Allison Beatty (yxn127)
Ian Scarff (iie728)
DA6223 - Data Analytics Tools and Techniques
December 10, 2019

Animal Disease Outbreak Analysis

INTRODUCTION

For this report, we analyzed various disease outbreaks in animals through time. This is a topic of importance because these outbreaks pose a health, agricultural, and economic risk for the areas affected. The data used for this analysis comes from the Emergency Prevention System (EMPRES) Global Animal Disease Information System, which is maintained by the Food and Agriculture Organization of the United Nations. This dataset was found on the website Kaggle.com under the page *EMPRES Global Animal Disease Surveillance*. This report will examine where, when, and what types of diseases affect different animals, discuss the real-world impacts of these outbreaks, and provide recommendations to different groups on what they can do to prevent these outbreaks.

DATA & DATA PREPARATION

This dataset contains approximately 17,000 observations, where each row is an incident of a disease outbreak in animals. Some of the information provided in the data includes the location of the outbreak, the date in which it occurred, the specific disease that was present, the species of the animals that were affected, and how many animals were infected. A full list of the variables used in this analysis and a data dictionary is provided in the Appendix. In addition to these variables, other variables were created during the data preparation phase of this analysis. These variables include the month and year in which the outbreak occurred, the season (based on latitude) in which the outbreak occurred, indicator variables for whether the animal was domestic, wild, or captive, and what were the main animals affect. The main animals were determined by the frequency their species description appears in the data. Based on this, the main animals were summarized into the following categories: 1) “Cattle,” 2) “Sheep,” 3) “Pig/Swine,” 4) “Wild Boar,” 5) “Chicken,” 6) “Duck,” and 7) “Unknown Bird.” The remaining species descriptors were placed into a category called “Other.”

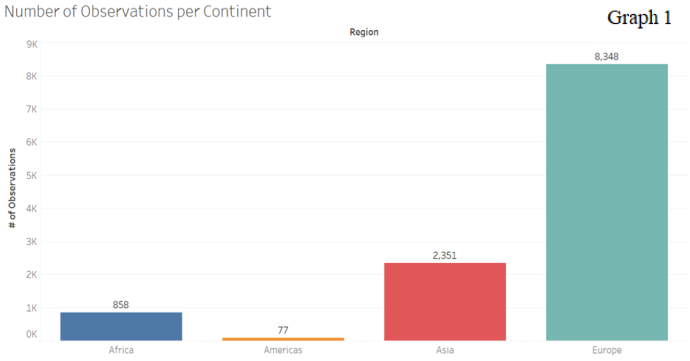
FOCUS OF ANALYSIS

For this analysis, we examined outbreaks from 2015 to 2017 where we knew the exact location of the outbreak and the incident had a confirmed status. Using this data, we aimed to answer the following questions:

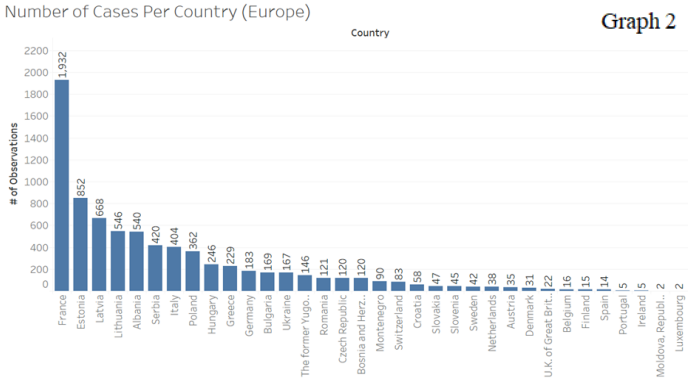
1. What are the most prevalent animal diseases?
2. What countries are most at risk for these types of diseases?
3. Are these diseases harmful to humans?
4. Which animal (domestic, wild, or captive) poses the greatest risk of spreading these diseases?

ANALYSIS

The first step in the analysis was to see where these outbreaks in the data come from. Graph A in the Appendix shows a map of the countries listed in the data. Graph 1 expands on this further by examining how many observations by continent. From this bar graph we can see that a large majority of observations come from Europe. Because of this, the analysis was split into two parts. The first part takes a focused look at disease outbreaks in Europe, while the second part examines certain diseases from a global perspective.



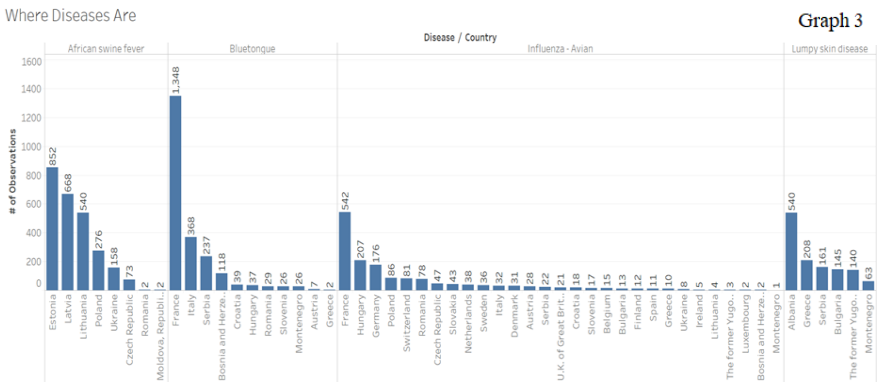
Graph 1



Graph 2

Continuing with the first part of the analysis, we next examined the number of observations in each European country, as shown in Graph 2. However, Russia was excluded from this exploration because in the data it is classified as a part of Europe, but in reality, it is also a part of Asia. This way we would not include any diseases that possibly only exist in Asia into the analysis of Europe. In Graph 2 we can see that the countries with the highest number of outbreaks are France and the Baltic States (Estonia, Latvia, Lithuania). The next geographical area that has a large amount of observations is Eastern Europe (Hungary, Ukraine, Serbia, etc.). The geographic area with the least number of outbreaks is Central and Northern Europe (Germany, Sweden, Austria, etc.).

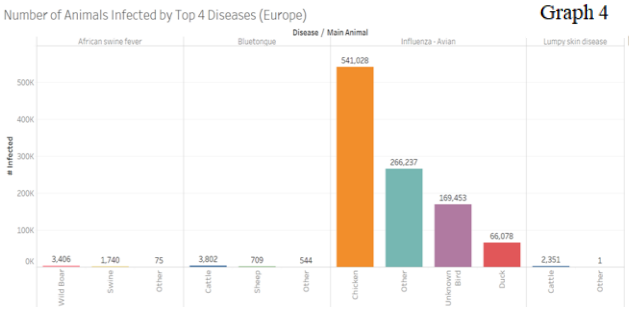
Next, we examined are the most prevalent diseases in Europe. Graph B in the Appendix shows that the top four diseases in terms of observations are African Swine Fever, Bluetongue, Avian Influenza, and Lumpy Skin Disease. Because the frequency of these diseases in the data set are greater than all the other disease types combined, our analysis was further focused on these four diseases. Graph 3 shows how these diseases are distributed across each European country. We see that a majority of Bluetongue is located in France, a majority of African Swine Fever is located in the Baltic State, Lumpy Skin Disease is concentrated in Southeastern Europe, and Avian Influenza is spread out all across Europe.



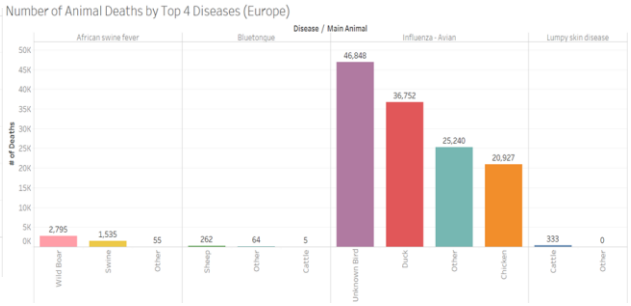
Graph 3

Next, we examined how each of the top four disease affects different animals. Graph 4 shows the total number of animals infected animals for each disease by animal type. Graph 5 shows the total number of animal deaths for each disease by each animal. We can see that certain diseases only affect certain animals. African Swine Fever only infects species of swine and wild boar, Bluetongue mainly affects cattle and sheep, Avian Influenza only affects birds, and Lumpy Skin Disease also affect cattle. Based on these graphs, Avian Influenza by far is the most contagious and the deadliest when compared to the other diseases. It has the highest number infected cases and the highest number of deaths.

Graph 4



Graph 5

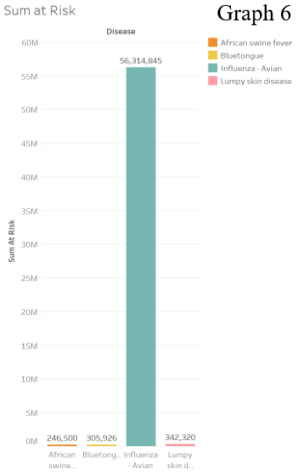


Finally, we combined all of our findings into two interactive maps in Tableau. Both maps can be found in the file named "Disease_Ian.twb." The first map, called *Top 4 Diseases Through Time (Europe)*, shows the exact location (by latitude and longitude) of each observation in the data, played through time. The diseases are color coded to distinguish them. Playing through time, we see at certain points there is a widespread outbreak of African Swine Fever in the Baltic States, a large concentration of Bluetongue in the middle of France and Northern Italy, and a concentration of Lumpy Skin Disease spread across Southeastern Europe. Avian Influenza appears all throughout the map, with a large concentration in Southwestern France.

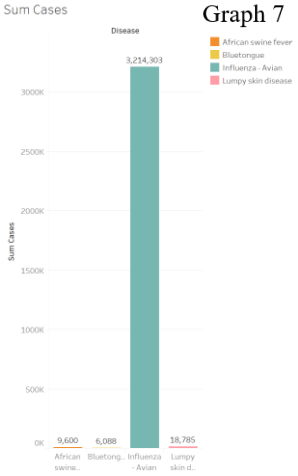
The second map in the file is called *Top Animals/Diseases Through Time (Europe)*. This map expands upon the first map by now color coding by animal types and coding the diseases by shapes. Playing through time, we can see that 1) the African Swine Fever outbreak in the Baltic States mainly infected species of swine, 2) the Bluetongue outbreak in France and Italy and the Lumpy Skin Disease outbreak in Southeastern Europe mainly infected cattle, and 3) the Avian Influenza outbreak in Southwestern France mainly infected chickens and ducks. The reason why we see these patterns in discussed later in this report.

Moving to the second part of the analysis, we wanted to take a more generalized view of these diseases. When looking at Graphs 6, 7, and 8 we can see that the sum at risk, sum of cases, and sum of deaths is significantly

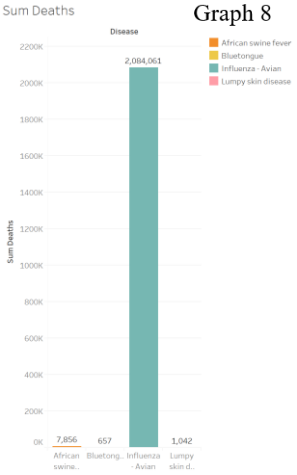
Graph 6



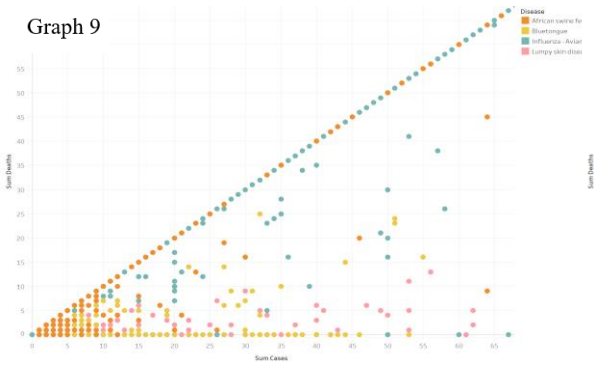
Graph 7



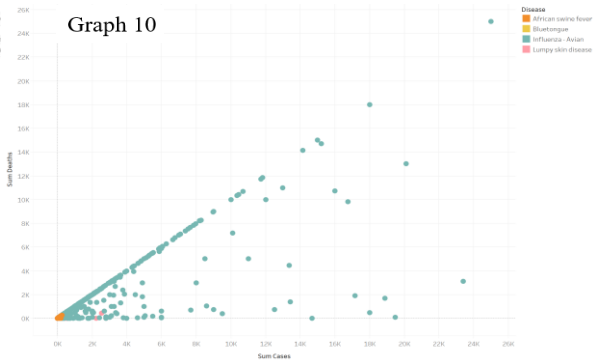
Graph 8

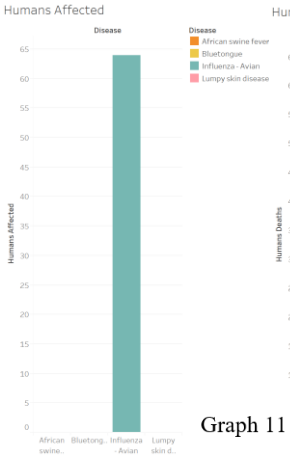


Graph 9

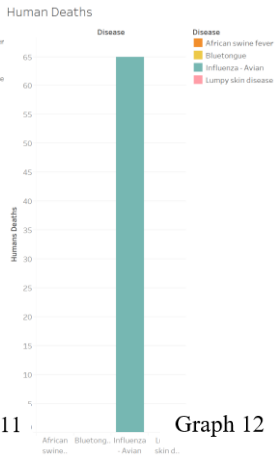


Graph 10





Graph 11



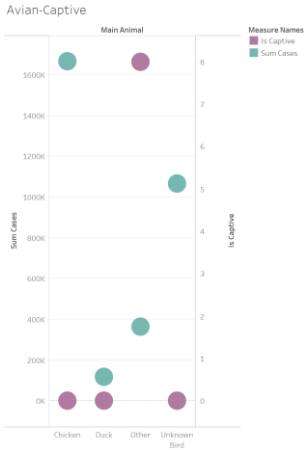
Graph 12

greater for Avian Influenza. According to Graphs 9 and 10, we can also see that for most outbreaks, there are minimal animal deaths; however, a few of the outbreaks have resulted in significant loss of animal life - most of those being from Avian Influenza.

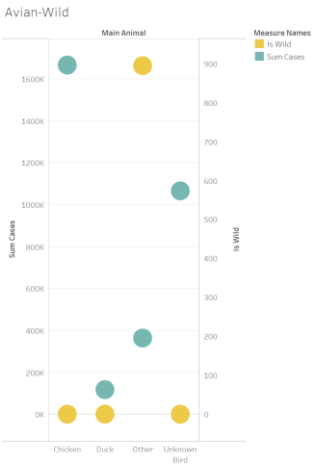
We then wanted to see which of these four diseases, if any, are harmful to humans. According to Graphs 11 and 12, we can see that only Avian Influenza is harmful to humans. It both affects and kills humans. After doing more research into these diseases, we learned that some strains of Avian Influenza are transferable to humans through direct contact with the infected animal (dead or alive). We also learned that Bluetongue, African Swine Fever, and Lumpy Skin Disease are not transferable to humans. This confirms the findings in our graphs.

Graph 13

Graph 14



Graph 15



Based on our generalized findings, we decided to then figure out which type of animal (captive, domestic, or wild) poses the greatest risk of spreading the disease (Avian Influenza specifically). When looking at Graphs 13, 14, and 15, we can see that there are no cases where the animals were categorized as being captive or wild (except for the "Other" animal). However, we can see that all the animal species had some cases where the animal was categorized as being domestic.

Our analysis also highlights the impacts these diseases are having in reality. Referring back to the Tableau maps, Graphs C and D in the Appendix tell us why we see disease concentrations in certain places. Graph C shows the density of grazing livestock in European Union countries. We see that there are very dense areas in the middle of France and in the Baltic States. Note that in the second Tableau map that the main animals in these areas were farm animals. This indicates why we see such high concentrations of these animals and these disease in those areas. Graph D further indicate this for France by showing where large concentrations of dairy production and meat livestock are. We see the concentration is highest in the middle of France, where we saw the Bluetongue outbreaks in cattle.

Doing further research into France and the Baltic States, we found that these diseases are taking an agricultural and economic toll. Based on a 2018 article in the Dairy Global, a farming industry news outlet, France has stopped compulsory Bluetongue vaccination. The disease there has been uncontrollable and there has been a shortage of vaccines. Another article in 2016 from the French Agency for Food, Environmental and Occupational Health & Safety discusses how Bluetongue affects cattle and how it has caused economic losses for farmers. In a 2018 article from Pig Progress, a pig production news outlet, the Baltic States had to slaughter 35,000 pigs due to African Swine Fever. This had a big impact for local meat-processing companies. These companies rather rely on local products than imported products. This culling of pigs would have an effect of supply. The economic impact in the Baltic States is further shown in a 2019 article from Pig Progress stating that the price of pork has risen over 40% from the start of the year.

CONCLUSIONS & RECOMMENDATIONS

Based on our analysis we have answered the four questions we original proposed:

1. What are the most prevalent animal diseases?
2. What countries are most at risk for these types of diseases?
3. Are these diseases harmful to humans?
4. Which animal (domestic, wild, or captive) poses the greatest risk of spreading these diseases?

We have found that the most prevalent diseases, in Europe at least, are African Swine Fever, Bluetongue, Avian Influenza, and Lumpy Skin Disease. We have seen that the countries most at risk in Europe for these diseases are France and the Baltic states because they have the highest variety of diseases. We have determined that the disease most harmful to humans is Avian Influenza. Finally, we have found that domestic animals pose the greatest risk of spreading these diseases.

From our analysis and conclusions, we can make the following recommendations:

- France and the Baltic States should focus research and development efforts on developing cheaper vaccines. This way they could help prevent economic losses from agriculture.
- Medical professionals, organizations, and the general public should be informed about the presence, causes, and effects of these diseases.
- Farmers should vaccinate their animals. If an outbreak occurs, they should put down all affected animals as soon as possible in order to decrease economic impact and possibly slow the spread of disease.

From our analysis and conclusions, we can make the following suggestions for future study:

- Our analysis should be expanded to include more species of animals.
- Pathogenic research should be undertaken to develop preventative measures and cures to stop outbreaks
- Pathogenic research should also be undertaken to develop early identification systems for outbreaks of zoonotic diseases (diseases that are transferable from animals to humans).

Works Cited

European Commission. *CAP context indicators – 2017*. 2018. <https://ec.europa.eu/agriculture/cap-indicators/context/2017_en>.

French Agency for Food, Environmental and Occupational Health & Safety. *Bluetongue (BT)*. 8 April 2016. <https://www.anses.fr/en/content/bluetongue-bt>.

Harrand, Rob. *EMPRES Global Animal Disease Surveillance*. 2019. <https://www.kaggle.com/tentotheminus9/empres-global-animal-disease-surveillance>.

Peys, Ruud. *France stops compulsory bluetongue vaccination*. 25 January 2018. <https://www.dairyglobal.net/Health/Articles/2018/1/France-stops-compulsory-bluetongue-vaccination-239742E/>.

Site Du Ministere De L'Agriculture Et De L'Alimentation. *Overview of French Agricultural Diversity*. 22 September 2015. <https://agriculture.gouv.fr/overview-french-agricultural-diversity>.

Vorotnikov, Vladislav. *Baltic processors faced with soaring pork prices*. 21 November 2019. <https://www.pigprogress.net/Health/Articles/2019/11/Baltic-processors-faced-with-soaring-pork-prices-503202E/>.

—. *Baltic states cull 35,000 pigs due to ASF*. 18 September 2018. <https://www.pigprogress.net/Health/Articles/2018/9/Baltic-states-cull-35000-pigs-due-to-ASF-335126E/>.

Appendix

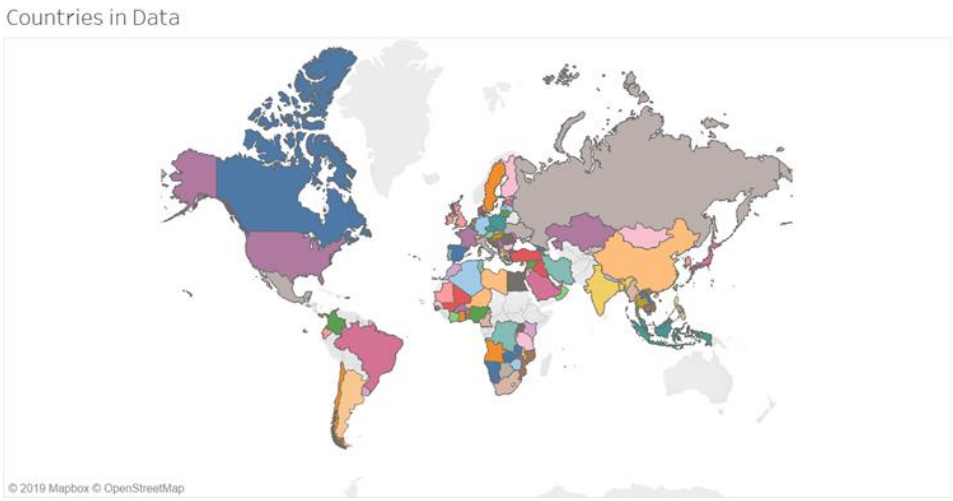
Data Dictionary

- Latitude
  - Latitude of the observation.
- Longitude
  - Longitude of the observation.
- Region
  - Continent where observation is located.
- Country
  - Country where observation is located.
- Admin1
  - The administrative region where observation is located.
- Locality Name
  - Locality where observation is located
- Locality Quality
  - Quality of the location.
- Observation Date
  - Date of observation.
- Month
  - Month of observation.
- Year
  - Year of observation.
- Season
  - Season of observation (based on latitude).
- Status
  - Status of observation.
- Disease
  - Name of disease.
- Species Description
  - Names of the species affected.
- Main Animal
  - The main animal affected.
- Is Domestic
  - Indicator for whether an animal is domestic or not.
- Is Wild
  - Indicator for whether an animal is wild or not.
- Is Captive
  - Indicator for whether an animal is captive or not.
- Sum At Risk
  - Sum of animal population at risk of contracting disease.
- Sum Cases
  - Number of animals that got infected.
- Sum Deaths
  - Number of animals that died from disease.
- Humans Affected
  - Number of humans that got infected

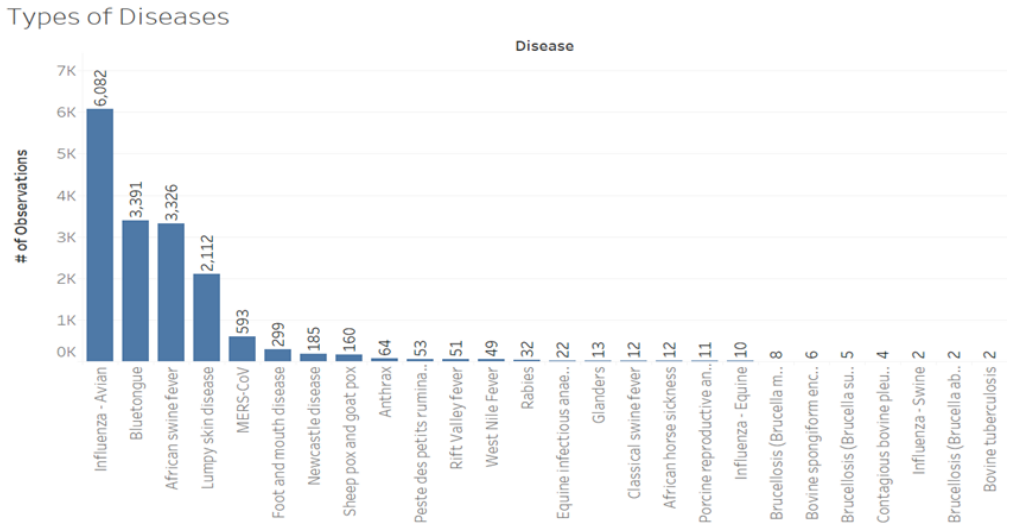
- Humans Deaths
    - Number of humans that died from disease.

Graphs

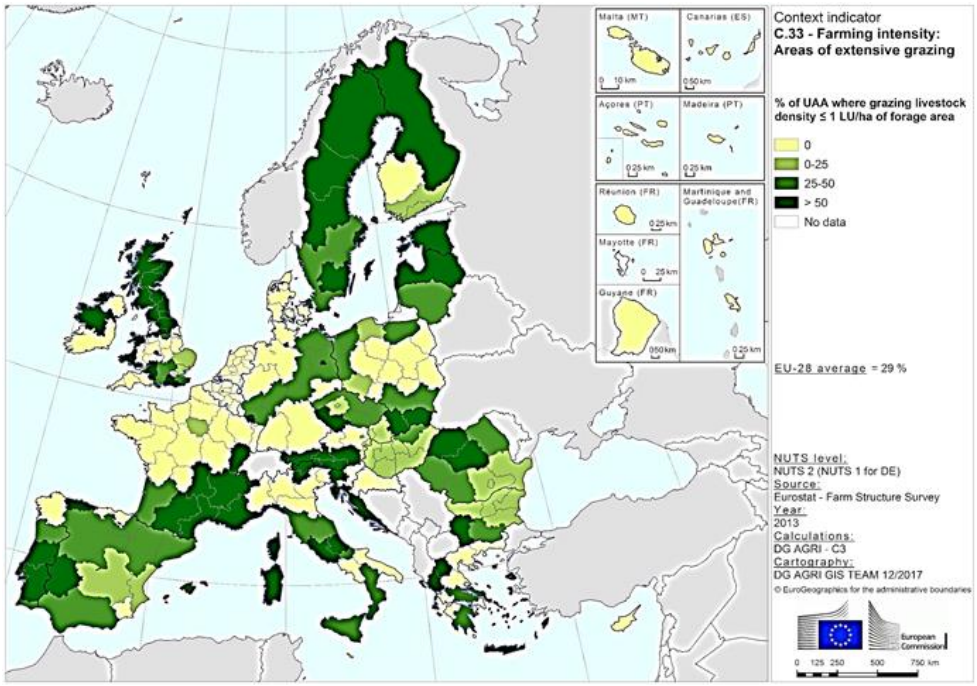
A



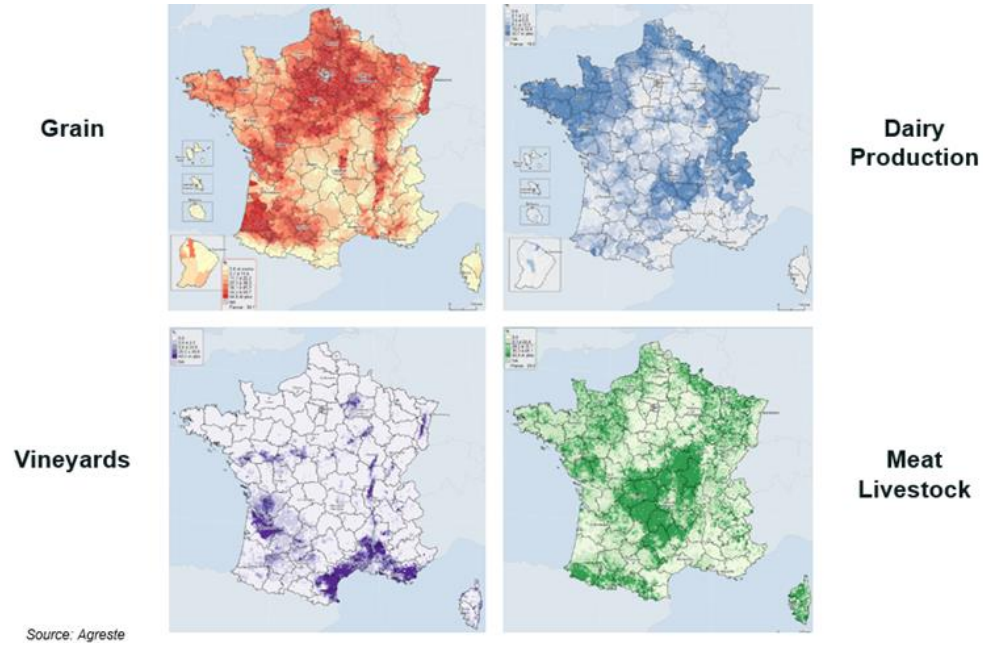
B



C

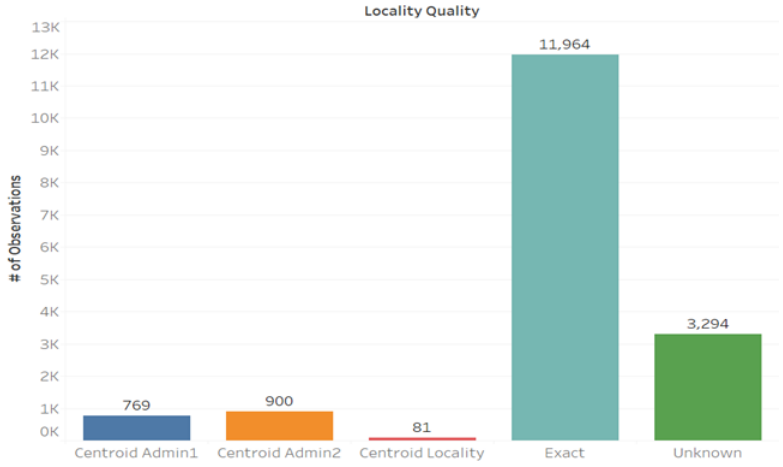

Malta (MT)
Canarias (ES)
Açores (PT)
Madeira (PT)
Réunion (FR)
Martinique and Guadeloupe (FR)
Mayotte (FR)
Guyane (FR)
Context indicator
C.33 - Farming intensity: Areas of extensive grazing
% of UAA where grazing livestock density ≤ 1 LU/ha of forage area
0
0-25
25-50
> 50
No data
EU-28 average = 29 %
NUTS level:
NUTS 2 (NUTS 1 for DE)
Source:
Eurostat - Farm Structure Survey
Year:
2013
Calculations:
DG AGRI - C3
Cartography:
DG AGRI GIS TEAM 12/2017
© EuroGeographics for the administrative boundaries
European Commission
0 125 250 500 750 km

D


Grain
Dairy Production
Vineyards
Meat Livestock
Source: Agreste

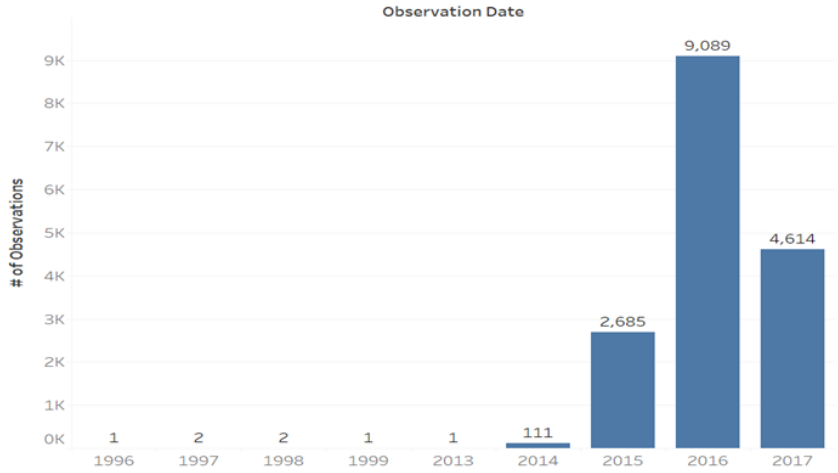
E

Location Quality



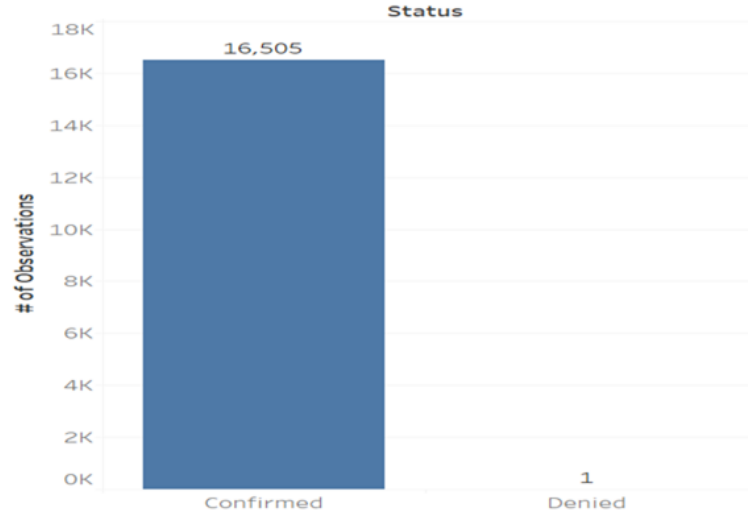
F

Years in Data



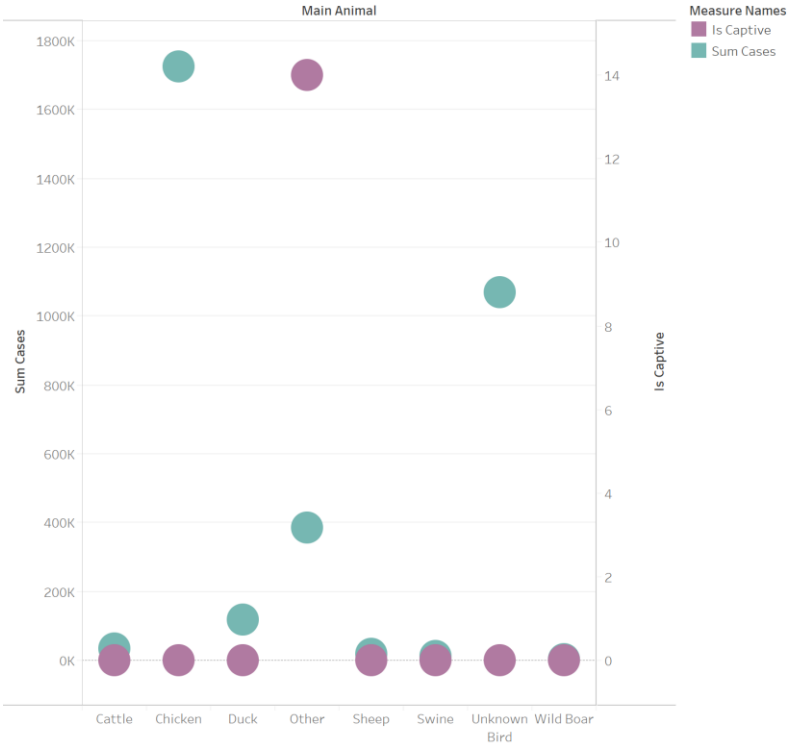
G

Status Level



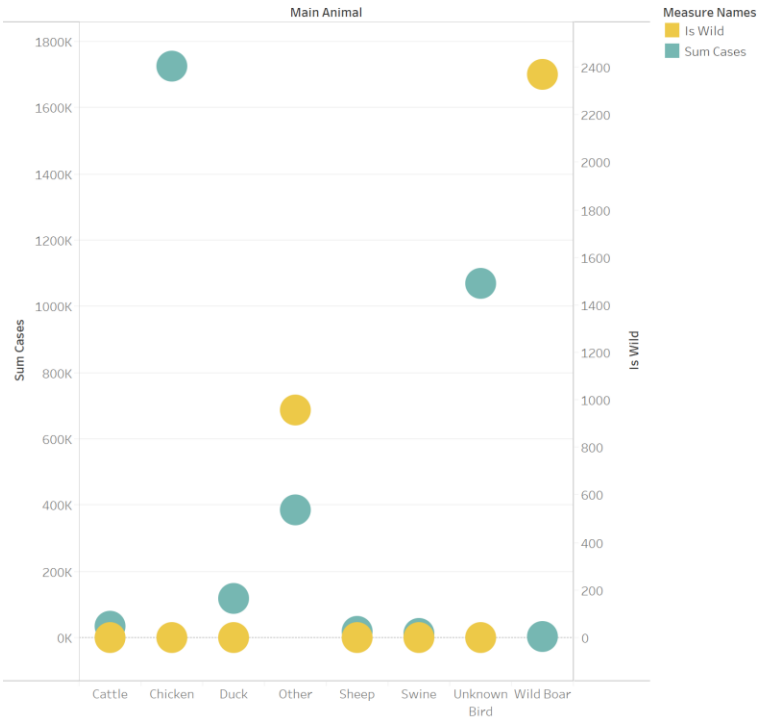
H

Captive



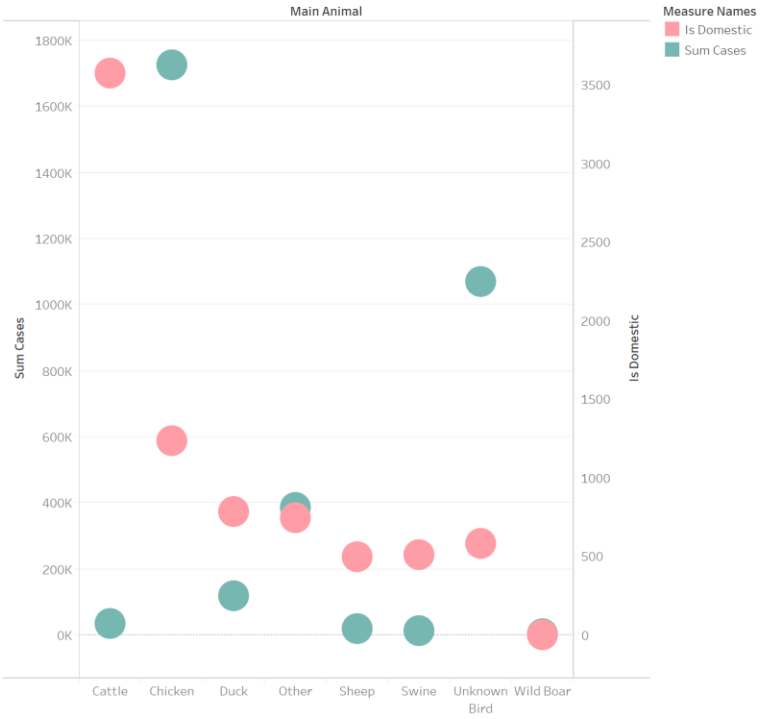
I

Wild



J

Domestic


Main Animal
Measure Names
Is Domestic
Sum Cases
Sum Cases
1800K
1600K
1400K
1200K
1000K
800K
600K
400K
200K
0K
Is Domestic
3500
3000
2500
2000
1500
1000
500
0
Cattle
Chicken
Duck
Other
Sheep
Swine
Unknown Bird
Wild Boar