Network Analysis for Food Safety: quantitative and structural study of data gathered through the RASFF system in the European Union

Abstract: This paper reports a quantitative and structural analysis of data gathered on the food issues reported by the European Union members over the last forty years. The study applies statistical measures and network analysis techniques. For this purpose, a graph was constructed of how different contaminated products have been distributed through countries. The work aims to leverage insights into the structure formed by the involvement of European countries in the exchange of goods that can cause problems for populations. The results obtained show the roles of different countries in the detection of sensitive routes. In particular, the analysis identifies problematic origin countries, such as China or Turkey, whereas European countries, in general, do have good border control policies for the import/export of food.

Keywords: Food Safety, Graph Theory, Network Analysis.

1 Introduction

We live in a globalized world where it is easy to find products from any country in any store. This is an advantage for consumers who, however, are often unaware of the dangers to which they are exposed. Risks arise due to the different policies and legislation that countries must comply with regarding consumer rights and feed and food safety. As a result, (Jongwanich, 2009) and (Timmis, 2017) conclude that regulations established by developed countries hold back food exports from developing countries because of the unique mistrust of these products, which are closely related to human health.

Since the benefits of food safety are directly related to the costs of health systems (Pal et al., 2015), there is a growing interest in determining how food safety policies in different countries can influence the number and severity of alerts detected in import/export operations.

The European Commission considers the highest possible food safety standards to be a key political priority, (Mutukumira & Jukes, 2003). To this end, the Rapid Alert System for Food and Feed (RASFF) was established in 1979 and now includes the Commission, the EU Member States, the European Food Safety Authority (EFSA), the European Surveillance Authority (ESA), Iceland, Liechtenstein, and Norway, (About RASFF, 2022). This tool enables the efficient exchange of information about food and feed issues that can be a risk to human health or can be considered fraudulent. The food issues are registered by contact points, which work at a national level, in an online system called iRASFF. The registration of notifications is shared between countries and organizations so that they can make fast and effective decisions.

The issues recorded in the RASFF are made up of a set of categorical characteristics that encode the countries involved and the types of contaminated or hazardous food involved. They describe the movement of a product between countries, and useful insights can be leveraged by taking advantage of the representation of import and export information in the form of a graph or network and its subsequent analysis. The process used to analyse graphs is called network analysis (NA). NA comprises a set of graph theory techniques that can be used to analyse networks formed by various actors. In this case, NA will be applied to countries and their trade of products; their behaviour can be analysed as a network. The results provide food safety patterns that could be used by food policy authorities and official organizations.

This research aims to provide a complete snapshot of the chain trade formed by the involved countries, the issued products, and their hazards, covering the food alerts stored in the RASFF records since their inception. To build a graph, RASFF data representing the countries involved (origin, distribution, and destination), the products issued, and their hazards were used. In this graph, nodes represent countries, while edges describe the flow followed by food alerts. Edge directions are set by the role of a country in the commercial chain: origin, distribution, or destination. Edges are also labelled with the issued product and its hazards. The detailed analysis of the graph allows the extraction of general statistical measures and NA metrics calculated from the general graph. These metrics can be used to recommend countries with good performance in terms of food policies, pay attention to others that have problems with contaminated products, increase the monitoring of certain products and their origin countries or find sensitive routes of trading. This is the first time that the complete RASFF dataset has been analysed as a graph using NA, and a complete set of metrics and statistics are reported and interpreted. Although there are papers studying the networks created by RASFF data, (Petroczi et al., 2011), (Bui-Klimke et al., 2014), (Popp et al., 2018) and (Petroczi et al.), 2010), they only comprise a particular period or are focused on a subset of issues. Another contribution is the free availability for the first time of the complete dataset. This fact has more weight as now the European Commission has decided to change its website and only provides data from 2019 with restrictions on some features.

2 Related works

Since RASFF import/export information can be configured as a graph, NA techniques are applied to extract and analyze the information contained in these data. NA is defined in general terms as a way to find patterns of relationships between social entities such as people, events, organizations, and other entities (Jamali & Abolhassani, 2006). (Mincer & Niewiadomska-Szynkiewicz, 2012) give a more accurate definition closer to that of the present research: NA provides a set of techniques used in graph theory that allow understanding networks formed by several actors. These techniques are often applied in different fields to understand the roles of different actors in a specific environment and how they interrelate.

In the case of food and feed safety, several papers make use of NA techniques. The model presented in (Wu & Guclu, 2013) uses NA in a use case of maize trade with data from 2000 to 2009. (Xu et al., 2014) present a framework for food traceability using data from various resources and applying the Internet of Things, social analytics, and mobile technologies. A network-based methodology for Hungarian cattle holding is presented in (Jozwiak et al., 2016). (Fair et al., 2017) use information from the global wheat trade between 1986 and 2011 to create a preferential attachment network model. Finally, (Wang, 2018) created a graph-based model for food safety monitoring based on a set of random graphs created artificially. Although the papers cited in this paragraph use NA techniques in the field of food and feed safety, they do not take advantage of RASFF data and, thus, are unable to analyze which countries cause more issues or which routes should be monitored more closely due to their risks.

Several papers make use of the RASFF dataset to perform statistical analyses; however, none benefit from the whole historical context. In these cases, studies focus on a subset of products or hazards or a particular region or period. For example, (Dada et al., 2021) studied microbiological hazards of products from the Asia Pacifica region from 2000 to 2020, and (Papapanagiotou, 2021) analysed food contact materials from 2012 to 2019. Similar work can be found in (Piglowski, 2020), (Kuchheuser et al., 2022), (Kępinska-Pacelik & Biel, 2021), (Alshannaq & Yu, 2021), (Caldeira et al., 2021), (Somorin et al., 2021), (De Leo et al., 2021), (Djekic et al., 2017), (Amico et al., 2018), (Kononiuk & Karwowska, 2017), (Duan et al., 2017), (Cınar et al., 2017), (Czepielewska et al., 2018), (Pennone et al., 2018), (Papapanagiotou, 2017), (Van Asselt et al., 2018), (Lüth et al., 2019) and (Tudela-Marco et al., 2017)].

Other works apply advanced computer science algorithms to RASFF data. For example, (Pigłowski, 2021) conducts a clustering analysis of RASFF data for the period of 1999-2018 alongside data from the Standard International Trade Classification. (Pigłowski, 2017) also applies cluster analysis to understand the dependencies between products and hazards. In (Pigłowski, 2017), tree clustering and k-means are applied to RASFF data for 2000 to 2015 to identify dangerous European countries. (Bouzembrak et al., 2018) applied Bayesian networks to determine the probability of contamination in species and herbs.

Only a few papers apply NA to the case of RASFF data and most are very specific. For instance, (Petroczi et al., 2011) focused only on mycotoxin contamination. (Bui-Klimke et al., 2014) studied how aflatoxin regulations have affected the global trade of pistachios. (Popp et al., 2018) also applied NA to the field of food safety but used a dataset from the FAO to study the honey trade network. (Petroczi et al., 2010) make a graph analysis but only consider a period from 2000 to 2009 and focused on the perspective given by notifier countries.

Finally, a couple of tools using the RASFF dataset have been identified. (Naughton et al., 2015) describe a prototype tool for NA using the RASFF, but it lacks metrics or quantitative results, and it is no longer available. (Robson et al., 2021) developed a tool to identify threats and risks of the beef supply chain.

Thus, although some studies have used NA with RASFF data, none work with the whole dataset except (Naughton et al., 2015), who describe the functionalities of a tool without providing NA metrics. Our work uses all the issues stored in the RASFF portal, instead of a subset, and reports a complete analysis of quantitative results that have not been found in previous literature. Since RASFF information can be modelled as a graph, NA techniques are applied to extract and analyze the information in the full RASFF dataset. We have also provided free access to the dataset which contains information that is inaccessible nowadays through the actual RASFF portal.

3 Materials and methods

This section describes the dataset used to build the graph and the techniques applied in this work.

3.1 Materials

The information stored in the RASFF cannot be downloaded all at once, as the website only allows users to compile 5,000 records in an XLS file of a total of 56,351 records registered in the RASFF portal from 1979 to the time of the research, 20th July of 2020. Additionally, the obtained file does not contain all the registered fields. For instance, it does not include information about distribution and destination countries that are necessary for the described research. In our case, we downloaded all the information needed to build a graph. To obtain the whole data history with all the features, we created a web scraper, (RASFF scrapper, 2022). This is a tool that uses the necessary techniques to generate structured data based on available unstructured data on the web, (Saurkar et al., 2018). After launching the scraper, a CSV file with all the information in RASFF Portal is obtained, (Nogales & García-Tejedor, 2022). We must highlight that this version of the RASFF portal changed in 2020. This fact supposes a limitation as no data from onwards is available in this format. However, as the available data now comprises from 2019 to the actual year (2022), we consider that having this dataset openly available is another contribution of the work.

Table 1 Features of the RASFF records used in this research.

Feature	Description
Notification country	Country registering the issue. The cardinality value is 31 (EU countries, Norway, Liechtenstein, and Iceland).
Product	The specific name of the product.
Hazard	The hazards or anomalies that have caused the issue.
Origin country	Country of origin of the product. Any country in the world.
Destination country	Country or countries of the product's destination. Any country in the world.
Distribution country	Countries within the transportation chain of the product between the origin and destination. Any country in the world or certain international regulatory organizations.

It must be noted that some issues have more weight (those having more instances) in the dataset, thus influencing the results. This could be related to stricter food policies adopted in the countries that detect them or to a larger number of issues related to a particular product in terms of productivity.

3.2 Methods

For this research, a two-stage analysis was carried out. First, a statistical analysis was performed on the scraped RASFF dataset to describe the characteristics of the different elements of the dataset, the countries, and the hazards reflected by registered food issues. The results can be used to identify potentially sensitive countries, countries with good food policies or products, and hazards that should be closely monitored. Second, we conducted an analysis based on network analysis techniques to understand the graph created from the information on the issues registered in the RASFF. This information reflects the number of issues reported among countries: the characteristics of the different actors, their behaviour and roles, and the structure of the graph itself.

3.2.1 Statistical analysis

This analysis is based on the frequency analysis of the data contained in the RASFF dataset. The aim is to obtain the following:

- From the information on countries and routes, the frequency distributions of the most involved countries, considering the role each country plays in the chain: origin, distributor, destination, and notifier. The information reflects the countries with the most restrictive or effective safety policies. Additionally, sensitive routes can be found. With this information, hypotheses about which countries should increase their food safety policies or which are the countries with good food controls that could be replicated in others could be established.

- Based on the information of the products affected by the reported issues, which products generate the most problems, or which are the most common dangers, as well as the combinations between them. Countries may use this information to pay more attention to specific products or hazards or to adjust the stringency of product safety policies in certain countries.

3.2.2 Graph NA

The next step involves applying NA techniques to study the graph formed by RASFF issues. Graph G is defined in (Schaeffer, 2007) as pair $G=(V, E)$, where V is a set of vertices or nodes, the number of nodes $n=|V|$ denotes the order of the graph, and set E contains the edges of the graph. The edges also have a weight denoted by w that represents the number of occurrences of the relation between the two nodes connected by an edge. We use NetworkX, (NetworkX, 2022), a Python library developed to deal with complex networks/graphs, to create a graph using the data extracted from the RASFF.

The first step is to represent the data in the form of a graph $G = (N, E)$, where N is the set of nodes represented by the countries, and E is the set of directed edges that represent the relationships between countries in each RASFF record and the character of the nodes as origins, destinations, or distributions (this was only taken into account to establish the direction of each edge). To include the number of issues between different countries, a weight is added to the edges, reflecting the number of occurrences of an issue between two given countries for a product contaminated by the same hazard. Thus, a weighted directed acyclic graph, $DAG=(N,E,W)$, is obtained, where W is the set of weights of the graph. This is a unique graph that includes all the registered issues recorded in the RASFF. In this case, the nodes denote countries, an edge and its direction denote how a product has travelled through countries (origin, distribution, and destination), and the label of an edge represents the product and hazard of the issue. Fig. 1 shows how a record is transformed into a part of the graph.

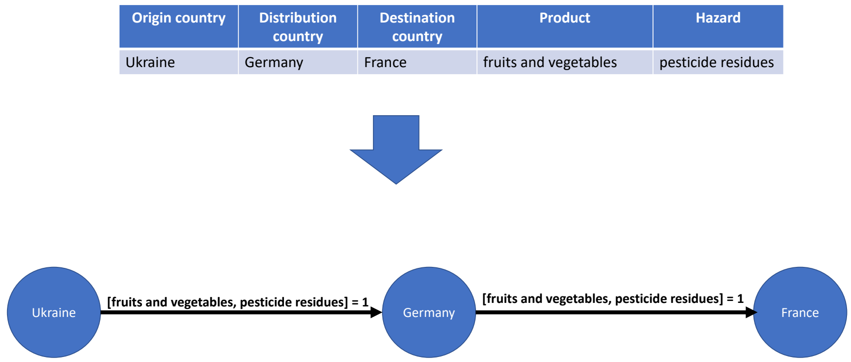

Fig. 1: An example of the RASFF record transformed into part of the graph.

NA processes RASFF data as a unit using a DAG formed by all previously scrapped registered issues considering the connections created by the transport of a contaminated product through different countries on any date. With such analysis, three types of results are obtained: general results, those related to connectivity, and those linked to centrality. General metrics describe some characteristics of the graph itself that will allow us to understand the structure formed by the different food issues. Connectivity metrics inform how the different countries are related, allowing for a better understanding of network performance. Finally, centrality provides information about the relative importance of each country in the network.

3.2.2.1 Definitions of general metrics

Density refers to the number of edges in the network relative to the number of potential edges, (Goldberg, 1984). This indicates how nodes are connected between them. The measure can be used to determine if most countries can spread a contaminated product to many countries. In Equation 1, m is the number of edges, and n is the number of nodes.

$$d = \frac{m}{n*(n-1)} \qquad (1)$$

Determining whether a network is heterogeneous or homogeneous is also relevant. In heterogeneous networks, most nodes have only a few edges; that is, there are only a few nodes with many edges. This describes how the connectivity of the countries is distributed. In contrast, for homogeneous nodes, most nodes are not highly connected and have approximately the same number of edges. This kind of network can be modelled by a normal distribution.

When a graph is heterogeneous, its degree distribution follows a power law. Power laws are defined in (Clauset et al., 2009) as forming a model of the relationship between two variables that are inversely related. The mathematical definition of a power law is as follows:

$$p(x) = Cx^{-\alpha} \text{ for } x > x_{min} \qquad (2)$$

where x corresponds to the quantity whose distribution is studied, and C and α are two constants. Then, x_{min} is a lower bound as the distribution diverges at zero, having a value greater than 0. To normalize the power law, conditions $\alpha>1$ and $x_{min}>0$ must be fulfilled.

$$C = (\alpha-1)x_{min}^{\alpha-1} \qquad (3)$$

Relationships obeying power laws appear as straight lines on log–log plots and they are “scale-free”, which implies that while most nodes in the network have a small number of links, there are a few selected nodes that act as highly connected hubs with many links. For spatial or geographical networks this has implications for failure tolerance, risk assessment, and production chains in general.

3.2.2.2 Definition of connectivity metrics

Definitions of connectivity metrics can be found in (Barnes, 1984). A connected component is defined as the largest subgraph that can be obtained from a graph where all pairs of nodes are connected by a path.

By obtaining the connected components, groups of nodes with similar behaviour between them but sufficiently different from others can be found.

A graph is strongly connected when for every pair of nodes, it is possible to find a (directed) path that connects them. The maximal strongly connected subgraphs are known as the strongly connected components (SCC). On the other hand, a graph is said to be weakly connected if there is not a (directed) path between any two pairs of vertices, but it is connected when considered as an undirected graph. In that case, a weakly connected component (WCC) is a subgraph of the original graph where all vertices are connected by some path, ignoring the direction of edges.

The diameter is related to the type of connectivity of the whole graph. It is defined as the maximum of the shortest paths between all pairs of nodes, (West, 1996). It is described by the following equation.

$$diam(G) = max_{x,y \in X} d(x,y) \qquad (4)$$

When the diameter is low, it is easier to reach the rest of the network from a node. For the RASFF case, the value represents the minimum number of connections needed to reach one country from another in a certain subgraph of the data.

The clustering coefficient of a node is a measure used to understand whether a node's neighbours are also linked, (Holland & Leinhardt, 1971). It is also denoted in Equation 5, where d_n and e_n correspond to the degree of the node and its number of edges.

$$C_n \begin{cases} 0 if d_n = 0 \\ \frac{e_n}{\binom{d_n}{2}} \quad if d_n \geq 2 \end{cases} \qquad (5)$$

When applied to the whole network, the value is called the average clustering coefficient. In the present case, it helps to determine whether there are sensitive routes between all countries, as the measure is useful for understanding how connected the nodes to each other are.

3.2.2.3 Definition of centrality metrics

Four different centrality measures are used: degree, closeness, betweenness, and eigenvector centrality. Degree centrality is related to nodes' degree (that is, the number of edges in a node), (Freeman, 1978). In the case of directed graphs, it is necessary to differentiate between the in-degree and out-degree; the first refers to the number of edges reaching a node, and the second is related to edges leaving a node. This metric can be used to understand which nodes are the most influential and their dependence on the network. In the case of an in-degree, this indicates countries that receive more products. An out-degree is used to identify countries that export or distribute the most. Both metrics are defined in Equations 6 and 7.

$$C_j^{IN} = \sum_{i=1}^{n} a_{ji} \qquad (6)$$

$$C_j^{OUT} = \sum_{i=1}^{n} a_{ij} \qquad (7)$$

Closeness centrality is calculated based on the number of edges between a node and each other node, (Bavelas, 1950). Equation 8 is formalized where *n* is the number of nodes and *d(v,u)* is the shortest path between this pair of nodes.

$$C(u) = \frac{n-1}{\sum_{v=1}^{n-1} d(v,u)} \quad (8)$$

This provides information about how easily an issued product can be spread through all countries. When the value is small, spreading information costs less, so the nodes are considered reference points in the network. In this paper, it is useful to determine how fast an issued product could reach many countries.

Betweenness centrality measures the importance of a node in connecting other nodes in a network, (Bonacich, 1972). The following equation describes this measure. *V* is a set of nodes, $\sigma(s,t)$ is the number of shortest (s,t) paths and $\sigma(s,t|v)$ is the number of paths passing through node *v* over *s, t*.

$$c_B(v) = \sum_{s,t \in v} \frac{\sigma(s,t \vee v)}{\sigma(s,t)} \qquad (9)$$

The measure can be used to determine the role of a node as an intermediary between nodes that are transmitting information, such as the distributing countries in this case.

Eigenvector centrality explains how a node influences the rest of the network. The eigenvector of a node is high when it is connected to many nodes, and they also have many connections (Bihari & Pandia 2015). Equation 10 describes this where $a_{v,t}$ is an element if the adjacency matrix and λ is a constant

$$x_v = \frac{1}{\lambda}\sum_{t\in M(v)} x_t = \frac{1}{\lambda}\sum_{t\in G} a_{v,t}x_t \quad (10)$$

It is useful to rank the nodes in terms of how they influence others in the network. Countries with high eigenvectors have a very strong influence on the rest of the network.

4 Results

The results are grouped following the techniques and methodologies described in Section 3 on the statistical analysis of the RASFF dataset and NA of the graph formed by all examined issues.

4.1 Statistical analysis of countries

The general analysis of the countries' behaviour is based on their roles in the notification system: origin, distribution, destination, and notifier. This information uses the number of times countries present certain issues to show which countries perform the most as origins, distribution destinations, or notifiers and to find sensitive routes.

Fig. 2 depicts the 35 countries with the biggest number of issues reported by the RASFF implicated as origin countries. As shown, three countries stand out from the rest: China, Turkey, and India. A few other countries also seem to produce an important number of issues. Nevertheless, most are concentrated in the lower part of the distribution.

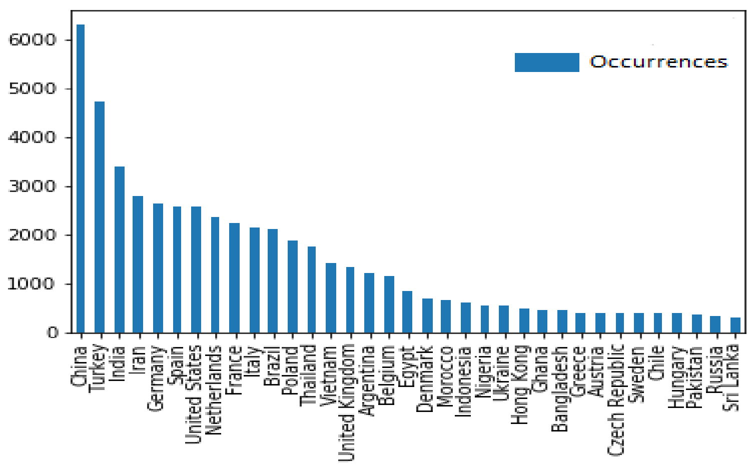

Fig. 2: Top 35 most frequent origin countries in the issues reported by the RASFF

Table 2 shows the top 5 countries that appear as origin countries. The table indicates the number of instances per country and its percentage. China is by far the most issued origin country. Additionally, Turkey and India stand out among the others. From the third position onwards, the percentage differences are less than 1%.

Table 2. Top 5 origin countries (n=56,531).

Country	Notifications	Percentage (%)
China	6,280	11.1%

Turkey	4,707	8.35%
India	3,375	5.98%
Iran	2,770	4.91%
Germany	2,628	4.66%

Fig. 3 lists the top 35 distribution countries. The diagram of bars shows that almost all countries belong to the EU. It is worth noting the use of the label “No country”, which refers to issues involving no distribution country. Then, fewer countries are found in the medium range. Most of the countries are distributed in the lower part of the graphic.

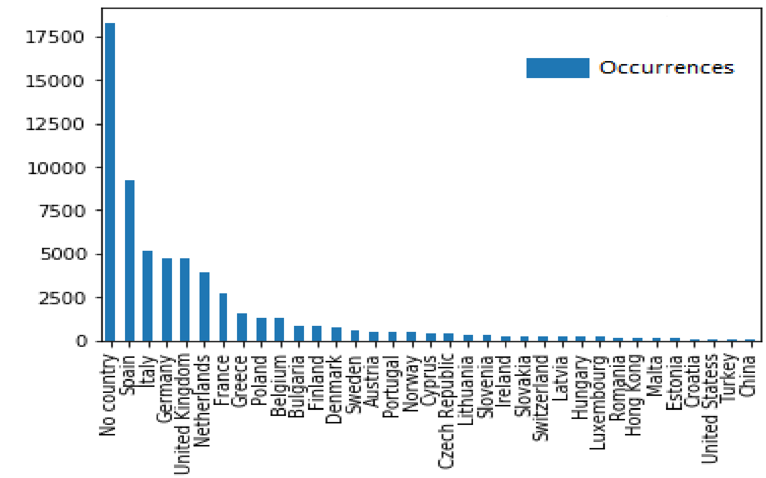

Fig. 3: Top 35 most frequent distribution countries in the issues reported by the RASFF

In Table 3, the top 5 distribution countries with their instances and the percentages they represent are shown. It is worth noting that more than 30% of the reported issues do not involve a distribution country. This shows that much of the trade is done directly between two countries (no intermediary country) or that many issues are already reported in the origin country. Another interesting point is that the rest of the countries belong to the EU, which is also plausible, as the RASFF is a tool of the European Commission.

Table 3. Top 5 distribution countries (n=56,531).

Country	Notifications	Percentage (%)
No country	18,240	32.36%
Spain	9,202	16.32%
Italy	5,138	9.11%
Germany	4,746	8.42%
United Kingdom	4,716	8.36%

Fig. 4 shows how the 35 top destination countries are distributed. Again, almost all countries belong to the EU, which could be expected due to the RASFF's nature. The bar with more occurrences also shows that many issues involve no destination country (or it is not properly reported). No country stands out, even though Germany and Italy appear to be more frequently involved. The rest of the distribution is like the one shown in Fig. 3.

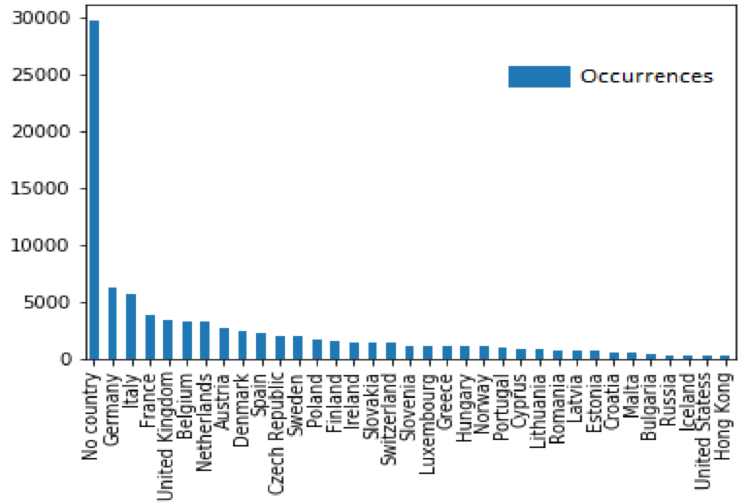

Fig. 4: Top 35 most frequent destination countries in the issues reported by the RASFF

The top 5 destination countries are shown in Table 4, again with instances and percentages. On the one hand, it should be noted that more than one destination country may be involved in a reported issue. On the other hand, the first position corresponds to issues involving no destination country. This occurs when a problem is found before a product leaves the origin country or at the distribution moment.

Table 4. Top 5 destination countries (n=56,531).

Country	Notifications	Percentage (%)
No country	29,612	52.54%
Germany	6,220	11.03%
Italy	5,598	9.93%
France	3,866	6.86%
United Kingdom	3,372	5.98%

The information related to the types of countries provided above can be complemented with information related to notification countries, which have identified product hazards. Notification is likely to occur in any of the stages of the trading chain. It may occur before a product leaves the origin country, when being distributed by another country, at the destination immediately before entering a country, or even when the product has been distributed to the market. For each issue, only one country can be a notifier. This makes it possible to find key countries in the logistics chain.

The distribution of the 35 leading notifier countries is shown in Fig. 5. Again, almost all countries belong to the EU and exhibit similar behaviour as that shown in Fig. 2. In this case, three countries can be highlighted: Italy, Germany, and the United Kingdom, serving as a set of countries that stand out enough to be studied. Last, most of the countries only show very few occurrences.

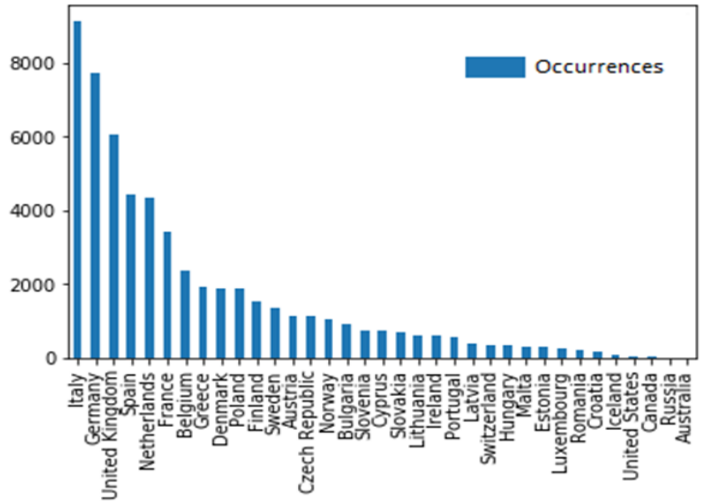

Fig. 5: Top 35 most frequent notifier countries in the issues reported by the RASFF

In Table 5, the top 5 notification countries are shown. Except for the Netherlands, the other four countries also appear among the most issued countries by distribution, origin, and destination.

Table 5. Top 5 countries (n=56,531).

Country	Notifications	Percentage (%)
Italy	9,111	16.16%
Germany	7,704	13.67%
United Kingdom	6,061	10.75%
Spain	4,410	7.82%
Netherlands	4,357	7.73%

To learn more about the role of notification countries, Fig. 6 relates this role to the others of the trading chain studied above. The figure shows the percentage of countries that are notifiers while at the same time performing any of the other roles: origin, distribution, destination, or all of them at the same time (goods that are traded within a country). “Does not apply” denotes cases with an error in records (i.e., it has no countries in the record). The figure shows that when a country reports an issue, it is often also a distributor or destination country.

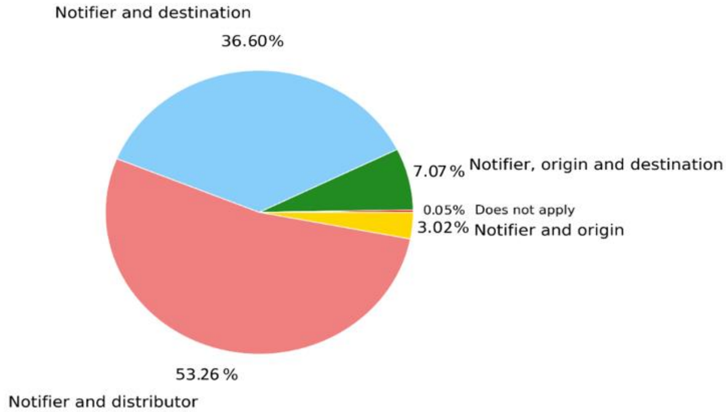

Fig. 6: Proportion of countries that are notifiers as well as performing any other role in the trading chain.

Table 6 shows the top 5 countries for each case presented in Fig. 6. The first column lists countries that are countries of origin and at the same time have notified the issue. The second column provides the same information for countries of distribution. The third column shows the top 5 destination countries that have found contaminated products. The last column contains information on countries that are the origin and destination of a product with the reported issues. Numbers in parentheses denote the percentages of issues reported by countries.

Table 6. Top 5 notifier countries vs. other roles.

Notifier and origin	Notifier and distribution	Notifier and destination	Notifier, origin and destination
France (256)	Italy (4,781)	Italy (4,154)	Italy (3,553)
Netherlands (247)	United Kingdom (4,127)	Germany (3,619)	Germany (3,006)
Germany (229)	Germany (3,853)	United Kingdom (1,773)	United Kingdom (1,414)
Belgium (215)	Spain (3,695)	France (1,626)	Denmark (1,264)
Italy (171)	Netherlands (2,818)	Denmark (1,394)	France (1,151)

It is also interesting to determine from which countries notifiers have found contaminated products. Two chord diagrams are used to represent the routes followed by the reported products, as shown in Fig. 7 and 8. The first category accounts for countries that are notifiers and distributors, while the second category refers to countries that are notifiers and destinations. In both diagrams, countries from Table 6 are shown in the upper half of the circle. Then, they are connected to the origin countries of a registered issue with a specific-coloured line. The width of a line depends on the number of issues identified with one country as the origin/destination and the other as a notifier.

Fig. 7 shows three country pairs with stronger connections than those of the rest of the considered countries: China-Italy, Iran-Germany, Spain-China, Germany-Turkey, and India-United Kingdom.

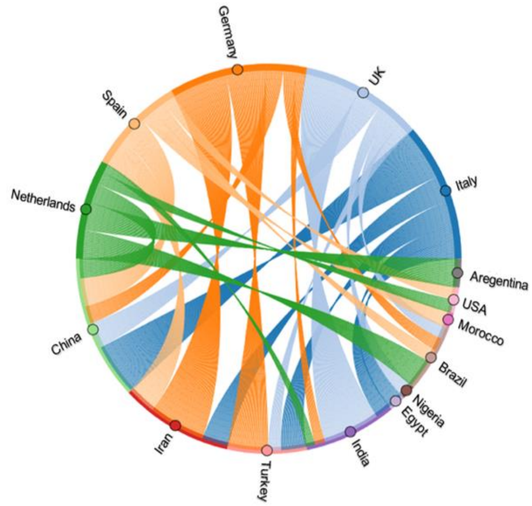

Fig. 7: Chord diagram. Top 5 notifiers and distributor countries with their origin countries.

Fig. 8 shows the relations between destination and notifier countries, and some cases can be highlighted. The strongest connection is from Spain to Italy. This is followed by a group with connections of almost the same level: Germany-the Netherlands, Germany-Turkey, Italy-France, or Germany-China.

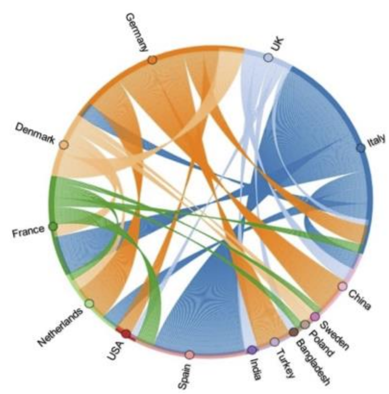

Fig. 8: Chord diagram. Top 5 notifiers and destination countries with their origin countries.

4.2 Statistical analysis of issued products

The following analysis focuses on which products are most frequently giving problems and which hazards are most common. This information is useful, as organizations can pay more attention to these products or increase analyses to detect these hazards.

Table 7 shows the most issued products, where the columns correspond to the numbers of instances and their percentages relative to the total amount of records in RASFF. Nuts are the products that generate the most problems. Problems with fruits and vegetables and fish and fish products are also common.

Table 7. Top 5 issued products.

Product	Instances	Percentage (%)
Nuts, nut products, and seeds	10,454	18.55
Fruits and vegetables	8,321	14.76
Fish and fish products	5,956	10.56
Meat and meat products (other than poultry)	3,067	5.44
Food contact materials	2,910	5.16

Table 8 shows the 5 hazards that are found in the issued products. The fact that mycotoxins (toxic metabolites produced by the fungi) are the most reported hazard is as expected since Table 7 shows that the most reported products are nuts and fruits and this toxin is a contaminant that could be found especially in these products, (Smith et al., 2016). Pathogenic microorganisms also appear to be the cause of many registered issues, as these bacteria can be found in every kind of food, (Pigłowski, 2019), in particular, the most found are salmonella and listeria monocytogenes. Pesticides create problems because some countries ban them while others do not, (Carvalho, 2006), so they can be used in products in one country that exports them to another country where they are forbidden. Considering (Pigłowski, 2018), heavy metals are related to fish and food contact materials, which are also reflected in Table 7, where fish appears among the most issued products. Finally, microbial contaminants (non-intended or accidental introduction of infectious material) have been reported in different products, such as processed foods and raw agricultural products (Jha, 2015).

Table 8. Top 5 identified hazards.

Hazard	Instances	Percentage (%)
Mycotoxins	12,068	21.41
Pathogenic microorganisms	9,629	17.08
Pesticide residues	6,459	11.46
Metals	4,666	8.28
Microbial contaminants (other)	4,490	7.96

Finally, to better understand the relation between products and their hazards, Fig. 9 is provided. The graph depicts a bubble diagram of the top products and hazards registered in the RASFF. To compile this information, all combinations and their frequencies were obtained. Then, on the X-axis we represented the 10 products with more instances, while the Y-axis lists the hazards that have been found more times in these products. The size of the bubbles corresponds to the frequency of combinations, and each colour is univocal. As is shown, three issues stand out from the rest. Nut products and seeds are contaminated by mycotoxins, which supports Tables 7 and 8, as they are in the leading positions. Additionally, fruits and vegetables where pesticides have been found and meat products with pathogenic microorganisms are also identified as expected as they are in high positions in these tables.

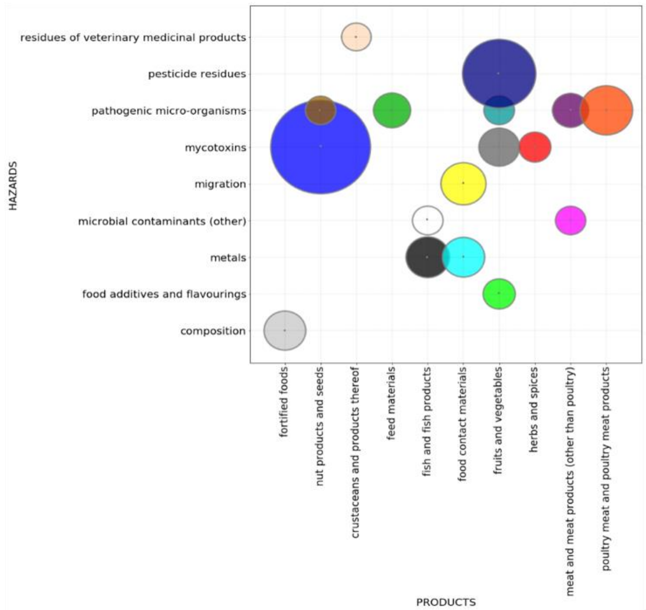

Fig. 9: Bubble diagram relating issued products to their hazards.

4.3 RASFF graph NA

The dataset was used to build a graph considering the different roles of the countries in the trading chain. A representation of the graph with RASFF information was obtained using the Basemap Matplotlib Toolkit, (Matplotlib, 2022), which allows plotting 2D data in maps (Fig. 10). In the figure, the size of a node represents the total number of issues that concern a country, and the thickness of an edge denotes the total number of issues involving two countries. The colour of a node represents a country's role as an importing (blue) and exporting country (red). Blue means that it has more edges entering the node and red that it has more edges leaving the node. Green nodes represent countries with a mixed role. The figure shows that China is the leading exporter of issued products, Spain and Germany as importers finding hazards, and France and Poland are balanced countries in terms of finding contaminated imports and distributing or exporting hazardous food.

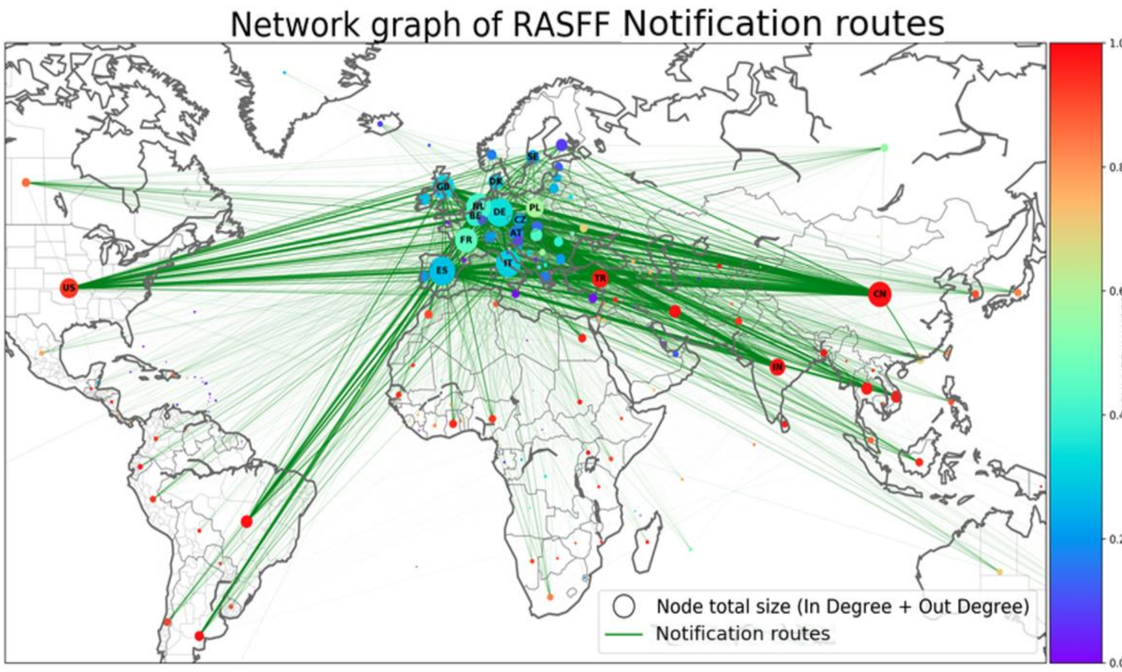

Fig. 10: Graph with all RASFF issues represented on a global map.

Fig. 11 shows a subgraph of Fig. 10 for Europe since the RASFF data come from EU countries. As shown in Fig. 11, Spain, the United Kingdom, Germany, and Italy receive the products with more issues in RASFF. These trends are consistent with the results obtained in the statistical study previously described in subsection 4.1.

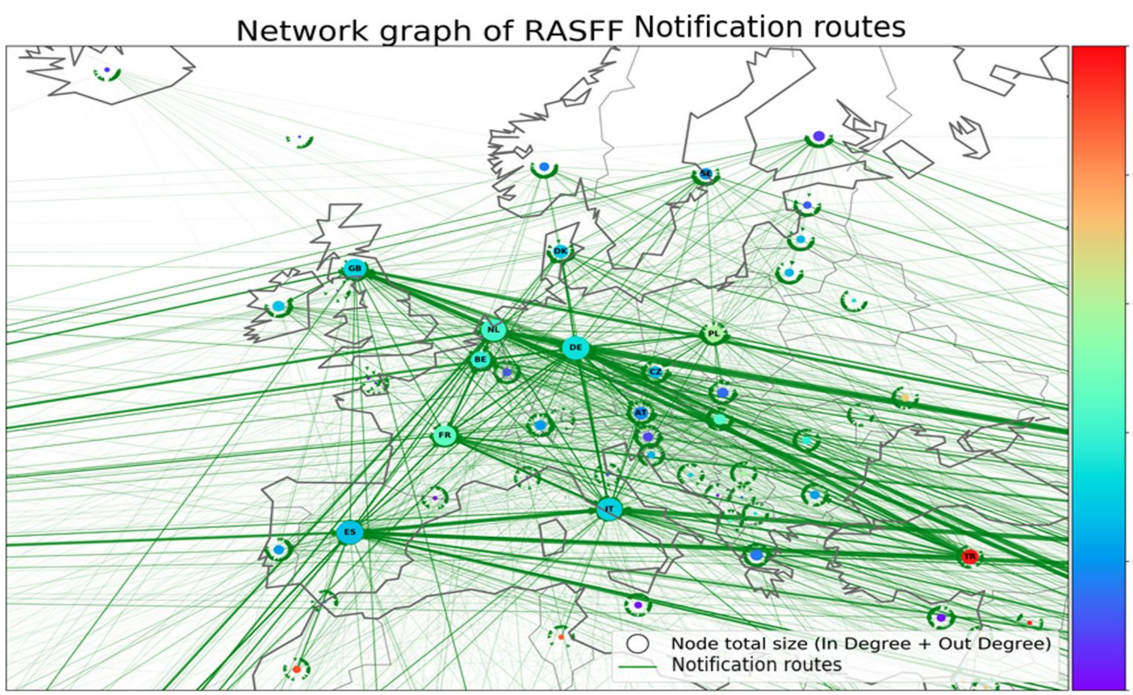

Fig. 11: Subgraph of Fig. 10 for EU countries.

4.3.1 General NA metrics

Table 9 shows the metrics for the RASFF graph depicted in Fig. 10. For all metrics, not only the global values are computed (considering the full RASFF graph) but also the metrics for each issue subcategory (alert, border rejection, and information). Alert represents a serious health risk on the market taking rapid action. Border rejection occurs when the product has been rejected outside the EU. Information is like an alert but without taking rapid action. The number of nodes corresponding to countries in the global network is 222. However, the number of countries in the world is lower. This mismatch is explained by two factors: territories such as Monaco and former countries such as the former Yugoslavia can also be found in the historical graph.

Table 9. General metrics.

Metrics	Total	Alerts	Border rejection	Information
Number of nodes	222	215	164	201
Number of edges	7,092	5,932	1,260	4,010
Density	0.14	0.13	0.05	0.10
Type of network	Heterogeneous	Heterogeneous	Heterogeneous	Heterogeneous

The number of edges that correspond to the total number of product movements is 7,092, and the RASFF graph density value is 0.14.

To study the heterogeneity of the RASFF network, the distribution of node degrees versus the number of nodes with these degrees was analysed. Fig. 12 shows that the graph seems to follow a power law. In this plot, a small number of instances are clustered in the left part, denoting a heterogeneous graph.

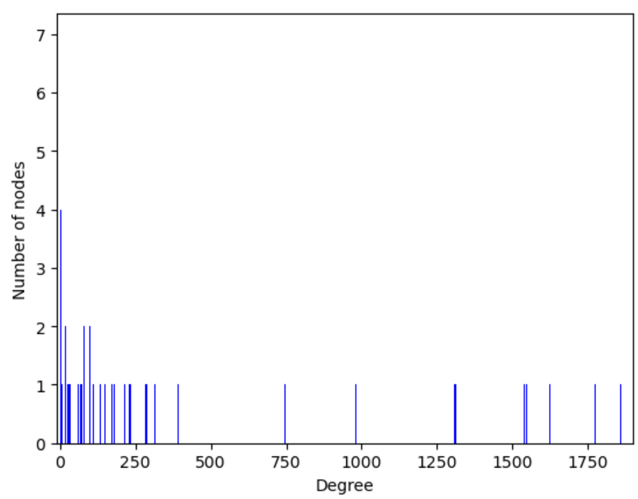

Fig. 12: Distribution of node degrees and their frequencies.

Confirmation was achieved using the power law Python library implemented in (Alstott & Bullmore, 2014). Using this library, the values for x_{min} and α *were calculated.* The obtained values were measured as 79 and 1.45, respectively. Since α is greater than 1, it can be concluded that the distribution follows a power law. Fig. 13 shows how the distribution fits the power law, showing that the RASFF graph presents a heterogeneous network.

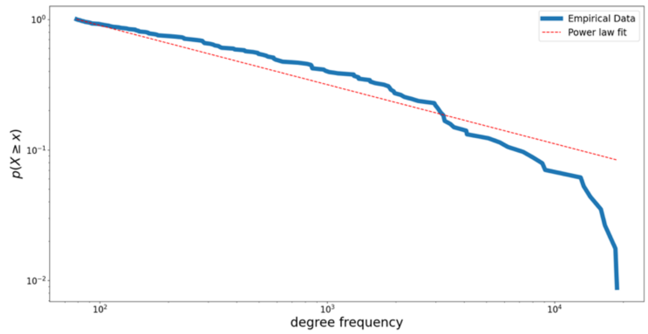

Fig. 13: Degree distribution fitting the power law.

4.3.2 Connectivity

Connectivity was also studied. Table 10 compiles these metrics, which can be used to understand how the different countries are connected due to the registered issues. The RASFF graph is a weakly connected graph with a single connected component that contains all 222 nodes. As the graph is not strongly connected, the diameter does not apply. The same results are found for each subcategory.

Table 10. Connectivity metrics.

Metrics	Global	Alerts	Border Rejection	Information
Strongly connected	False	False	False	False
Weakly connected	True	True	True	True
Number of WCCs	1	1	1	1
Diameter	Does not apply	Does not apply	Does not apply	Does not apply
Average clustering coefficient	0.78	0.79	0.24	0.77

However, the clustering coefficient should be analysed in depth to understand the role of a country as a possible spread vector for possible issues. The rationale behind these results is that there are countries with a clustering coefficient of 0 and others with a value of 1. Lower coefficients point to countries that play more critical roles in the flow produced by the alerts as they are the main connection with their neighbours who, in turn, tend to be isolated from each other. We study in detail the 5 countries with the highest and lowest clustering degrees excluding 0 and 1 in Tables 11 and 12, respectively.

Table 11 shows countries with a low clustering and small nonzero clustering coefficient.

Table 11. Top 5 countries with the lowest clustering coefficients.

Total	Alerts	Border rejection	Information
France (0.20)	France (0.21)	Italy (0.05)	Netherlands (0.21)
United Kingdom (0.20)	Netherlands (0.23)	United Kingdom (0.05)	Spain (0.22)
Netherlands (0.21)	Belgium (0.23)	Spain (0.06)	Italy (0.23)
Spain (0.22)	United Kingdom (0.28)	France (0.07)	France (0.24)
Italy (0.23)	Italy (0.28)	Malta (0.07)	Germany (0.24)

In Table 12, countries have high clustering levels that do not reach 1.

Table 12. Top 5 countries with the highest clustering coefficients

Total	Alerts	Border rejection	Information
French Polynesia (0.99)	French Polynesia (0.99)	Suriname (0.67)	Syria (0.99)
Yemen (0.99)	Burundi (0.99)	Paraguay (0.67)	Madagascar (0.99)
West Bank and Gaza Strip (0.98)	Togo (0.99)	Togo (0.53)	Paraguay (0.99)
Honduras (0.98)	Belize (0.99)	Yemen (0.5)	Ethiopia (0.98)
Burundi (0.98)	Gabon (0.99)	Honduras (0.5)	Sudan (0.98)

4.3.3 Centrality

The last set of metrics is related to centrality. In-degree and out-degree centrality values show how countries behave in terms of how a contaminated product move within the trading chain. In Tables 13a and 13b, the values of the top 5 countries of both metrics are obtained. By studying the in/out-centrality of the nodes, how countries within the network work as a distribution gateway for third-party products or whether they represent a focus of problems as countries of origin can be assessed.

Table 13a. Top 5 countries by in-degree centrality

In-degree centrality	Alerts	Border Rejection	Info.
Spain (0.71)	Netherlands (0.68)	Spain (0.67)	Spain (0.71)
France (0.65)	France (0.66)	United Kingdom (0.57)	Italy (0.675)
Italy (0.65)	Belgium (0.64)	Italy (0.53)	Netherlands (0.66)
United Kingdom (0.64)	Spain (0.63)	France (0.42)	United Kingdom (0.65)
Germany (0.60)	Germany (0.62)	Netherlands (0.40)	Germany (0.64)

Table 13b. Top 5 countries by out-degree centrality

Global	Alerts	Border Rejection	Info.
France (0.72)	France (0.76)	China (0.31)	Netherlands (0.59)
Netherlands (0.72)	Belgium (0.72)	India (0.26)	China (0.52)
Belgium (0.68)	Netherlands (0.69)	Turkey (0.21)	France (0.52)
United Kingdom (0.65)	United Kingdom (0.66)	United States (0.18)	Germany (0.51)
China (0.62)	Italy (0.64)	Egypt (0.18)	Spain (0.48)

In Table 14, the top 5 countries by closeness centrality are shown. The table measures how close a node is to the rest of the nodes in the graph.

Table 14. Top 5 countries by closeness centrality

Global	Alerts	Border Rejection	Info.
Spain (0.73)	Netherlands (0.74)	Spain (0.69)	Spain (0.74)

France (0.68)	France (0.73)	Netherlands (0.57)	Italy (0.72)
Italy (0.68)	Belgium (0.71)	Italy (0.52)	Netherlands (0.71)
United Kingdom (0.68)	Spain (0.71)	United Kingdom (0.52)	United Kingdom (0.70)
Germany (0.65)	Germany (0.70)	Germany (0.52)	Germany (0.70)

The values of betweenness centrality are gathered in Table 15 and are very low (close to zero). This metric shows which nodes behave as connectors in a network; apparently, there are no critical brokers.

Table 15. Top 5 countries by betweenness centrality

Global	Alerts	Border Rejection	Info.
United Kingdom (0.07)	France (0.13)	Turkey (0.15)	Netherlands (0.12)
Belgium (0.06)	Netherlands (0.10)	Portugal (0.12)	Spain (0.10)
France (0.05)	Belgium (0.09)	Morocco (0.12)	United Kingdom (0.09)
Italy (0.05)	United Kingdom (0.07)	China (0.12)	France (0.08)
Netherlands (0.05)	Italy (0.06)	Spain (0.10)	Italy (0.08)

Information on the eigenvector centrality of the top 5 countries was obtained and is shown in Table 16. The measures obtained are close to the average value or lower (eigenvector centrality values of 0 to 1). This metric provides information about how a node influences the rest of the network.

Table 16. Top 5 countries by eigenvector centrality

Global	Alerts	Border Rejection	Info.
Germany (0.41)	Netherlands (0.17)	Spain (0.48)	Italy (0.19)
Italy (0.41)	Spain (0.17)	Bulgaria (0.39)	Spain (0.19)

France (0.28)	France (0.16)	Netherlands (0.36)	Netherlands (0.19)
Netherlands (0.26)	Belgium (0.16)	Poland (0.32)	Germany (0.19)
Spain (0.26)	Germany (0.16)	Germany (0.29)	United Kingdom (0.18)

5 Discussion

From the results compiled in the previous section, we can draw several conclusions.

Data related to the origin countries show the countries that exhibit recurrent problems with their products or more intense export activity compared to others. This indicates that countries should pay attention to the products imported from these areas. It should be stated that most of the hazardous products reported in RASFF come from Asia (China, Turkey, India, and Iran), which is confirmed by (Pigłowski, 2020). However, as most of the issues are categorized as border rejections, this is a problem that seems to be under control because these products have not entered other countries. Germany is the only EU member among the countries having problems with products at origin, and this should be studied in-depth as it could be a consequence of detecting issues of products that come from the same country, so such food policies could be replicated by others. As noted by (Pigłowski, 2017), Germany presents many issues in origin, perhaps due to problems with meat products. As this type of issue could be a problem for the free market of the European territory without the possibility of border rejections, market controls should be reinforced for this product category.

From the data of distribution countries, we can see that most of the issues do not involve a distributor. Additionally, we found a set of European countries (Spain, Italy, Germany, and the United Kingdom) that seems to have good control policies when distributing products, Table 3. It seems that food quality policies with products entering Europe are strong. Spain appears above the rest of the countries in Table 3, probably due to the country's large number of ports and its central role as a hub of the distribution chain. This is confirmed by centrality studies and by (Caldeira, 2021), who describes Spain as the main gateway for fishery products in the EU.

Table 4 interestingly shows that three of the countries are the same as those shown in Table 3 (distribution countries). It could be that these countries import more products or that they are very concerned about foreign products entering their borders.

To summarize our study regarding the countries' role in the trading chain, two main conclusions are drawn. First, most of the origin countries are not in the EU, perhaps due to the different food policies of different countries. Second, Germany, Italy, and the United Kingdom are the leading distributors and destinations, which makes them more active in the trading chain than the rest according to RASFF data.

Regarding notifier countries, these countries notify more than others probably due to their food safety policies being stronger. This makes sense when comparing the results in Table 5 to those in Tables 3 and 4. In this case, Spain, Italy, Germany, and the UK are the top distribution and destination countries, demonstrating their high levels of activity. If this information is complemented with Table 5 (notifier countries), we can consider them good countries finding products with hazards.

From Fig. 6, we can make two possible interpretations. First, countries, in general, may receive more products than they export, creating more possibilities to identify issues or food policies in countries in the EU (most of the countries are member states as shown in Tables 3 and 4), which are more restrictive. Second, countries may be more concerned about the products they import, as food policies in other countries could be different.

An in-depth analysis of the notifier countries and the other roles are shown in Table 6. First, Germany and Italy are among the four top countries. This can be interpreted as suggesting that these countries have good food safety controls. After analyzing each column separately and comparing the information provided in Tables 2, 3, and 4, the following points can be made. Germany appears aware of not exporting issued products, as it is also at the top of Table 2 (origin countries). From a comparison to Table 3 (distribution countries), countries in the second column (notifier and distribution) are concerned with not distributing contaminated products. The same conclusions, related to imported products, can be drawn when comparing the third column (notifier and destination) to Table 4 (destination countries), except in the case of Denmark. The final column (notifier, origin, and destination) lists countries that identify food problems internally, so their food policies should be replicated.

Fig. 7 and Fig. 8 show related countries by trade. In Fig. 8, countries maintain a trade relationship but have different food policies, denoting sensitive routes. As India is a former colony of the United Kingdom, trading occurs regularly between these two countries. (Mukherjee et al., 2019) note that even though India is an interesting supplier for the United Kingdom, limited use of technology in the food safety field presents problems. Therefore, we propose that countries with the possibility of establishing profitable food trading chains should create funding programs to introduce new technologies that can improve food safety. Fig. 9 shows that the route between Spain and Italy should be revised and that Germany has a very restrictive food safety policy or is importing many products. This should be combined with trade balances, enforcement, and controls.

Concerning the results obtained for products and hazards, we can make the following proposals. The appearance of mycotoxins in nuts is difficult to prevent, although recommendations can be made for the entire harvesting cycle. However, as they are directly related to weather patterns, predictive models of regions with similar weather can be implemented and tested. The decrease in mycotoxin issues is something that should be addressed, as it is directly related to the level of trade (Nes & Schaefer, 2018). The appearance of pathogenic microorganisms in meat products was reported by (Piglowski, 2021), who showed a strong correlation between the number of kilograms traded among EU countries and the alerts reported by them. As this happens in a free distribution market, the best way to reduce these issues is to improve the traceability policies for these products. Finally, there are many issues with fruits and vegetables containing pesticides. In this case, new and effective integrated pest management (IPM) processes should be implemented.

Table 9 shows an important figure that has consequences for the rest of the metrics. The number of border rejections is significantly lower than that of other types of notifications. With this information, we can conclude that some products posing high risks for the population are not entering the EU.

Summarizing the results, we can guess that China and Turkey are sensitive countries, as the products imported from these countries create many issues, so special effort should be put into analyzing products coming from there. European countries' policies seem to be good and consistent, as these countries report many issues when they act as distributors. Most of these issues have no destination registered as they have been found before reaching the expected destination. In this sense, Italy, Germany, the United Kingdom, and Spain report more cases and could be considered safe countries in this way. From the analysis of the roles played by different countries (origin, distribution, or destination) when reporting a problem (notifier), we can conclude which European countries are more stringent with food controls. In particular, Germany and Italy always lead in this regard. Regarding hazards and products, nuts, nut products, and seeds contaminated with mycotoxins and fruits and vegetables contaminated with pesticides are the most reported.

In terms of country management, we recommend the following. The EU should establish agreements with big exporters like China. Due to the high level of alert notifications controls at the border should be increased. In this way, based on historical data resources should be focused on analyzing particular products coming from particular countries.

A density value, Table 9, of close to zero could be interpreted that in the case of a food crisis, the product has not been distributed to many countries. This could be due to the prompt intervention of the mechanisms to analyze the products. In other words, trading between all countries' networks is slow so it is possible to have good responses on time.

The heterogeneity of the network in Table 9 implies that there are a few countries that dominate the network as they have many connections with the rest. Countries such as Spain, France, Italy, the UK, and Germany should be studied in more detail, as they either create more problems or have better food policies that should be replicated.

The average clustering coefficient from Table 10 denotes strong interconnections among certain origin, distribution, and destination countries. These areas are relevant as particular parts of the network could become quickly affected by an issue originating in one of them due to their cohesion.

In this research and concerning RASFF data we find different groups of countries with clustering coefficients of which two of them have values of 0 and 1. The first group includes non-EU countries that do not make many exports, and their destination countries do not trade between them. For the second group, we find countries that seem to not make many exports, but in this case, their distribution or destination countries are highly connected between them. Apart from these groups, we have studied countries with high and low coefficients different from 0 and 1. Countries with low clustering coefficients, compiled in Table 11, trade with many countries, and the possibilities of them being connected are fewer. Finally, there are countries with high clustering coefficients in Table 12. As the network is very large, countries that do not export heavily (no matter the reason) have fewer trade partners, increasing the likelihood that at a particular moment all trading links between them can be established.

From the analysis of the clustering coefficient, countries can be classified into four groups. The first group has a clustering coefficient of zero. This group seems to include countries with few exports whose distribution or destination countries are not connected. In this case, a food issue will not spread to other countries. The second group includes countries that trade heavily but whose neighbours are not highly connected. In this case, the problems are the same as those of the previous use case. The third group comprises countries with high degrees of coefficient clustering that seem to be small or to make few exports. In these cases, food issues are not common but could be spread to a few countries. Finally, countries with a clustering coefficient of 1 are countries that are not much into exports but whose distribution or destination countries are very connected.

Regarding in-degree centrality, EU countries such as Spain appear as difficult-to-save ports in terms of food policy, reflecting the strict character of their safety standards. Regarding out-degree centrality in Table 13a, the EU countries that appear are those with more relaxed regulations on exporting products. It should be noted that the UK and France, in turn, also impose high standards on the products they receive, possibly reflecting an imbalance in their food safety standards. In the case of out-degree centrality from Table 13b, China is the non-European country with more problems with its exports, so Europe probably applies stricter control over its products. It should also be noted that all the countries listed in the border rejection column are non-European countries.

From the closeness centrality results in Table 14, the identified countries (Spain, Germany, Italy, the United Kingdom, or the Netherlands) should be considered major sources of the rapid spread of contaminated products. This means that these countries should implement strong food policies, as they represent sensitive points in the network.

Regarding betweenness centrality found in Table 15, it can be concluded that none of the countries acts this way in the graph. This means that there are no countries whose removal could cause a disconnection of the network. In terms of issued products, there are no main distributors in the chain.

Table 16 shows eigenvector centrality, the most important countries play this role. However, as the values are less than 0.5, such countries are important but do not significantly influence the rest of the graph. The RASFF network shows that these sensitive routes are not critical.

Two other separate analyses were also conducted:

- Distinguishing between the three main issues: notification, alerts, and border rejection.
- Analyzing each decade separately.

In the first case, the results for each issue are consistent with the global ones; from some metrics, relative differences are found (e.g., some countries seldom appear among the top countries or change in relative position), but most of the same EU countries monopolize the network measures (the table considering the clustering coefficient shows the only exception). Other isolated cases could be highlighted: Belgium always leads in terms of alerts, and there are some exceptions, such as Bulgaria, Malta, or Morocco. These cases can be studied in future work. In the second case, the only relevant observation (though arguably obvious) is that the number of alerts grows decade by decade, which reflects an increase in the adoption of the RASFF alert system by trade countries and the number of network members.

6 Conclusions

This paper presents a statistical and structural analysis of historical information on food issues previously obtained by scraping the online RASFF portal. First, a statistical analysis of raw data was carried out to obtain significant metrics shown in tables and graphics. We then constructed a graph from the most relevant characteristics of each record: origin country, distribution country, destination country, product, and hazard. The result is a weighted DAG where nodes represent countries and edges describe trade flows and issued products plus hazards. The directions of edges denote the roles of countries and how products are transported (origin, distribution, and destination country). Weights are computed by summing the total number of times an issue involving the same pair of countries occurs. Using this graph, we can characterize the chain trade formed by the countries involved, the issued products, and their hazards by applying NA techniques. The result is several descriptive metrics of the graph, including general, connectivity, and centrality metrics. The dataset used in the work is also a contribution as it is the first time the complete RASFF Portal information is freely available. Although the data stopped being scrapped during 2020, the actual version of the RASFF Portal only allows access for a period that starts in 2019.

In future work, an in-depth analysis of the countries highlighted in this study seems necessary. Future work should explore the specific reasons why a country stands out, whether due to importing more products or adopting safe food policies plus enforcement and controls. The behaviour of trading products should also be studied as a time series. An in-depth study of seasonality could then be carried out via time series analysis to identify seasonal products that are related to the month of the year and the hemisphere in which they are produced. This would provide information on how countries' behaviour changes depending on the season and the products that are traded at a given time. For instance, issues with figs mostly originate in Turkey in September. An isolated study of networks formed by countries and their trade will be useful in obtaining metrics such as clustering coefficients to measure the risk of propagating contaminated products. Another interesting study would consist of comparing results to the global trade network so we could leverage the results based on proportionality. This will let us confirm which recommendations from this paper are nearer to reality. This would allow us to better understand the conclusions drawn in this paper, as we could compare the number of issues reported to the total amount of exports and imports. Additionally, a socioeconomic study of the relations obtained in this paper could be used to propose solutions to the origins of some problems. Finally, we can benefit from using graph convolutional neural networks, an artificial intelligence technique used in the field of deep learning that is currently one of the most effective techniques for managing a large amount of data.

Conflicts of interest

The authors have no conflicts of interest to declare.

References

[1] Alshannaq, A., & Yu, J. H. (2021). Analysis of EU Rapid Alert System (RASFF) Notifications for Aflatoxins in Exported US Food and Feed Products for 2010–2019. *Toxins*, *13*(2), 90.

[2] Alstott, J., & Bullmore, D. P. (2014). power-law: a Python package for analysis of heavy-tailed distributions. PloS one, 9(1).

[3] Amico, P. D., Nucera, D., Guardone, L., Mariotti, M., Nuvoloni, R., & Armani, A. (2018). Seafood products notifications in the EU Rapid Alert System for Food and Feed (RASFF) database: Data analysis during the period 2011–2015. *Food Control*, *93*, 241-250.

[4] Barnes, J. A., & Harary, F. (1983). Graph theory in network analysis. *Social networks*, *5*(2), 235-244.

[5] Bavelas, A. (1950). Communication patterns in task-oriented groups. *The journal of the acoustical society of America*, *22*(6), 725-730.

[6] Bihari, A., & Pandia, M. K. (2015, February). Eigenvector centrality and its application in research professionals' relationship network. In *2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE)* (pp. 510-514). IEEE

[7] Bonacich, P. (1972). Factoring and weighting approach to status scores and clique identification. *Journal of mathematical sociology*, *2*(1), 113-120.

[8] Bouzembrak, Y., Camenzuli, L., Janssen, E., & Van der Fels-Klerx, H. J. (2018). Application of Bayesian Networks in the development of herbs and spices sampling monitoring system. *Food Control*, *83*, 38-44.

[9] Bui-Klimke, T. R., Guclu, H., Kensler, T. W., Yuan, J. M., & Wu, F. (2014). Aflatoxin regulations and global pistachio trade: insights from social network analysis. *PloS one*, *9*(3), e92149.

[10] Caldeira, A. J. R., Alves, C. P. P., & Santos, M. J. (2021). Anisakis notification in fish: an assessment of the cases reported in the European Union Rapid Alert System for Food and Feed (RASFF) database. *Food Control*, *124*, 107913.

[11] Carvalho, F. P. (2006). Agriculture, pesticides, food security and food safety. *Environmental science & policy*, *9*(7-8), 685-692.

[12] Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. SIAM review, 51(4), 661-703.

[13] Czepielewska, E., Makarewicz-Wujec, M., Rozewski, F., Wojtasik, E., & Kozłowska-Wojciechowska, M. (2018). Drug adulteration of food supplements: A threat to public health in the European Union?. *Regulatory Toxicology and Pharmacology*, *97*, 98-102.

[14] Cınar, S., Yılmaz, S. N., Aydın, E., & Yorulmaz, A. (2017). Rapid Alert System for Food and Feed (RASFF) 2009-2016 Turkey Report. *Turkish Journal of Agriculture-Food Science and Technology*, *5*(8), 873-882.

[15] Dada, A. C., Somorin, Y. M., Ateba, C. N., Onyeaka, H., Anyogu, A., Kasan, N. A., & Odeyemi, O. A. (2021). Microbiological hazards associated with food products imported from the Asia-Pacific region based on analysis of the rapid alert system for food and feed (RASFF) notifications. *Food Control*, *129*, 108243.

[16] De Leo, F., Coluccia, B., Miglietta, P. P., & Serio, F. (2021). Food contact materials recalls and international trade relations: An analysis of the nexus between RASFF notifications and product origin. *Food Control*, *120*, 107518.

[17] Djekic, I., Jankovic, D., & Rajkovic, A. (2017). Analysis of foreign bodies present in European food using data from Rapid Alert System for Food and Feed (RASFF). *Food Control*, *79*, 143-149.

[18] Duan, M., Li, Q., Liu, W., Hao, J., Liu, P., Han, B., & Wang, Y. (2017). Analysis on the safety of metallic food contact materials and products in China based on RASFF. *Journal of Food Safety and Quality*, *8*(12), 4865-4869.

[19] Fair, K. R., Bauch, C. T., & Anand, M. (2017). Dynamics of the global wheat trade network and resilience to shocks. *Scientific reports*, *7*(1), 7177.

[20] Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, *1*(3), 215-239.

[21] Goldberg, A. V. (1984). *Finding a maximum density subgraph*. Berkeley: University of California.

[22] Holland, P. W., & Leinhardt, S. (1971). Transitivity in structural models of small groups. *Comparative group studies*, *2*(2), 107-124.

[23] Jamali, M., & Abolhassani, H. (2006, December). Different aspects of social network analysis. In 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2006 Main Conference Proceedings) (WI'06) (pp. 66-72). IEEE.

[24] Jha, S. N. (2015). Rapid Detection of Food Adulterants and Contaminants: Theory and Practice. Academic Press.

[25] Jongwanich, J. (2009). The impact of food safety standards on processed food exports from developing countries. *Food Policy*, *34*(5), 447-457.

[26] Jozwiak, A., Milkovics, M., & Lakner, Z. (2016). A network-science support system for food chain safety: a case from Hungarian cattle production. *International Food and Agribusiness Management Review*, *19*(1030-2016-83145), 1-26.

[27] Kępinska-Pacelik, J., & Biel, W. (2021). Microbiological Hazards in Dry Dog Chews and Feeds. *Animals*, *11*(3), 631.

[28] Kononiuk, A. D., & Karwowska, M. (2017). Meat and meat products–analysis of the most common threats in the years 2011-2015 in Rapid Alert System for Food and Feed (RASFF). *Roczniki Państwowego Zakładu Higieny*, *68*(3).
[29] Kuchheuser, P., & Birringer, M. (2022). Pesticide residues in food in the European Union: Analysis of notifications in the European Rapid Alert System for Food and Feed from 2002 to 2020. *Food Control*, *133*, 108575.
[30] Lüth, S., Boone, I., Kleta, S., & Al Dahouk, S. (2019). Analysis of RASFF notifications on food products contaminated with Listeria monocytogenes reveals options for improvement in the rapid alert system for food and feed. *Food Control*, *96*, 479-487.
[31] NetworkX. (2022). *Matplotlib basemap toolkit*. Retrieved from https://matplotlib.org/basemap/. Accessed July 12, 2022
[32] Nogales, A., & Garcia-Tejedor, Alvaro J. (2022). Food and feed health risk notifications in the European Union: a historical dataset of the issues registered in the Rapid Alert System for Food and Feed portal., Mendeley Data, V4, doi: 10.17632/yxkm4gs7zf.4
[33] Mincer, M., & Niewiadomska-Szynkiewicz, E. (2012). Application of social network analysis to the investigation of interpersonal connections. *Journal of Telecommunications and information technology*, 83-91.
[34] Mukherjee, A., Satija, D., Sinha, S., & Sarma, A. P. (2019). Food Imports in India: Prospects, Issues and Way Forward. *Journal of Economic Science Research*, *2*(3).
[35] Mutukumira, A., & Jukes, D. J. (2003, May). The Development of National Food Safety Control Systems in Sub-Saharan Africa-Issues and Opportunities. In *Food Africa, Congress centre/palais de congress, Yaoundé, Cameroon* (pp. 5-9).
[36] Naughton, D. P., Nepusz, T., & Petroczi, A. (2015). Network analysis: A promising tool for food safety. *Current Opinion in Food Science*, *6*, 44-48.
[37] Nes, K., & Schaefer, K. A. (2018). *Hazard Mitigation in EU Food Trade* (No. 2229-2019-1917).
, and C etc.
[38] NetworkX. (2022). *Network analysis in Python*. Retrieved from https://networkx.org/. Accessed May 11, 2022
[39] Pal, M., Gerbaba,T., Kumar, A. & Shukla,P.K. (2015).Impact of food safety on One Health. Beverage and Food World 42.
[40] Papapanagiotou, E. P. (2017). Foodborne norovirus state of affairs in the EU rapid alert system for food and feed. *Veterinary sciences*, *4*(4), 61.
[41] Papapanagiotou, E. P. (2021). Serious Notifications on Food Contact Materials in the EU RASFF. *Veterinary Sciences*, *8*(4), 56.
[42] Pennone, V., Lehardy, A., Coffey, A., Mcauliffe, O., & Jordan, K. (2018). Diversity of Listeria monocytogenes strains isolated from Agaricus bisporus mushroom production. *Journal of applied microbiology*, *125*(2), 586-595.
[43] Petroczi, A., Taylor, G., Nepusz, T., & Naughton, D. P. (2010). Gate keepers of EU food safety: Four states lead on notification patterns and effectiveness. *Food and Chemical Toxicology*, *48*(7), 1957-1964.
[44] Petroczi, A., Nepusz, T., Taylor, G., & Naughton, D. (2011). Network analysis of the RASFF database: a mycotoxin perspective. World Mycotoxin Journal, 4(3), 329-338.
[45] Pigłowski, M. (2017). Product categories and hazard categories in the RASFF notifications: Dependences between chosen variables. *Quality Assurance and Safety of Crops & Foods*, *9*(3), 335-344.
[46] Pigłowski, M. (2017). Notifications to food from European Union countries in the rasff. *Studia i Prace WNEiZ US*, (47/3), 345-358.
[47] Pigłowski, M. (2018). Heavy Metals in notifications of rapid alert system for food and feed. *International journal of environmental research and public health*, *15*(2), 365.
[48] Pigłowski, M. (2019). Pathogenic and Non-Pathogenic Microorganisms in the Rapid Alert System for Food and Feed. *International journal of environmental research and public health*, *16*(3), 477.
[49] Pigłowski, M. (2020). Food hazards on the European Union market: The data analysis of the Rapid Alert System for Food and Feed. *Food science & nutrition*, *8*(3), 1603-1627.
[50] Pigłowski, M. (2021). The Intra-European Union Food Trade with the Relation to the Notifications in the Rapid Alert System for Food and Feed. *International journal of environmental research and public health*, *18*(4), 1623.
[51] Popp, J., Kiss, A., Olah, J., Mate, D., Bai, A., & Lakner, Z. (2018). Network analysis for the improvement of food safety in the international honey trade. *Amfiteatru Economic*, *20*(47), 84-98.
[52] About RASFF. (2022). *About the RASFF Portal*. Retrieved from https://food.ec.europa.eu/safety/rasff-food-and-feed-safety-alerts_en. Accessed September 06, 2022

[53] RASSF Scrapper. (2022). *About the RASFF Portal*. Retrieved from https://github.com/ufvceiec/RASFFportal_scrapping. Accessed July 12, 2022
[54] Robson, K., Dean, M., Haughey, S. A., & Elliott, C. T. (2021). The identification of beef crimes and the creation of a bespoke beef crimes risk assessment tool. *Food Control*, *126*, 107980.
[55] Saurkar, A.V., Pathare, K.G., & Gode, S.A. (2018). An Overview On Web Scraping Techniques And Tools, International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRSCE), Volume 4 Issue 4, PP: 363 – 367.
[56] Schaeffer, S. E. (2007). Graph clustering. *Computer science review*, *1*(1), 27-64.
[57] Smith, M. C., Madec, S., Coton, E., & Hymery, N. (2016). Natural co-occurrence of mycotoxins in foods and feeds and their in vitro combined toxicological effects. *Toxins*, *8*(4), 94.
[58] Somorin, Y. M., Odeyemi, O. A., & Ateba, C. N. (2021). Salmonella is the most common foodborne pathogen in African food exports to the European Union: Analysis of the Rapid Alert System for Food and Feed (1999–2019). *Food Control*, *123*, 107849.
[59] Timmis, H. (2017). The Impact of Standards on Developing Country Exports.
[60] Tudela-Marco, L., Garcia-Alvarez-Coque, J. M., & Martí-Selva, L. (2017). Do EU member states apply food standards uniformly? A look at fruit and vegetable safety notifications. *JCMS: Journal of Common Market Studies*, *55*(2), 387-405.
[61] Van Asselt, E. D., Banach, J. L., & Van Der Fels-Klerx, H. J. (2018). Prioritization of chemical hazards in spices and herbs for European monitoring programs. *Food Control*, *83*, 7-17.
[62] Wang, L. (2018). Design Optimization of Food Safety Monitoring System with Social Network Analysis. IEEE Transactions on Computational Social Systems, 5(3), 676-686.
[63] Wu, F., & Guclu, H. (2013). Global maize trade and food security: Implications from a social network model. *Risk analysis*, *33*(12), 2168-2178.
[64] West, D. B. (1996). *Introduction to graph theory* (Vol. 2). Upper Saddle River, NJ: Prentice hall.
[65] Xu, J., Zhao, V. P., Shan, L., & Huang, C. X. (2014). A framework for developing social networks enabling systems to enhance the transparency and visibility of cross-border food supply chains. *GSTF Journal on Computing (JoC)*.