World Trade Analysis

Brendan Fruin

brendan@cs.umd.edu

Introduction

With the vast amount of data being collected and made publicly available, individuals from all walks of life have been able to provide analysis, intuitive visualizations and spot trends with the help of tools such as the Microsoft Excel NodeXL template. Using this tool, I have been able to find noteworthy headlines concerning international trade, a few of which are outlined below.

For my analysis, I used the trade data from the Correlates of War Project ([1], [2]) which I verified with the IMF’s dataset ([3]).

The above graph shows my initial dataset using NodeXL where nodes are different countries and edges are imports or exports between the two countries. I quickly noticed that both data for Hong Kong and Germany were missing. Since Germany had the fourth largest GDP in 2009 ([3]), I believed that it was necessary to find at least its major trading partners ([4]) in order to give more accurate results.
Annual exports exceeding $25 billion occur between countries in close proximity and/or with international trade hubs

Above is a map displaying all of the countries that exported or imported at least $25 billion dollar in (consumer goods) from at least one other country in 2009. 29 countries are displayed and labeled at their relative locations and directed edges are used to show the flow of goods over $25 billion from one country to another. Edge thickness, transparency and color are used to show the total number of goods being exported. Thicker, more opaque and blue lines represent the larger export connections between countries where the thinner, more transparent and orange lines show smaller export connections between countries.

On first observation, the countries of China and the United States are quickly noticed and will be termed in this example as international trade hubs. The international trade hubs trade heavily regionally as is seen in the United States in its connections with Mexico and Canada or China’s connections in Asian countries. More importantly in terming these countries international trade hubs, is the fact that these countries trade at least $25 billion with countries across the world. This is quickly seen by China and the United States trading with most of the displayed European countries, South American and Middle Eastern countries along with North America for China and Asia for the United States.
By removing the map underlay and applying the Harel-Koren Fast Multiples mapping algorithm with continent color coding, it can be seen that countries from the same continent are grouped and placed close together. This shows that countries in the same continent mostly do trading exceeding $25 billion within their respective continents or with an international trade hub which are placed in the center of the map due to their high degree.

**The 10 Countries with the Lowest GDP Trade More with the 10 Countries with the Highest GDP than each other**
The trade dataset was merged with the IMF 2009 GDP dataset. The resulting set was then filtered to only show the 10 countries with the highest GDP excluding Germany for which data was missing and the 10 countries with the lowest GDP. The 20 countries were assigned their flag as their node representation and sized by their GDP where the larger GDP results in a larger flag. The countries were then grouped by GDP for which the increments can be seen in the graph.

After some study of the graph, it can be seen that groups with low GDPs have strong connections with groups with high GDPs. Furthermore, there exists only one connection between the smaller GDP countries which can be seen by the small edge from the “GDP < $1 billion” group to the “GDP < $.5 billion” group. This result may be due to insufficient reporting of trade between lower GDP countries or missing data, but it could also be taken as an indicator that lower GDP countries may not have the resources or infrastructure to organize trade between themselves.

It is not just China where the United States sees a large trade deficit

The graph shown in this map is the result of calculating the balance of trade (exports-imports) between the United States and all countries in the dataset. The results are then filtered by relations in which the United States has a trade deficit (imports > exports) exceeding $2 billion and these countries are placed on the map at their respective locations. An edge is drawn from the United States to each of these countries with the total deficit in millions of 2009 USD displayed on the edge. To better visualize this, the edge thickness is varied from a very large deficit and a thick edge as is seen in the China or Canada.
connection to a thin edge as is seen in Ecuador or Finland. The color is also telling of the deficit where red edges show large deficits and orange edges show smaller deficits.

To more clearly see the United States trade deficits, the map underlay has been removed and a spiral graph has been used to show countries in order of the size of the trade deficit starting with Ecuador and ending in China. Continents have been color coded. {Africa is pink, Asia is orange, Europe is blue, North America is purple and South America is brown} As can be seen from this spiral map, there is a huge trade deficit between China and the United States almost 4 times the amount of the next largest deficit seen in Mexico. In the ten countries with the largest trade deficits, three are Asian, two are North American, three are European, one is African and one is South American. This shows that the United States has trade deficits exceeding two billion dollars in all of the continents of the world excluding Australia.
Critique of NodeXL

Positive

- By being an Excel template, it is easy to use for anyone familiar with Excel and does not require any programming experience.
- Can quickly filter data, change edge properties and clean up your data.
- Multiple graph algorithms are built in to the system which allows for different representations of your data. The resulting graph can then be manipulated as you please my moving edges, changing vertex properties or even applying further dynamic filtering.
- Very fast graph updates for smaller datasets.
- Although unused in the above example, automatic network download of social networks such as Twitter.

Negative

- Without dual displays hard to change the underlying Excel spreadsheet and view the graph at the same time.
- Not all processes report the progress of the operation to the user. When I attempted to bundle edges, my computer began processing and appeared unresponsive after time passed and there was no feedback. A progress report is given for newer features such as the grouping.
- Nasty error messages are propagated to the user which should be replaced by a friendly descriptive message. For example, the following error was displayed when a stray character was in the Image column of the Vertex sheet:

![Error message from NodeXL](image)
References


