, from the beginning of their production to the final disposal to the consumer (Lawley, R. et al., 2013).

In order to prevent, eliminate or only reduce the hazards that may contaminate the food during its production process and subsequent distribution, requirements which, once contemplated, have led to the effective application of the HACCP system should be taken into account ('The food safety hazard guidebook', 2013).

The prerequisites control the hazards associated with the surrounding environment in the food production process, while the HACCP system controls the hazards associated with the production process(Stankovic, 2015): Structures and Equipment; Sanitation Plan; Pest Control; Water supply; Collection of waste; Materials in contact with food; Personal Hygiene; Training.

Economic and Food Safety Authority is a criminal police body specialized in the areas of food security and economic inspection in Portugal (ASAE, 2018b) (Mil-Homens, 2007). It is responsible for the evaluation and communication of risks in the food chain, as well as for legal compliance with economic activities in the food and non-food sectors (ASAE, 2018c). The Economic and Food Safety Authority 's mission is to supervise and prevent compliance with the regulatory legislation for the exercise of economic activities in the food and non-food sectors, as well as the evaluation and communication of risks in the food chain (ASAE, 2018c). Economic and Food Safety Authority has followed, with special attention, the growing of technological innovation and today is created the need to have capacity to control and store the data produced in the daily tasks of all its stakeholders, with the aim of improving public health and automation of inspector's tasks. Economic and Food Safety Authority generates a large and complex volume of data in the various inspections it performs, in the reception of complaints, in its classification, in its registration and in the monitoring until the end of the process. This factor enhances the use of specialized storage tools capable of transforming data into information that positively influences decision making on inspection days (Lele, 2019). According to this expectation and with the predominant characteristic of the large volume of data, the techniques and tools for the analysis of these arise as a response, more concretely, through Data Analysis (Provost and Fawcett, 2013). To achieve formal declarations, the organization intends to use Data Analysis techniques, since it recognizes its capacity to respond to large amount of data (IRMA, 2016). Moreover, Economic and Food Safety Authority is the only organization in the world that brings economic security together with food security. Another aspect about Economic and Food Safety Authority it is the national connection body with its counterparts at European and international level (ASAE, 2018a). In fact, Economic and Food Safety Authority cooperates with other similar organizations around the world (ASAE, 2018a). As a criminal police body receives tens of thousands of complaints per year that convert into a lot of data. It is important to analyze the data available in the datasets and evaluate how can be an improvement in public health. Economic and Food Safety Authority generates a large and complex volume of data in the various inspections, in the reception of information requests and complaints, in its classification, in its registration and in the monitoring until the end of the process. This factor enhances the use of specialized storage tools capable of transforming data into information that positively influences decision making on inspection days (Regulamento (CE) n.º 852, 2004). According to this expectation and with the predominant characteristic of the large volume of data, the techniques and tools for the data analysis are related with Knowledge Discovery in Database and Data Analysis (Han, Kamber and Pei, 2012). To understand from the many data provided by Economic and Food Safety Authority those that are related to non-compliance with the HACCP system and to analyze its recurrence is the motivational factor of this study.

The remaining of the paper is structured as follows: section 2 presents the complaints flow, infractions classification and methodology used to understand and analyze data. Section 3 presents the results and discussion of the obtained result. The last part of this paper presents the major conclusions, and some directions for future work.

2. METHODOLOGY

This section presents the methods used to acquire more information about complaints that reach Economic and Food Safety Authority and which are related with non-compliance of HACCP system. First, a complaint is a formal complaint in order to expose an irregularity/non-compliance of economic group with the legislation (DL194/, 2012). All reports of the datasets were made online by filling in the form available on the Economic and Food Safety Authority website or sent by email. Then a decision circuit is made by inspectors to classify complaints in many aspects, such as competence (as can be seen in Figure 1) or type of complaints.
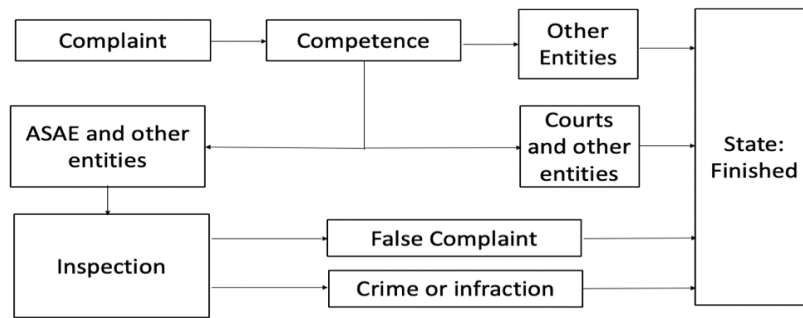
Figure 18. Complaint competence in decision circuit.

In terms of competence, there are many complaints that should be treated by other entities such as courts or other institutional entities. In this work, only complaints that are from Economic and Food Safety Authority competence will be analyze. The competence complaints have code assigned in order to classify the kind of infraction is presented.

Figure 2 presents an example of classification codes and frequencies in the year 2014. In this example the code 1.C.31 has digits and letters that represent different information: 1 represents "Segal", an abbreviation for "Segurança Alimentar" that means food security; C represents hygiene and food safety; 31 represents "Failure to comply with general and specific hygiene requirements".

```
'2.R.19 ': 102,
' 1.B.00.9 ': 217,
'1.L.10 ': 238,
'1.C.31 ': 1577,
'1.B.00.8 ': 42,
'2.G.02.9 ': 211,
'3.ZA.04.1 ': 8,
' 1.C.31 ': 2109,
```

Figure 19. Frequencies and infractions codes.

In this example, the most frequent infraction in 2014 was a HACCP compliance. Also, in this example there are different format related with the deep of detail in the code with two or three points dividing the fields of information.

2.1 Data and Ethical Issues

The data was made available by Economic and Food Safety Authority in xlsx format. Data are from 2014 to 2018 and are divided into two datasets: data on complaints and information requests and data related to type of complaints associated with infractions. Statistical analysis of the complaints file and information requests, a comparison was made between the number of complaints and information requests, including complaints and pending information requests per year. This statistical analysis was made using Python libraries, Pandas and NumPy (Mckinney, 2018). The first descriptive analysis was applied to five excel files related to the years of 2014 to 2018 with approximately 30MB each with a total of 120661 entries of complaints and information requests. This analysis was simple and with no setbacks because there was not missing values or incorrect data in any of the files. In 2018 the available data was only until October. Concerning the file related to denounced entities a 670mb file containing 2882629 reported entities was analyzed. The file's variables include:
- RNPC - National Registry of Legal Entities;
- Relation - If belongs to a head office or to an establishment;
- NIPC / NIF - Identification Number of Collective Person/ Tax Identification Number;
- Type - Type of Entity;
- Name - Entity Name;
- Postal Code;
- Community/Region;
- Address;

- District / County / Town;
- Legal Nature – RNPC;
- Main CAE – Main Code of Economic Activity;
- Secondary CAE – Secondary Code of Economic Activity.

The descriptive analysis about these variables focus on the type of entity most reported/denounced and in which districts most of the complaints were made. The Economic and Food Safety Authority has an epigraph file containing the meaning of each letter or code number associated with each infraction to make possible the analyses of the dataset related to type of complaints associated with infractions. Analyzing this file, it was verified that infractions beginning with the number "1" are related to HACCP. Each year between 2014 and 2018 corresponds an excel file with each report made, including the infringement code assigned to each report. After importing each excel file, per year, into the Jupyter Notebook (Mckinney, 2018) an initial observation was made to identify the columns of interest. Given the amendment to the General Data Protection Regulation, the directive and legislation on data protection have been replaced and all those involved in these works have signed terms in accordance with the new legislation to ensure the right of secrecy of both the complainant and of the complaint.

2.2 Data Analysis Methods

Descriptive data analysis was made using Pandas, a software library written for the Python programming language for manipulation and data analysis (DataCamp, 2001). It was also used NumPy that provides the data structures, algorithms, and it is needed for most scientific applications involving numerical data in Python(McKinney, 2018).

2.2.1 KDD Methodology

The data volume, generated and stored in the course of any activity, exceeds the capacity of human analysis and makes it impossible to extract knowledge from the same data, without resorting to a system that automates this process (Fayyad, Piatetsky-Shapiro and Smyth, 1996). This context justifies the existence of the research area of Knowledge Discovery in Databases (KDD), generally defined as "the non-trivial process of identifying valid and potentially useful patterns, perceptible from the data" (Fayyad, Piatetsky-Shapiro and Smyth, 1996), in which the principles associated to it are influenced by the areas of Artificial Intelligence, Automatic Learning, Pattern Recognition, Statistics, Database, Information Systems, among others(Maimon and Rokach, 2010). The method used to search the patterns in the data is called Data Mining and is considered the core step of the process of Knowledge Discovery in Databases, which is developed in several phases, such as: selection, preprocessing, transformation, data mining, interpretation and evaluation (Witten, Frank and Hall, 2005). These patterns may or may not represent useful knowledge, being one of the phases that normally requires the participation of the user (Fayyad, Piatetsky-Shapiro and Smyth, 1996). Data Analysis phase represents 20% in the KDD process and is better supported by software than the other phases. All other phases, from the selection of data to the interpretation of the patterns found, are more a matter of "art" than a routine that can be automated(Jothi, Rashid and Husain, 2015).

2.2.2 Data Pre-Processing

The data has some problems that were tackle using pre-processing. These problems were related with missing/incomplete values as can be analyzed in Table 1.

Table 12. Variables vs Percentage Invalid/Null data information

Variables	Invalid/Null (%)	Variables	Invalid/Null (%)
RNPC	< 0,01	Address	0,53
Relation	< 0,01	District	2,40
NIPC/NIF	< 0,01	County	6,05
Type	0,01	Town	90,74
Name	< 0,01	Legal Nature RNPC	0,60
Postal Code	73,23	Main CAE	41,20
Community/Region	0,43	Secondary CAE	89,73

In terms of missing values were identified the following problems:
- 3-digit suffix of the postal code missing;
- District / County / Town" field the information for "County" and "Town" to be filled.
- In terms of incorrect and/or outdated values:
- Town as region;
- Incomplete / ambiguous values;
- Data corruption;
- CAEs missing values.
- There were also inappropriate separators and delimiters:
- Mixture of separation by spaces, by "," and by ";".

The majority of data considered invalid or null is in the "CAEs", the "Postal Code" and the "Town". The missing data were filled matching it with online data bases with valid information. In case of unavailable information from online datasets, the lines with missing values were excluded from analysis.

From another dataset the column with the infractions codes is called "Infracções" as shown in Figure 3. In this column we have information about the infraction code, the type of infraction and a description that explains the meaning of the infraction code. All the following examples are from 2014 dataset.

Figure 3.

2.2.3. Transformed Data

As the goal is to study the compliance analysis of Economic and Food Safety Authority's Hazard Analysis and Critical Control Point system the focus of study will be in the column selected in Figure 3.

Each complaint is associated with an NID, which is a code assigned to identify each complaint. In each complaint, more than one infraction can be identified, as demonstrated in Figure 4, and it is necessary to separate the different infractions within each complaint.

Figure 4. NID and Infractions columns.

As each infraction of the same complaint is separated by paragraph, the split is done by the paragraph as can be seen in Figure 5 the output.


Figure 5. Output from split by paragraph.

Since the infractions are separated it is necessary to split each code so that we can proceed to the count.

Since the infraction code, the type of infraction, and the description of the infraction are separated by commas in the column "Infrações", it is used the comma split. This split creates a new column (Fig.6) named 'Codes' that

contains the infraction code that represents the infraction in each complaint. This column can be added to the original dataset for future studies in the database.

Using unique, one of the commands of the Pandas library, it turns out that split created a space at the end of each code. This space was removed so that it is possible to count the different infraction codes. In order to understand the meaning of each code it is necessary to analyze the file of epigraphs made available. By analyzing this file, it is verified that all the complaints started by the number '1' are related to the HACCP system. In order to know how many complaints are related to the HACCP system a new split was made. The output of this new split gives a 26083 total number of rows in the dataset for the year 2014. In order to know the number of 9726 complaints with the infringement code related to the HACCP system was applied a count of these entries. The same procedure described for the year 2014 was be applied for the years 2015, 2016 and 2018, trying to verify, through data analysis by descriptive statistics if there is any pattern over the years.

3. RESULTS
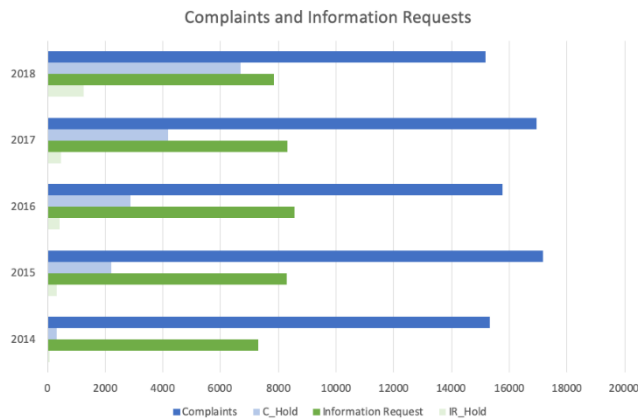
Data Analysis, is a superset of Data Mining that involves extracting, cleaning, transforming, modeling and visualization of data with an intention to uncover meaningful and useful information that can help in deriving conclusion and take decisions (Abzalov, 2016). This Section presents the results obtained by analyzing the complaints and information requests and the complaints related to HACCP system from 2014 to 2018.

3.1 Descriptive analysis about infractions related to complaints and information requests

From a total of 120661 entries between 2014 and 2018, 80358 of them are complaint and 40303 are information requests (see Figure 6).

Figure 6. Complaints and Information Requests frequencies.

"Complaints" represent the solved and archived complaints. "C_Hold" represent complaints on hold to be solved. "Information Requests" is related to information that the sender needs to be enlightened. "IR_Hold" represent information requests that have not yet been answered. The number of complaints has not decreased over the years. Complaints on hold (C_Hold) have not yet been solved. 2018 is the year with more "C_Hold" which is expected because there was no time for response yet. From a total of 120661 entries of denounced entities nearly 100% represent Economic Agents. To try to solve missing data related to entities addresses was used information in georeferenced websites. As a result of this cross-referencing of information 97,33% of data became complete.

3.2 Descriptive analysis about infractions related to HACCP system

The number of rows of all original datasets increased after the split (Fig. 6). Once each line represents a complaint and to each complaint is associated an infraction code, it means that there are complaints with more than one infraction code. In Figure 7 it is possible to compare the number of original rows with the number of rows created after the split from 2014 to 2018.
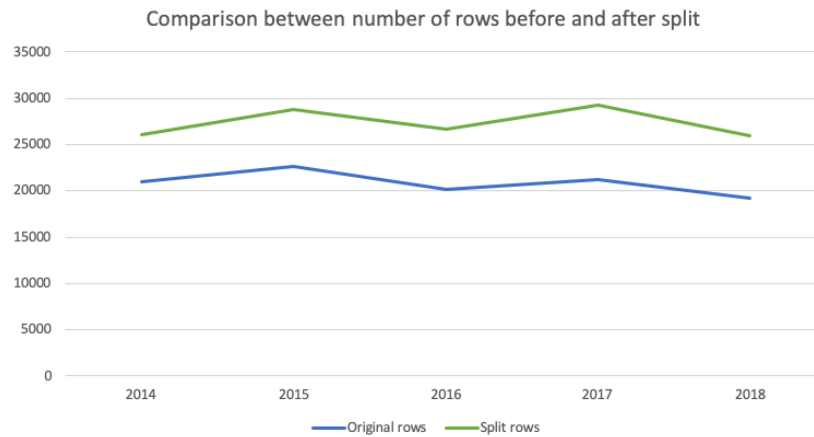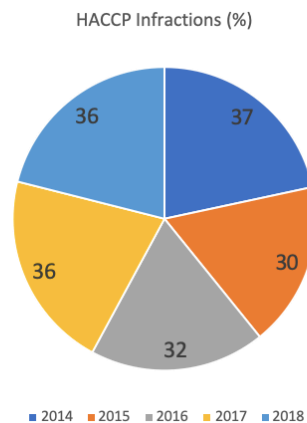

Figure 7. Comparison between the number of rows before and after split.

As can be seen (Figure 7), in all datasets from 2014 to 2018 there are complaints with more than one infraction committed. As the objective is the analysis of compliance with the HACCP system, it is important to analyze

the infractions committed related to this system in each year under study (Figure 8).

Figure 8. Percentage of infractions related to non-compliance with HACCP per year.

The percentages of non–compliance of the HACCP system were rounded to the units. As can be seen, the variation between the interval of years was not accentuated. From 2014 to 2015 the compliance improved by 7%, however in the remaining years this percentage of compliance was disappearing.

4. CONCLUSIONS AND FUTURE WORK

Food and Economic Security Authority receives more complaints than information requests. It is important to try to understand where law enforcement is failing so that there is not a significant decrease in complaints over the years. Incomplete, incorrect and omitted data relating to the entities complained may be justified since the complainant does not always know the correct and complete data of the entity it intends to report.

From the point of view of compliance with the HACCP system, values between 30% and 37% are relatively high since the Food and Economic Security Authority is responsible for food and economic security and not only for food security and hygiene. Although, as mentioned above, there was an improvement between the first two years under study, followed by worse results in the following years, the values related to the non-compliance rate did not have marked differences.

As future work, it would be interesting to evaluate the density of distribution of the complaints by regions. Once this has been done, and in an attempt to improve public health, it is necessary to associate them to the HACCP system related complain. Then try to understand what the most common complaint is and see if the problem is related to understanding the legislation or if it is related to the difficulty of complying with the legislation because it is too restrictive.

ACKNOWLEDGEMENT

REFERENCES

Abzalov, M., 2016. *Exploratory data analysis*, in Modern Approaches in Solid Earth Sciences. doi: 10.1007/978-3-319-39264-6_15.

ASAE, 2018a. ASAE-Cooperação Internacional Bilateral, https://www.asae.gov.pt/cooperacao/internacional/bilateral/medio-oriente.aspx.

ASAE, 2018b. Como Actua a ASAE. Available at: https://www.asae.gov.pt/inspecao-fiscalizacao/como-atua-a-asae.aspx (Accessed: 30 May 2019).

ASAE, 2018c. Missão, Visão e Valores. Available at: https://www.asae.gov.pt/asae20/missao-visao-e-valores.aspx (Accessed: 30 May 2019).

DataCamp, 2001. *Python for Data Science*, Cheatsheet. doi: 10.1097/00007890-200105270-00005.

DL194/ (2012) 'Decreto-Lei no 194/2012 ASAE', Diário da República 1a série, pp. 4641 – 4646

Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P., 1996. *The KDD process for extracting useful knowledge from volumes of data*. Communications of the ACM, 39(11), pp.27-34.

Han, J., Kamber, M. and Pei, J., 2012. *Data Mining: Concepts and Techniques*. doi: 10.1016/C2009-0-61819-5.

IRMA, I. R. M. A., 2016. *Big Data: Concepts, Methodologies, Tools, and Applications*. Information Resources Management Association ISBN-13: 978-1466698406

Jothi, N., Rashid, N. A. and Husain, W., 2015. *Data Mining in Healthcare – A Review*, in Procedia Computer Science. doi: 10.1016/j.procs.2015.12.145.

Lawley, R.; Curtis, L.; Davis, J., 2013. *The food safety hazard guidebook*, Choice Reviews Online. doi: 10.5860/choice.46-3282.

Lele, A., 2019. *Big data*, in Smart Innovation, Systems and Technologies. doi: 10.1007/978-981-13-3384-2_9.

Maimon, O. and Rokach, L., 2010. *Introduction to Knowledge Discovery and Data Mining*, in Data Mining and Knowledge Discovery Handbook. doi: 10.1007/978-0-387-09823-4_1.

McKinney, W., 2018. *Python for Data Analysis Data Wrangling with Pandas, NumPy, and IPython*, Climate Change 2013 – The Physical Science Basis. doi: 10.1017/CBO9781107415324.004.

Mil-Homens, S., 2007. *ASAE - HACCP*. Available at: http://www.asae.gov.pt/pagina.aspx?back=1&codigono=54105579AAAAAAAAAAAAAAAA (Accessed: 19 January 2019).

Provost, F. and Fawcett, T., 2013. *Data Science and its Relationship to Big Data and Data-Driven Decision Making*. Big Data, 1(1), pp.51–59.

Regulamento (CE) n.o 852, 2004. Regulamento (UE) no 852/2004 do parlamento europeu e do conselho', *Jornal Oficial da União Europeia*, 139, pp. 54.

Stankovic, I., 2015. Codex Alimentarius, in *Encyclopedia of Food and Health*. doi: 10.1016/B978-0-12-384947-2.00180-X.

Wallace, C. A. and Mortimore, S. E. (2016) 'HACCP', in Handbook of Hygiene Control in the Food Industry: Second Edition. doi: 10.1016/B978-0-08-100155-4.00003-0.

Witten, I. H., Frank, E. and Hall, M. A., 2005. *Data Mining Practical Machine Learning Tools And Techniques*, Data Mining. doi: 0120884070, 9780120884070.