### International Cheese Awards - Exploratory Data Analysis

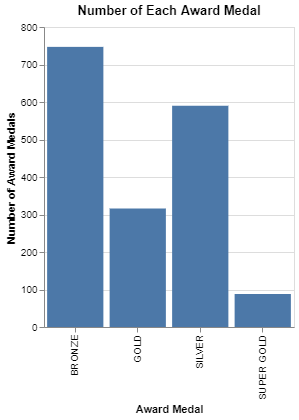
### Kasha Muzila, Karl Eirich

### W209 - Section Number 1

**The following images were generated via Google Colab using Altair and Python libraries. Please reference this Google Colab notebook for Data joining and cleaning as well as additional visualizations that were not mentioned in this document due to contextual purposes.**

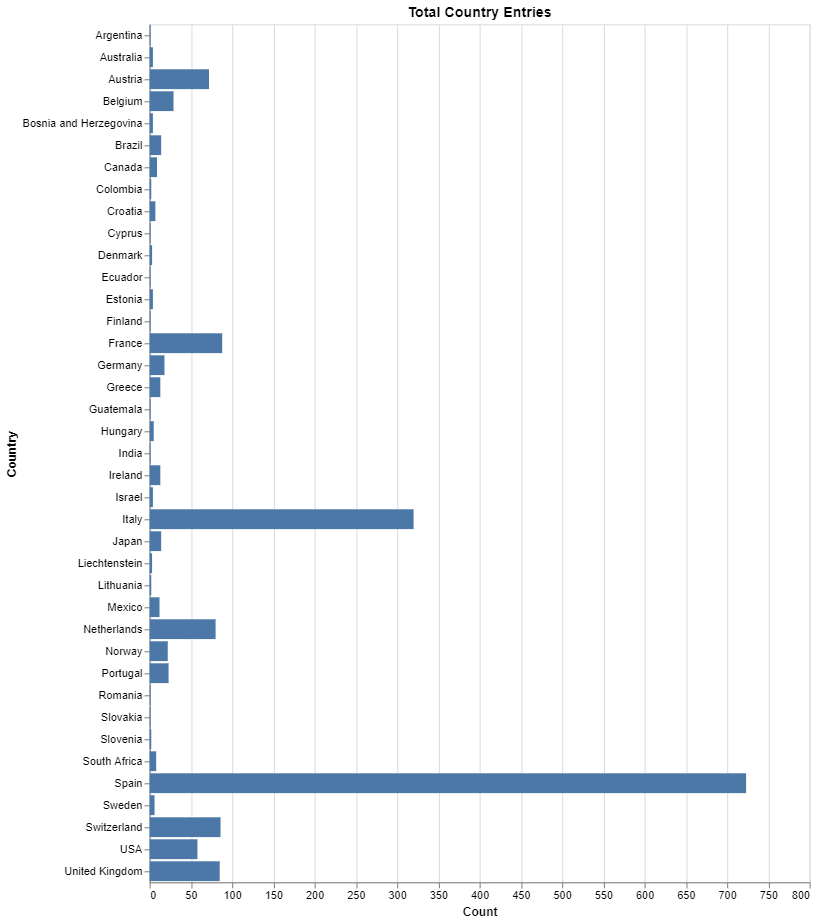
[**https://colab.research.google.com/drive/1dhr8k\_w7M4JHARAqjvDWfz5r7hjMDG-h?usp=sharing**](https://colab.research.google.com/drive/1dhr8k_w7M4JHARAqjvDWfz5r7hjMDG-h?usp=sharing)

**Hypothesis 1:** The top 3 countries with the most awards are France, Switzerland, and Italy, with bronze being the most award type.



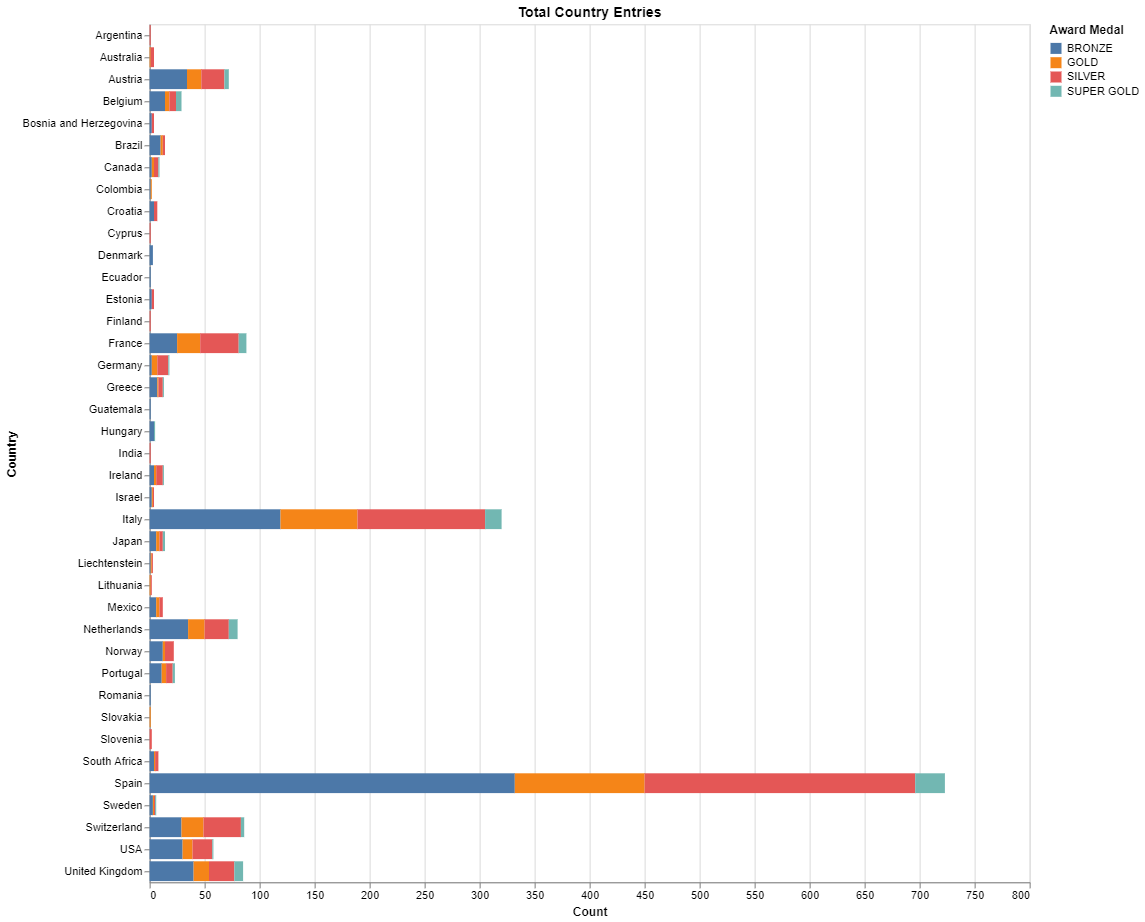
**What's informative about this view:** This view shows the number of each type of award that was given out. This view is helpful in knowing the quantity of the type of awards for the international competition to see which type of award was easier to win. We can see in this graph that the most awards to the least awards were Bronze, Silver, Gold, and Super Gold.

**What could be improved about this view:**  A total amount of awards given to each country would give us better insight into which countries participated and would answer the first part of our hypothesis.



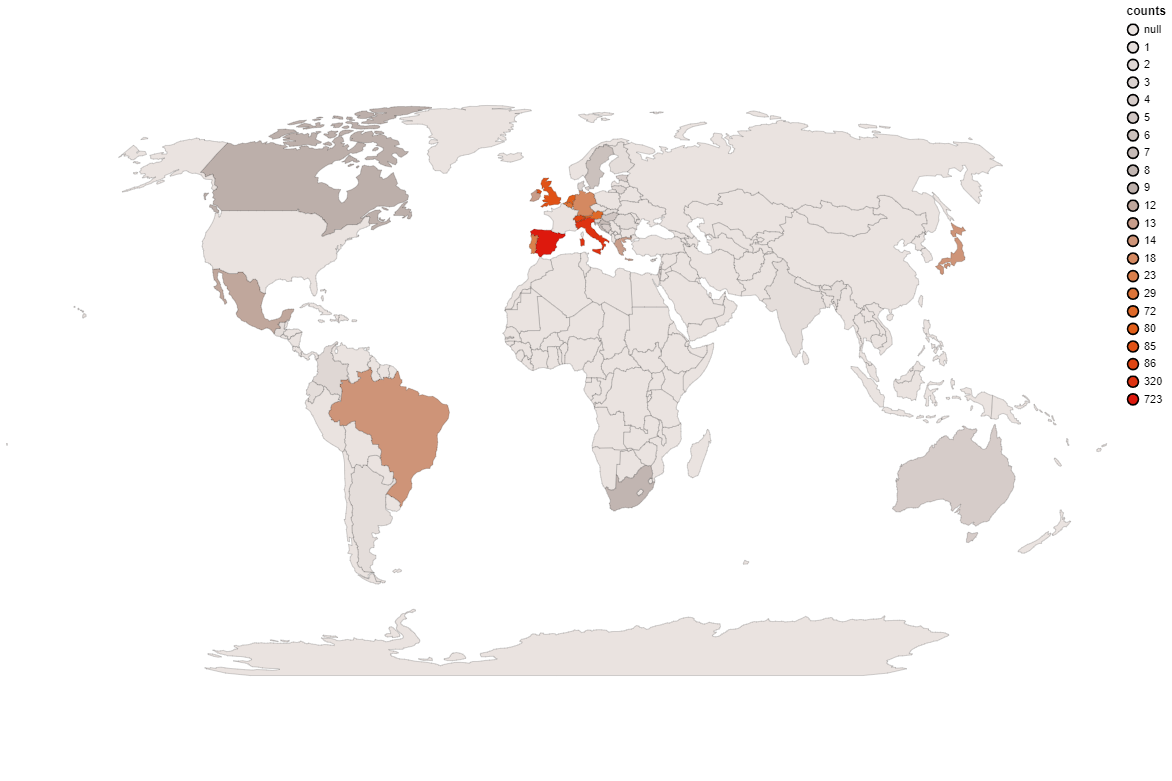
**What's informative about this view:** This view explores all of the countries that participated in the international cheese awards and the number of awards that were given to each country. It is helpful by seeing which country was given the most awards.

**What could be improved about this view:** Knowing how much of each type of award that was given per country would be an improvement and would answer the second part of our hypothesis.



**What's informative about this view:** This view explores all of the countries that participated in the international cheese awards, the total number of awards that were given to each country, and the amount of each type of award. It is helpful by seeing how much each award was given to each country that participated in the international cheese awards.

**What could be improved about this view:** Seeing the total amount of awards given to each country on a geographical map would give a better visual for those that aren’t familiar with where these countries are located.

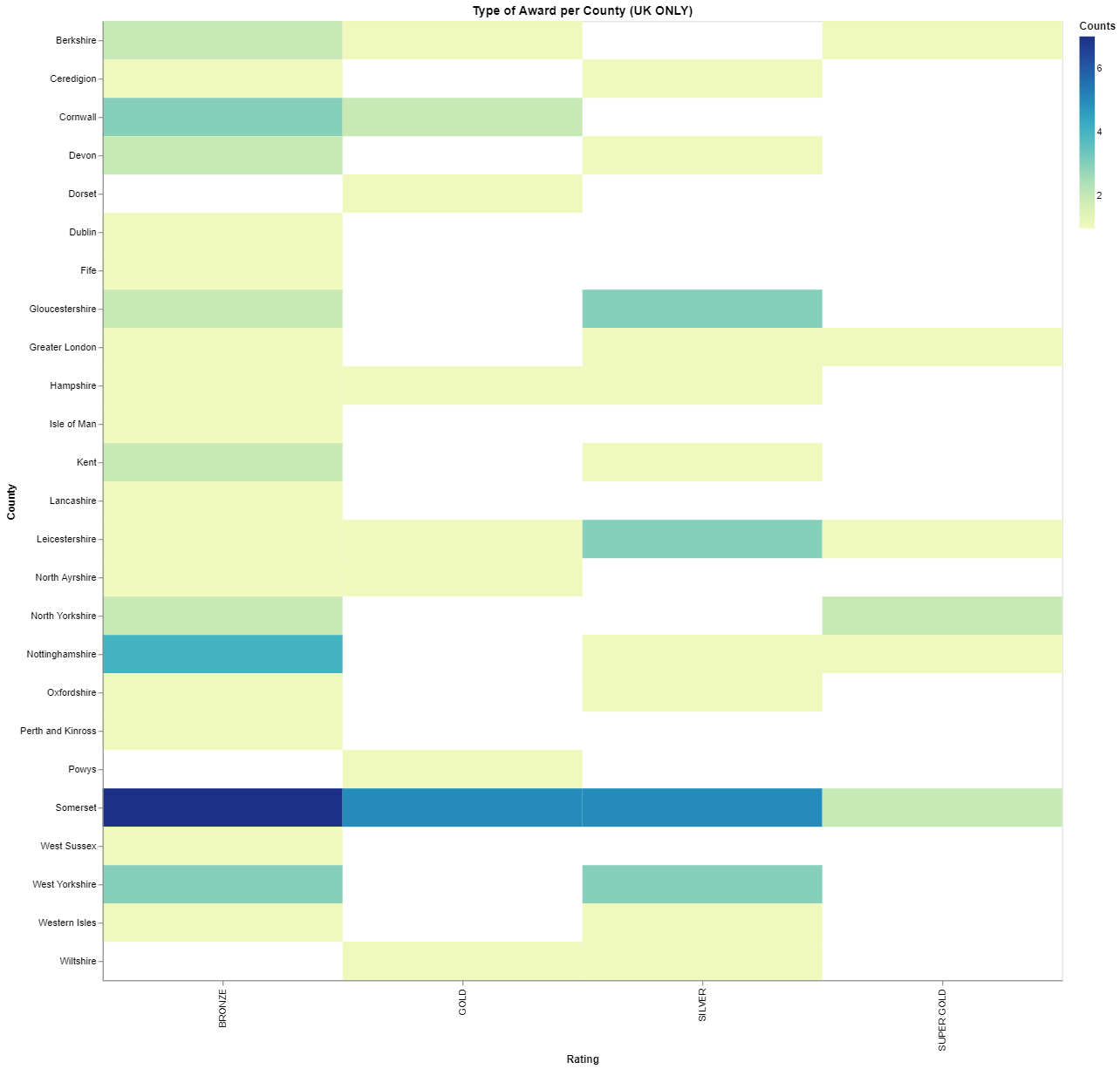
****

**What's informative about this view:** This view explores the total amount of awards given to each country that participated in the international cheese awards on a world map. It is helpful by seeing where in the world countries participated the most during the competition.

**What could be improved about this view:** We already answered our hypothesis at this point. However, it would be interesting if we could zoom in on the countries and see more details about the companies that participated in each country (if there were more than one company that participated). This view was challenging to make as there weren’t a lot of resources to help create this map. This map is also interactive when hovering over each country. It shows the country name and total number of awards. This view can also be improved by having the countries that didn’t participate show it’s name and a 0 for the total amount of awards, or that they didn’t participate. So far, the countries that didn’t participate say “null” when hovering over the country.

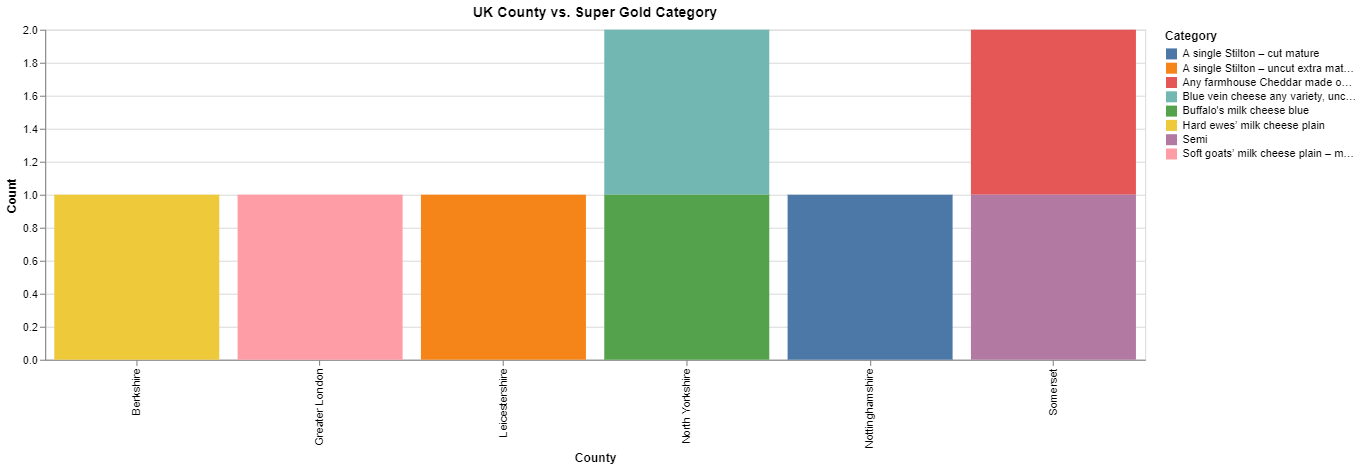
**Conclusion** These data views suggest that Spain, Italy, and France were the top 3 countries given the most awards with Spain and Italy given mostly Bronze awards, and France given mostly Silver awards. It was surprising to see how involved Spain was with the international competition with a total of 723 awards and Switzerland only having 86 total awards.

**Hypothesis 2:** The most awarded county in the United Kingdom was Greater London with cow’s milk being the best category of cheese in Greater London.



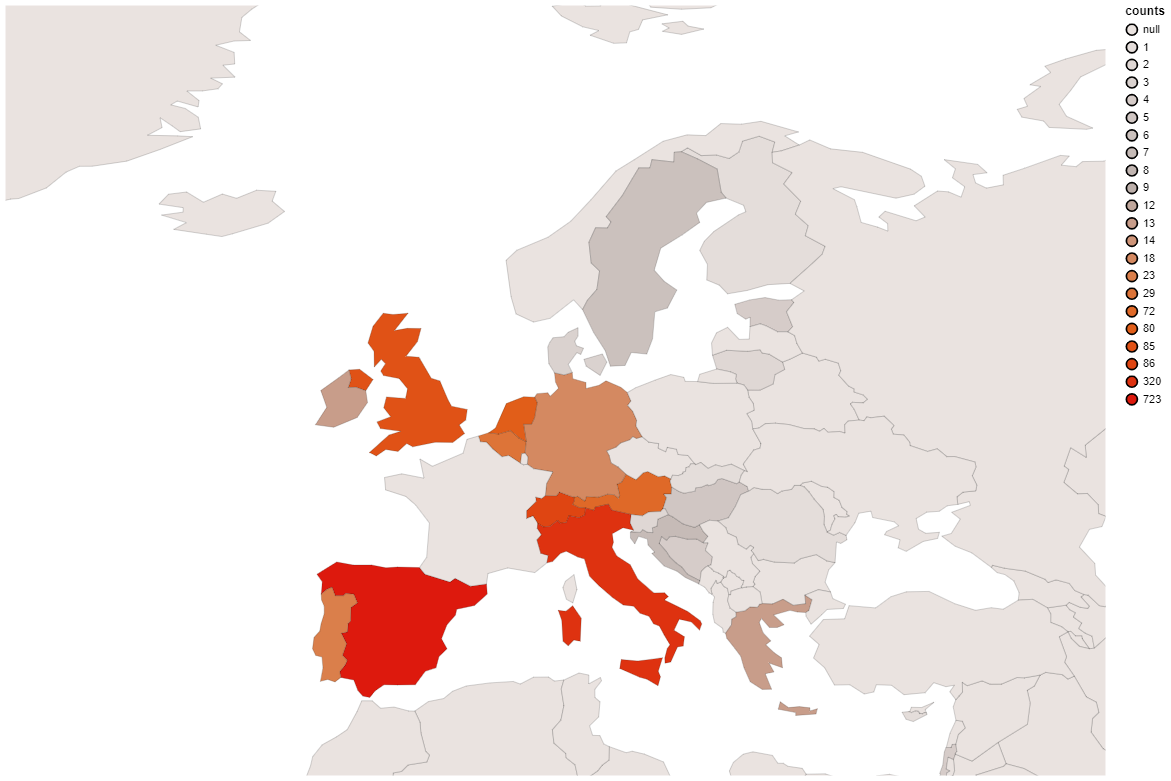
**What's informative about this view:** This view explores all of the counties in the United Kingdom that won awards in the international cheese competition and shows a legend of the amount of each award. It is helpful by showing that the first part of our hypothesis is false. It is Somerset County that was awarded the most in the United Kingdom.

**What could be improved about this view:** Seeing only the Super Gold awarded to each county and what category they won the Super Gold award would help answer the second half of our hypothesis.

****

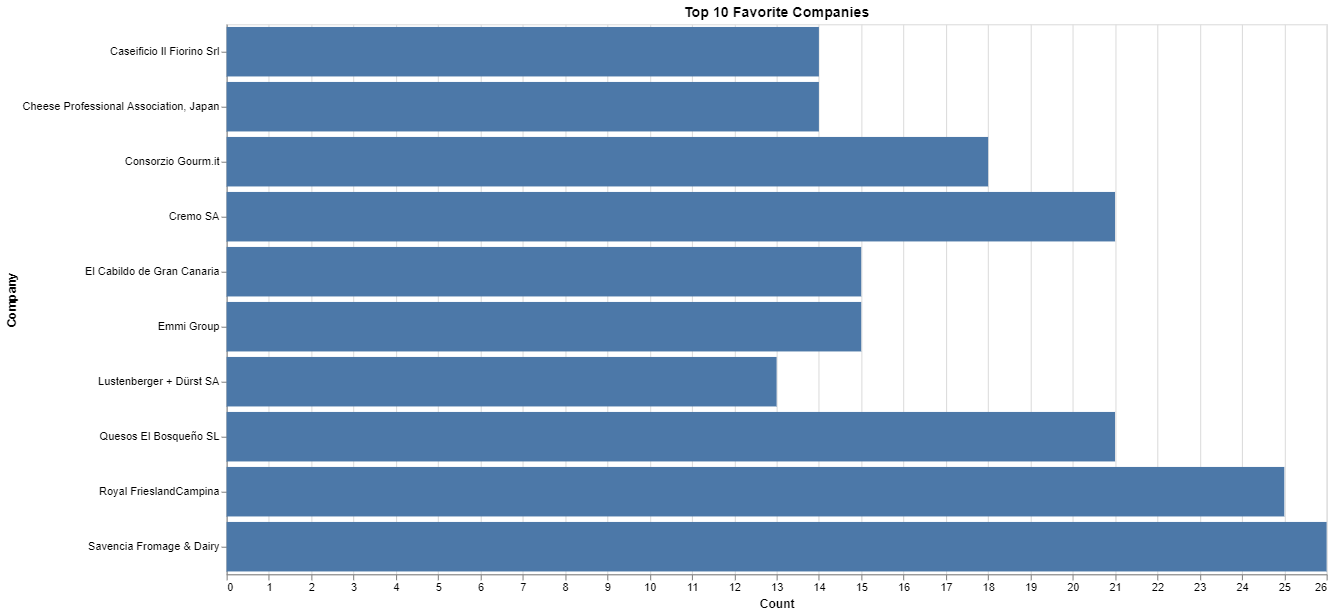
**What's informative about this view:** This view explores all of the Super Gold awards given out to each United Kingdom county. It is helpful by seeing the amount or Super Gold awards given out and which category each county won a Super Gold. We can see that our second part of our hypothesis is false. It was Any Farmhouse Cheddar and Semi being the best category of cheese in Somerset.

**What could be improved about this view:** Being able to see the different counties on a United Kingdom geographical map would be an improvement. However, we have only been able to figure out how to zoom in on all of Europe but not be able to show the counties. This is because the original dataset does not have county identification numbers (i.e., latitude, longitude, county number, etc.). We would have to find a separate dataset that has these identification numbers in order to plot a geographical map. We could not look into other country counties because the original dataset only has United Kingdom counties available. Again, we would need to find a separate dataset to include other country counties. Below, is a screenshot of how we were able to zoom in on a geographical map, but could not get the counties:



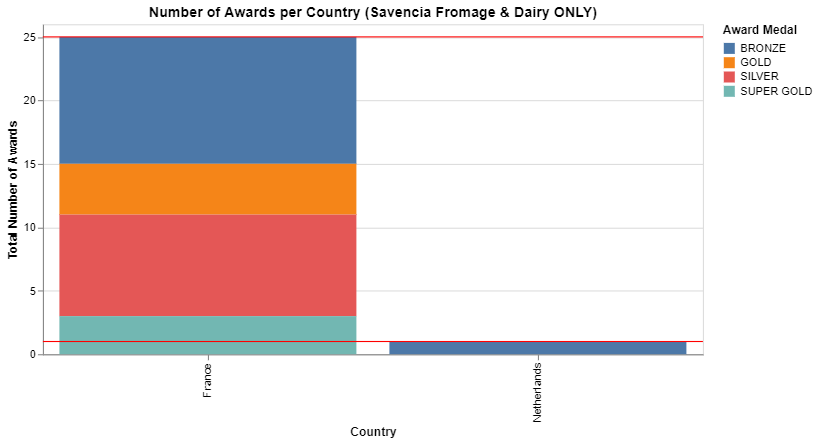
**Conclusion** These data views suggest that Somerset County was the most awarded United Kingdom county in the international cheese awards with its best cheese being in the Any Farmhouse Cheddar and Semi category. It was surprising to see that mixed cheese categories were the best awarded as typically, the most common cheese made it cow’s milk.

**Hypothesis 3:** The company that won the most awards was Savencia Fromage & Dairy which is located in the Netherlands with goat’s milk being its best category.



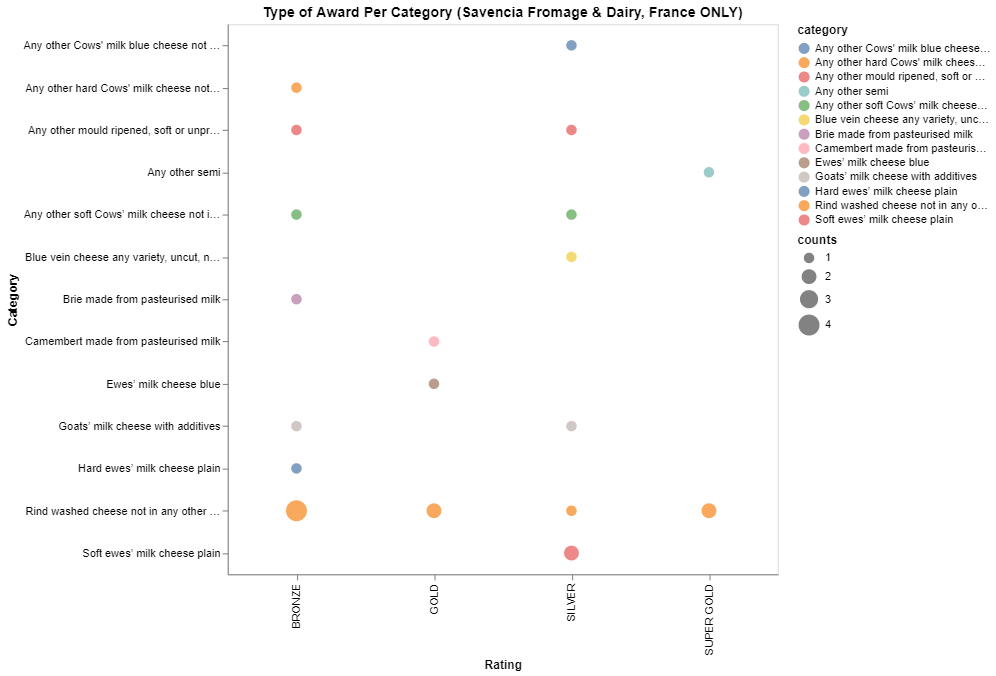
**What's informative about this view:** This view explores the top 10 most awarded companies. It is helpful by showing the total number of awards that the top 10 companies won with Savencia Fromage & Dairy being the most awarded company. This shows that the first part of our hypothesis is true.

**What could be improved about this view:** With this graph, we don’t know what country Savencia Fromage & Dairy is located in or if it’s in multiple countries. Generating a new graph that only looks at the countries Savencia Fromage & Dairy is located in would help answer the second part of our hypothesis.



**What's informative about this view:** This view explores more information about the company Savencia Fromage & Dairy by looking at the countries they reside in and what awards they have won. It is helpful because this shows that Savencia Fromage & Dairy is located in the Netherlands and France with France winning most of the awards. We can also see that the highest award France won was Super Gold. This view shows that our hypothesis of the Savencia Fromage & Dairy location was false.

**What could be improved about this view:** To finish off our hypothesis, we need a graph that will show more details about the companies that won awards within the France location of Savencia Fromage & Dairy.



**What's informative about this view:** This view gives a detailed explanation on how many of each type of award was won for each category. It is helpful because this shows that Savencia Fromage & Dairy’s France location won 2 Super Gold awards in the “Rind washed cheese not in any other class” category and answers that the last part of our hypothesis is false.

**What could be improved about this view:** Again, it would be interesting to see more of this information in a geographical map but we would need a separate dataset that could find the latitude and longitude information of each company so we can plot it on a county basis. We were not able to accomplish this by the time this EDA analysis was due. However, we will look into it further within our project.

**Conclusion** These data views suggest that Savencia Fromage & Dairy company within the France location won the most awards with their best cheese category being “Rind washed cheese not in any other class.” It was surprising that the common cheeses such as cow’s, goat’s, sheep’s, milk was not the highest awarded and was the uncategorized or mixed cheeses during this project’s exploratory data analysis. We hope to incorporate more latitude and longitude data to gain better insight on company locations further into this project.