Graph Methods and Network Analysis Final Project

Rutgers Business School, Spring 2024

Port Insights: Analyzing Trade Activity for USA

Project By –

JIBRAN SIRAJ NAKHWA

SMITKUMAR JAGDISHBHAI PATEL

SUVAPNIL DUBEY

Index -

1. Introduction

1.1 Introduction of Project

This project delves into the intricate web of trade networks, focusing on the United States' involvement in global trade through sea, land, and air routes. By analyzing trade flow patterns, port connectivity, and dependencies on trade partners, we aim to unravel the dynamics shaping the nation's commerce landscape.

1.2 Aim of the project

The primary aim of this project is to capture the essence of trade analysis for the United States. Through comprehensive exploration of sea, land, and air networks, we seek to understand trade patterns and dynamics, identifying major trade routes and corridors connecting ports domestically and globally.

Major objective we wish to accomplish at the end of the project are –

- Trade flow patterns - What are the major trade routes and corridors connecting ports within the USA and with ports globally?
- Port Connectivity and Centrality - Which ports have the highest degree of connectivity within the trade network?
- Trade Partners and Dependencies – Which countries or regions are the primary trading partners of ports in the USA.

1.3 Data Description

Our datasets contained variable like Port name, Imports, Location, Trade volume, Trade partners and other variable as well. We used the data from 2015 to 2022 for the analysis of trade of USA with different countries through different ports.

1.4 Data Source

To accomplish our objectives, we've gathered data from various reputable sources:

- ustradenumbers.com/ports

- trade.gov/maritime-services-trade-data

- census.gov/foreign-trade/reference/products/catalog/port.html

- census.gov/foreign-trade/balance/c0004.html

These sources provide a wealth of information crucial for analyzing trade patterns, port connectivity, and trade dependencies. Gathering data from diverse sources enables us to compile a comprehensive dataset essential for our analysis.

Our methodology involves meticulous data aggregation and analysis, ensuring accuracy and reliability in our findings. By integrating insights from multiple sources, we aim to provide a holistic understanding of the trade landscape, facilitating informed decision-making for stakeholders involved in trade policy, logistics, and economic development.

2. Data Pre-processing

2.1- Data Cleaning Process –

The project commenced with a meticulous data cleaning process aimed at ensuring the integrity and accuracy of our analysis. We initiated this phase by identifying and addressing null values within the dataset. Additionally, thorough scrutiny was applied to detect any discrepancies or anomalies in the data that could potentially skew our results.

2.2 Processing the data-

Upon initial examination of the dataset, it became apparent that certain columns held the potential for transformation, offering the opportunity to extract more insightful results. Recognizing this, we embarked on a process to enhance the data and elevate the quality of our analysis. One significant aspect of this enhancement was the transformation of specific data fields to optimize their utility in our analysis.

Our efforts extended beyond mere data transformation. We delved into identifying the top trade partners, a crucial aspect of understanding the United States' trade dynamics. To enrich our analysis further, we leveraged additional datasets to merge quarterly data, enhancing the temporal dimension of our analysis. This integration allowed us to capture seasonal variations and trends, providing a more comprehensive understanding of trade dynamics over time.

The amalgamation of various datasets was a meticulous process aimed at synthesizing a holistic view of the trade landscape. We meticulously curated and consolidated disparate sources of information, ensuring that no relevant variables were overlooked. By harnessing the full spectrum of available data, we endeavored to extract actionable insights that would inform strategic decision-making and policy formulation.

In our pursuit of thorough analysis, we prioritized the utilization of all available variables. This approach enabled us to gain a nuanced understanding of the data, uncovering hidden patterns and correlations that might otherwise have remained obscured. By maximizing the richness of the dataset, we aimed to provide comprehensive insights into the intricacies of trade networks within the United States, facilitating informed decision-making for stakeholders across various sectors.

```
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Port                 47 non-null     object
 1   Trade Volume - 2015  47 non-null     int64
 2   Trade Volume - 2016  47 non-null     int64
 3   Trade Volume - 2017  47 non-null     int64
 4   Trade Volume - 2018  47 non-null     int64
 5   Trade Volume - 2019  47 non-null     int64
 6   Trade Volume - 2020  47 non-null     int64
 7   Trade Volume - 2021  47 non-null     int64
 8   Trade Volume - 2022  47 non-null     int64
```

```
Data columns (total 5 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Trading Partners - 1  47 non-null     object
 1   Trading Partners - 2  47 non-null     object
 2   Trading Partners - 3  47 non-null     object
 3   Trading Partners - 4  47 non-null     object
 4   Trading Partners - 5  47 non-null     object
dtypes: object(5)
memory usage: 2.0+ KB
```

```
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   Port       47 non-null     object
 1   Location   47 non-null     object
 2   Port Type  47 non-null     object
dtypes: object(3)
memory usage: 1.2+ KB
```

```
Column Names: Index(['Port', 'Location', 'Port Type', '15-Mar', '15-Jun', '15-Sep', '15-Dec',
       '16-Mar', '16-Jun', '16-Sep', '16-Dec', '17-Mar', '17-Jun', '17-Sep',
       '17-Dec', '18-Mar', '18-Jun', '18-Sep', '18-Dec', '19-Mar', '19-Jun',
       '19-Sep', '19-Dec', '20-Mar', '20-Jun', '20-Sep', '20-Dec', '21-Mar',
       '21-Jun', '21-Sep', '21-Dec', '22-Mar', '22-Jun', '22-Sep', '22-Dec',
       'Trading Partners - 1', 'Trading Partners - 2', 'Trading Partners - 3',
       'Trading Partners - 4', 'Trading Partners - 5', 'Trade Volume - 2015',
       'Trade Volume - 2016', 'Trade Volume - 2017', 'Trade Volume - 2018',
       'Trade Volume - 2019', 'Trade Volume - 2020', 'Trade Volume - 2021',
       'Trade Volume - 2022', 'Trade Volume - Partner 1 - 2015',
       'Trade Volume - Partner 1 - 2016', 'Trade Volume - Partner 1 - 2017',
       'Trade Volume - Partner 1 - 2018', 'Trade Volume - Partner 1 - 2019',
       'Trade Volume - Partner 1 - 2020', 'Trade Volume - Partner 1 - 2021',
       'Trade Volume - Partner 1 - 2022', 'Trade Volume - Partner 2 - 2015',
       'Trade Volume - Partner 2 - 2016', 'Trade Volume - Partner 2 - 2017',
       'Trade Volume - Partner 2 - 2018', 'Trade Volume - Partner 2 - 2019',
       'Trade Volume - Partner 2 - 2020', 'Trade Volume - Partner 2 - 2021',
       'Trade Volume - Partner 2 - 2022', 'Trade Volume - Partner 3 - 2015',
       'Trade Volume - Partner 3 - 2016', 'Trade Volume - Partner 3 - 2017',
       'Trade Volume - Partner 3 - 2018', 'Trade Volume - Partner 3 - 2019',
       'Trade Volume - Partner 3 - 2020', 'Trade Volume - Partner 3 - 2021',
       'Trade Volume - Partner 3 - 2022', 'Trade Volume - Partner 4 - 2015',
       'Trade Volume - Partner 4 - 2016', 'Trade Volume - Partner 4 - 2017',
       'Trade Volume - Partner 4 - 2018', 'Trade Volume - Partner 4 - 2019',
...
Trade Volume - Partner 5 - 2020    0
Trade Volume - Partner 5 - 2021    0
Trade Volume - Partner 5 - 2022    0
```


Figure 1 Description of data sets used and combined together

3. Exploratory Data Analysis

3.1 Quarterly Trade volume

```
plt.figure(figsize=(25, 25))
plt.plot(data_2['Year'], data_2['Imports'], marker='o', linestyle='-')
plt.title('Imports Over Time')
plt.xlabel('Year')
plt.ylabel('Imports')
plt.grid(True)
plt.show()
```

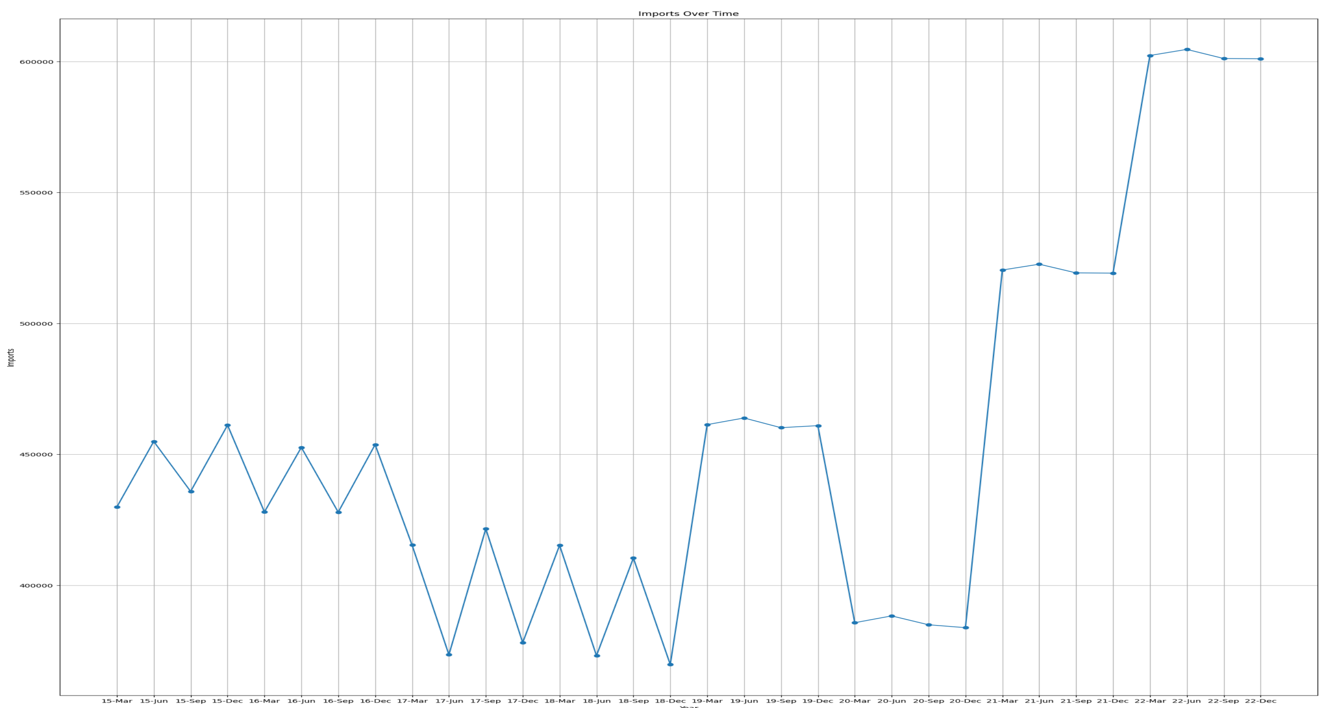

Figure 2 The quarterly imports volume for U.S

The graph depicts the trade volume in terms of value ($) across the different types of ports in United states. The graph particularly depicts the **trade volume of imports** for the United States. The data shows a clear dip in the year 2022 since it was the peak covid period which resulted in decline in trade.

3.2 Trade volume over the years with respect to port type

```
df = pd.DataFrame(data_port_type)
df.set_index('Year', inplace=True)
import matplotlib.pyplot as plt

plt.figure(figsize=(10, 6))
plt.plot(df.index, df['Air'], marker='o', label='Air')
plt.plot(df.index, df['Land'], marker='o', label='Land')
plt.plot(df.index, df['Sea'], marker='o', label='Sea')

plt.xlabel('Year')
plt.ylabel('Trade Volume (In trillions)')
plt.title('Trade Volume Over Years')
plt.legend()
plt.grid(True)


plt.tight_layout()
plt.show()
```

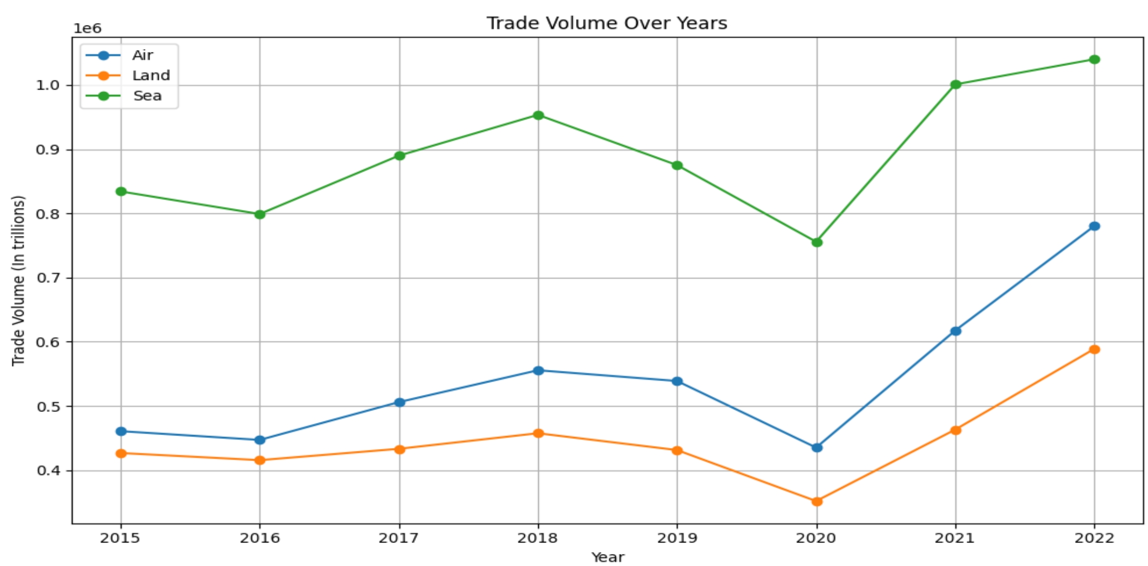

Figure 3 Trade volume of the United States in terms of Value from different sources.

Import through sea has always been higher as compared to land and air. Although post covid we see an increase in trade through land and air as well.

3.3 Types of Ports

```
Data_connected.groupby('Port Type').size().plot(kind='barh',
color=sns.palettes.mpl_palette('Dark2'))
plt.gca().spines[['top', 'right',]].set_visible(False)
```

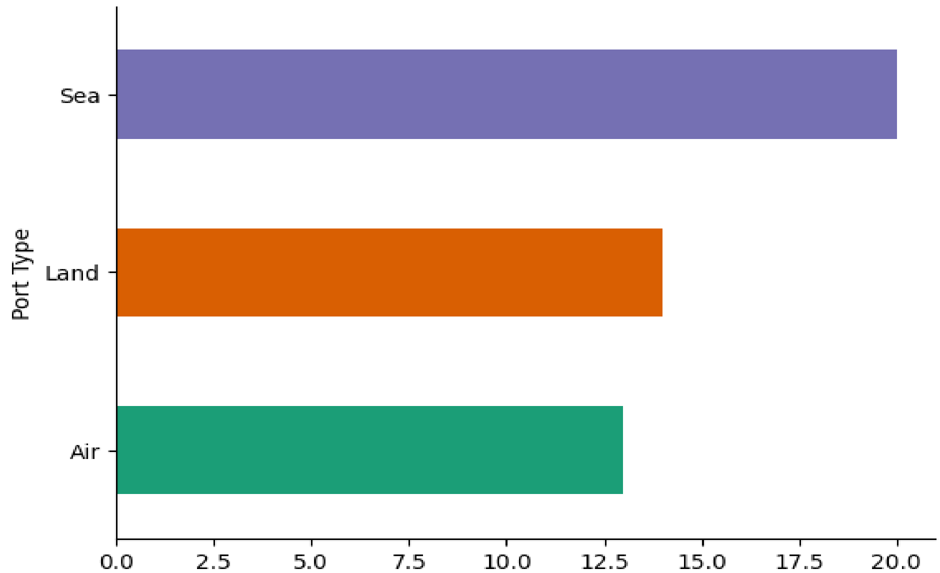

Figure 4 Sea has the highest number of ports.

3.4 Major Trade partner

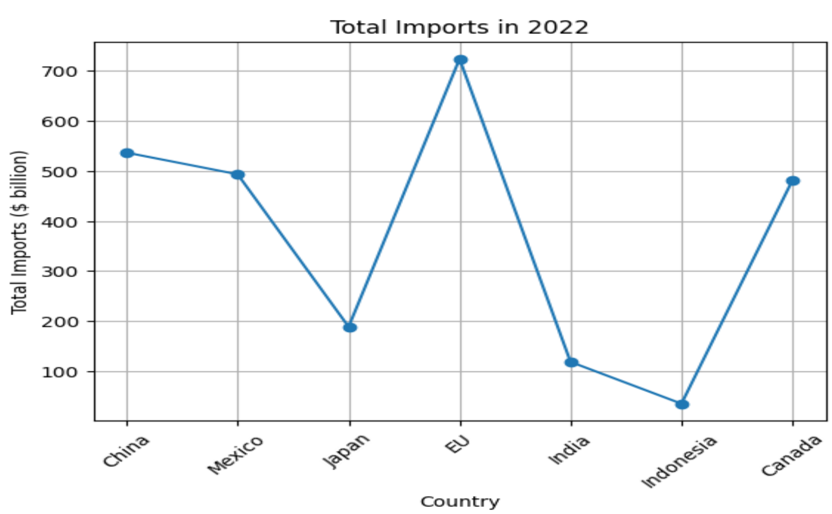

Figure 5. Major Import partners for USA in 2022

3.5 Port Locations

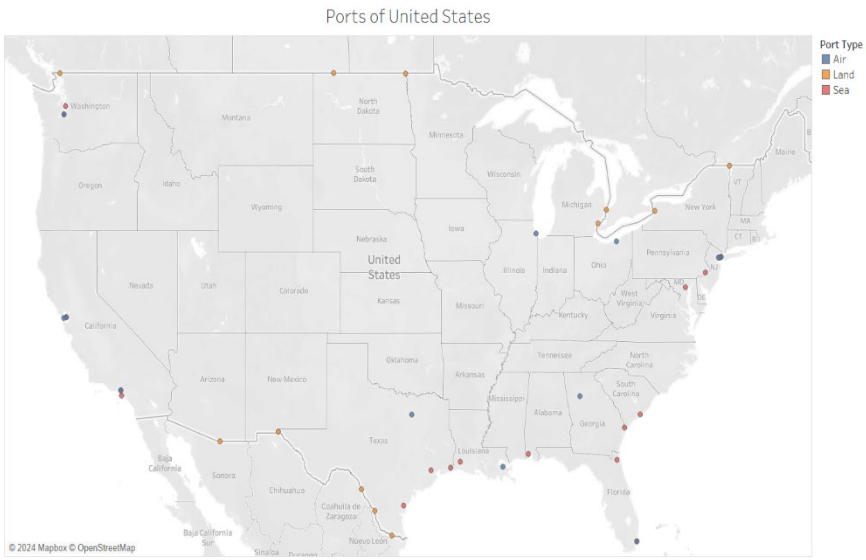

Figure 6 Locations of ports.

A graphical representation vividly illustrates the primary entry points for trade. Predominantly, land-based entry points are concentrated along the border with Mexico, while major seaports cluster predominantly along the West Coast.

3.6 – Top 3 trading partners and Trade type

```
plt.subplots(figsize=(8, 8))
df_2dhist = pd.DataFrame({
    x_label: grp['Trading Partners - 1'].value_counts()
    for x_label, grp in Data_connected.groupby('Port Type')
})
sns.heatmap(df_2dhist, cmap='viridis')
plt.xlabel('Port Type')
_ = plt.ylabel('Trading Partners - 1')
```

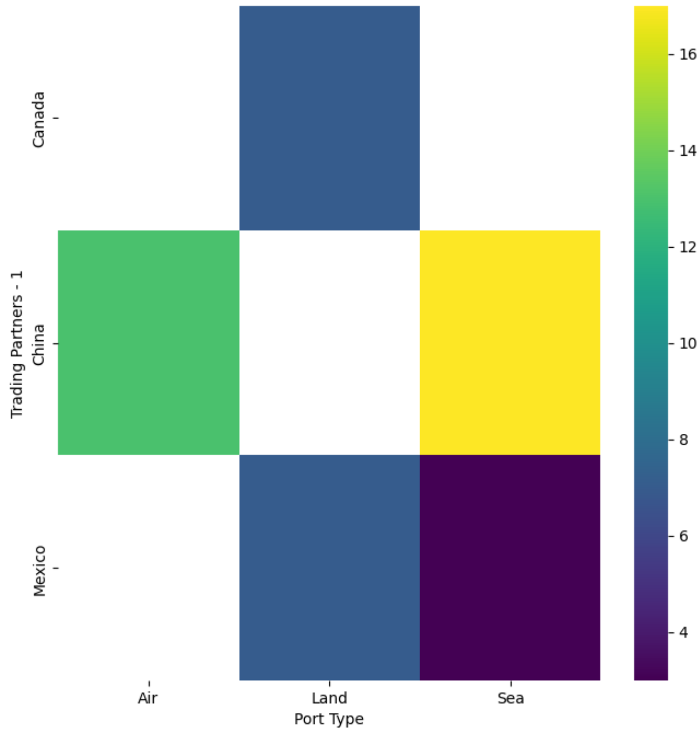

Figure 7 Top 3 major trading partners.

Top 3 import partners for US are China, Mexico and Canada. Mexico and Canada being at the borders of US use more land ports than any other for trade, while trade with China relies heavily upon Sea and Air routes.

4. Network Analysis

The subsequent phase of the project involved the establishment of a network and the creation of graphical representations. This entailed delineating connections and constructing graphs to visually depict the intricate relationships within the trade landscape.

The ensuing stage of the intricate endeavor necessitated the meticulous establishment of an interconnected network coupled with the adroit crafting of graphical representations. This multifaceted undertaking entailed the precise delineation of intricate connections and the construction of sophisticated graphs, thereby enabling a visually compelling depiction of the labyrinthine relationships interwoven within the dynamic tapestry of the trade landscape. The overarching objective was to distill the complexity inherent in this domain into a cohesive and readily comprehensible format, facilitating a deeper understanding and more informed decision-making processes.

4.1 Network Construction

4.1.1 Initial Network graph

The initial graph was constructed to find out the underlying patterns and understand the connection between different nodes and edges.

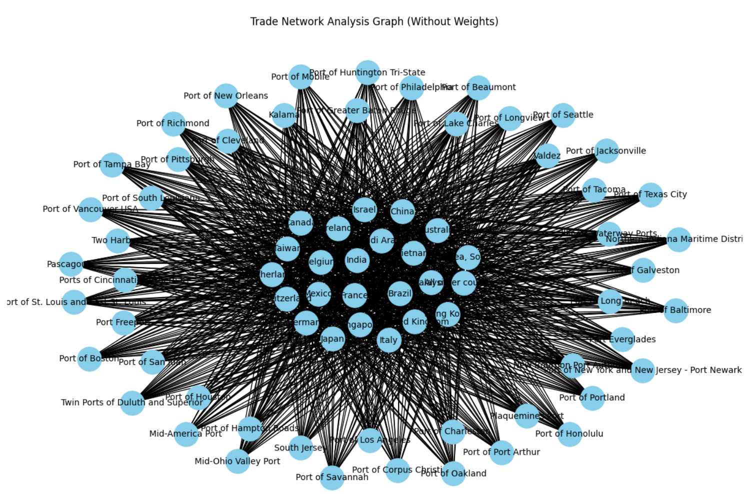

Figure 8 Initial Graph of the data

In the initial dataset, we began with information on the trade relationships between various countries and ports. Here's a breakdown of the initial data:

- Total Countries: 24

This indicates that the dataset covers trade interactions involving 24 different countries.

- Total Ports: 50

This denotes the total number of ports included in the dataset. Ports serve as crucial hubs for international trade, facilitating the movement of goods between countries via maritime routes. Each port in the dataset represents a node in the trade network.

- Total Edges: 1200

In the context of network analysis, edges refer to the connections or links between nodes (in this case, ports). The total number of edges indicates the total number of trade relationships or connections between pairs of ports in the dataset. For example, if goods are traded between Port A and Port B, there exists an edge between these two ports in the trade network. The total number of edges being 1200 suggests that not all possible trade relationships between ports are present in the dataset, which could be due to various factors such as data availability or trade volume.

In summary, the initial data provides insights into the extent of trade relationships captured in the dataset, including the number of countries involved, the ports facilitating trade, and the connections between these ports. This information forms the basis for further analysis, such as network visualization and identification of trade patterns and dependencies.

4.1.2 Pruning the Graph

In our analysis, we encountered a graph with a multitude of nodes and edges, presenting a complex web of trade relationships. To enhance clarity and focus on the most pertinent connections, we embarked on a process of pruning the graph. Our objective was to distill the essential trade relationships by removing non-essential ports and countries where the trade value was not significant.

By selectively removing nodes and edges deemed less critical to our analysis, we refined the graph to better highlight the key trade pathways and dependencies. This strategic pruning allowed us to streamline the visualization and concentrate on the most relevant trade relationships.

Subsequently, we reconstructed the graph, incorporating only the ports and countries that played a significant role in the trade network. This revamped graph provided enhanced clarity and insight into the essential trade dynamics, enabling us to discern patterns and relationships with greater precision.

Overall, this iterative process of graph pruning, and reconstruction yielded a refined representation of the trade network, facilitating a deeper understanding of the critical relationships and driving forces shaping global trade.

```
import networkx as nx
import matplotlib.pyplot as plt

G = nx.Graph()
ports = Data_connected['Port'].unique()
G.add_nodes_from(ports)
for _, row in Data_connected.iterrows():
    port = row['Port']
    partners = row[['Trading Partners - 1', 'Trading Partners - 2', 'Trading
Partners - 3', 'Trading Partners - 4', 'Trading Partners - 5']].values
    for partner in partners:
        if pd.notna(partner):
            G.add_edge(port, partner)

plt.figure(figsize=(12, 8))
nx.draw(G, with_labels=True, node_color='skyblue', node_size=1000,
edge_color='grey', linewidths=1, font_size=10)
plt.title("Network of Ports and Trading Partners")
plt.show()
```

Constructed Graph –

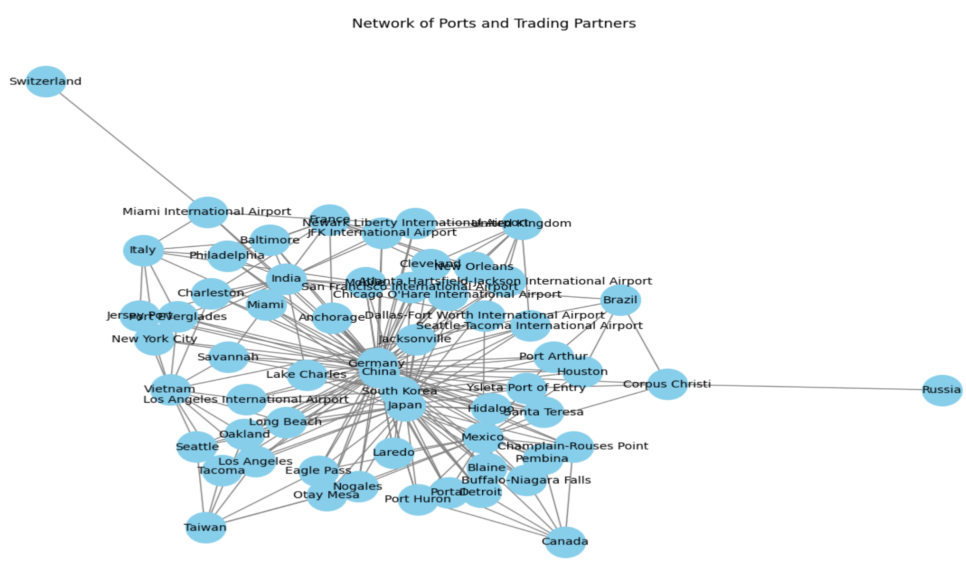

Figure 9 The constructed graph of countries and ports.

4.1.3 – Creating a circular graph with Countries and Ports

```
ports = data['Port'].unique()
trading_partners = pd.unique(data[['Trading Partners - 1', 'Trading Partners -
2', 'Trading Partners - 3', 'Trading Partners - 4', 'Trading Partners -
5']].values.ravel('K'))
G = nx.Graph()
G.add_nodes_from(ports, layer='outer')
G.add_nodes_from(trading_partners, layer='inner')


for index, row in data.iterrows():
    port = row['Port']

    for i in range(1, 6):
        trading_partner = row[f'Trading Partners - {i}']
        if pd.notna(trading_partner):

            trade_volumes = row[[f'Trade Volume - Partner {i} - {year}' for year
in range(2015, 2023)]].sum()
            G.add_edge(port, trading_partner, weight=trade_volumes)


outer_circle = nx.circular_layout([node for node in G.nodes if
G.nodes[node]['layer'] == 'outer'], scale=1)
inner_circle = nx.circular_layout([node for node in G.nodes if
G.nodes[node]['layer'] == 'inner'], scale=0.5)
pos = {**outer_circle, **inner_circle}
plt.figure(figsize=(15, 15))
nx.draw_networkx_nodes(G, pos, nodelist=ports, node_color='blue', node_size=100,
alpha=0.7, label='Ports')
nx.draw_networkx_nodes(G, pos, nodelist=trading_partners, node_color='red',
node_size=50, alpha=0.7, label='Trading Partners')
nx.draw_networkx_edges(G, pos, edge_color='grey', width=0.5, alpha=0.5)
nx.draw_networkx_labels(G, pos, font_size=8, font_weight='bold',
font_color='black')

plt.title('Trade Network Graph')
plt.legend()
plt.axis('off')
plt.show()
print("Number of nodes (ports and trading partners):", G.number_of_nodes())
print("Number of edges (trade relationships):", G.number_of_edges())
```

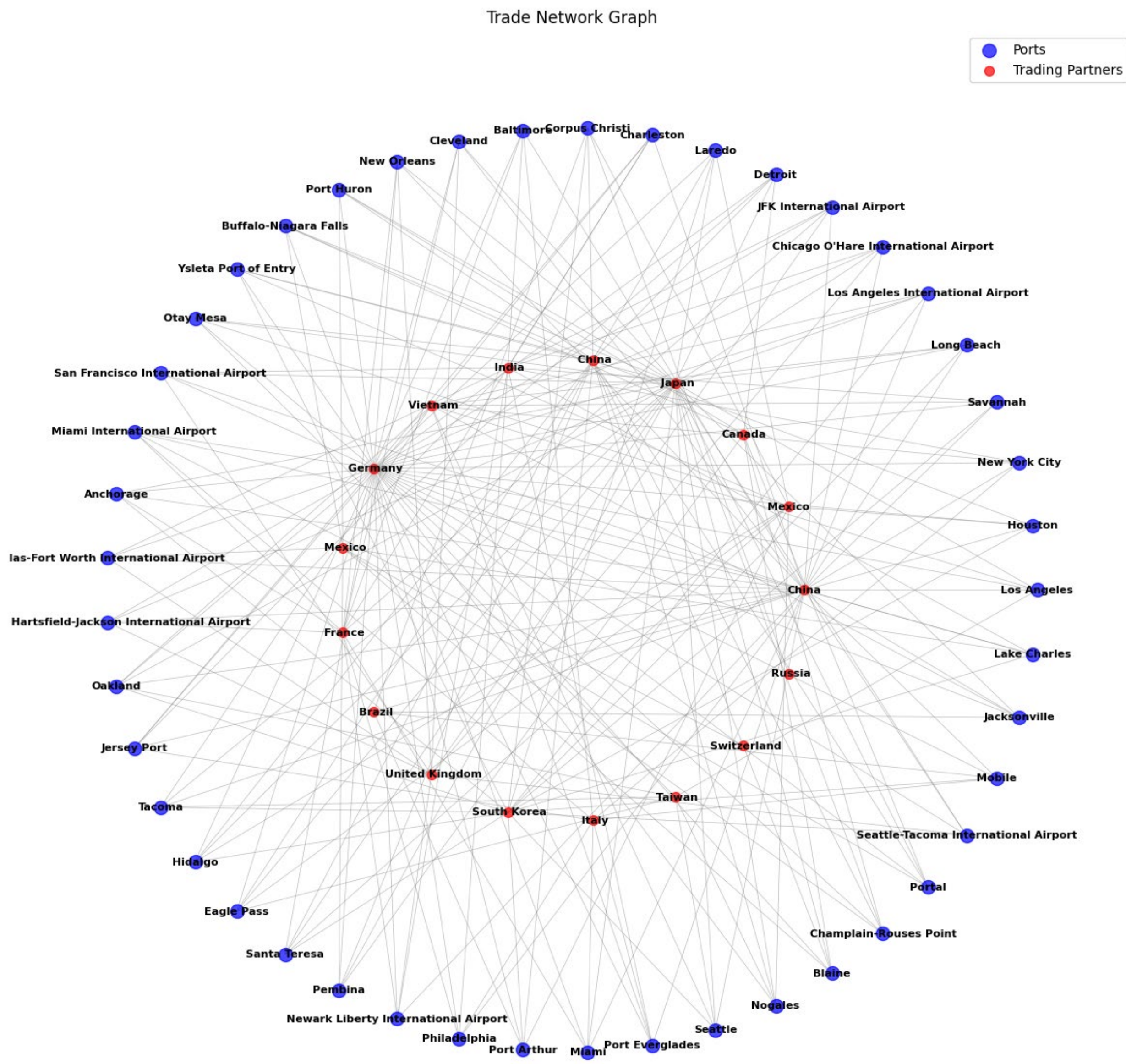

Figure 10 - The constructed graph.

This graph represents a trade network, where the blue nodes represent ports, and the red nodes represent trading partners or countries. The graph shows the complex interconnections and relationships between different ports and their trading partners in a circular form.

The ports are typically major international airports or maritime ports located in various cities and countries across the world. Some of the prominent ports shown in the graph include Los Angeles

International Airport, San Francisco International Airport, Miami International Airport, JFK International Airport, and Newark Liberty International Airport.

The red nodes represent the trading partners or countries that these ports engage with for trade and commerce. Countries like China, Japan, India, Canada, Mexico, Germany, France, Russia, United Kingdom, South Korea, Italy, and others are shown as trading partners.

The edges or lines connecting the ports and trading partners indicate the existence of trade relationships between them. The more connections a port or country has, the more integrated it is within the global trade network.

Number of Nodes, i.e **ports and trading partners is now reduced to 61** and the edges, **i.e trade relationships are changed to 226 in the graph.**

4.1.4- Creating a circular graph with Countries, Ports and Port locations

```
ports = data['Port'].unique()
trading_partners = pd.unique(data[['Trading Partners - 1', 'Trading Partners - 2', 'Trading Partners - 3', 'Trading Partners - 4', 'Trading Partners - 5']].values.ravel('K'))
locations = data['Location'].unique()

# Initialize your graph
G = nx.Graph()

# Add nodes for ports, trading partners, and locations
G.add_nodes_from(ports, layer='outer')
G.add_nodes_from(trading_partners, layer='inner')
G.add_nodes_from(locations, layer='location')

# Add edges with weights based on trade volumes
for index, row in data.iterrows():
    port = row['Port']
    location = row['Location']
    # Add an edge between the port and its location
    G.add_edge(port, location)
    for i in range(1, 6):
        trading_partner = row[f'Trading Partners - {i}']
        if pd.notna(trading_partner):
            # Sum the trade volumes for the years 2015 to 2023
            trade_volumes = row[[f'Trade Volume - Partner {i} - {year}' for year in range(2015, 2023)]].sum()
            # Add an edge with the cumulative trade volume as weight
```

```python
            G.add_edge(port, trading_partner, weight=trade_volumes)

# Create the layout for the nodes
outer_circle_ports = nx.circular_layout([node for node in G.nodes if
G.nodes[node]['layer'] == 'outer'], scale=1)
inner_circle_trading_partners = nx.circular_layout([node for node in G.nodes if
G.nodes[node]['layer'] == 'inner'], scale=0.5)
outer_circle_locations = nx.circular_layout([node for node in G.nodes if
G.nodes[node]['layer'] == 'location'], scale=1.5)

# Combine the positions into one dictionary
pos = {**outer_circle_ports, **inner_circle_trading_partners,
**outer_circle_locations}

plt.figure(figsize=(20, 20))
nx.draw_networkx_nodes(G, pos, nodelist=ports, node_color='blue', node_size=100,
alpha=0.7, label='Ports')
nx.draw_networkx_nodes(G, pos, nodelist=trading_partners, node_color='red',
node_size=50, alpha=0.7, label='Trading Partners')
nx.draw_networkx_nodes(G, pos, nodelist=locations, node_color='green',
node_size=50, alpha=0.7, label='Locations')
nx.draw_networkx_edges(G, pos, edge_color='grey', width=0.5, alpha=0.5)
nx.draw_networkx_labels(G, pos, font_size=8, font_weight='bold',
font_color='black')

plt.title('Trade Network Graph with Locations')
plt.legend()
plt.axis('off')
plt.show()

num_nodes = G.number_of_nodes()

num_edges = G.number_of_edges()

print(f"The graph has **{num_nodes} nodes** and **{num_edges} edges**.")
```

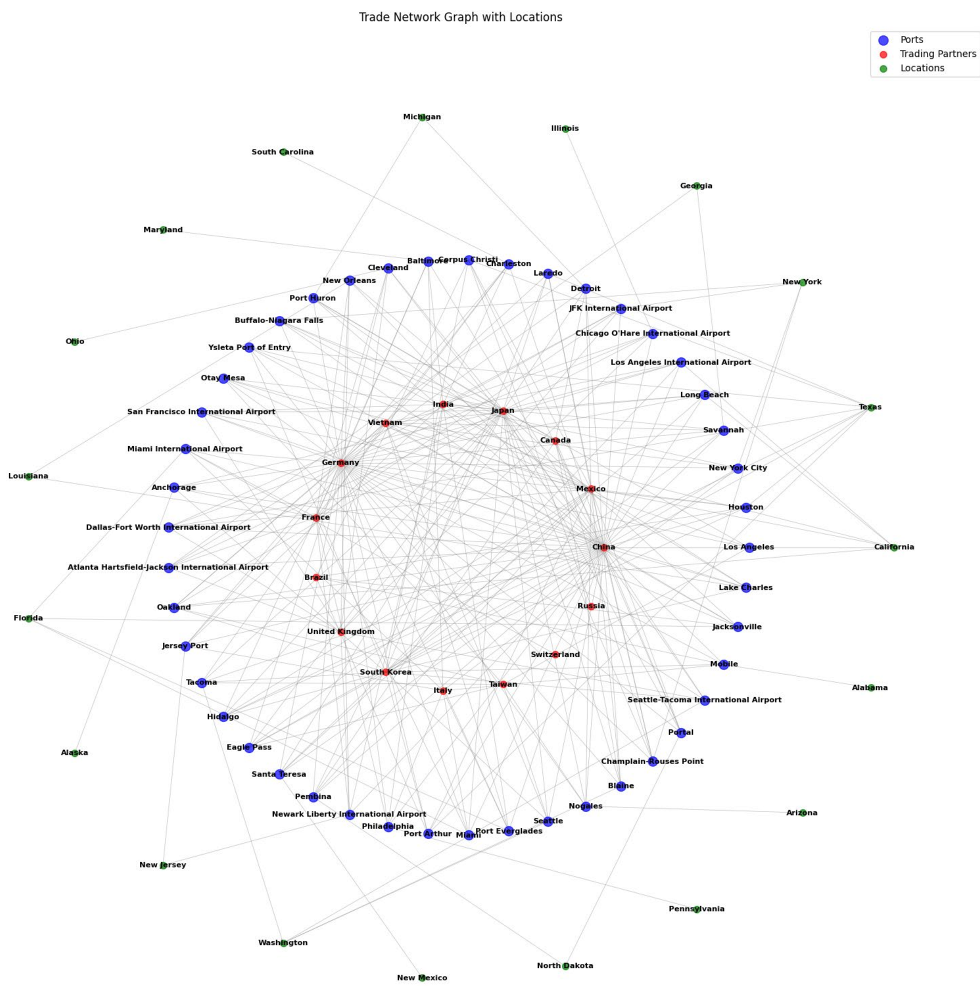

Figure 11 Network graph of Ports, Ports location and Trading Partners.

The expanded trade network graph now **comprises 79 nodes (ports, ports location and trading partners)** and **271 edges (connections between them)**. By incorporating additional port locations into the analysis, new significant nodes have been identified. These newly uncovered key nodes can facilitate a deeper understanding of the relationships and interdependencies among the various nodes within the trade network.

4.2 Network Analysis Metrics

4.2.1- Degree Centrality:

China (0.7627):

China has the highest degree centrality in the network, which means it has the highest number of direct connections to other ports or nodes. A high degree centrality indicates that China plays a central role in the network, potentially acting as a major hub or gateway for trade, transportation, or other interactions within the network.

Germany (0.7119):

Germany has the second-highest degree centrality, suggesting that it is also highly connected within the network. Like China, Germany likely serves as a significant hub, facilitating numerous direct connections and interactions with other ports or countries.

Japan (0.4915):

Japan ranks third in terms of degree centrality, indicating a substantial number of direct connections to other nodes in the network. This high degree centrality suggests that Japan is an important player in the network, potentially due to its strong trade and economic ties with various countries and ports.

Mexico (0.3220):

With the fourth-highest degree centrality, Mexico maintains a considerable number of direct connections within the network. This could be attributed to its geographic location, as well as its trade relationships with other countries, particularly those in North America.

South Korea (0.2881):

South Korea rounds out the top 5 ports based on degree centrality, indicating a relatively high number of direct connections within the network. This could be influenced by South Korea's strong export-oriented economy and its trade partnerships with various countries in the region and beyond.

4.2.2- Betweenness Centrality:

China (0.3477):

China has the highest betweenness centrality in the network, indicating that it lies on a significant number of shortest paths between other ports or nodes. This means that a substantial amount of flow or traffic within the network likely passes through China, making it a crucial bridge or gatekeeper for the exchange of resources, information, or goods.

Germany (0.2671):

Germany ranks second in terms of betweenness centrality, suggesting that it also acts as an important bridge or gateway within the network. A high betweenness centrality implies that Germany facilitates the flow of resources or information between other ports, potentially due to its strategic geographic location or its well-established trade and transportation networks.

Japan (0.0992):

With the third highest betweenness centrality, Japan plays a significant role in connecting other ports or nodes within the network. While its betweenness centrality is lower than China and Germany, Japan still serves as a bridge or intermediary for the flow of resources or information between various parts of the network.

Mexico (0.0423):

Mexico's betweenness centrality indicates that it acts as a bridge or gatekeeper, facilitating the flow of resources or information between other ports or nodes in the network. This could be attributed to its geographic location, serving as a link between North and Central America, as well as its trade relationships with various countries.

Corpus Christi (0.0391):

Corpus Christi, a port located in Texas, has the fifth highest betweenness centrality in the network. Despite being a specific port rather than a country, its relatively high betweenness centrality suggests that it plays a crucial role in connecting other ports or nodes within the network, potentially due to its strategic location or its role in facilitating trade or transportation.

Ports with high betweenness centrality are essential for the efficient flow of resources, information, or goods within the network. They act as bridges or intermediaries, facilitating connections between other ports or nodes that may not have direct links. These ports can potentially exercise control or influence over the flow within the network, making them strategically important for maintaining connectivity and ensuring the smooth operation of the network as a whole.

4.2.3. Closeness Centrality:

China (0.8082):

China has the highest closeness centrality in the network, indicating that it has the shortest average distance to all other ports or nodes. This means that China is closely connected to most other ports, either directly or through a small number of intermediaries. A high closeness centrality suggests that China can efficiently reach or influence other ports within the network.

Germany (0.7284):

Germany ranks second in terms of closeness centrality, implying that it has a relatively short average distance to other ports in the network. This could be attributed to Germany's central geographic location and well-developed transportation and trade infrastructure, allowing it to maintain close connections with various ports across the network.

Japan (0.5413):

With the third-highest closeness centrality, Japan is also relatively close to other ports in the network. While its closeness centrality is lower than China and Germany, Japan still maintains a reasonable proximity to other nodes, facilitating efficient communication, transportation, or exchange of resources within the network.

Oakland (0.4876):

Oakland, a port located in California, has the fourth-highest closeness centrality in the network. This suggests that Oakland is well-connected to other ports, with relatively short average distances separating it from other nodes. This could be due to its strategic location on the West Coast of the United States and its role as a major shipping hub.

Los Angeles (0.4797):

The port of Los Angeles ranks fifth in terms of closeness centrality, indicating that it has a relatively short average distance to other ports in the network. Like Oakland, its proximity to other nodes could be attributed to its location on the West Coast and its significance as a major port for international trade and transportation.

Ports with high closeness centrality are well-positioned within the network, with efficient access to other ports or nodes. This could translate to faster and more direct communication, transportation, or exchange of resources within the network. Closeness centrality is particularly important for ports that need to maintain close connections with various partners or destinations, as it can facilitate timely and efficient operations.

4.2.4- Community Detection:

The analysis identified four distinct communities within the network:

Community 0: This community includes major ports such as Los Angeles, New York City, Savannah, Long Beach, and JFK International Airport, along with countries like China, Vietnam, Taiwan, and India.

Community 1: This community consists of ports like Houston, Laredo, Corpus Christi, and Anchorage, as well as countries like Japan, Mexico, Brazil, South Korea, and Russia.

Community 2: This community includes ports like Chicago O'Hare International Airport, Miami International Airport, and Dallas-Fort Worth International Airport, along with countries like Germany, Italy, United Kingdom, France, and Switzerland.

Community 3: This community comprises ports located near the Canada-US border, such as Detroit, Port Huron, Buffalo-Niagara Falls, and Pembina, along with Canada itself.

These communities may indicate stronger connections or interactions among the ports and countries within each group, potentially driven by factors such as geographic proximity, trade agreements, or transportation networks.

4.2.5 – Eigenvector centrality:

Germany (0.3872):

Germany has the highest eigenvector centrality in the network, indicating that it is connected to other highly connected and influential nodes. This suggests that Germany is a central and important player within the network, likely due to its strong economic and trade ties with other major ports and countries.

China (0.3741):

China ranks second in terms of eigenvector centrality, meaning that it is also connected to many other well-connected and influential nodes in the network. This reflects China's significant role in global trade and its strong connections with various major ports and countries.

Japan (0.2540):

With the third-highest eigenvector centrality, Japan is linked to other influential nodes within the network. This could be attributed to Japan's position as a major economic and trading power, maintaining strong ties with other important ports and countries.

South Korea (0.2105):

South Korea's eigenvector centrality indicates that it is connected to other influential nodes in the network, likely due to its strong export-oriented economy and trade relationships with various countries and ports.

Mexico (0.1610):

Mexico ranks fifth in terms of eigenvector centrality, suggesting that it is connected to other well-connected nodes within the network. This could be influenced by its geographic location and trade relationships, particularly with other countries in North America.

Nodes with high eigenvector centrality are considered influential or important within the network because they are connected to other well-connected and influential nodes. This metric considers not only the number of connections a node has but also the importance or influence of the nodes it is connected to. They maintain strong connections with other influential ports and countries. These nodes may play a pivotal role in shaping the overall structure and dynamics of the network.

4.2.6 – Separate measures for ports and trading partners

Now explaining the top 5 ports and countries for degree centrality and betweenness centrality separately, highlighting the reason for analyzing these metrics separately for ports and trading partners.

Top 5 USA Ports by Degree Centrality:

1. Oakland (0.0847)

2. Los Angeles International Airport (0.0678)

3. Los Angeles (0.0678)

4. New York City (0.0678)

5. Jersey Port (0.0678)

Degree centrality for ports indicates the number of direct connections a port has within the network. Analyzing this metric for ports specifically helps identify the major hubs or gateways within the United States that have numerous direct connections to other ports or nodes in the network.

Top 5 Countries by Degree Centrality:

1. China (0.5254)

2. Germany (0.5085)

3. Japan (0.3559)

4. South Korea (0.2881)

5. Mexico (0.2373)

Degree centrality for countries highlights the nations that have the highest number of direct connections or trade partnerships within the network. Examining this metric for countries separately helps identify the most well-connected and potentially influential trading partners in the network.

Top 5 Ports by Betweenness Centrality:

1. Miami International Airport (0.0400)
2. Corpus Christi (0.0373)
3. Jacksonville (0.0165)
4. Mobile (0.0164)
5. Oakland (0.0146)

Betweenness centrality for ports indicates the extent to which a port lies on the shortest paths between other nodes in the network. Ports with high betweenness centrality act as bridges or gateways, facilitating the flow of resources or information between other ports. Analyzing this metric for ports specifically helps identify the key intermediaries or transit points within the network.

Top 5 Countries by Betweenness Centrality:

1. China (0.2971)
2. Germany (0.2786)
3. Japan (0.1305)
4. South Korea (0.0992)
5. Mexico (0.0471)

Betweenness centrality for countries highlights the nations that serve as bridges or intermediaries, facilitating the flow of resources or information between other countries or nodes in the network. Examining this metric for countries separately helps identify the trading partners that play crucial roles in connecting different parts of the network.

By analyzing these metrics separately for ports and trading partners, we can gain insights into the specific roles and importance of individual ports within the United States, as well as the influence and interconnectedness of different countries within the broader trade network. This separate analysis allows for a more nuanced understanding of the network dynamics and the relative positions of ports and trading partners within the overall structure.

5. Conclusion

The goal of this project was to answer questions about the trade dependencies of USA and what ports hold the most importance. On the basis on the analysis, we have the following analysis

1. **Trade Flow Patterns**: The analysis indicates several major trade routes and corridors connecting ports within the USA and globally. Ports like Los Angeles, New York City, and Oakland emerge as significant hubs within the USA, likely due to their strategic locations and extensive connections with other ports globally. Additionally, ports in communities like Community 0 (which includes Los Angeles, New York City, Savannah, Long Beach, etc.) and Community 1 (which includes Houston, Laredo, Corpus Christi, etc.) highlight important trade corridors within the USA and with countries like China, Vietnam, Japan, Mexico, Brazil, and South Korea.

2. **Port Connectivity and Centrality**: In terms of port connectivity and centrality, ports like Oakland, Los Angeles, and New York City emerge as highly connected hubs within the USA, based on degree centrality. These ports have numerous direct connections to other ports or nodes in the network, making them crucial for trade and transportation within the USA. Moreover, ports like Miami International Airport and Corpus Christi stand out for their high betweenness centrality, indicating their role as key intermediaries or transit points facilitating the flow of resources or information between other ports.

3. **Trade Partners and Dependencies**: The primary trading partners of ports in the USA include countries like China, Germany, Japan, South Korea, and Mexico, based on degree centrality and betweenness centrality metrics. These countries have the highest number of direct connections or trade partnerships within the network, highlighting their importance as trading partners for ports in the USA. Additionally, the identified communities within the network also suggest strong connections or dependencies between ports within the USA and their counterparts in other countries or regions, emphasizing the interconnected nature of global trade networks.

In conclusion, the analysis highlights the major trade routes and corridors connecting ports within the USA and with ports globally, identifies the ports with the highest degree of connectivity within the trade network, and highlights the primary trading partners of ports in the USA. This information is essential for understanding the dynamics of global trade networks and can help stakeholders make informed decisions regarding trade, transportation, and economic policies.