Jordan Page

Final Project Proposal

**What is the problem you are attempting to solve?**

* I will attempt to predict the number of meat products or live stock being exported worldwide from the US during the next five years.
* Given some of the events that has happened this year in the world such as the fires in the Amazon and more demand for farming, as well as African swine fever that spread across Europe, we want to explore what reasons contributed these events.
* Look into the meat products and livestock data exported from the US to the world from 1990 to present.

**How is your solution valuable?**

* I think my solution would be valuable because consumers, government, environmentalists, and meat farmers would benefit from this. These groups of people will be able predict future demands and supplies of meat products and livestock. I believe this would help these groups of people to make better decisions on meat consumption, land usage, and impact climate to change.

**What is your data source and how will you access it?**

* The last couple of years I have been trying to learn more about the relationship between agriculture and global warming. Also, I have been trying to learn more about the climate change crisis. I would like to explore the data and see what the trend of meat exports from the US is. With this I can see if the meat exporting industry in the US is increasing or decreasing.
* The dataset I obtained is from data.gov which is a website of the US government’s data for public use.
* The variables I am going to use are the amount of livestock exported outside of the US in kilograms, the countries each livestock or meat product they are being exported to

**What techniques from the course do you anticipate using?**

* I will use unsupervised machine learning techniques such as neural network techniques to model the amount of meat products and livestock exported from the US and project it further out.
* I will employ machine learning natural language processing techniques to classify each commodity description to a simplified commodity.
* I will use ARIMA modeling, which is from the specialization topic I chose time series analysis, to model the number of meat products and livestock exported from the US and project it further it.

**What do you anticipate to be the biggest challenge you’ll face?**

* Feature selection would be a challenge as the dataset only has the amount and the year as a measurement factor.
* When dealing with big data and complexity of the dataset that includes a large amount of countries and a complex commodity description, I would need to try different methods and techniques to aggregate data.

**The goals that I aim to achieve with this project, are to answer these questions.**

* Can we predict the amount of livestock or meat products being exported worldwide from the US over the next 3 years?
* Based on what we learn, what does this prediction mean for the environment?
* Given recent events such as the fires in the Amazon and more demand on land cultivation, can we find any correlation between those events and the export of meat?