MERCURY GROUP, PROJECT 3

CRISTIAN SALGUEIRO, HARSHA MAHAJAN, SAHIL KUMAR

NOVEMBER 12, 2020

ECE 59500 SOCIAL NETWORKS W/ MACHINE LEARNING

INDIANA UNIVERSITY PURDUE UNIVERSITY INDIANAPOLIS

The purpose of this project is to demonstrate the impact of day of the week on new COVID-19 cases by country, compare COVID-19 trends using a 7-day moving average, and demonstrate “friendliness” among countries using trade data.

Figure 1 shows the day-of-week trends among the top 7 countries with Active cases. Starting with Sunday the number of active cases reduced on Monday for the US, Spain, Philippines. However, the daily active cases increase throughout the week and then declines from a peak on Friday and Saturday.

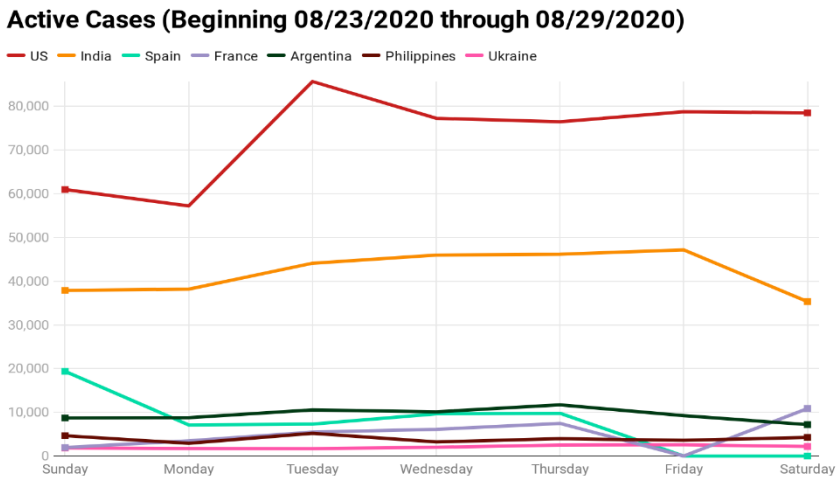

Figure 1 Day of the Week Trends Among Countries

Figure 2(a) shows the graph of the number of tests versus the number of active cases in the United States with a 7-day moving average line. The graph shows that the moving average is not a persuadable measure because it can be seen that there is a significant difference between the average number of active cases and the actual number of active cases reported for a day.

Figure 2(b) shows the graph of the number of tests with the ratio of positive cases in the United States. The graph shows that the average positive ratio does not differ very much from daily positive ratio. Therefore, it can be used as a persuadable measure to notice the weekly positive ratio of COVID cases.

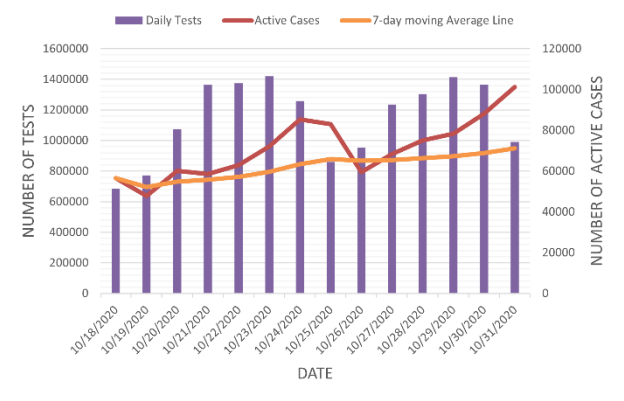

Figure 2(a). Daily Tests and Active Cases

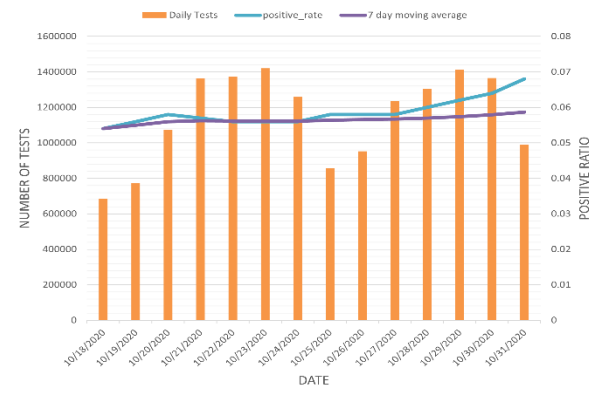

Figure 2(b). Daily Tests and Positive Ratio

Figure 3 below shows a trade network of the top 10 countries for total COVID-19 cases. Russia is not included due to lack of data available from the World Trade Organization (WTO). The link color represents maximum import value between two countries. The network demonstrates the significant economic relationship the United States, which has the most COVID-19 cases of any country, has with the other top 10 countries. It can also be seen that Brazil's economic presence may be a factor in driving COVID-19 spread in South America. Despite the comparatively low value trade interactions Argentina and Colombia have with the other top 10 countries, they both share a significant economic relationship with Brazil.

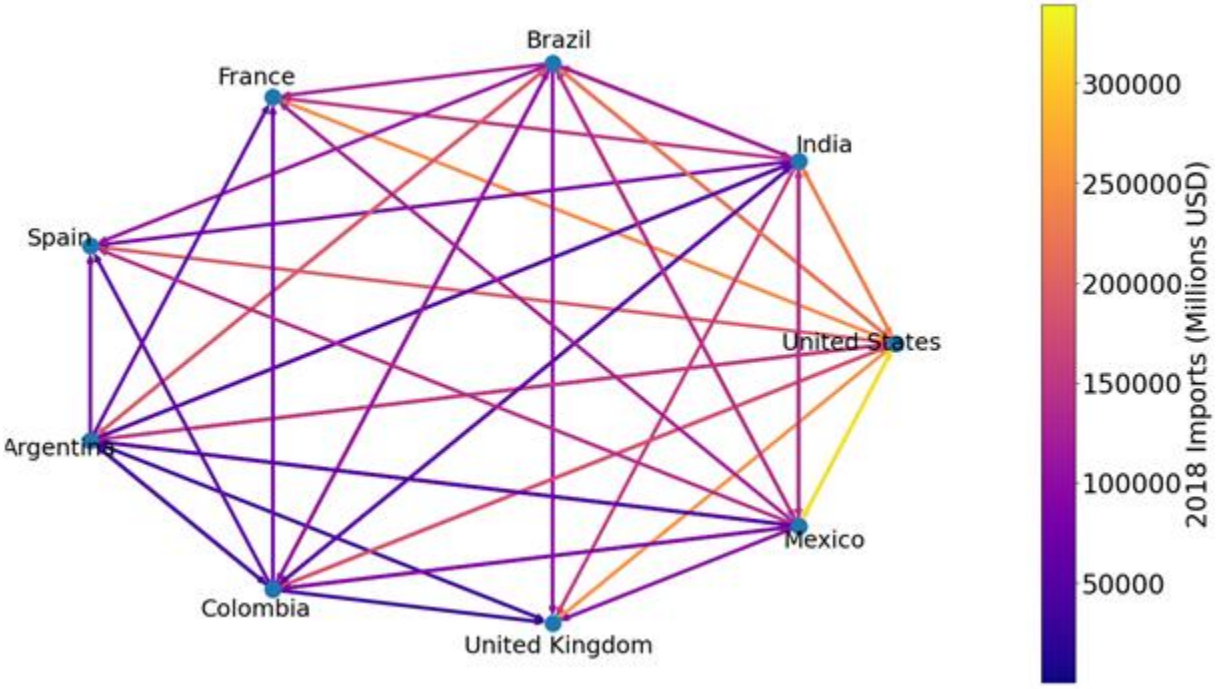

Figure 3. International Trade Network for Top COVID-19 Countries

References:

1. https://www.wto.org/english/res_e/statis_e/trade_datasets_e.htm
2. https://ourworldindata.org/coronavirus-source-data

Code:

```
import pandas as pd
import networkx as nx
import matplotlib.pyplot as plt
import numpy as np
import math
data = pd.read_csv("WTOdata.csv",encoding="latin-1")
top10 = ["United States of
America","India","Brazil","Russia","France","Spain","Argentina","Colombia","United
Kingdom","Mexico",]#"Peru","South
Africa","Italy","Iran","Chile","Germany","Iraq","Bangladesh","Indonesia","Phillippines"]
data = data[data['Reporting Economy'].isin(top10)]
data = data[data['Partner Economy'].isin(top10)]
data = data.loc[(data['Year'] == 2018) & (data['Product/Sector Classification'] == 'Harmonized System')]
data.to_csv("top20Trade.csv")
colname = ['Reporting Economy','Partner Economy', 'Import Value']
out = pd.DataFrame(columns = colname)
for i in top10:
   for j in top10:
      trade = sum(data[(data['Reporting Economy']==i) & (data['Partner Economy']==j)]['Value'])
      trade = int(round(trade/1000000)) #final value in millions USD
      out = out.append({'Reporting Economy':i,'Partner Economy':j,'ImportValue':trade},
ignore_index=True)
out = out.loc[out['Import Value'] !=0]
out.to_csv("Network_Data.csv")
plt.figure(figsize = (15,8))
G = nx.from_pandas_edgelist(out,source = 'Reporting Economy',target = 'Partner
Economy',edge_attr='Import Value',create_using=nx.DiGraph())
pos = nx.circular_layout(G,scale = .5)
nx.draw(G,pos)
label_pos = {}
countries = list(pos.keys()) #create list of countries to offset label position
for k in countries:
label_pos[k] = pos[k]*1.12 #node label offset
nx.draw_networkx_labels(G,label_pos)
imports = nx.get_edge_attributes(G,'Import Value').values()
cval = np.log(np.array(list(imports))) #create logscale for color values
vmin = min(cval)
vmax = max(cval)
cmap = plt.cm.plasma
edges = nx.draw_networkx_edges(G,pos,width = 4,edge_color = cval,edge_cmap =
cmap,vmin=vmin,vmax=vmax)
sm = plt.cm.ScalarMappable(cmap = cmap,norm = plt.Normalize(vmin = math.exp(vmin),vmax =
math.exp(vmax))) #take exponential to map colorbar to real values
clb = plt.colorbar(sm)
clb.set_label('2018 Imports (Millions USD)')
plt.show()
```