In [1]:
```
import pandas as pd
import numpy as np
import plotly.graph_objs as go
from plotly.offline import download_plotlyjs, init_notebook_mode, plot, ipl
init_notebook_mode(connected=True)
import holoviews as hv
import plotly.graph_objects as go
import plotly.express as pex
hv.extension('bokeh')
import matplotlib.pyplot as plt
import seaborn as sns
from collections import Counter
import networkx as nx
```

Q1A: Trade between top five exporters and importers (Sankey)

In [2]:
```
def detail_trade(country_name, year):
    # Read all information in the sheet
    country = pd.read_excel('Q1.xlsx', sheet_name = country_name)
    # Read specific columns in the sheet
    country_export = pd.read_excel('Q1.xlsx', sheet_name = country_name, us
    # Choose specific year (There are two years in the sheet, including 201
    country_export = country_export.loc[country_export['Year'] == year]
    # Sort from highest to lowest
    country_export = country_export.sort_values(by=['Export (US$ Thousand)'
    # Get the first ten countries with highesting numbers
    country_export = country_export[:10]
    # Rename column names
    country_export_new = country_export.rename(columns={'Export (US$ Thousa
    # Reset the index
    country_export_new = country_export_new.reset_index()
    # Add new column, named "Measure"
    country_export_new['Measure'] = ['Export' for i in range(len(country_ex

    # Do the same thing for import
    country_import = pd.read_excel('Q1.xlsx', sheet_name = country_name, us
    country_import = country_import.loc[country_import['Year'] == year]
    country_import = country_import.sort_values(by=['Import (US$ Thousand)'
    country_import = country_import[:10]
    country_import_new = country_import.rename(columns={'Import (US$ Thousa
    country_import_new = country_import_new.reset_index()
    country_import_new['Measure'] = ['Import' for i in range(len(country_im

    # Combine two table including the exporting and importing for specific
    total = country_export_new.append(country_import_new)
    # Reset index
    total = total.reset_index()
    # Drop columns
    total = total.drop(columns=['level_0', 'index', 'Reporter Name', 'Year'
    # Reorder columns
    total = total[['Measure', 'Country', 'Value']]

    return total
```

In [3]:
```
total_usa = detail_trade('United States', 2019)
```

In [4]:
```
# Draw the graph
hv.Sankey(total_usa)
sankey1 = hv.Sankey(total_usa, kdims=["Measure", "Country"], vdims=["Value"

sankey1.opts(cmap='Colorblind',label_position='left',
                                edge_color='Country', edge_line_width=0,
                                node_alpha=1.0, node_width=40, node_sort=T
                                width=800, height=600, bgcolor="snow",
                                title="Trade in United States in 2019")
```

Out[4]:

In [5]:
```
total_china = detail_trade('China', 2019)
```

In [6]:
```
# Draw the graph
hv.Sankey(total_china)
sankey1 = hv.Sankey(total_china, kdims=["Measure", "Country"], vdims=["Valu

sankey1.opts(cmap='Colorblind',label_position='left',
                                    edge_color='Country', edge_line_width=0,
                                    node_alpha=1.0, node_width=40, node_sort=T
                                    width=800, height=600, bgcolor="snow",
                                    title="Trade in China in 2019")
```

Out[6]:

In [7]:
```
exporters = ['China', 'United States', 'Germany', 'Japan', 'Korea, Rep.']
exporters_name = ['China_ex', 'United States_ex', 'Germany_ex', 'Japan_ex',
importers = ['United States', 'China', 'Germany', 'United Kingdom', 'France
importers_name = ['United States_im', 'China_im', 'Germany_im', 'United Kin
```

In [8]:
```
exporter = pd.read_excel('Q1_new.xlsx', sheet_name = 'Exporter_2019')
exporter = exporter.rename(columns={'Reporter Name': 'Exporter', 'Partner N
importer = pd.read_excel('Q1_new.xlsx', sheet_name = 'Importer_2019')
importer = importer.rename(columns={'Reporter Name': 'Importer', 'Partner N
importer = importer[['Exporter', 'Importer', 'Value']]
```

In [9]:
```
total_trade = exporter.append(importer)
total_trade = total_trade.dropna(subset=['Value'])
```

In [10]:
```
total_trade['Exporter'] = total_trade['Exporter'].replace(exporters, export
total_trade['Importer'] = total_trade['Importer'].replace(importers, import
```

In [11]:
```
for index, row in total_trade.iterrows():
    if row['Exporter'] not in exporters_name:
        total_trade['Exporter'] = total_trade['Exporter'].replace([row['Exp

for index, row in total_trade.iterrows():
    if row['Importer'] not in importers_name:
        total_trade['Importer'] = total_trade['Importer'].replace([row['Imp
```

In [12]:
```
sankey1 = hv.Sankey(total_trade, kdims=["Exporter", "Importer"], vdims=["Va

sankey1.opts(cmap='Colorblind',label_position='left',
                                    edge_color='Importer', edge_line_width=0,
                                    node_alpha=1.0, node_width=40, node_sort=T
                                    width=1000, height=600, bgcolor="snow",
                                    title="Trade in 2019 for Top Five Exportin
```

Out[12]:

Q1B: Trade between top 10 gdp countries (Weighted graph)

This is the other way to represent the data of exports and imports. However, we didn't show in the presentaion and report result, since in the future we would like to improve the presentation of these two weighted graphs by adding the arrows and having clear explanation of the amount of either importing or exporting for these seleced countres.

In [13]:
```
data = pd.read_excel('top10.xlsx')
datanew = data[["Country","ExportCountry","Export"]]
datanew
```

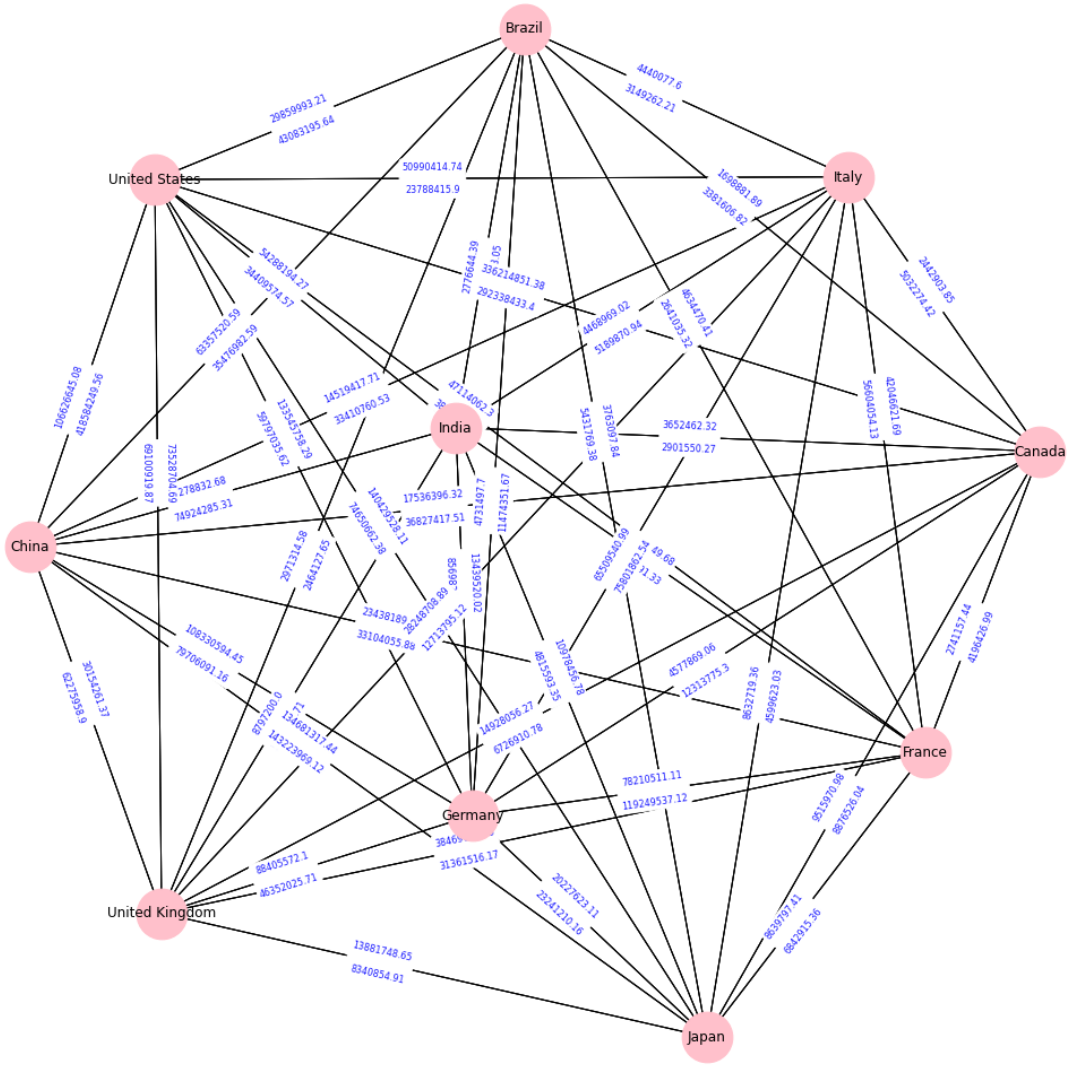
Out[13]:

	Country	ExportCountry	Export
0	Brazil	Canada	3381606.82
1	Brazil	China	63357520.59
2	Brazil	France	2641035.32
3	Brazil	Germany	4731497.70
4	Brazil	India	2776644.39
...	...	...	...
85	United States	Germany	59797035.62
86	United States	India	34409574.57
87	United States	Italy	23788415.90
88	United States	Japan	74650662.38
89	United States	United Kingdom	69100919.87

90 rows × 3 columns

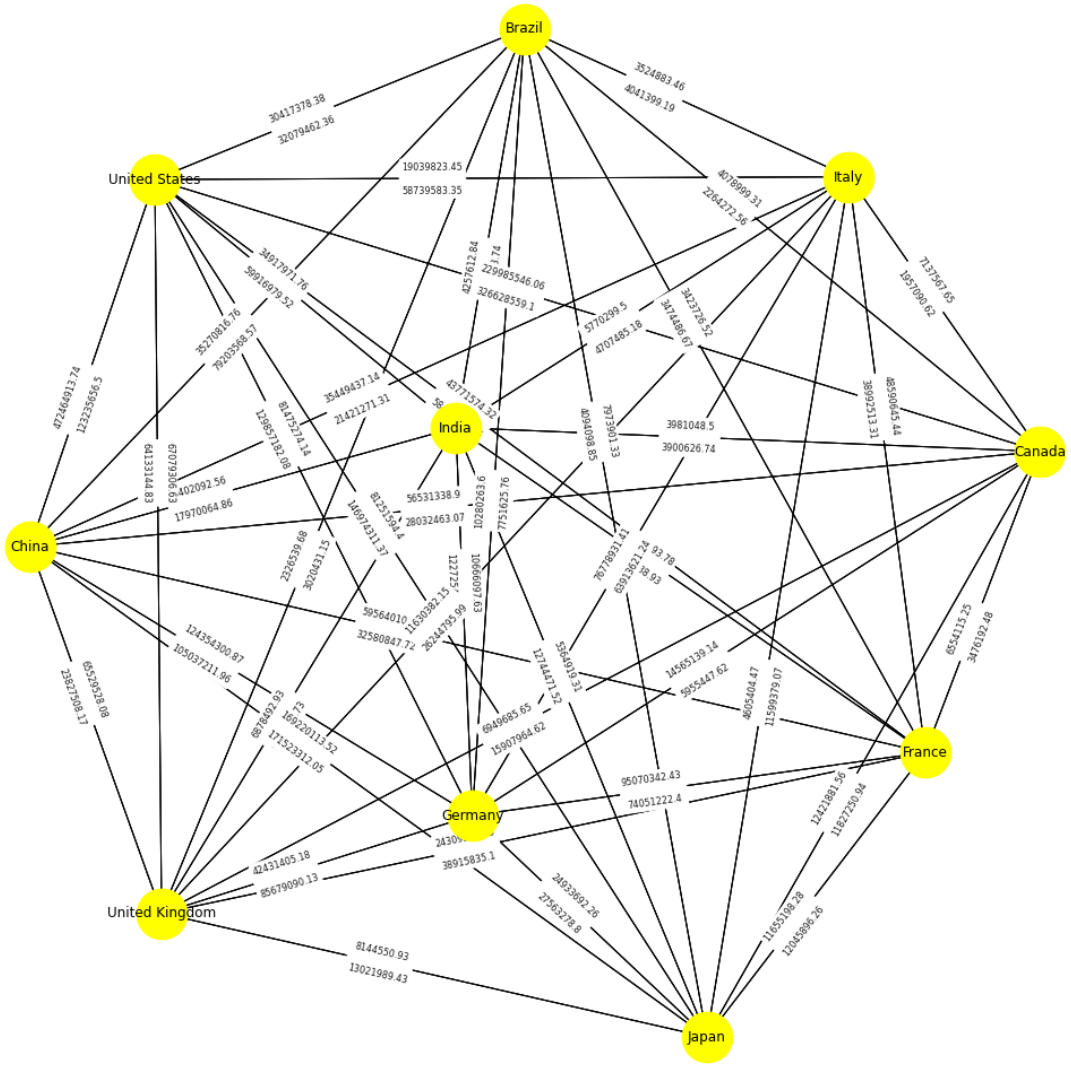
In [14]:
```
G = nx.DiGraph()
for index, row in datanew.iterrows():
    G.add_node(row['Country'])
    G.add_node(row['ExportCountry'])
    G.add_edge(row['Country'], row['ExportCountry'], length = row['Export']
```

In [15]:
```
plt.figure(figsize=(15,15))
pos = nx.spring_layout(G,seed=20)
nx.draw(G, pos, with_labels=True, connectionstyle='arc3,rad=0')

edge_labels = dict([((u, v,), f'{d["length"]}\n\n{G.edges[(v,u)]["length"]}
                for u, v, d in G.edges(data=True) if pos[u][0] > pos[v][0]]



nx.draw_networkx_edge_labels(G, pos, edge_labels=edge_labels, font_color='b
nx.draw_networkx_nodes(G, pos, node_size=2000, node_color = 'pink')
nx.draw_networkx_edges(G,pos,alpha=0.5,edge_color='black')




#plt.tight_layout()
plt.title("Top 10 GDP Countries' Export Trade ")
plt.show()
```



In [16]:
```
data = pd.read_excel('top10.xlsx')
Import = data[["Country","ExportCountry","Import"]]

G = nx.DiGraph()
for index, row in Import.iterrows():
    G.add_node(row['Country'])
    G.add_node(row['ExportCountry'])
    G.add_edge(row['Country'], row['ExportCountry'], length = row['Import']
```

In [17]:
```
plt.figure(figsize=(15,15))
pos = nx.spring_layout(G,seed=20)
nx.draw(G, pos, with_labels=True, connectionstyle='arc3,rad=0')

edge_labels = dict([((u, v,), f'{d["length"]}\n\n{G.edges[(v,u)]["length"]}
                for u, v, d in G.edges(data=True) if pos[u][0] > pos[v][0]]



nx.draw_networkx_edge_labels(G, pos, edge_labels=edge_labels, font_color='b
nx.draw_networkx_nodes(G, pos, node_size=2000, node_color = 'yellow')
nx.draw_networkx_edges(G,pos,alpha=0.5,edge_color='black')




#plt.tight_layout()
plt.title("Top 10 GDP Countries' Import Trade ")
plt.show()
```



Q2: Analyze the value of imports & exports, 16 categories, 10 countries

In [18]:
```
df2017 = pd.read_excel("Q2_2017.xlsx")
df2018 = pd.read_excel("Q2_2018.xlsx")
df2019 = pd.read_excel("Q2_2019.xlsx")
```

In [19]:
```
df2017.head()
```

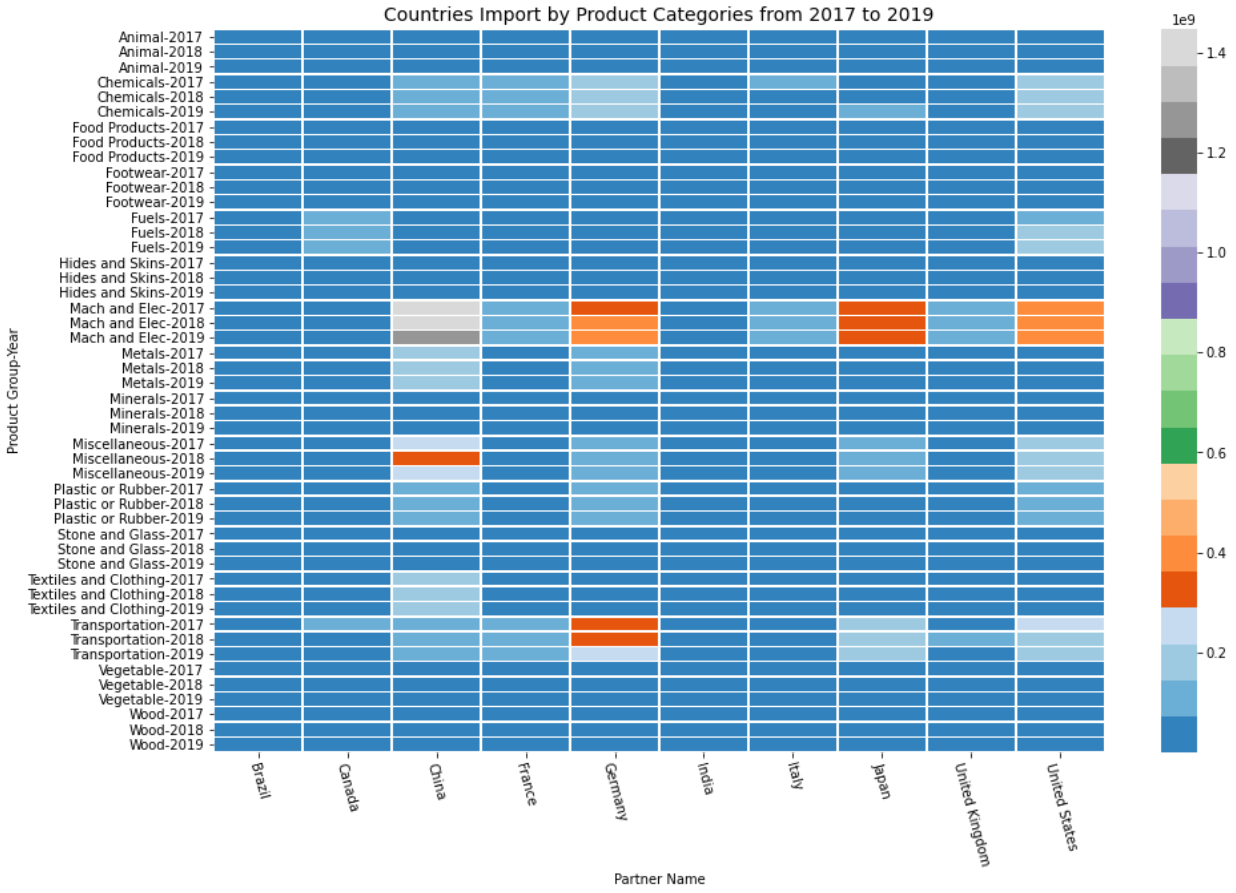
Out[19]:

	Reporter Name	Partner Name	Year	Product Group	Export (US$ Thousand)	Import (US$ Thousand)
0	World	United States	2017	Animal	3.271923e+07	2.776733e+07
1	World	United States	2017	Chemicals	2.688639e+08	1.903602e+08
2	World	United States	2017	Food Products	8.013343e+07	4.503055e+07
3	World	United States	2017	Footwear	3.087877e+07	1.185737e+06
4	World	United States	2017	Fuels	1.558189e+08	1.307280e+08

Heatmap: Changes and Trends

In [20]:
```
# Countries import from 2017-2019 heatmap
df_2017_2018_2019 = pd.concat([df2017,df2018,df2019])

df_country_prod_imp = df_2017_2018_2019.pivot_table(index=["Product Group",

plt.figure(figsize=(15,10))
sns.heatmap(df_country_prod_imp, cmap="tab20c",linewidths=.5)
plt.xticks(rotation=-75)
plt.title("Countries Import by Product Categories from 2017 to 2019",fontsi
```

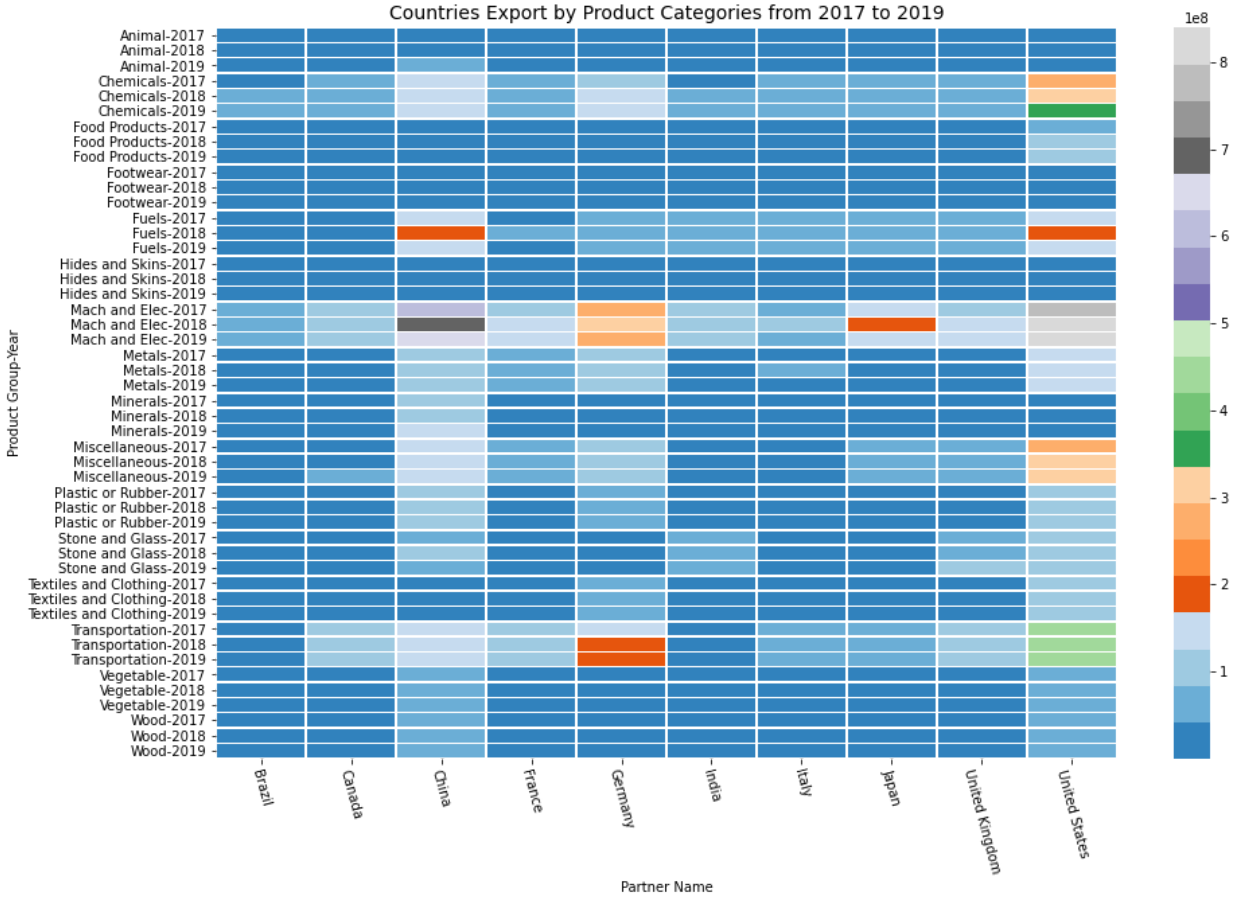
Out[20]:
```
Text(0.5, 1.0, 'Countries Import by Product Categories from 2017 to 201
9')
```




```
- Import
- Trend & influence of tariff
- Categories of Mach & Elec and transportation should be a metter of
concern
- Mach & Elec: China: downward trend, highest
               Germany: upward,highest
- Transportation: Germany: down
-------------------------------------------------------------------------
Why heatmap: remarkable in color: which categories and countries
compared to line chart and bar chart
line chart: too many lines, hard to tell
bar chart: differencies in
```

In [21]:
```
# Countries export from 2017-2019 heatmap
df_2017_2018_2019 = pd.concat([df2017,df2018,df2019])

df_country_prod_exp = df_2017_2018_2019.pivot_table(index=["Product Group",

plt.figure(figsize=(15,10))
sns.heatmap(df_country_prod_exp, cmap="tab20c",linewidths=.5)
plt.xticks(rotation=-75)
plt.title("Countries Export by Product Categories from 2017 to 2019",fontsi
```

Out[21]:
```
Text(0.5, 1.0, 'Countries Export by Product Categories from 2017 to 201
9')
```




```
- Export
- Categories:
    Chemical: USA: increase, highesr
    Mach & Elec: China: up then down, higher then most of the countries
                 USA: up,highest
    Miscellaneous: USA:up, highest
    Transportation: Germany: up
                    Usa: highest
```

Comparison of export among countries in 2018

and 2019

In [22]:
```
df2018_exp = df2018[["Partner Name","Year","Product Group","Export (US$ Tho
df2019_exp = df2019[["Partner Name","Year","Product Group","Export (US$ Tho
df2019_exp["Export (US$ Thousand)"] = df2019_exp["Export (US$ Thousand)"]*(

df2018_2019_exp = pd.concat([df2018_exp,df2019_exp ])
df2018_2019_exp.head()
```

Out[22]:

	Partner Name	Year	Product Group	Export (US$ Thousand)
0	United States	2018	Animal	3.319013e+07
1	United States	2018	Chemicals	3.190755e+08
2	United States	2018	Food Products	8.796040e+07
3	United States	2018	Footwear	3.298597e+07
4	United States	2018	Fuels	1.848361e+08

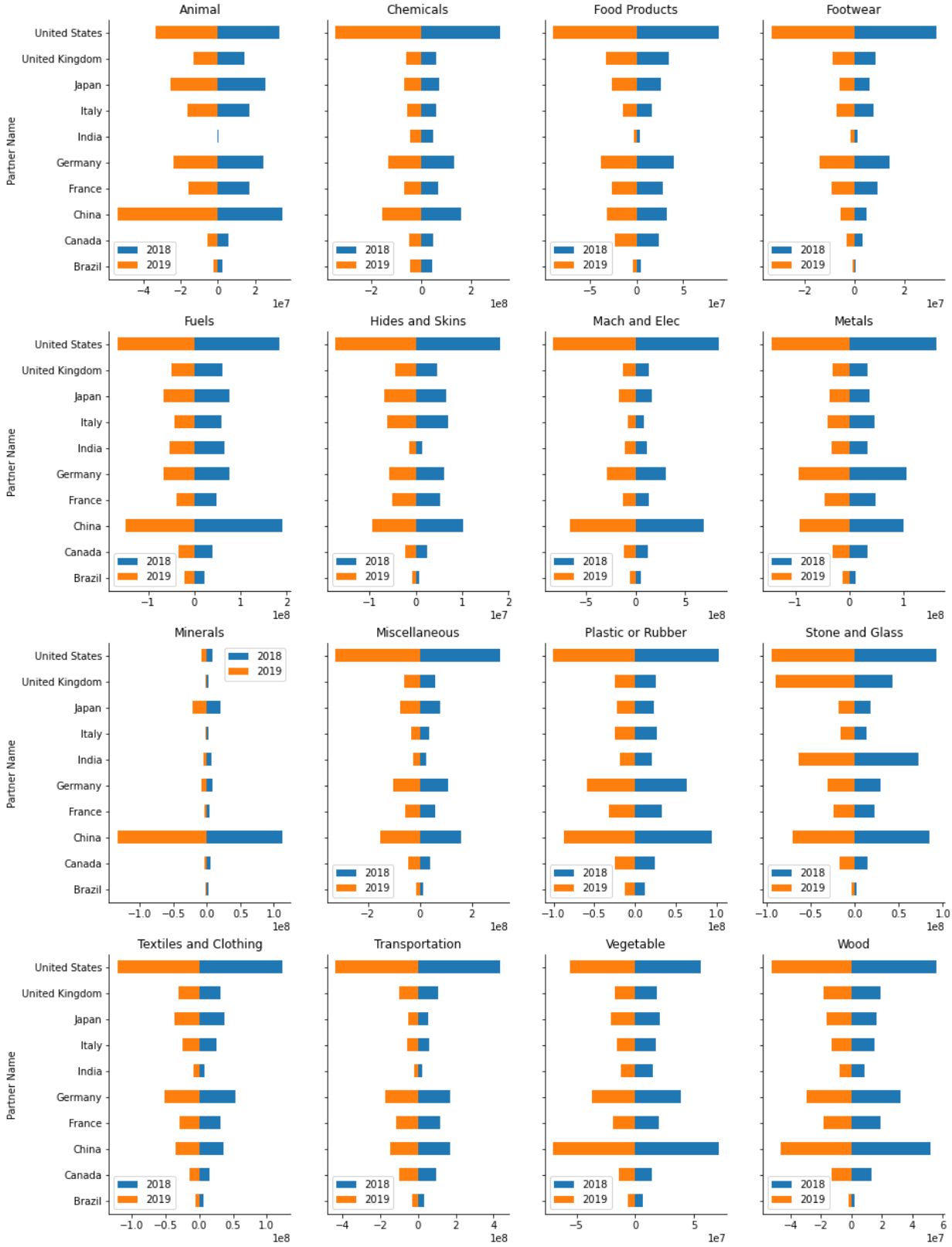
In [23]:
```
df2018_2019_exp_piv = df2018_2019_exp.pivot_table(index="Partner Name",colu
df2018_2019_exp_piv
```

Out[23]:

Product Group	Animal		Chemicals		Food Products		
Year	2018	2019	2018	2019	2018	2019	
Partner Name							
Brazil	2506623.24	-2400707.78	4.245714e+07	-4.275616e+07	4591824.79	-4199421.86	64
Canada	5503902.79	-5613622.18	4.862393e+07	-4.865184e+07	24099363.35	-23787571.66	309
China	34788272.26	-54192877.17	1.610071e+08	-1.587524e+08	32184724.86	-31656304.90	480
France	16735662.79	-15797783.42	6.772859e+07	-6.702447e+07	27894631.80	-27139176.83	908
Germany	24567795.22	-23811105.69	1.319325e+08	-1.327650e+08	39845411.13	-39143404.74	1421
India	204622.45	-207026.28	4.687901e+07	-4.548054e+07	3125548.59	-2918894.34	121
Italy	16976949.03	-16440973.24	5.973384e+07	-5.830004e+07	16097890.80	-15432784.08	781
Japan	25702199.07	-25325328.80	7.029119e+07	-6.762307e+07	25759121.67	-27173445.97	602
United Kingdom	14393798.08	-13295555.74	6.121933e+07	-5.897131e+07	34152790.05	-32834074.66	851
United States	33190126.28	-33847179.13	3.190755e+08	-3.456493e+08	87960396.35	-89719182.63	3298

10 rows × 32 columns

In [24]:
```
titles = [x[0] for x in df2018_2019_exp_piv.columns][::2]

fig,axs = plt.subplots(4,4,figsize=(15,22))
axs = axs.flatten()

for i in range(0,16):
    tmp_df = df2018_2019_exp_piv.iloc[:,i*2:(i+1)*2].copy()
    tmp_df.columns = ["2018","2019"]
    tmp_df.plot(kind="barh",sharey=True,
                                                stacked=True,
                                                legend=False,
                                                title=titles[i],ax=axs[i],
                                                )
    sns.despine()
    axs[i].legend(fontsize=10)
```




```
Remarkable categories and countries:
Strength and weakness
- China: Mineral, prominent
         Vegetables
         Animal
         wood
- USA: large in each categories
        Textile
```

Comparison of import among countries in 2018 and 2019

In [25]:
```
df2018_imp = df2018[["Partner Name","Year","Product Group","Import (US$ Tho
df2019_imp = df2019[["Partner Name","Year","Product Group","Import (US$ Tho
df2019_imp["Import (US$ Thousand)"] = df2019_imp["Import (US$ Thousand)"]*(

df2018_2019_imp = pd.concat([df2018_imp,df2019_imp ])
df2018_2019_imp.head()
```

Out[25]:

	Partner Name	Year	Product Group	Import (US$ Thousand)
0	United States	2018	Animal	2.725376e+07
1	United States	2018	Chemicals	1.992497e+08
2	United States	2018	Food Products	4.543039e+07
3	United States	2018	Footwear	1.027420e+06
4	United States	2018	Fuels	1.866005e+08

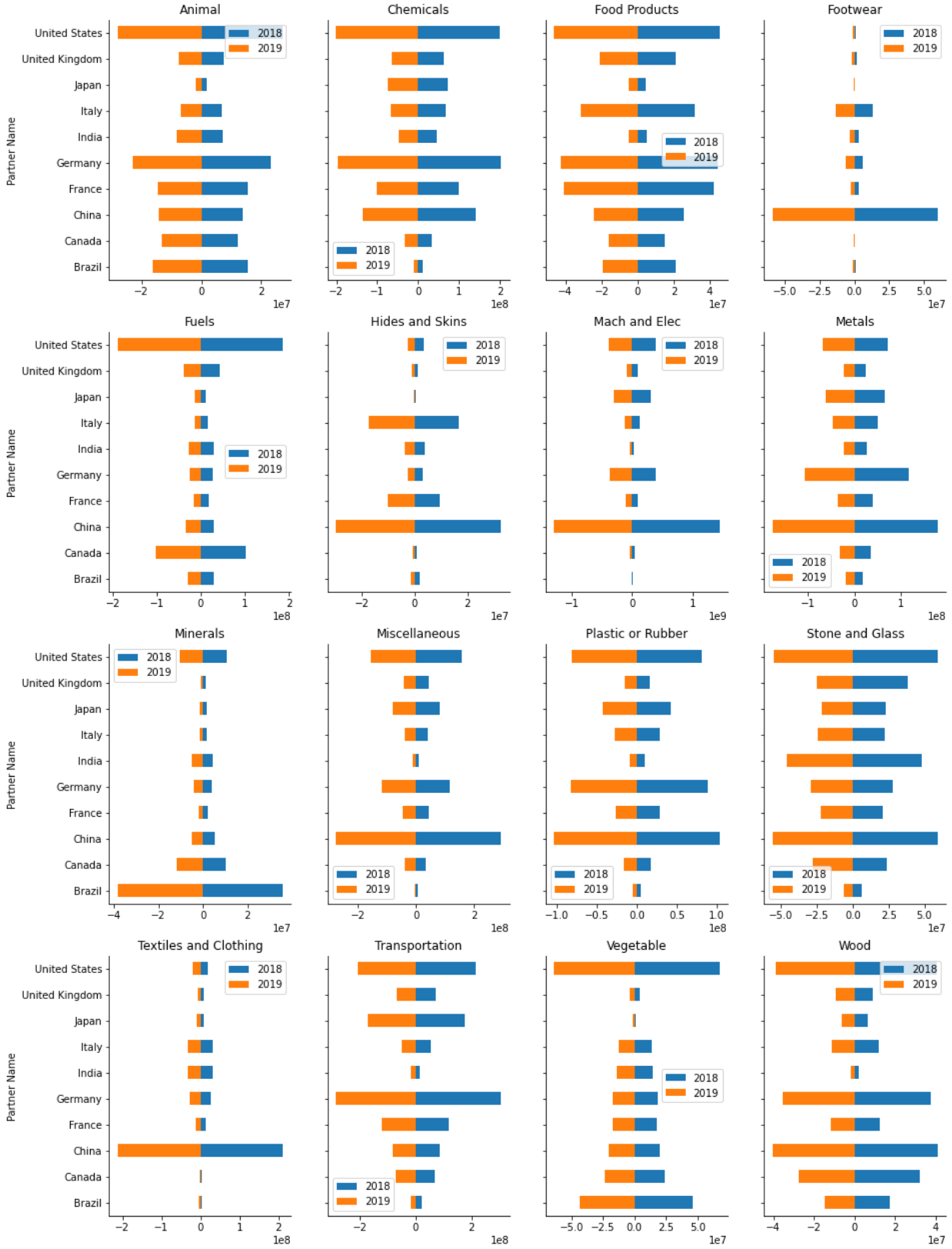
In [26]:
```
df2018_2019_imp_piv = df2018_2019_imp.pivot_table(index="Partner Name",colu
df2018_2019_imp_piv
```

Out[26]:

Product Group	Animal		Chemicals		Food Products		
Year	2018	2019	2018	2019	2018	2019	
Partner Name							
Brazil	15606572.62	-16315503.88	1.224826e+07	-1.131585e+07	21167159.01	-19531441.12	124
Canada	12311063.03	-13290351.48	3.263724e+07	-3.289952e+07	15249785.52	-16081800.84	13
China	13906857.69	-14117630.79	1.418938e+08	-1.346791e+08	25401544.51	-24090965.84	6037
France	15357902.72	-14513304.61	9.810821e+07	-1.020266e+08	42050243.12	-40984708.23	303
Germany	23056377.95	-22920338.51	2.030625e+08	-1.969241e+08	44291652.31	-42456662.09	605
India	7085532.08	-8360596.95	4.455538e+07	-4.768041e+07	5240238.69	-5166647.45	355
Italy	6794929.11	-6868738.41	6.803527e+07	-6.800381e+07	31751588.33	-31415604.66	1305
Japan	1758446.32	-1793970.15	7.214194e+07	-7.478562e+07	4347000.48	-4705111.86	27
United Kingdom	7463848.87	-7514938.62	6.343358e+07	-6.337713e+07	21273807.85	-20915778.98	179
United States	27253762.22	-28000361.18	1.992497e+08	-2.011104e+08	45430388.25	-46271482.93	102

10 rows × 32 columns

In [27]:
```
titles = [x[0] for x in df2018_2019_imp_piv.columns][::2]

fig,axs = plt.subplots(4,4,figsize=(15,22))
axs = axs.flatten()

for i in range(0,16):
    tmp_df = df2018_2019_imp_piv.iloc[:,i*2:(i+1)*2].copy()
    tmp_df.columns = ["2018","2019"]
    tmp_df.plot(kind="barh",sharey=True,
                                                stacked=True,
                                                legend=False,
                                                title=titles[i],ax=axs[i],
                                                )
    sns.despine()
    axs[i].legend(fontsize=10)
```




```
Import
Strength and weakness
Remarkable categories and countries:
- China: footwear, Hides and Skins, Mach and Elec, Metals, Miscellaneous,
Textiles
- Brazil: Minerals
- USA: Vege, Fueks
```

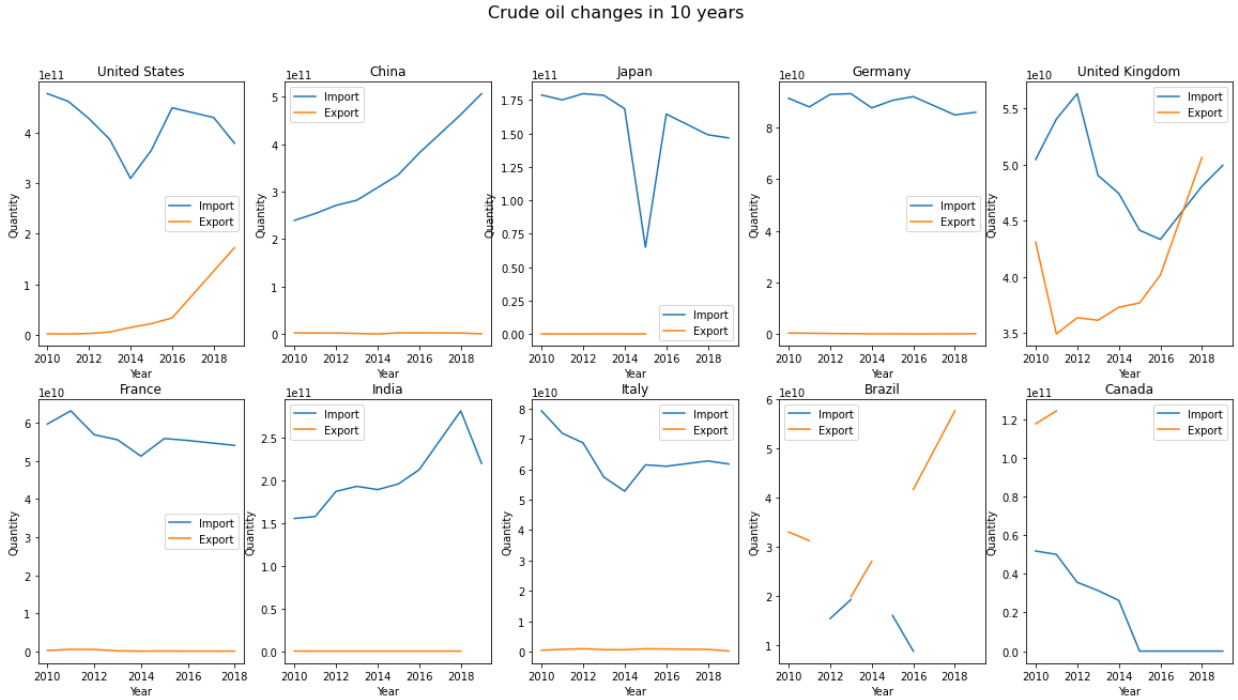
Comparison between import and export in one country in 2018 and 2019

In [28]:
```
# export
df2018_exp=df2018[["Partner Name","Year","Product Group","Export (US$ Thous
df2019_exp=df2019[["Partner Name","Year","Product Group","Export (US$ Thous
df2019_exp["Export (US$ Thousand)"]=df2019_exp["Export (US$ Thousand)"]*(-1

df2018_2019_exp=pd.concat([df2018_exp,df2019_exp ])
df2018_2019_exp_piv=df2018_2019_exp.pivot_table(index="Product Group",colum

# import
df2018_imp=df2018[["Partner Name","Year","Product Group","Import (US$ Thous
df2019_imp=df2019[["Partner Name","Year","Product Group","Import (US$ Thous
df2019_imp["Import (US$ Thousand)"]=df2019_imp["Import (US$ Thousand)"]*(-1

df2018_2019_imp=pd.concat([df2018_imp,df2019_imp ])
df2018_2019_imp_piv=df2018_2019_imp.pivot_table(index="Product Group",colum

# prepare title
titles=[x[0] for x in df2018_2019_exp_piv.columns][::2]


for i in range(0,10):
    fig,axs=plt.subplots(1,2,figsize=(6,5))

    # plot export
    tmp_df_exp=df2018_2019_exp_piv.iloc[:,i*2:(i+1)*2].copy()
    tmp_df_exp.columns=["2018","2019"]
    tmp_df_exp.plot(kind="barh",sharey=True,
                                                stacked=True,
                                                legend=False,
                                                title="{} Export".format(t
                                                )

    sns.despine()
    axs[0].legend(fontsize=8,loc="upper right")

    # plot import
    tmp_df_imp=df2018_2019_imp_piv.iloc[:,i*2:(i+1)*2].copy()
    tmp_df_imp.columns=["2018","2019"]
    tmp_df_imp.plot(kind="barh",sharey=True,
                                                stacked=True,
                                                legend=False,
                                                title="{} Import".format(t
                                                )

    sns.despine()
    axs[1].legend(fontsize=8,loc="upper right")
    plt.show()
```

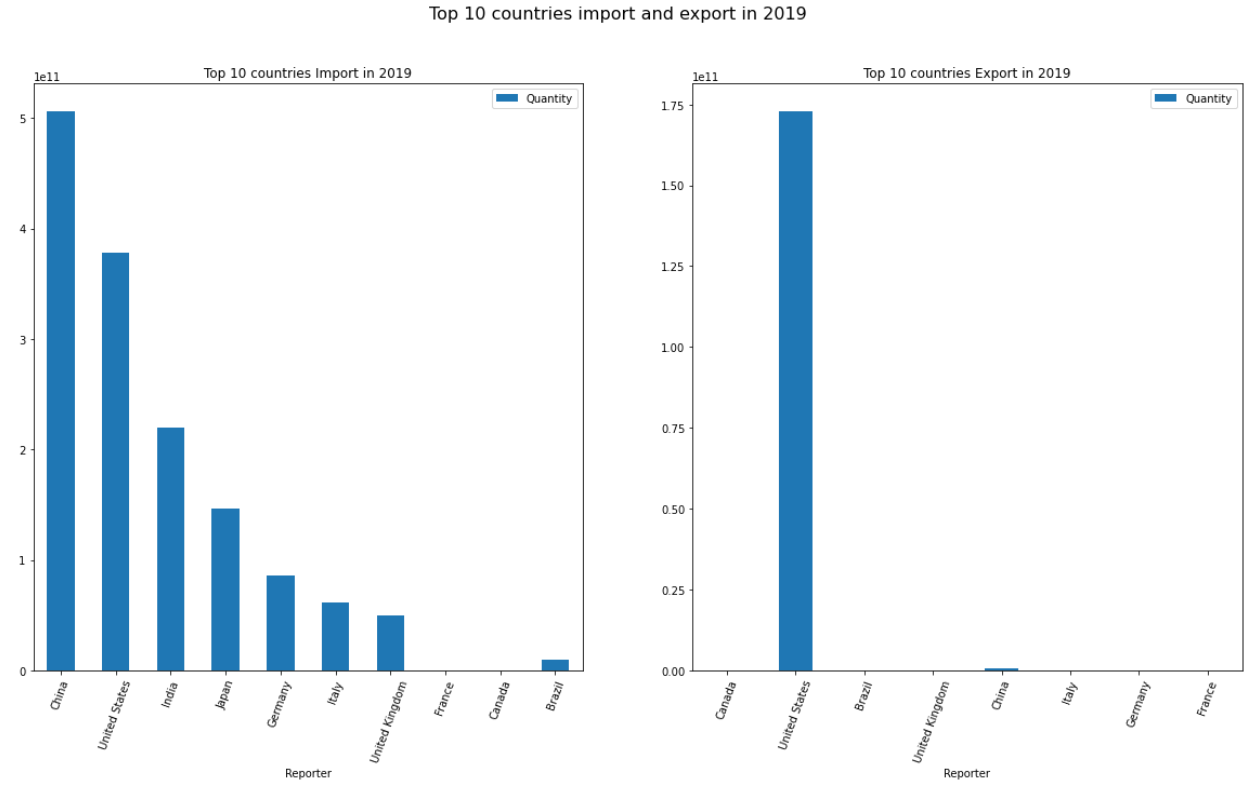

```
Brainstorm for this part:
- Export vs import in the same country, also does the countries rely on
import? Features?
- relationship between export and import
- difference between developed countries and developing countries
Attention: quantity difference
```

Q3: Analysis Of The Trade For Crude Oil Among Ten Years

In [29]:
```
plt.figure(figsize=(20,10))


# use pandas to read the excel file into a dataframe
data = pd.read_excel("Q3.xlsx", engine="openpyxl", sheet_name="By-HS6Produc

country = ["United States","China","Japan","Germany","United Kingdom","Fran

# Line Plot for each country
num = 1
for c in country:
    # first determine import and export quantity over years for each countr
    # find rows for each country, separate by import and export, then extra
    Import = data.loc[(data["Reporter"] == c) & (data["TradeFlow"] == "Impo
    Export = data[(data["Reporter"] == c) & (data["TradeFlow"] == "Export")

    # get currect axis, draw import and export on the same graph
    plt.subplot(2, 5, num)
    ax = plt.gca()
    Import.plot(kind = 'line', x = "Year", y = "Quantity", ax = ax, label="
    Export.plot(kind = 'line', x = "Year", y = "Quantity", ax = ax, label="
    ax.set_ylabel("Quantity")
    ax.set_title(c)
    num += 1
plt.suptitle("Crude oil changes in 10 years", size=16)
plt.show()
```



In [30]:
```
# Bar chart to compare all countries' import and export quantities in 2019
# find all rows where year is 2019, separate by import and export
data_2019_import = data[(data["Year"] == 2019) & (data["TradeFlow"] == "Imp
data_2019_export = data[(data["Year"] == 2019) & (data["TradeFlow"] == "Exp

fig, axes = plt.subplots(nrows=1, ncols=2, figsize=(20,10))

data_2019_import.plot.bar(x = "Reporter", y = "Quantity", rot = 70, title =
data_2019_export.plot.bar(x = "Reporter", y = "Quantity", rot = 70, title =
plt.suptitle("Top 10 countries import and export in 2019", size=16)
plt.show()
```



World Trade

Jiawen Liang, Jiayang Liu, Qiyue Cui, Yitian Liang
Northeastern University, Boston, MA, USA

Introduction

One news about the blockage of the Suez Canal in March this year caught our attention on global supply chains. This blockage lasted for around six days which triggered the decrease of global annual trade growth by approximately 0.2% to 0.4% [1]. Since one single container ship resulted in the drawback of the whole global trade, our group formed the hypothesis on the recent trade policies (Trade War) and pandemic (Covid-19) impact on world trade.

Even with the non-existence of recent pandemic and its related policies, world trade can be considered as an intriguing and profound subject to evaluate. World trade represents the correlation and collaboration between every single country. Based on the principle of mutual benefit and reciprocity, import and export countries simultaneously promote economic growth, development and poverty reduction by trade exchanges. Therefore, through this project, our group created the hypothesis that the total value of global trade between exporters and importers would have a positive impact on the gross domestic product (GDP) of every single country, and each category of global trade will be altered independently under big events.

To evaluate world trade as comprehensively as possible, our group aimed at three perspectives. First perspective is to generally find the top five importers and exporters based on total values, take a deeper exploration on the trade of the largest import and export country and visualize it by Sankey diagram. Second perspective is to visualize the trend of global trade of all sixteen categories during different years by using heatmap and stacked bar charts. The final visualization is to study the global trade of crude oil specifically since crude oil is always considered as one of the most significant fuel resources, and is deeply applied into a wide area.

Data Sources and Methods

Our dataset is gathered from World Integrated Trade Solution software [2], an application developed by the World Bank. WITS is a professional, authoritative and effective platform, allowing users to retrieve official wholescale information on trade and tariff data. By adopting the WITS databases, we have access to select our ideal data by setting country, year, types of trade flow and product categories during the process of extracting data. In this project, with the attempt to analyze world trade by product in the past decade among top GDP countries, we extract both import and export data in 16 categories of product of countries including USA, China, Japan, Germany, UK, France, India, Italy, Brazil, Canada in the analytical databases [3].

The research targets for the first perspective are the top five exporters and importers, which are the United States, China, Germany, Japan, South Korea, France, and the United Kingdom. All these separate files downloaded from the website [2] are Excel format. Due to the

limitation of extraction for particular columns, there are parts of integration done in Excel, such as removing the unnecessary columns and combining all files into one. Next, Function "detail_trade" is the place doing all rest integration work. For example, Pandas read_excel() method is to read an excel file as a DataFrame. After creating a DataFrame in Python, many Pandas functions were used to change the original input table, like sorting the values for specific columns by using DataFrame.sort_value, and renaming, adding, dropping as well as reordering columns. More importantly, to build the Sankey diagram, three columns are required, corresponding to three important factors of the Sankey diagram, which are the left side, the connection line, and the right side. For instance, for Exhibit 1, the three columns are exporters, importers, value respectively. The left side presents exporting countries and the importers would be set on the right side. The thickness of the line connecting two sides was determined by the amounts of the value. For the detailed graph about the trades in the United States (Exhibit 2) and China (Exhibit 3), the three columns are measures (Export/Import), countries and values. For a picked country, which measures it used to connect with the countries on the right sides, which are the top ten connected countries based on the value for different measures.

For the second perspective of our project, the dataset for analyzing the trends and changes of trade flow is basically three separate excel files covering information for each year from 2017 to 2019. We clean the data in this section by adopting the pandas module, specifically, the embedded functions of concat and pivot tables. Concat provides an easy way to concatenate these three DataFrame objects into one integrated DataFrame, while pivot tables provide a more flexible and logical way for us to compile our desired DataFrame. By setting parameter such as product group and year as index, partner name as columns, and trading amount as values, we finally put our data in the certain order, which later we are able to use partner name, which are countries, to represent x-axis, and product group and year to represent y-axis in the heatmap (Exhibit 4, 5). Three bars for three years for each category arranged together in the heatmap is significant, which helps us recognize the changes in values of trade directly through the wide-range changing colour palette over time. Otherwise, if we made it into three separate heatmaps for each year, we are not able to compare the 3-year data in a single graph and there is a limitation to learn the output of significant product categories and countries in a one-sided view.

In addition, we would like to compare the trading value of each product category among 10 countries in both 2018 and 2019 through horizontal stacked bar charts in each subplot. To realize the visualization that the left side for 2018 data and right side for 2019 data, we had to process and multiply the 2019 data by -1 in the first step. Then, we set the parameter of sharey to be true so that the 2-year data from the same country can be stacked together (Exhibit 6, 7). From the vertical perspective, we are able to compare the trading values across countries as well as recognize those countries with outstanding values. The processing method of data is also applied in the next section of the second part, with a different attempt to compare import and export data in the same country in 2018 and 2019, we are supposed to change the setting of some

parameters, including index, columns, which finally enable us to study the value of 16 product categories of our interested countries in depth (Exhibit 8, 9).

For the last part, we'll look at the trend of changes in the number of imports and exports of specific products. So we utilized ten nations as subplots in a line chart that best represents the change in amount (Exhibit 10). This is a good way to indicate changes in amount. We use the line-chart in Python to construct and adjust the title after importing data from ten nations from 2010 to 2019. The excel file is firstly read into a dataframe by using pandas. Secondly, for each nation, we generate a Line Plot and determine the number of imports and exports for each nation over a period of years. Thirdly, for each nation, locate rows, partition them by import and export, and extract the amount. Finally, we draw import and export on the same graph with the correct axis. In 2019, a bar chart will be used to compare the import and export volumes of all nations. Then we search all rows with 2019 as the year, and sort them by import and export.

Analysis and Results

As for the first goal of this project, by looking at Exhibit 1, it is an overview for the trades among those countries. Using the Sankey diagram, it is easy to find the top exporter and importer by seeing the number near each node which is the total value of export or import for each country in a specific year. It concluded from Exhibit 1 that the top exporting country is China with the amount $3.2205 trillion and the top importing country is the United States with the amount $3.3337 trillion. Also, it is convenient to go through the amount of trading between two countries in a specific year by clicking on the connection line. This time, except top five exporters and importers, other countries in the files were set as others listing both sides, which indicates the value top five exporting countries exports to the rest of the world except top five importers and the value top five importing countries imports from the rest of the world except top five exporters. Moreover, after knowing the top importer and exporter, Our group tried to look at some detailed connections for them. Then, Exhibit 2 and Exhibit 3 are two charts for the top importer United States and top exporter China respectively. From the graph for the United States, it provided some detailed information about the United States's trade. It can conclude that the United States exports and imports more to and from Canada, Mexico, and China in 2019. For the graph associated with China, it shows that China has relatively close cooperation with the United States, Hong Kong, Japan, South Korea for exporting and importing products in 2019.

As for the second goal of this project, we would like to learn what categories these countries trade, who are the biggest players and what are the trends and features in world trade. For the first section, the heatmap (Exhibit 4, 5) gives us an insight into how each category of each country varies over time and where to focus on, directing people towards outstanding areas of visualizations that matter most. Unlike common heatmaps, these two visualizations may be unique when seeking for interpretations, which should be read as a group of 3 years in 3 bars for each product category. By digesting information in this way, we can easily find out the remarkable value changes of significant categories in some outstanding countries. For example, when looking at Exhibit 4 as if we were looking across a blue ocean, we can position our eyes

precisely in the machinery and electronic category which are filled in both grey and orange colors, representing higher values in import trade then other categories. In other words, other categories like mach and elec with differentiated colors such as transportation, are significantly important in world trade. More importantly, we can figure out that China, which is with the lightest grey, heavily relies on the import of mach and elec, however, this situation tends to decline over time. Then the United States and Germany are second largest import countries in the mach and elec category. In exhibit 5 regarding the export trends, the mach and elec as well as transportation categories are still most remarkable. With the lightest grey blocks, the United States is recognized as the biggest exporter in the mach and elec category with an upward trend over time. Then China is the second biggest exporter in this field with volatile trends, which firstly rose then descended. This may indicate an underlying factor having an influence in world trade among countries.

In the next section (Exhibit 6, 7), we plot stacked bar charts in a format of 4 by 4 subplot to have an overview of countries import and export in each category and make comparisons across the 10 countries. In some subplots with extreme situations, which one country has large value while others are small, they easily raise our awareness, including categories of footwear, fuels, mach and elec, minerals, textile, hides, and vegetables. Most of these categories are related to the manufacturing sector, and we find China is their largest importer, which highlights the fact that China is strongly based around the manufacturing and further processing sectors. As for exhibit 7, subplots are about countries exporting information in each category. We can easily recognize that the United States almost accounts for the largest trades of export in each category, except minerals. This highlights the prominent trade status of the USA.

Based on the former sections, the trade flows of both China and USA are significantly outstanding. We would like to look at them in depth in this section through exhibit 8 and 9. Compared to other categories, in mach and elec category of China, both import and export values are large, but import amount is much larger than export when considering the difference in units in the graph denoted. It is not a favorable position in world trade, causing a huge trade deficit. While looking at the USA(exhibit 9) , it is the major country that exports mach and elec, where it is also larger than the value of its import. We can boldly assume that the USA will take advantage of its remarkable categories in world trade or even trade war.

In the last portion, Exhibit 10, we utilize a linear graph(line chart) to depict the top ten nations' oil imports and exports from 2010 to 2019. At the start of the project, the ten nations were selected to symbolize the global economy, and crude oil was chosen to represent the economies of all countries. Following on from the previous 9 exhibits, as everything we looked at was connected to economic growth, we wanted to look into the changes in the amount of significant imports and exports over the course of the decade, as well as the variables that drove those fluctuations. We concentrated our study on COVID-19 and the US-China trade war, which started in 2013. When we look at the line-chart 10 years later, we can see that export commerce in the United States, Japan, and India has all decreased significantly.After 18 years of line chart analysis, we can see that US imports have declined dramatically since 2018, while exports have

been steadily increasing since 2016. China's exports have been very low, while imports have increased since the trade war began in 2018.The United States is the greatest importer, as shown in exhibits 6 and 7, and a drop in imports from the United States would place a financial strain on China. Simultaneously, additional employment will be created in the United States, boosting the country's economy.One of the components that the US cannot afford is China's retaliatory tariffs. China's exports have been steadily increasing, indicating that the trade war has had little impact on China's export commerce. In addition, retaliatory tariffs have resulted in higher taxes for Americans.

Conclusions

According to the first visualization, there are several repeated importing and exporting countries appearing on both the top 10 GDP countries ranking and the top 5 countries related to the total value of global trade, such as the United States, China and Germany. Hence, we can conclude that the total value of global trade has an impact on increasing the GDP of each country. Furthermore, not only can countries focus more on their domestic trade, but also they can involve more trades with other countries to improve GDP as well as the living standards of citizens.

Take a deeper analysis on evaluating various categories of global trade between countries, there are 16 major categories of global trade. By visualizing the heatmap, we found out the total global trade value of machinery and electronic equipment conquered the most among all categories. Undoubtedly, the result represents that technology, which consists of machinery and electronic equipment, achieves a high status in exploration and development globally. Countries around the world place a higher value on trade related to technology, since technology is used in all fields of living. For developed countries which take advantage of the development of technology with either financial support or talents, they could exploit more resources to develop and innovate, and these countries are more likely to export technical goods actively. Simultaneously, countries who import technician goods will benefit from the latest technique.

For developing countries with comparative disadvantages on researching and developing technology, they could have an opportunity to be involved in natural resources trade, which opens another way to exchange globally. The conclusion gives rise to our view on analyzing one of the most valuable natural resources, crude oil. Among all countries, the United Arab Emirates (UAE), has an absolute advantage on mining oil, and it plays a key role in exporting oil to the rest.

To sum up, it is a good signal to see the increased total value of global trade year by year. There are a lot of expected advantages among countries who open the door for global trade. Expanding global trade and expanding more categories of global trade should be encouraged.

Future addition (Improvements)

There are two ways to visualize the first perspective, which are Sankey diagrams and weighted directed graphs. For the Weighted directed graph, to visualize the global trade between

the top 10 GDP countries, our group regarded the top 10 countries as vertices, and import and export value as weighted edges. However, the graph edges did not fully show the trade directions between countries, and the graph labels, representing import and export value, are not clear, since the import value from one country to another is not equal to the export between two countries, which resulted in two different weighted edges between two countries. In order to get a clear weighted directed graph, we need to explore more accurate data and adjust the edges and labels to visualize the trade relationship between countries.

In addition, since we are passionate to see how Trade War and Covid 19 impact global trade and its detailed categories, we need the updated data for Year 2020 and 2021. By having the updated information, we are more likely to see the shifted trade transactions and subtle differences of each category. Also, we missed the global trade data of Italy, Canada, and Brazil for some years. Therefore, we plan to explore and do more research on these countries' global trade.

Author Contributions

Qiyue Cui: I provide ideas to choose the topic and contribute different perspectives on analyzing global trade. To visualize the global trade between top 10 GDP countries, I extracted data through World Integrated Trade Solution software, created a new csv file including import, export and balance value of each country. I tried to use a weighted directed graph to show the global trade among 10 countries. I regarded the top 10 countries as vertex, and import and export value as weighted edges.

Jiayang Liu: I helped generate the idea about what we plan to do related to these three perspectives/graphs. To have a more convenient job with data in Jupyter, I did some integration in Excel after downloading the data. I took the first perspective by figuring out how to use Sankey diagrams to present.

Jiawen Liang: I mainly take responsibility for figuring out the visualization of heatmaps and stacked bar charts. Based on the second goal of the project proposal, I continuously brainstormed and developed the idea as well as collected and cleaned data to make it work, hoping to help draw out more meaningful and significant analyses on the aspect of world trade.

Yitian liang: I am primarily responsible for the line chart, which depicts crude oil import and export. Based on the information provided by the current charts, I examined the effect of COVID-19 and the trade war on crude oil exports and arrived at a highly solid result. I outlined a portion of the project, as well as its viability and future prospects.

Exhibit 1: Trade in 2019 for Top Five Exporting and Importing Countries

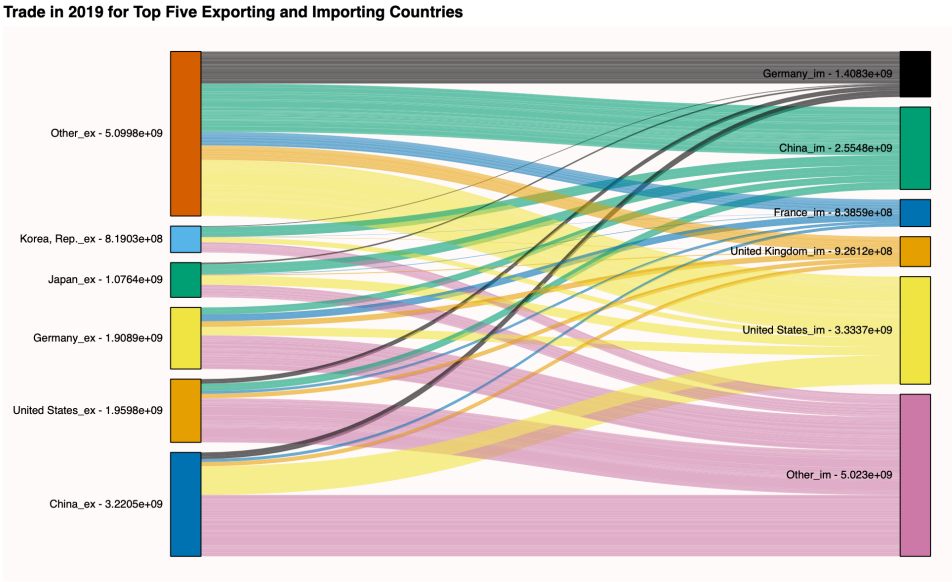


Exhibit 2: Trade in United States in 2019

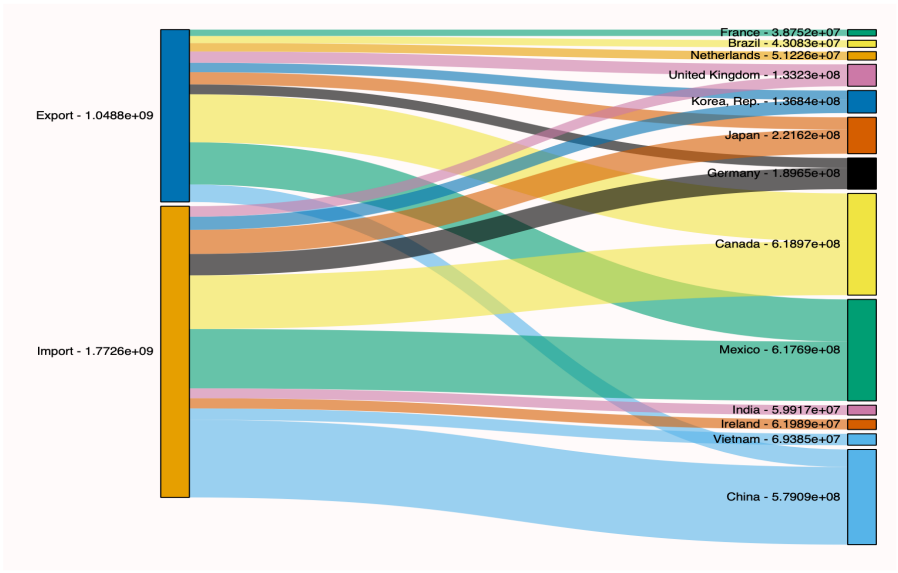


Exhibit 3: Trade in China in 2019

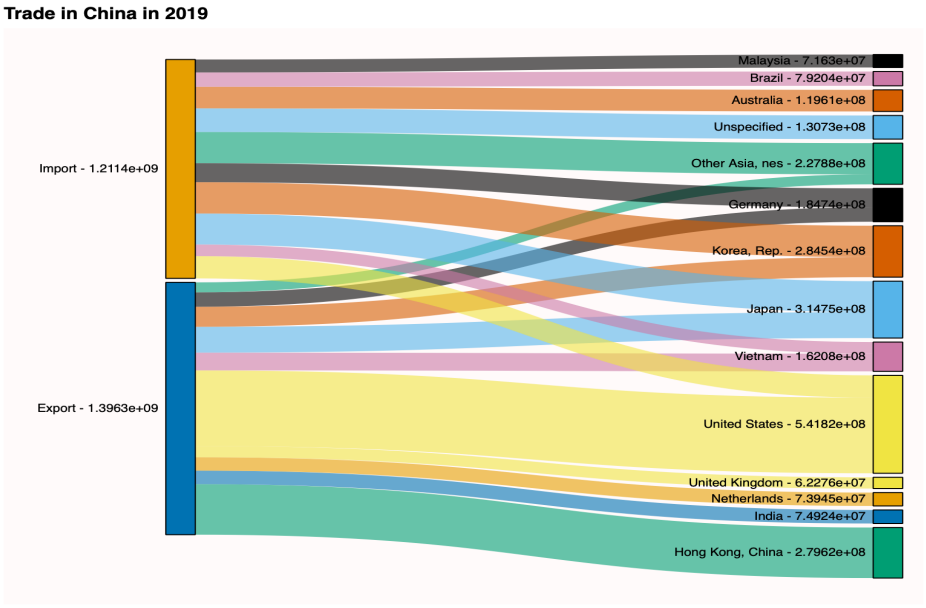


Exhibit 4: Changes and Trends - Import

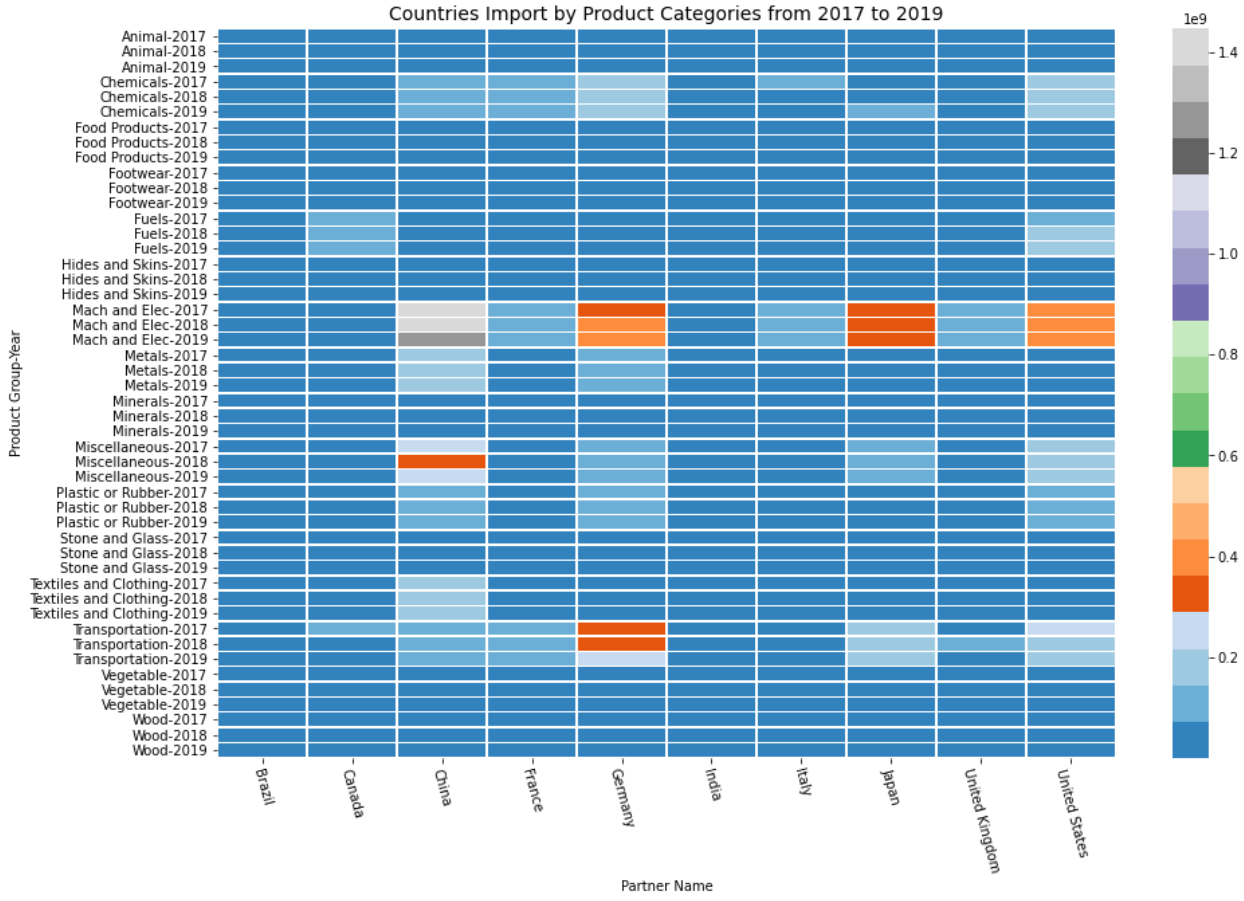


Exhibit 5: Changes and Trends - Export

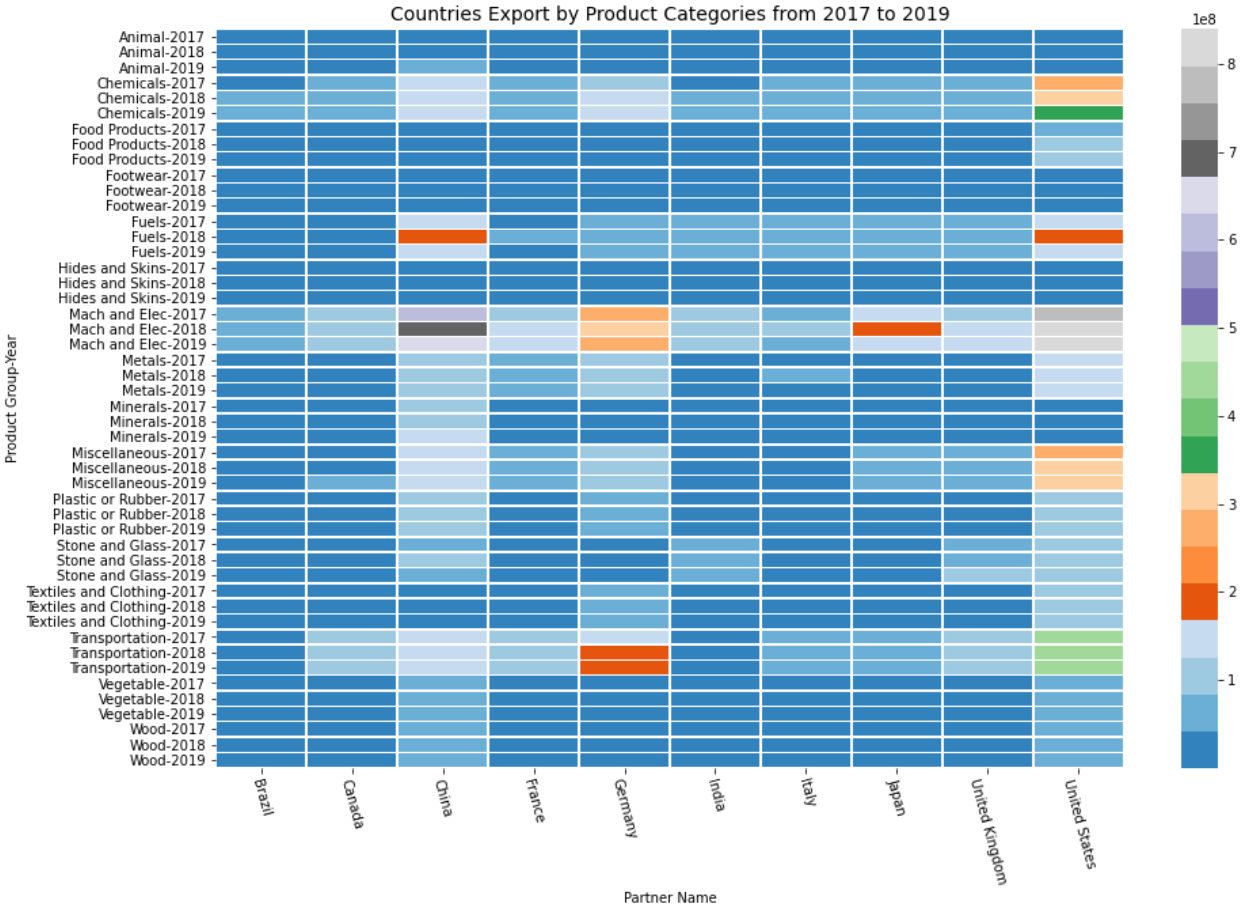


Exhibit 6: Comparison of import among countries in 2018 and 2019

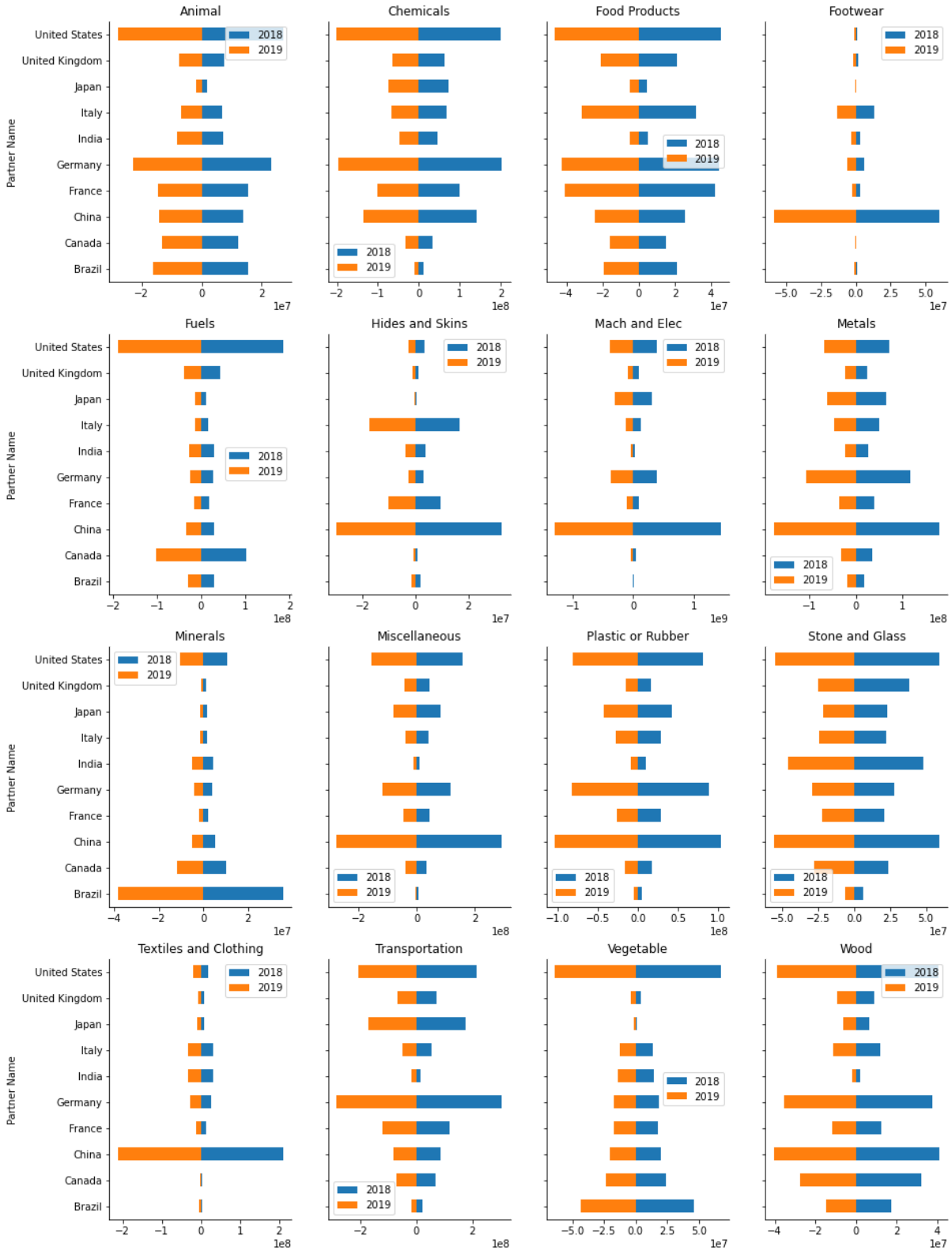


Exhibit 7: Comparison of export among countries in 2018 and 2019

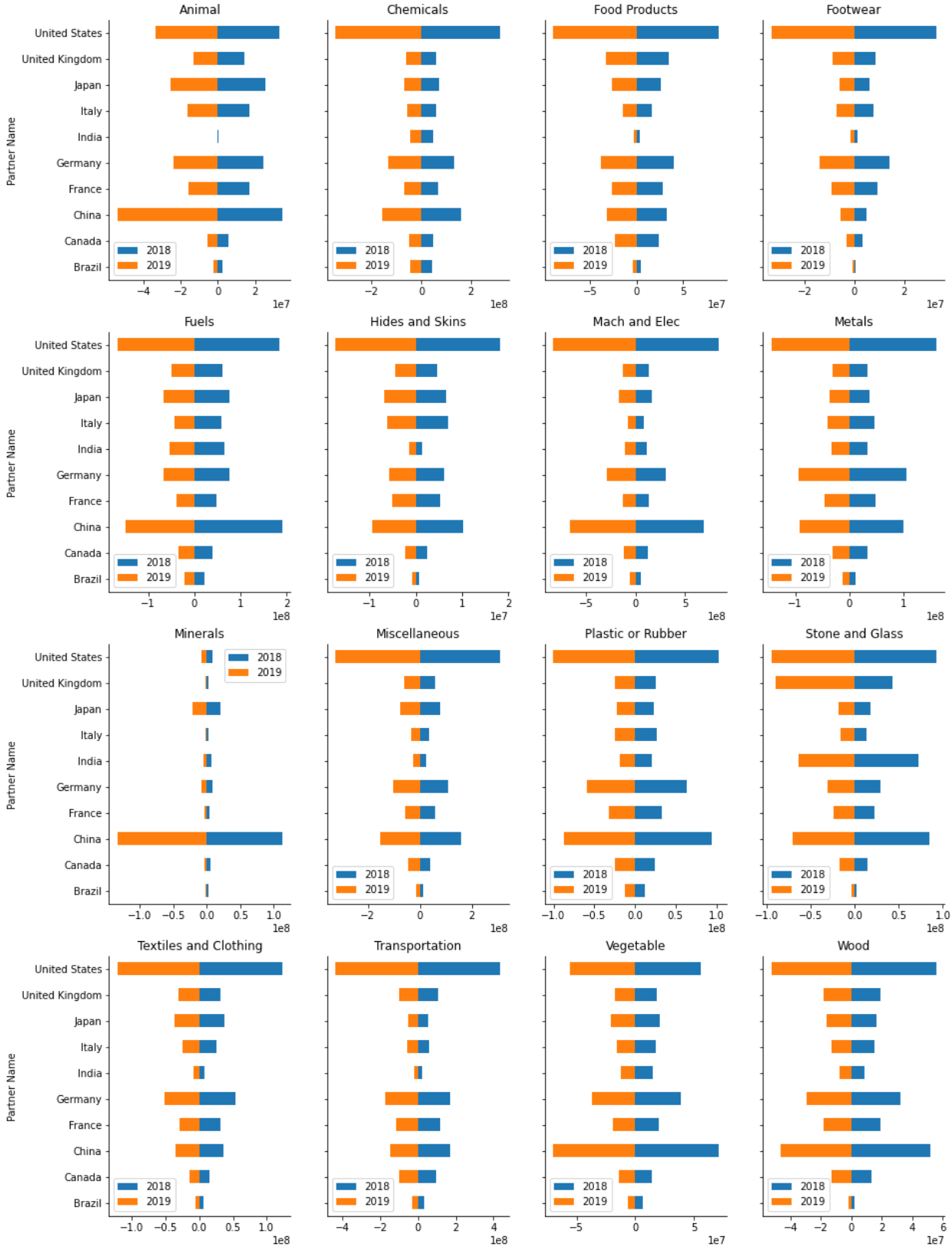


Exhibit 8: Comparison between import and export - China

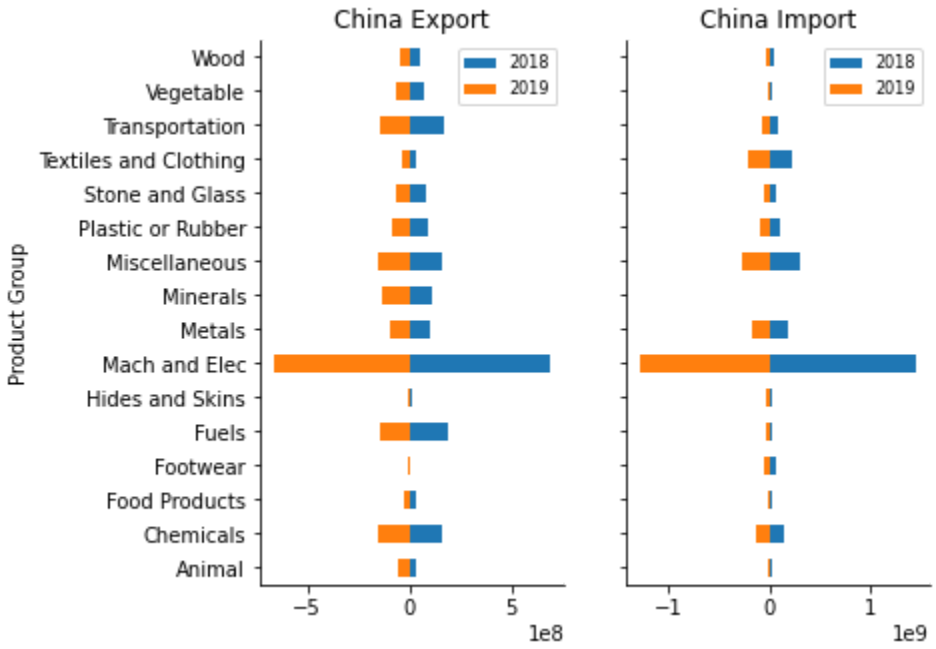


Exhibit 9: Comparison between import and export - United States

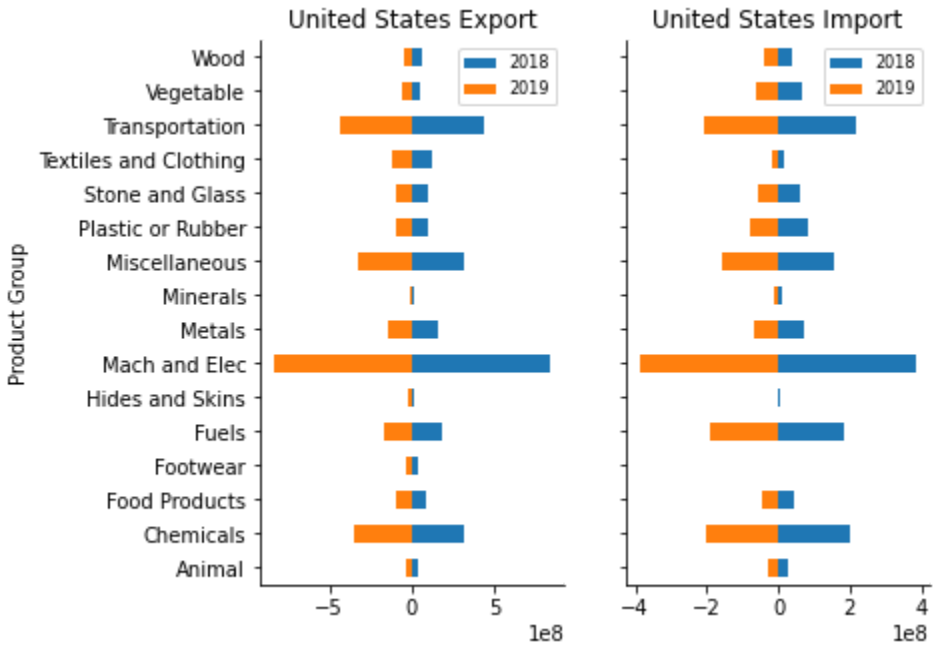
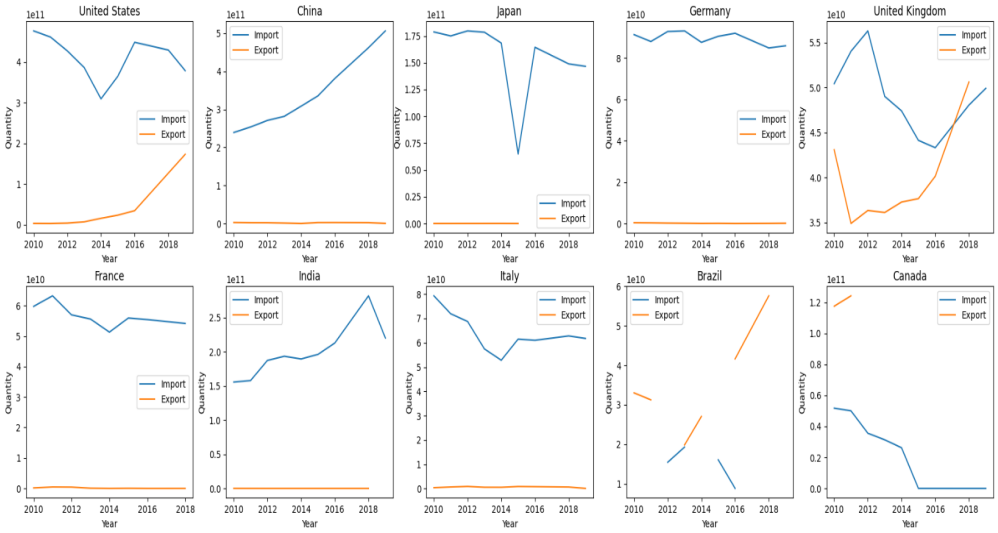


Exhibit 10: From 2010 to 2019, trends in crude oil imports and exports for ten nations.



References

1. New Zealand Ministry of Foreign Affairs and Trade. (n.d.). *The importance of the Suez Canal to global trade - 18 April 2021*. New Zealand Ministry of Foreign Affairs and Trade. Retrieved December 3, 2021, from https://www.mfat.govt.nz/ru/trade/mfat-market-reports/market-reports-middle-east/the-importance-of-the-suez-canal-to-global-trade-18-april-2021/.
2. *World Integrated Trade Solution (WITS)*. World Integrated Trade Solution (WITS) | Data on Export, Import, Tariff, NTM. (n.d.). Retrieved December 2, 2021, from https://wits.worldbank.org/.
3. *GDP Ranked by Country 2021*. (2021). World Population Review. https://worldpopulationreview.com/countries/countries-by-gdp