Group 6: Arvind Suriakanth, Calvin Moore, Ramesh Gururajan, Roy Ulrichiii

1) Introduction to Problem or Opportunity (include supporting domain information):
    a) Domain: Other → Food Distribution
    b) Problem Statement: With the global population increasing with life expectancy, class increase, and land available; the need for food is increasing.
    c) Project Description: In order to cope with the need for food, we need to devise a way to balance supply and demand. For any country, the production of food is influenced by two main factors, natural resources and institutional support. Natural resources include fertile land, forest, water, etc. Institutional resources include local and national government bodies. Our project is going to examine the demand for certain main food sources such as milk, wheat, and soy. from different countries within a given population. Then, we are going to merge this data with expected population growth to see if the countries can sustain using that resource for the growing population or should change to something better suited to the resources available.
2) Project Repository Github Link: https://github.com/ramkguru97/Group_6_6162

3) Research Question(s) of Interest:

a) What are the top 5 foods produced by each continent from 1961?
b) Which food items from 1990-2013 are known to increase production of feed and food in Asia? How does this affect food consumption globally?
c) Which top 3 countries produce the most food globally and is the rate of food production inversely or directly proportional to its growing population?
d) Given the volume of dairy production in Europe, which countries are the largest importers and exports, from 2000 - 2013?
e) What are the top 5 food ingredients (milk, wheat, soy, etc) for each continent in the world? What is the percent of demand of those given food compared to the total demand of all food ingredients for that continent?
f) Are African countries able to lessen their demand for one food source for maybe a more sustainable one? Do they have the resources to grow better resources than they already are or do they have arable land that is not being used?
g) What percentage of countries have more demand for a given food resource than supply? What continent do these countries reside in and what factors contribute to this inability to supply with the given demand?
h) Clustering → We intend to cluster by continent to showcase the food distribution and its population distribution.

4) Data Resources: This dataset is acquired from Kaggle and is related to worldwide food production, specifically focusing on comparison between food produced for human consumption and feed produced for animals. The data is collected from The Food and Agriculture Organization of the United Nations for over 245 countries from 1961 to 2013 and consists of 63 columns and 21478 rows. We are also in the process of taking a look at different datasets for different types of

foods and countries around the world's demand for that food. Datasets on the growing population of the world are being looked at as well.
https://www.kaggle.com/dorbicycle/world-foodfeed-production

5) Future Work:
-Due to globalization, some countries might have more imported goods and produce more than what they are exporting. The rate of imported goods and exported goods might shed light onto how food distribution is shifting along with supply/demand. With that being said, we apply clustering methods to dive deep into this and get a better understanding of the impacts in food distribution.