Exploratory Data Analysis

Don Irwin

U.C. Berkeley MIDS

W209 - Spring 2022 Thursday 4 P.M.

Background:

MIDS students taking W209, Data Visualization, are expected to demonstrate an ability to source datasets, analyze datasets, and present visualizations in the form of web-based applications, to convey insights from those datasets to users.

Our team will communicate trade trends over the past 6 years, for the world's top 20 trading nations.

Data Sources, and description of those data:

Our data is sourced from the World Trade Organization, and Bloomberg. These data consist primarily of datasets which contain total trade, total imports, and total exports, time-series, by any one of the top 20 trading nations with another. These data also contain trade categories, of each of these nations with the other.
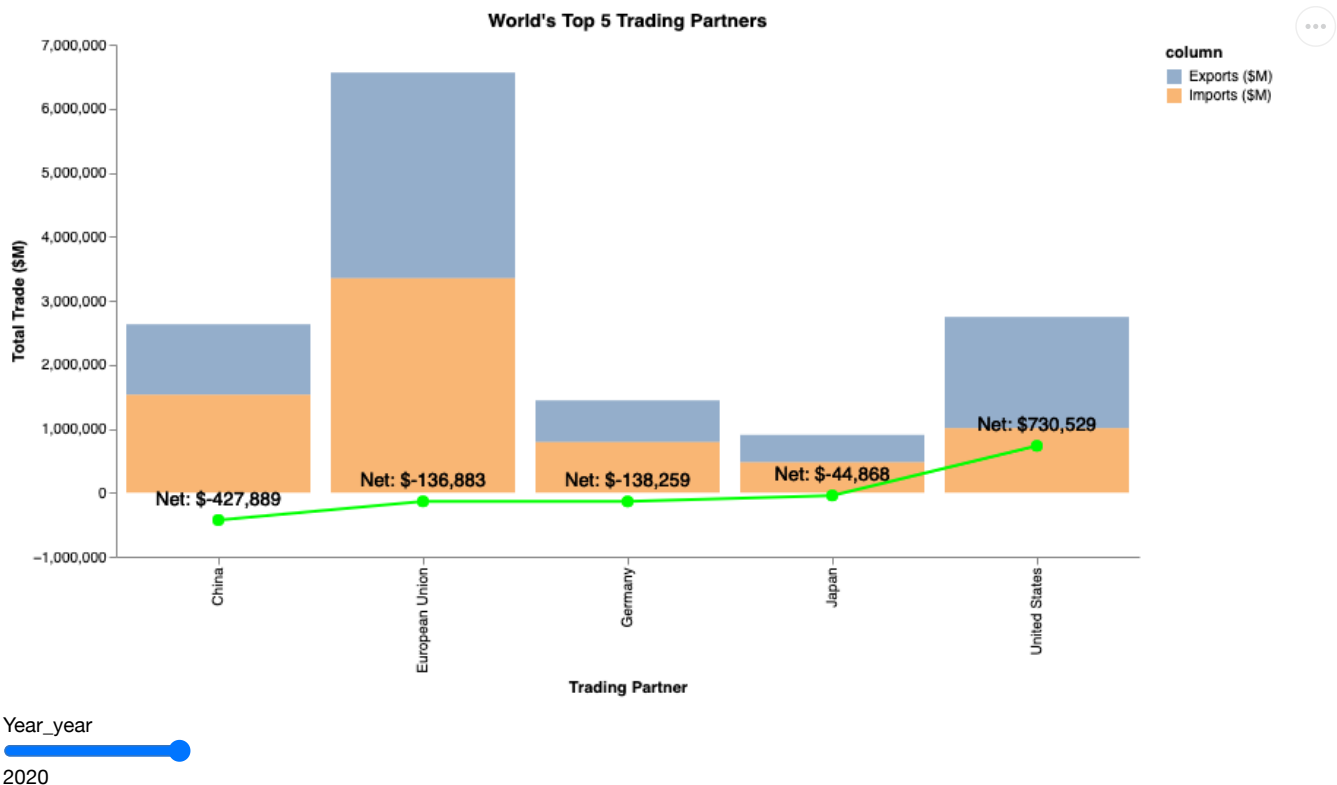
Limitations of data set:

This data set is limited to 20 countries for 6 years. This is due to the difficulty of harvesting, and pre-processing data. These data are sufficiently feature rich to demonstrate hypothesis. However, they lack longtitudinal depth require to draw conclusions about world trade trends over a long term. They do, however, provide a view into trade between the world's 20 largest trading nations over the past six years.

Hypothesis 1:

While not a country, the European Union, has been the world's largest trading block. It has been so for the past 6 years.

Out[2]:

Year_year

2020

What is informative about this view:

The view above shows the top 5 trading nations or blocks in the world with a slider for the years between 2014 and 2020.

Utilizing the slider, a user can see that for each of the years the European Union is consistently the top trading block.

While not directly related to the hypothesis, we see that nations with net negative trade values (related to the world) are nations with which the world is importing more things from. We see that the world is consistently importing more things from China, while consistently exporting more things to the United States. This means china is a net retrainer of trade dollars, while the United States is a net exporter, so to speak, of trade dollars.
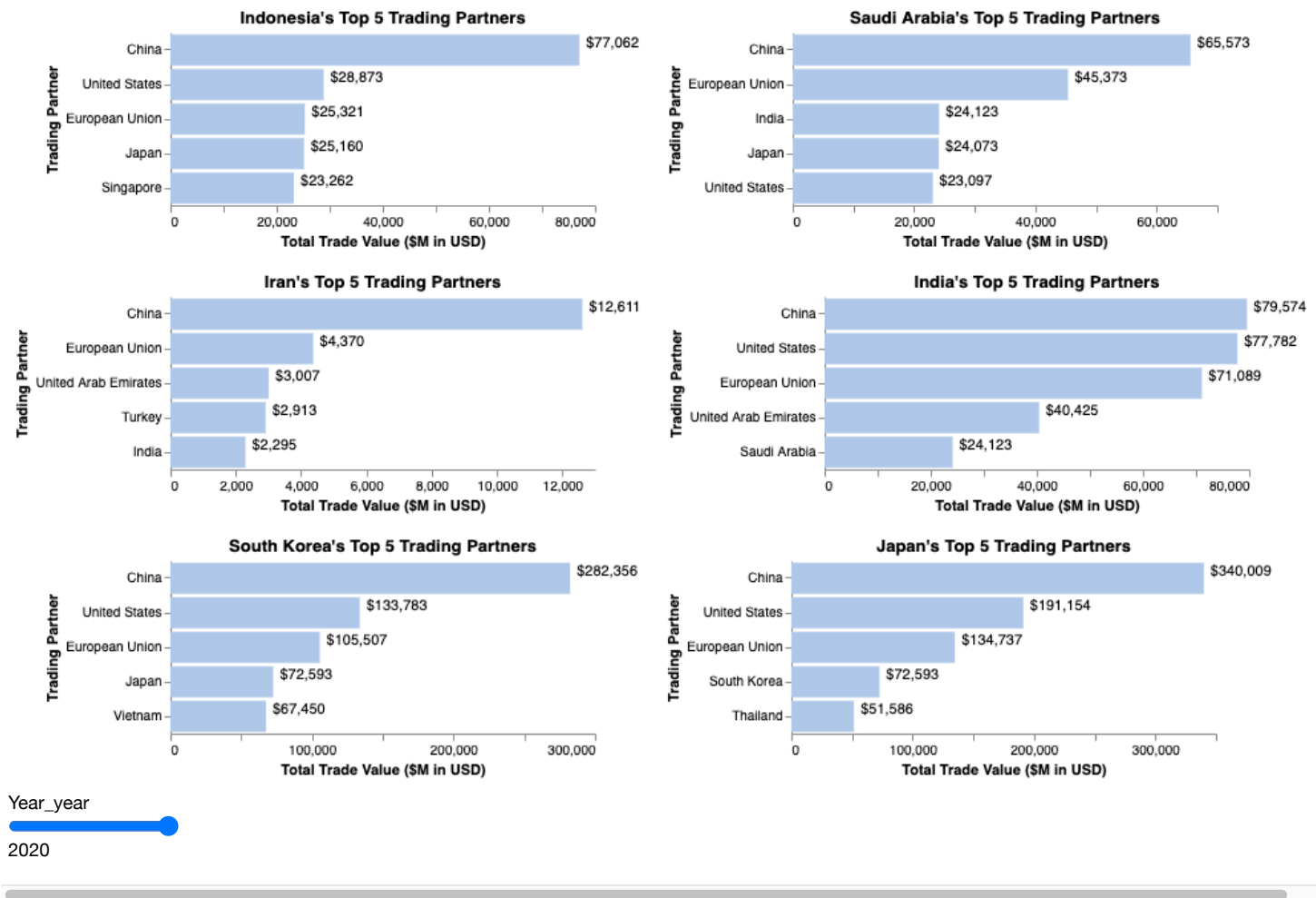
What could be improved about this view:

This view is bordering on being a bit busy, and may be better as a stacked bar chart, or ordered from highest trade, to lowest trade.

Hypothesis 2:

China dominates trade with the Middle East, South Asia, Southeast Asia, and Asia Pacific nations. China has done so for the past 6 years with, although, moving the slider, below will show that the European Union was the top trading partner of Saudi Arabia and India for a few of the years prior to 2016.

Out[3]:

What is informative about this view:

The view above shows 6 nations in Asia, from the Middle East, to the Far East, and it provides a slider to compare different years. We see, that for almost all of these nations, for almost all years in the visualization, China is the top trading partner. In the case of India and Saudi Arabia, we see that in 2014, their leading trading partner block was the European Union, but it switched to China.
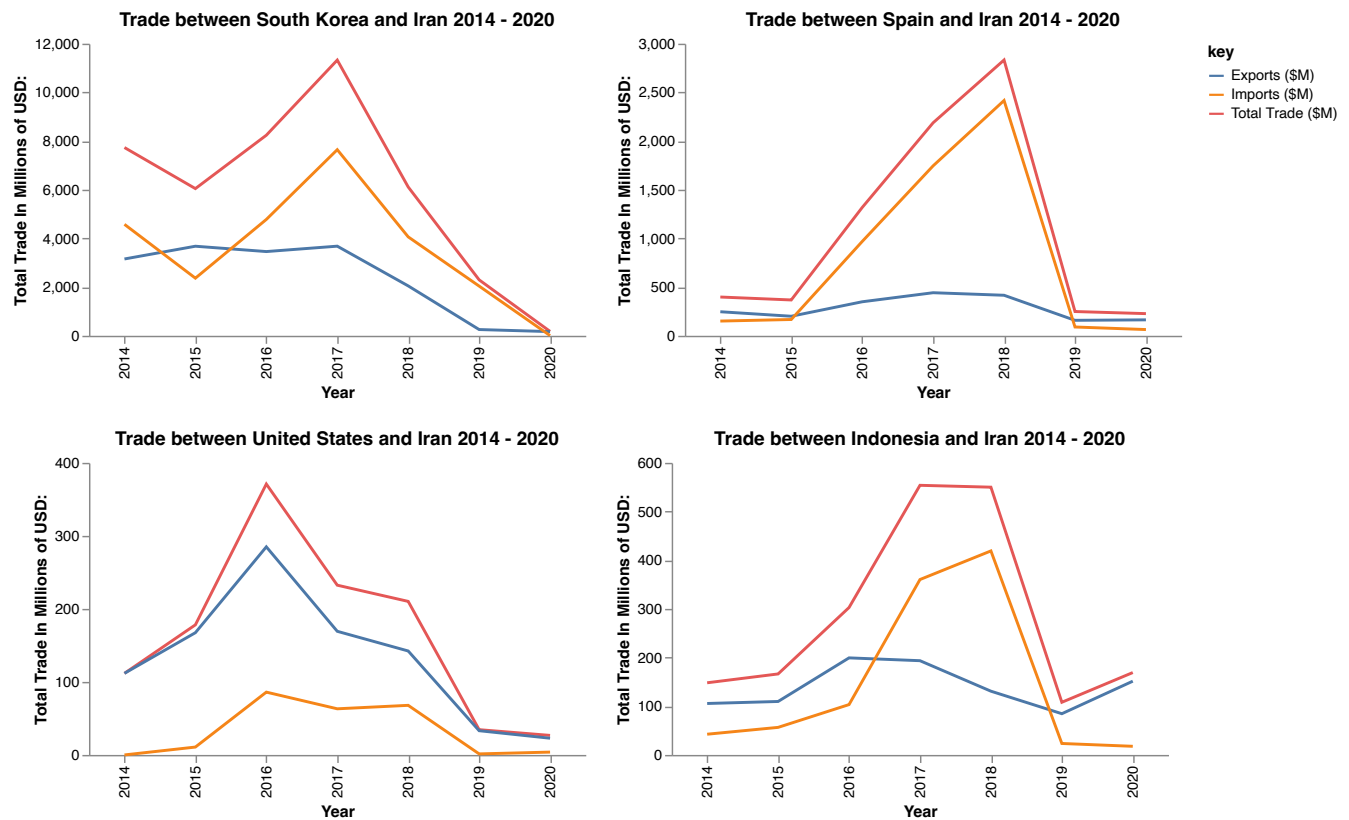
What could be improved about this view:

This view sometimes "wiggles" because of the labels pushing out from side to side. This could be improved.

Hypothesis 3

Geopolitical events can have a drastic impact on a nation's trade, and that nation's ability to buy and sell with the rest of the world.

Trade between South Korea and Iran 2014 - 2020

Trade between Spain and Iran 2014 - 2020

key
— Exports ($M)
— Imports ($M)
— Total Trade ($M)

Trade between United States and Iran 2014 - 2020

Trade between Indonesia and Iran 2014 - 2020

What is informative about this view:

The View above clearly shows the trade impact on Iran as the result of the Trump Administration's withdrawl from the JCPOA, also known as the "Iran Nuclear Deal".

We can see that generally from 2014 (the time of the JCPOA) until 2016-2017 trade between Iran and other nations was increasing drastically.

Once Trump imposed secondary sanctions on people trading with Iran, there was a massive crash in trade with Iran.

What could be improved about this view:

Lines should be thicker, and the tool tip should be improved. The charts could be changed to indicate in some fashion, the date of the cancellation of the JCPOA.

Conclusion:

The data do appear to support the various hypothensis presented.

The challenge with this project is to create a Summary-Zoom-Detail paradign within the application that can communicate effectively to a user.

At this stage, we are still struggling with how to represent different datasets and datum effectively.

Thank you for your instruction during this project.

Best regards,

Don Irwin