**Process Book**

# **Overview and Motivation**

**Overview**

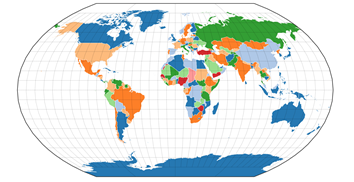
The world’s population is growing even faster from 7.3 billion today to 9.7 billion in the year of 2050. How to provide enough food for all those people globally would be a challenge and difficult task to each country and its government. As people’s requirement on diet is also processing there are tons of animals feeded by farmers to be served as the meat for meals, this is something that not only on agriculture but also on dieting culture, habits and farming business.

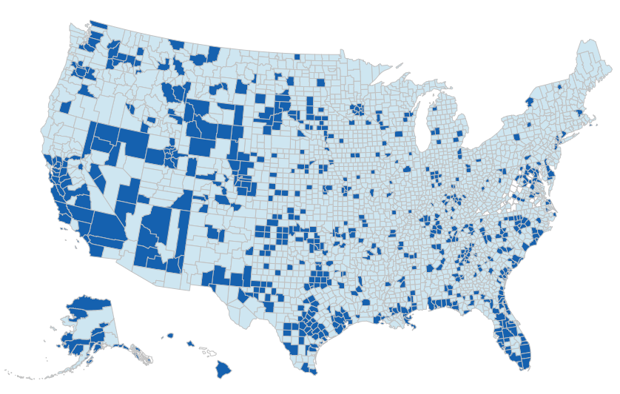
**Motivation**

We would like to demonstrate some visualizations regarding worldwide food and feed amount, and this should always indicate something that happened in different regions. By having a deep insight on our worldwide food and feed production and the comparison we can move forward more on human consumption and feed produced situation for animals.

# **Related Work**

By using GeoJSON or TopoJSON to generate map based visualization is always a good solution to demonstrate the differences and comparison all over the world. The geometry information about cities, states and countries are indexed in the JSON file and all of their shapes are encoded in polygon coordinates. With some interesting projection methodologies we can even create visualizations in different styles which could highlight and attract the viewers to see what we want them to focus. Below is a snapshot showing about it:





Here are some links that inspired us to further design and implement our most fun visualization with Geomap:

a. Symbol Map by Mike Bostock:

<https://bl.ocks.org/mbostock/4342045>

b. How to convert and prepare TopoJSON files for interactive mapping with d3

<https://hackernoon.com/how-to-convert-and-prepare-topojson-files-for-interactive-mapping-with-d3-499cf0ced5f>

# **Questions**

a. Does the amount of food consumed by a country indicate anything such as the political environment, economic development and even population fluctuation of this country? The best way to draw this is by using the bar chart to analyze with Tableau. By using Tableau we set up a visualization with country as the filter and its fluctuation bars of consumed food across the year of 1992 to 2013. Considering some news from TV and newspaper we tried with several countries as the options and finally decide to put Ukraine for the demonstration.

b. How about the amount of food and feed of a country compared to others? In order to compare the amount between different countries and clear see the difference we need involve in treemap. We can build a hierarchical dataset beforehand, and rank them based on amount. Additionally, we divided the countries by different continent, so these different groups can tell us more stories.

c. Lots of food was consumed to provide food to people nowadays, but we also have tons of food feeded to animals which grow them as the meat and meal for people. Due to the various economic development and regional religion belief across different areas and countries the ratio of comparison between food feeded to animals or as the meal for people also varies a lot. How to get a balance on this? Should we save more wheat and grain to the starving people rather than feeding animals to provide meat for the others? This topic inspired us to draw a geo map for further insight and analysis.

**Data**

We collected the data from Kaggle competition of Who eats the food we grow and it has the worldwide food\feed production and distribution data across Year 1961-2013. Here’s the column metadata:

|  |  |  |
| --- | --- | --- |
| **Column name** | **Column Description** | **Column type** |
| Area Abbreviation | Country name abbreviation | String |
| Area Code | Area Code | Numeric |
| Area | Country name | String |
| Item Code | Food item code | Numeric |
| Item | Food item | String |
| Element Code | Food or Feed code | Numeric |
| Element | Food or Feed | String |
| Unit | Unit of Measurement | String |
| Latitude | Latitude | Numeric |
| Longitude | Longitude | Numeric |
| Y1961 | Measurement of Year 1961 | Numeric |
| Y1962 | Measurement of Year 1962 | Numeric |
| …... | …... | …... |
| Y2013 | Measurement of Year 2013 | Numeric |

The Food Balance sheet's data was relatively complete. A few countries that do not exist anymore, such as Czechoslovakia, were deleted from the database. Countries which were formed lately such as South Sudan were kept, even though they do not have all full data going back to 1961.In addition, data aggregation for the 7 different continents was available as well, but was not added to the dataset. Food and feed production by country and food item from 1961 to 2013, including geocoding. Y1961 - Y2011 are production years that show the amount of food item produced in 1000 tonnes.

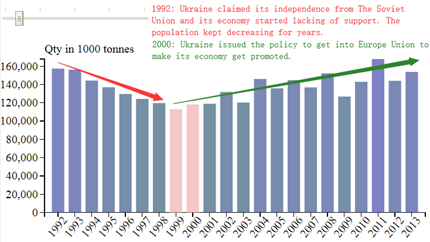
Consider to have better performance in data visualization via d3.js we decide to do some substantial data cleanup, filtering and aggregation on columns of Areacode, Years and Elements. We plan to extract from the raw dataset and use python code to manipulate the data.

final\_vis1\_barchart.ipynb is the jupyter notebook that we used to pre-handle the data for Visualization 1.

# **Exploratory Data Analysis**

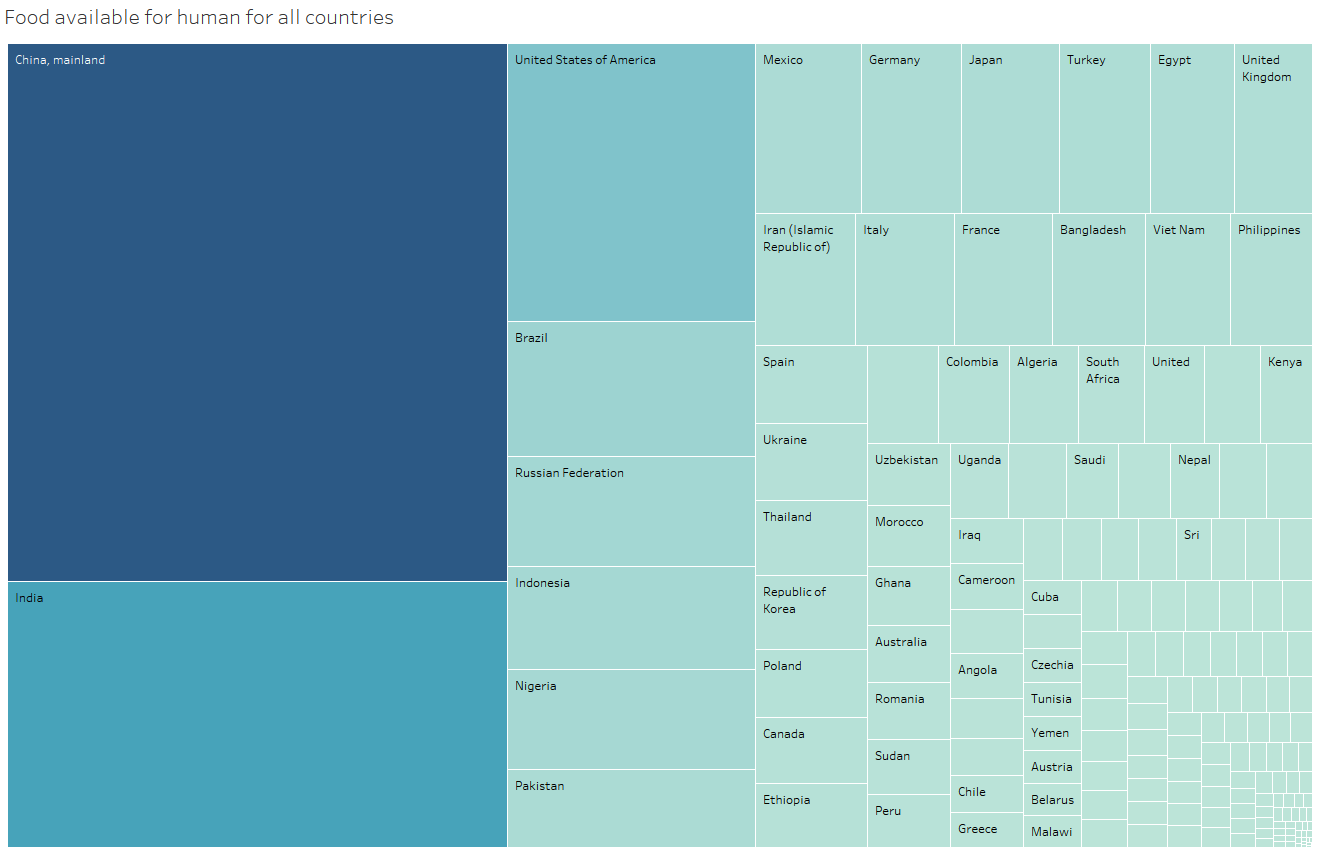
**Visualization 1: Trend of Food Consumed**

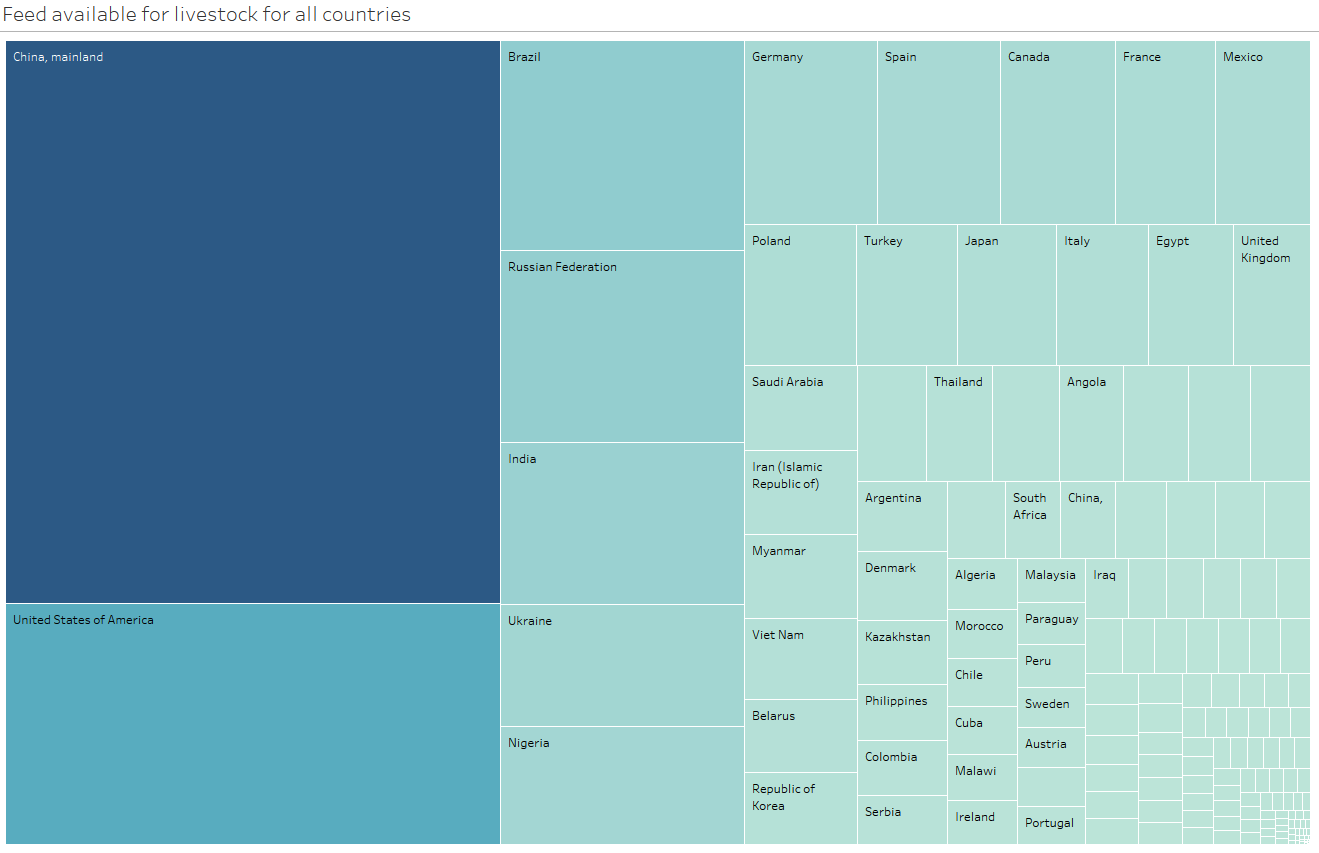
As we know Ukraine derived from the Soviet Union in 1992 and its political environment kept unstable since then. Losing the political support from the Soviet Union its population kept decreasing for many years. In 2000 the government of Ukraine issued the policy of joining the Europe Union to promote its economy so its population got boosted and kept increasing since then. The food consumed by Ukraine could also be seen as a good indication about what was happening with a representation way of bar charts.



**Visualization 2: Food and Feed Available among all Countries**

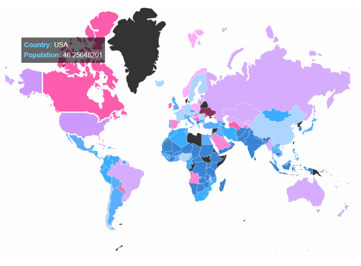
From two treemap(Feed treemap and Food treemap) we got, we can clearly see the amount comparison from the area of rectangle. The comparison can be affected by its religion and economic growth of a country. For example, China is the No.1 among two maps. India definitely has much more Food amount than Feed, since India called a vegetarian country with less meat consumption etc.





**Visualization 3: Worldwide Feed and Food Ratio Comparison**

From the geomap of ratio of Food and Feed, generally, we cannot say that the developed country has higher ratio, since this indicator can be affected by people eating diet, religion, population, animal agriculture, economic growth etc. Based on these factors, we should move forward and conduct the next analysis.



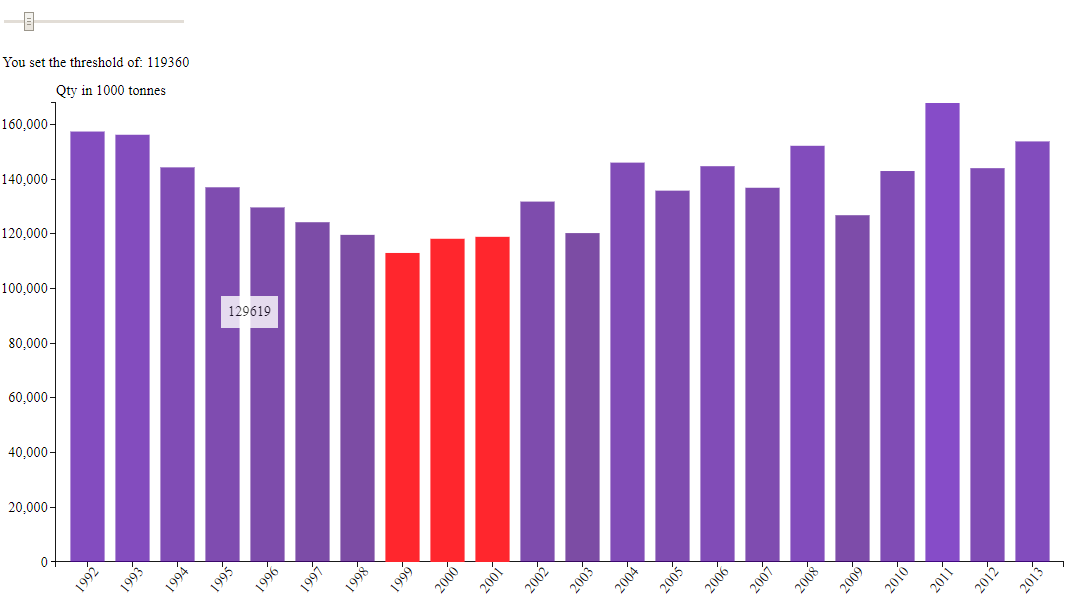
# **Design Evolution**

In order to have a deep insight of this dataset, we want to explore the trend of food consumed, the variety of food for human and feed for livestock. At first we want to use pie chart to visualize the variety and display the percentage of each part, but after we tried to visualize in Tableau, we found there are almost 50 different types of food and feed, and the final visualization would be hard to tell us the variety even we add legend and other interactions. So after that we decided to just pick top 15 of them on this pie chart, so we wrote python script and did map-reduce to generate the final data we need. After we submit our project proposal, we think it really broke the integrity of our dataset and pie chart are not useful in this aspect, they're too easy to screw up, and they don't accomplish the one thing we actually use charts for — to make information visually informative, finally we chose treemap to visualize the feed and food amount among all countries.

# **Implementation**

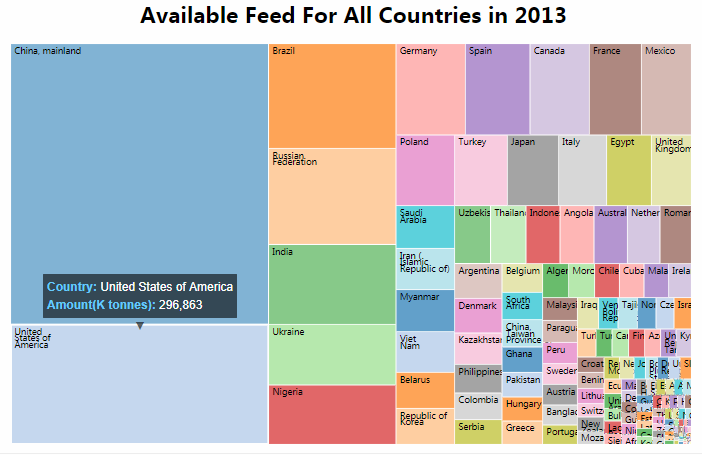
## **Visualization 1: Trend of Food Consumed**

* Add the mouseover event for displaying the label and value in opacity level of 0.8
* Use the web element of “input” to indicate the threshold that we would like to set to filter on the bars, and once user scrolls on the slider the threshold value would appear to assist
* Define the change event for the web element of “input” to implement the filter out of bars based on the threshold we set
* Implement the transition effect on the appearance and disappearance of bars
* Any bars with actual value smaller than the threshold viewer sets will turn to red



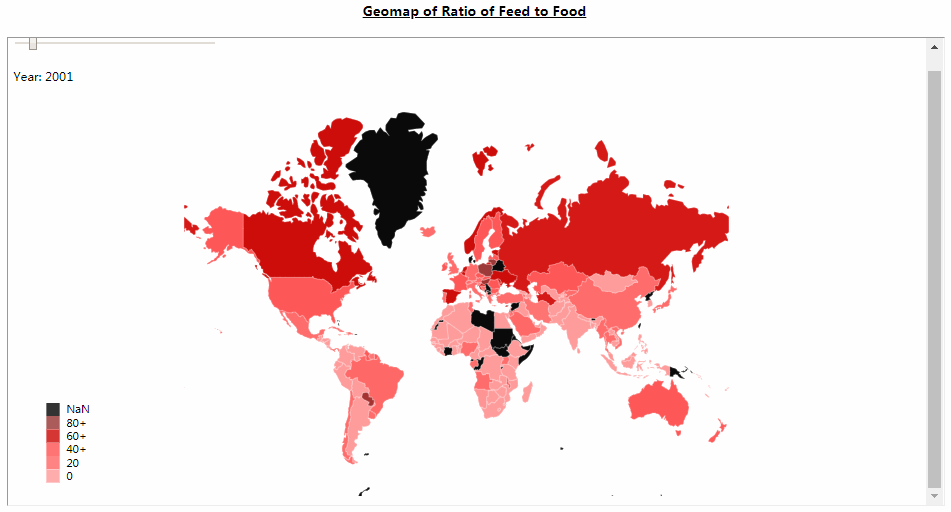
## **Visualization 2: Food and Feed Available among all Countries**

* Displayed a large amount of hierarchical data using nested rectangles of varying size and color. The total area of a treemap reflects the sum of its parts, which consist of inner rectangles or nodes.
* Prepared a metric set beforehand, convert dataset to hierarchy structure. For example, create a new metric set by dragging two measures (Feed amount and Country) and a hierarchy/dimension to the canvas.



## **Visualization 3: Worldwide Feed and Food Ratio Comparison**

* Preprocessing Data set and generate the ratio of Feed for livestock over Food for human, stored the ratio data into world\_population.tsv.
* Collect the world map geometry coordinates data into world\_countries.json.
* Join .tsv and .json with country id to visualize the ratio level in color onto world map.



# **Evaluation**

From the above visualizations, we can clearly see the global population is continually increasing In our current world hence the argument that producing feed for livestock conflicts with feeding hungry people is likely to continue for some years. It is clear that we need to improve our strategies to balance the by-product feeds to improve the efficiency of mitigating the environmental and economic impact of animal agriculture.

Furthermore, some import conclusions we cannot just make base on these visualization, because many other factor may affect conclusion. In order to have a deeper insight on our worldwide food and feed production, we still need to collect more related data from people eating diet, religion, population, economic growth from different country, we can do more analysis towards the ratio of feed to food and compute the correlation between ratio and other factors, and visualize them use Heatmap.