

# How the Maritime Shipping Influences Global Economy

Chayanggoon Hipkaeo (e-mail: [hipkaeo.chayanggoon@stud.hs-fresenius.de](mailto:hipkaeo.chayanggoon@stud.hs-fresenius.de))

HS-Fresenius: Data Science for Business

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Definitions</b>	<b>2</b>
<b>3</b>	<b>Data Sources</b>	<b>3</b>
<b>4</b>	<b>Maritime Market Analysis and Segments</b>	<b>3</b>
<b>5</b>	<b>Types of Cargo Ships (Cargo Vessels)</b>	<b>4</b>
<b>6</b>	<b>Leaders of The World's Major Fleets</b>	<b>8</b>
<b>7</b>	<b>The Global Port Operations</b>	<b>11</b>
<b>8</b>	<b>Conclusion</b>	<b>22</b>
<b>9</b>	<b>A Journey of R-Studio Project</b>	<b>23</b>
<b>10</b>	<b>Affidavit</b>	<b>27</b>
<b>11</b>	<b>References</b>	<b>28</b>

Rendered at 13 February, 2023

Word count: 5987

# 1 Introduction

Shipping is the cornerstone of international trade. It is connected to numerous industries. Shipping is used to deliver food, medicine, and technological equipment, among other things. Especially in the developing world, cost-effective shipping mechanisms are used to promote growth and sustainable development. There is no country in the world capable of surviving without maritime and shipping activities (direct or indirect activities). Due to the fact that trade and commerce are fully dependent on maritime shipping, it has become a crucial factor in the global economy. As raw materials or completed goods, our products for daily use are delivered by sea.

This project, therefore, focuses on the study of maritime shipping, beginning with the topic of how the maritime shipping industry impacts the global economy. Consequently, further study was undertaken to present the proposal. Nonetheless, as the study was being conducted, other questions arose, such as “What are the major components of maritime transportation?”, “Which region is the leader of the world’s major fleets?”, “Which countries are listed as the top five global port operations?”, “Who is the most powerful country in the world maritime industry?”, etc.

Also, the purpose of this project report is to express the use of “R Studio” with coding, processing, and using tools in order to present the database and statistics of the world economy regarding maritime transport operations. All of the above will be discussed, as well as the impact of the shipping industry on the global economy.

Following an explanation of several aspects relating to the marine transportation industry is provided as a result of the above:

- Maritime Market Analysis and Segments
- Types of Cargo Ships (Cargo Vessels)
- Leaders of the World’s Major Fleets
- The Global Ports Operations

## 2 Definitions

**Motor Vessel (or Motor Ship):** “A vessel that employs a marine diesel engine as its propulsion method. The fuel efficiency of diesel engines causes them to operate more cost effectively than steam turbine plants, which were used in earlier vessel designs” (Miller 2013).

**Dead-Weight Tons (dwt) and Gross Tons (gt):** According to UNCTAD (2022), the unit dead-weight tons (dwt) is used to characterize a vessel’s cargo-carrying capacity, whereas gross tons (gt) is used to measure its size. The latter is used to measure shipbuilding and recycling activity, while the former is used to quantify freight transit capacity.

**Hong Kong SAR:** The acronym SAR stands for “Special Administrative Regions.” As integral parts of the country, the special administrative regions (SAR) of the People’s Republic of China are one of the provincial-level administrative divisions directly under the supervision of the Central People’s Government (State Council) (WIKIPEDIA 2022).

**UN Code (United Nation Code):** UN Code refers to 'the identification of particular locations is frequently required in information interchange in international trade and transport, to direct the movement of goods, e.g., in addresses, shipping marks, and data elements identifying ports of call, ports or places of loading and unloading, ports or places of transshipment and destination, etc. The International Maritime Organization advises that ports and other sites adopt the United Nations code. The United Nations Code for Trade and Transport Locations has been known as UN/LOCODE for a number of years ([UNECE 1981](#)).

### 3 Data Sources

Since maritime transport is complicated, data sets are used to present the project beginning with 7 observations and 2 variables collect particular data in Statista 2022. This is used to display the current number of ships of different categories in existence. The next part utilizes UNCTAD data to begin the topic of shipping fleets, with 42 observations and 6 variables spanning the various fleet types from 1981 to 2022. In addition, 31 observations and 6 variables are utilized to characterize the flag states from 1992 to 2022. While the most frequently used data set in the study is global port operations, it also compiles a list of 480 ports comprised of 12 variables that are used to investigate the top five countries (top twenty port operations) with port operations and ship traffic density. The mentioned data sets, however, can be accessed from the list below.

- [Table 1: Number of ships in the world merchant fleet as of January 1, 2022](#) (Excel format)
- [Fleet Expansion \(2011 - 2021\)](#) Retrieved from: Figure 1. World fleet by principal vessel type
- [Table 2: Fleet market by region of beneficial ownership, 2022](#) Retrieved from: Figure 2. Fleet market by region of beneficial ownership, 2022
- [Registration of Flag State \(based on DWT\)](#) Retrieved from: Figure 3. Vessels capacity in top five registries
- [The Global Port Operations](#) (Port\_Data.csv)

### 4 Maritime Market Analysis and Segments

According to ModorIntelligence ([2022a](#)), during the forecast time frame, it is anticipated that the maritime freight transport business will grow by around 4%. The traditional port environment is changing as demographic, technological, and sustainability forces impact daily operations and shape a variety of substantial modifications. The global marine industry is predicted to face a number of challenges, including geopolitical issues such as the US-China trade negotiations and Brexit.

As one of the world's largest trading partners, China has a tremendous impact on the marine industry. The Chinese container shipping industry faces probable early-year disruptions as

a result of COVID-19. Even though bunker prices have increased between 35 and 45 percent since January 2020, principally owing to changes in fuel rules on the use of Low Sulphur Fuel by IMO 2020, this effect has been negated by the decline in oil prices caused by COVID-19 (ModorIntelligence 2022b).

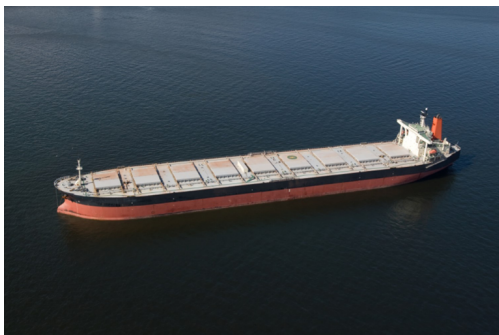
By type (containerized and non-containerized) and geography, the maritime freight transport industry is split (North America, Europe, Asia-Pacific, the Middle-East and Africa, and South America) (ModorIntelligence 2022b).

## 5 Types of Cargo Ships (Cargo Vessels)

As mentioned maritime transport activities and segments in previous chapter, it can be further discussed types of vessel in this section. Based on OneOcean (2022), cargo ships are classified according to their function, size, cargo type, etc. which can be mainly classified as follows:



- ***Container Ships*** : As the name suggests, a container vessel is a ship designed primarily to carry large quantities of cargo packed into various types of containers.



- ***Bulk Carriers*** : Bulk carriers are a type of ship that transports cargo in bulk amounts (often dry cargo). Such ships convey “loose cargo,” i.e., cargo that lacks specific packaging, and typically transport cereal grains, ores, coals, and even cement.



- ***Tanker Ships*** : Tanker ships are specialized vessels designed to transport enormous quantities of liquid cargo. Depending on the cargo they transport, tankers are further subdivided into distinct types. Oil tankers, liquid gas carriers, and chemical and product carriers are the most common types of tankers.



- ***Roll-on Roll-off Ships*** : Roll-on/roll-off is abbreviated as Ro-Ro. This category includes the Pure Car Carrier (PCC) and the Pure Car and Truck Carrier (PCTC), among others, which are designed to transport wheeled freight.



- ***Passenger Ships*** : As their name suggests, passenger ships are primarily used for transporting passengers, such as ferries, cruise ships, etc.



- **Offshore Vessels** : Offshore vessels mostly aid in oil exploration and maritime construction. A variety of offshore vessels exist, for instance, supply vessels (a.k.a. support vessels), accommodation barges, crew boats, etc.



- **Fishing Vessels** : Boats and ships designed for fishing purposes at sea are known as “fishing vessels.”

## Types of Ship in the World Merchant Fleet

Statista (2022) reported that as of January 1, 2021, over 15,000 of the approximately 55,000 merchant ships engaged in international trade were general cargo ships. With this breakdown, general cargo ships made up about 27% of the global merchant fleet, followed by bulk cargo carriers, crude oil tankers, Ro-Ro/Passenger ships, chemical tankers, container vessels, and liquefied natural gas carriers.

In light of this, the data set generated from such institute can be structured into tables to describe the mentioned information as follows.

```
#Steps of creating a data frame and table
types <- c("General cargo ships", "Bulk cargo carriers", "Crude oil tankers",
          "Ro-Ro/ Passenger ships", "Chemical tankers", "Container ships",
          "Liquefied natural gas tankers")
numbers <- c("15.106", "12.258", "7.350", "7.027", "5.664", "5.307", "2.031")
types <- data.frame(types, numbers)
knitr::kable(types, "pipe", col.names = c("Type", "Number(in thousand)"),
              align = c("l", "c"))
```



Type	Number(in thousand)
General cargo ships	15.106
Bulk cargo carriers	12.258
Crude oil tankers	7.350
Ro-Ro/ Passenger ships	7.027
Chemical tankers	5.664
Container ships	5.307
Liquefied natural gas tankers	2.031

Table 1: Number of ships in the world merchant fleet as of January 1, 2022

Source: Statista (2022)

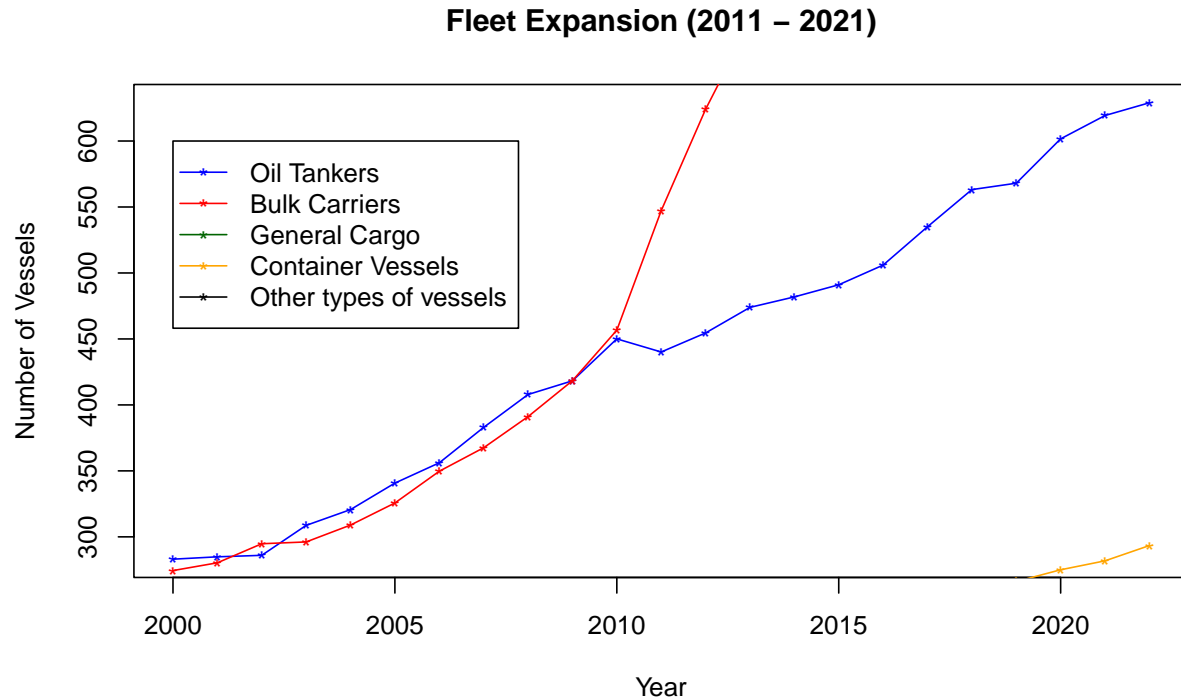
On another side, UNCTAD (2022) has looked at the fleet growth from a different angle than Statista, it measures fleet expansion by size with regard to the total mass in dwt of the different types of ships that have been built over time resulting in significant differences as follows; the world fleet's carrying capacity increased by 63 million dwt (dead-weight tons) in January 2021, reaching 2.1 billion dwt. In recent years, tonnage has expanded significantly across all market groups with the exception of general cargo carriers. Bulk carriers had a particularly rapid increase. From 2011 to 2021, their part of overall carrying capacity increased from 39% to 43%, but the share of oil tankers decreased from 31% to 29% and the share of general cargo decreased from 6% to 4%. All of these can be illustrated in the below line graph.

```
#Import and specify range of data
typeofvessels <- read.csv("~/Downloads/R-Project/HBS2022_5_2fig1 (1).csv",
                          dec="," , nrow = 42)

#Create data frame, and specify the range of data (the top five vessels)
a <- typeofvessels[c(20:42),]

#Plot data
plot(a$Category, a$`Oil.tankers`, type="o", col="blue", pch="*",
     main = "Fleet Expansion (2011 - 2021)",
     xlab="Year", ylab="Number of Vessels", lty=1)
points(a$Category, a$`Bulk.carriers`, col="red", pch="*")
lines(a$Category, a$`Bulk.carriers`, col="red",lty=1)
points(a$Category, a$`General.cargo`, col="dark green", pch="*")
lines(a$Category, a$`General.cargo`, col="dark green",lty=1)
points(a$Category, a$`Container.ships`, col="orange", pch="*")
lines(a$Category, a$`Container.ships`, col="orange",lty=1)
points(a$Category, a$`Other.types.of.ships`, col="black", pch="*")
lines(a$Category, a$`Other.types.of.ships`, col="black",lty=1)
legend(2000,600,legend=c("Oil Tankers","Bulk Carriers","General Cargo",
                        "Container Vessels","Other types of vessels"),
```

```
col=c("blue","red","dark green","orange","black"),
pch=c("*","*","*","*","*"),lty=c(1,1,1,1,1), ncol=1)
```



## 6 Leaders of The World's Major Fleets

By 2020, China, the Republic of Korea, and Japan would have dominated the global shipbuilding industry. These three nations were responsible for 94% of shipbuilding in terms of gross tonnage. Bangladesh and India together accounted for 71% of ship recycling, while Pakistan contributed an extra 17% (UNCTAD 2022).

As of January 2022, 53% of the world's fleet tonnage was owned by the top five ship-owning economies. Greece owned 18% of the market, which was followed by China (13%), Japan (11%), Singapore (6%), and Hong Kong SAR (6%). (5 percent). Asian companies owned half of the world's tonnage. Three-quarters of the owners were from Europe, and six percent were from North America. Africa, Latin America, and the Caribbean each had just over 1% of the market, and Oceania had just under 1% (UNCTAD 2022).

According to the different regions of the world's fleet, the statistics can be classified into six areas for commercial ship fleets and maritime transport operators. Asia has the greatest fleet, comprising nearly half of the global fleet. Europe, North America, Latin America, Africa, and Oceania come next, in that order, as displayed in the following table.



```

#Import data, create data frame, and specify the range of data
library(readxl)
Fleet_Ownership <- read_excel("~/Downloads/R-Project/HBS2022_5.xlsx",
  range = "A2:B8")

#Create new table for specific data
knitr::kable(Fleet_Ownership, "pipe", col.names = c("Regions", "Millions DWT"),
  align = c("l", "c"))

```

Regions	Millions DWT
Asia	1128.08064788556
Europe	840.883971237214
Northern America	128.418795910813
Latin America and the Caribbean	26.187948274988
Africa	25.7777428605973
Oceania	6.3004778883438

Table 2: Fleet market by region of beneficial ownership, 2022

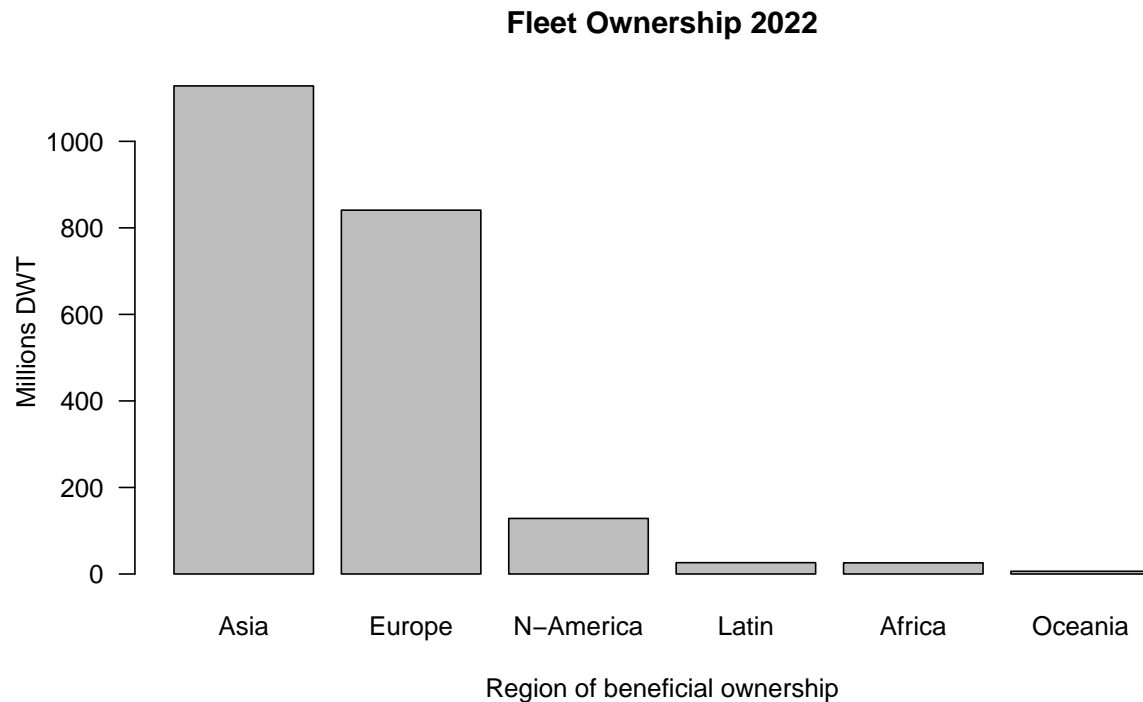
Source: UNCTAD (2022)

Based on the data in the table illustrating the world's fleet according to the world's regions, when turned to a bar graph as displayed. We can obviously see that Asia and Europe possess some of the greatest fleets in the world, demonstrating the importance of marine transport in global industry, trade, and commerce.

```

#Utilize bar plot format
fleetowner2022 <- barplot(c(1128.08064788556, 840.883971237214,
  128.418795910813, 26.187948274988, 25.7777428605973,
  6.3004778883438),
  names.arg = c("Asia", "Europe", "N-America", "Latin",
    "Africa", "Oceania"),
  main = "Fleet Ownership 2022", las=1,
  ylab = "Millions DWT", las=1,
  xlab = "Region of beneficial ownership")

```



## Flag States

However, numerous commercial ships are registered under a flag that does not correlate with the owner's nationality. More than half of all ships owned by Japanese businesses were registered in Panama at the start of 2021, while 25% of ships owned by Greek entities were registered in Liberia and 22% in the Marshall Islands whereas China - Hong Kong SAR and Singapore flags are ranked as fourth and fifth respectively ([UNCTAD 2022](#)).

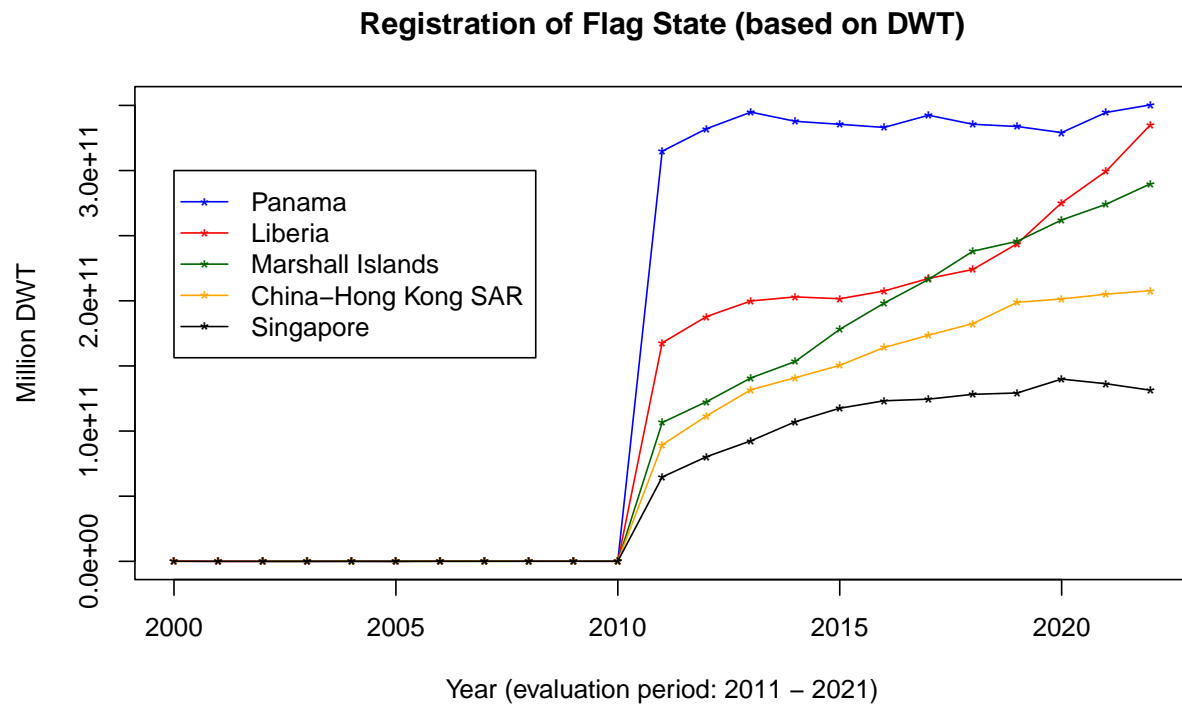
```
#Import data and specify range of data
library(readxl)
HBS2022_5_2_3 <- read_excel("~/Downloads/R-Project/HBS2022_5.2.3.xlsx",
  range = "A2:F25")

#Plot data
plot(HBS2022_5_2_3$Category, HBS2022_5_2_3$Panama, type="o", col="blue", pch="*",
  main = "Registration of Flag State (based on DWT)",
  xlab="Year (evaluation period: 2011 - 2021)", ylab="Million DWT", lty=1)
points(HBS2022_5_2_3$Category, HBS2022_5_2_3$Liberia, col="red", pch="*")
lines(HBS2022_5_2_3$Category, HBS2022_5_2_3$Liberia, col="red", lty=1)
points(HBS2022_5_2_3$Category, HBS2022_5_2_3`Marshall Islands`,
  col="dark green", pch="*")
lines(HBS2022_5_2_3$Category, HBS2022_5_2_3`Marshall Islands`,
  col="dark green", lty=1)
points(HBS2022_5_2_3$Category, HBS2022_5_2_3`China, Hong Kong SAR`,
  col="orange", pch="*")
```

```

lines(HBS2022_5_2_3$Category, HBS2022_5_2_3$`China, Hong Kong SAR`,
      col="orange",lty=1)
points(HBS2022_5_2_3$Category, HBS2022_5_2_3$Singapore, col="black", pch="*")
lines(HBS2022_5_2_3$Category, HBS2022_5_2_3$Singapore, col="black",lty=1)
legend(2000,3.0e+11,legend=c("Panama","Liberia","Marshall Islands",
                             "China-Hong Kong SAR","Singapore"),
      col=c("blue","red","dark green","orange","black"),
      pch=c("*","*","*","*","*"),lty=c(1,1,1,1,1), ncol=1)

```



According to the above line graph, it is consistent with the data derived from the UNCTAD (2022) explained that the top three registration flags belonged to Panama (344 million dwt), Liberia (300 million dwt), and the Marshall Islands (274 million dwt) followed by Hong Kong Special Administrative Region and Singapore, respectively. The Marshall Islands experienced the largest increase in registrations among these five countries during the past decade.

## 7 The Global Port Operations

Consideration of global port operations is essential when analyzing each region's maritime transport business. It can be seen that, in addition to the demand for cargo ships, the time of cargo operation must be factored into port operations components starting from the arrival

of the ship, the currently berthed ship, the departure from the port, and the estimated time of arrival (ETA) at the port, so that authorities can handle the ship and port operations.

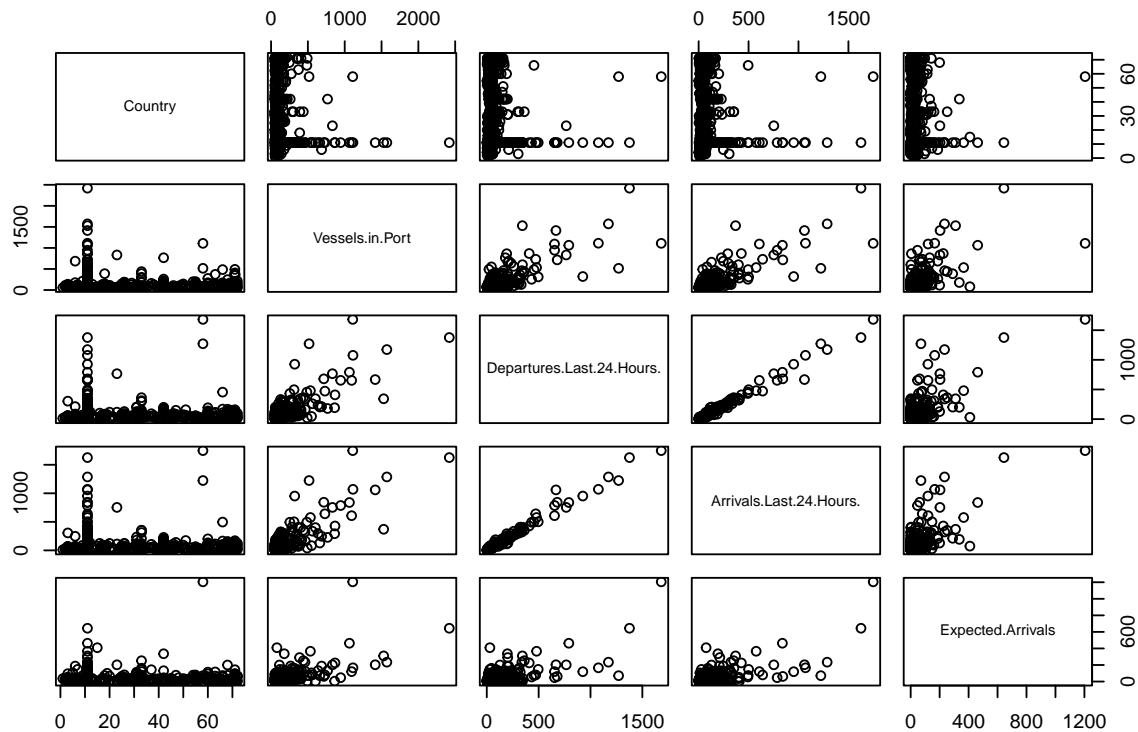
As marine operations have improved and stakeholders' needs have shifted, the question has evolved. A port serves different services today than it did decades ago. The sector has been drastically altered by seismic shifts, such as China's emergence as a manufacturing behemoth. The biggest difference, however, is the importance that seaports now play in the global economy as a whole. Once upon a time, a seaport was merely a site for loading and unloading cargo (Naik 2022).

```
#Import data and create data frame (480 obs. of 12 variables)
Port_Data <- read.csv("~/Downloads/Port_Data.csv", dec=",")
#Choosing specific columns
Column <- Port_Data[c(2, 5, 6, 7, 8)]
summary(Column)
```

```
##      Country      Vessels.in.Port Departures.Last.24.Hours.
## Length:480      Min.      : 51.0      Min.      : 0.00
## Class :character 1st Qu.: 63.0      1st Qu.: 26.75
## Mode  :character Median : 86.0      Median : 48.00
##                      Mean      : 153.3      Mean      : 98.98
##                      3rd Qu.: 144.0      3rd Qu.: 102.00
##                      Max.      :2420.0      Max.      :1682.00
## Arrivals.Last.24.Hours. Expected.Arrivals
## Min.      : 1.0      Min.      : 0.00
## 1st Qu.: 30.0      1st Qu.: 3.00
## Median : 56.0      Median : 16.00
## Mean      : 108.7      Mean      : 39.23
## 3rd Qu.: 106.2      3rd Qu.: 40.25
## Max.      :1748.0      Max.      :1205.00
```

```
#Summarize the whole data set of "Column" (480 obs. of 5 variables)
plot(Column, main = "The Global Port Operations in 24 Hours (480 Ports obs.)")
```

## The Global Port Operations in 24 Hours (480 Ports obs.)



The global port management relationship is represented by the scatter plot, which statistically averages it over a 24-hour period for each port. The number of ships that are currently berthing, arriving and waiting to come alongside, leaving the port, and the number of ships that report their expected time of arrival. These are important components of the port management. In addition, the clustering of the data reveals that many ports have fewer vessels entering than other areas that are not part of the conglomerate.

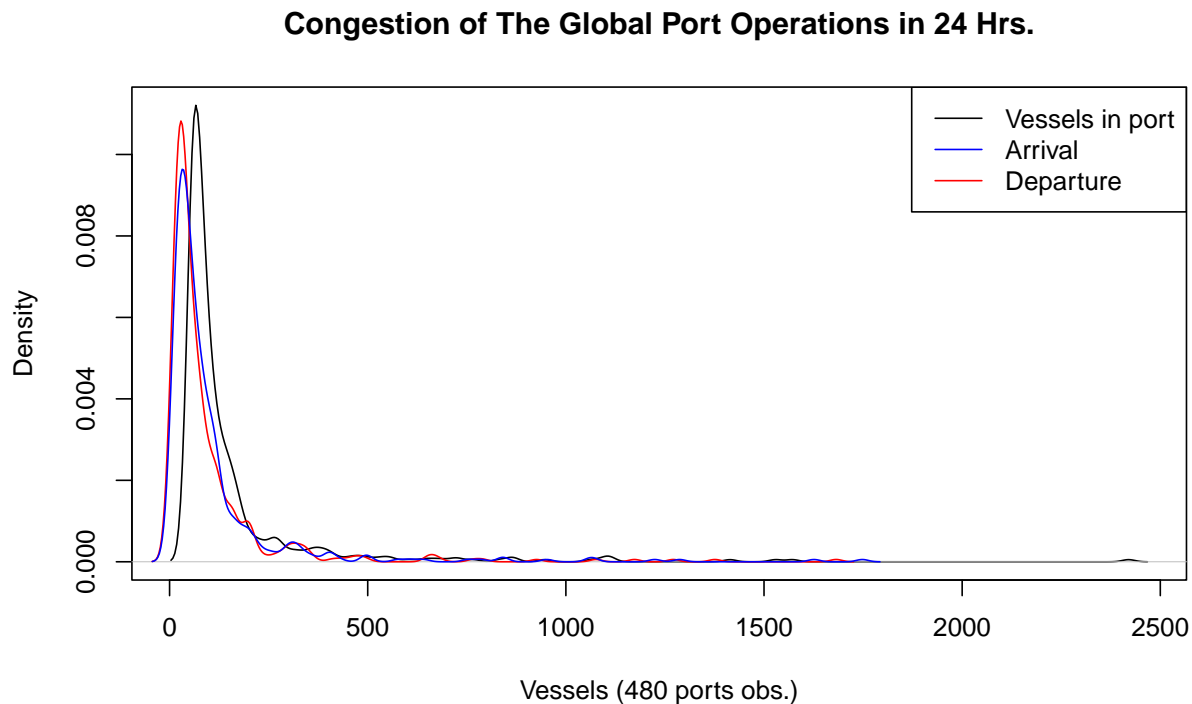
While multiple line graphs show the operational density of a port. A port density is defined mostly by the total number of ships affiliated with each harbor. The black line graph represents the berthed ships, the red line graph depicts ships at port states and waiting to berth, and the blue line graph depicts ships departing port. These three line graphs are interconnected since ships waiting to berth must be informed before they can come alongside.

```
Column1 <- (Column$Vessels.in.Port)
Column2 <- (Column$Departures.Last.24.Hours.)
Column3 <- (Column$Arrivals.Last.24.Hours.)

plot(density(Column1),
     main = "Congestion of The Global Port Operations in 24 Hrs.", xlab = "Vessels (480
lines(density(Column2), col = "red")+
  lines(density(Column3), col = "blue")
```

```
## integer(0)
```

```
legend("topright",  
      legend = c("Vessels in port", "Arrival", "Departure"),  
      col = c("black", "blue", "red"), lty = 1)
```



The objective of port operations is to make it possible for ships and their cargo to move smoothly from one terminal to another. Traveling by sea is possible in many different ways, and each kind of vessel has particular advantages. Almost all types of watercraft, including commercial ships, ferries, barges, and pleasure boats, rely on port operations to complete a voyage passage. A number of factors, such as topographical characteristics, the terminal's architectural style, and the ease of navigation, will affect how ports operate. Port authorities have developed municipal and state regulations that must be followed in addition to these operational issues. Depending on how port operations are managed, the security, safety, and success of a cargo or leisure shipping route may be affected. Port authorities are often in charge of a variety of duties, including maritime logistics, navigational support, licensing and regulatory contributions, and monitoring of the marine environment, though every port and terminal will be unique ([Daily\\_Logistics 2022](#)).

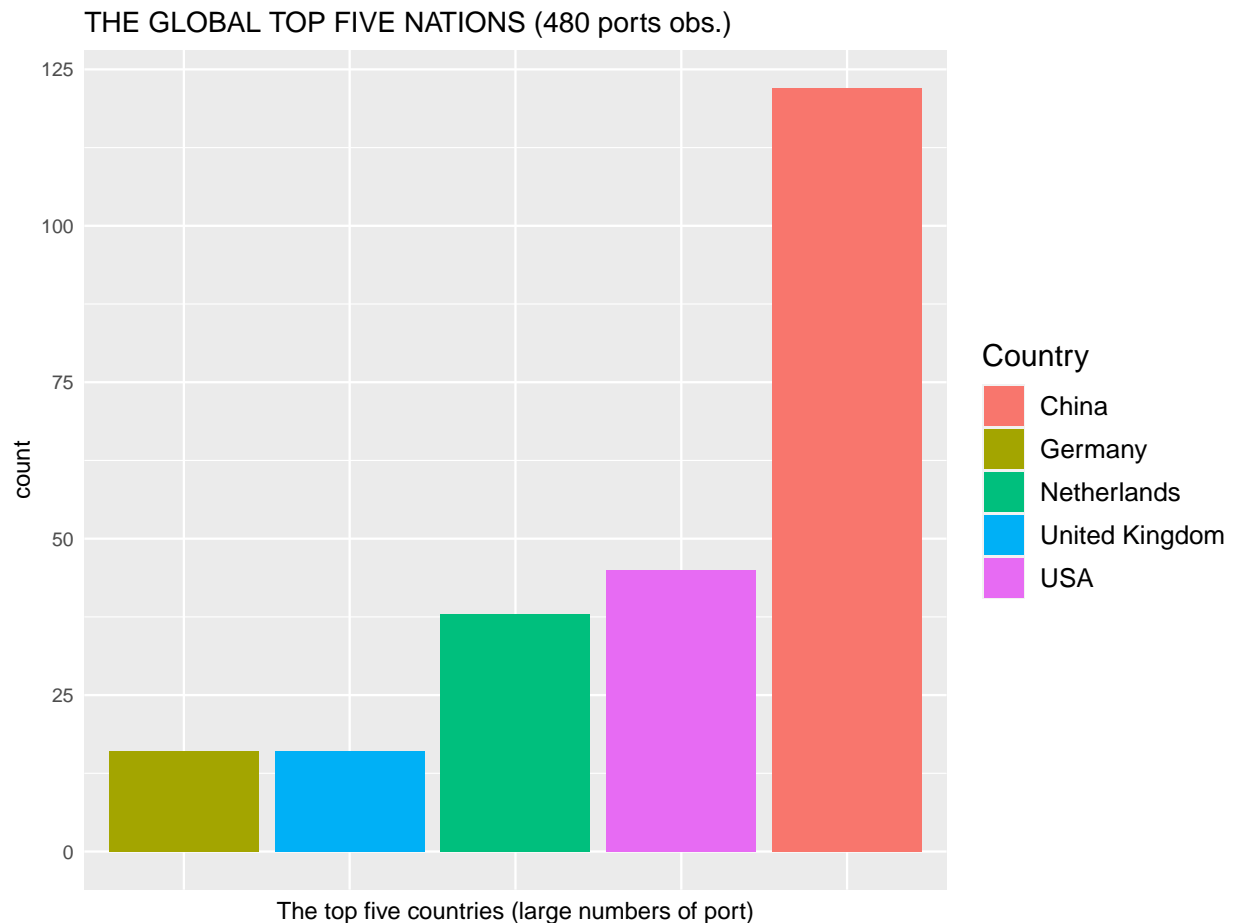
```
# Pipe Operators  
topfive <- Column %>%  
  select(Country) %>%  
  group_by(Country) %>%
```

```

summarize(count = sum(n()))>%
  arrange(desc(count))>%
  head(5)

ggplot(topfive, aes(reorder(Country, (count)), count, fill = Country))+
  geom_bar(stat = 'identity')+
  labs(title = "THE GLOBAL TOP FIVE NATIONS (480 ports obs.)")+
  xlab(label = 'The top five countries (large numbers of port)')+
  theme(axis.text.x = element_blank(), axis.ticks = element_blank())+
  theme(legend.key.size = unit(0.7,'cm'),
        legend.text = element_text(size = 12),
        legend.title = element_text(size = 14))

```



The above bar chart constructed from a list of 480 global ports reveals that the country with the most ports is among the top five, starting with China, where there are the most port operations globally, followed by Germany, the Netherlands, the United Kingdom, and the United States, in that order.

However, the port does not reflect the frequency with which ships arrive or depart. If you consider each type of ship individually, different cargo operation periods will be required.



Another aspect that indicates the frequency and density of each sort of cargo is the type of ship that enters and exits each port.

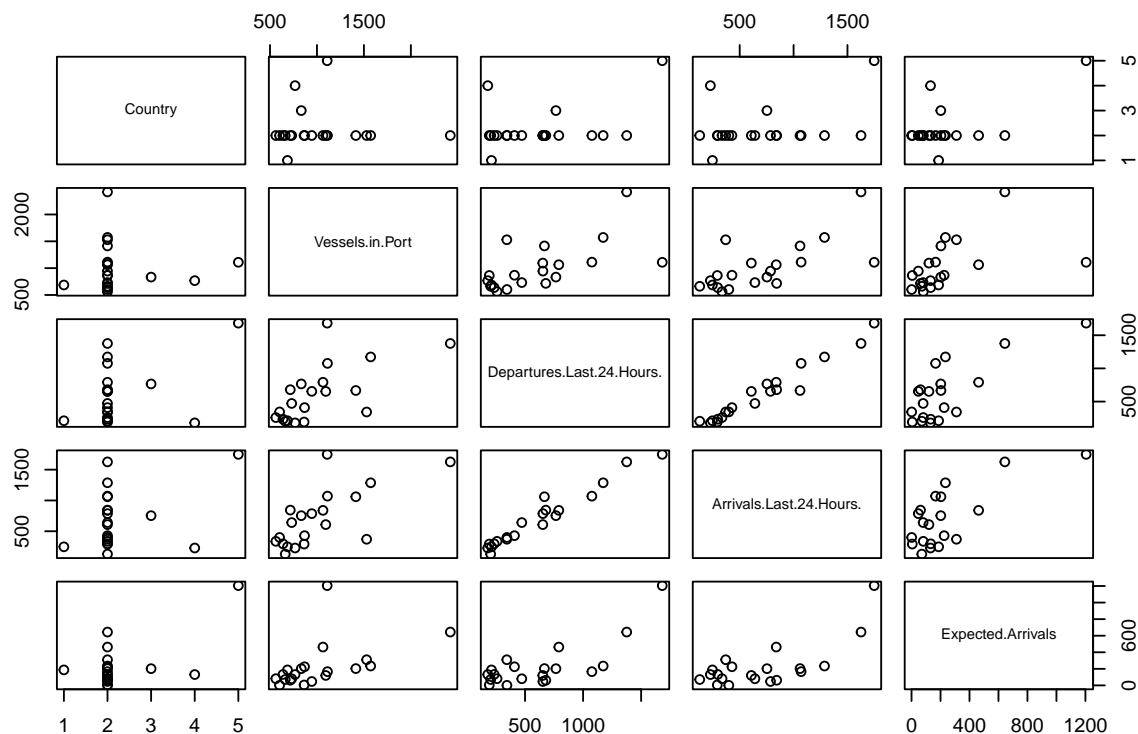
A port state with a large number of commercial ships is indicative of the importance of its commercial hub. Incorporating the vast economic potential of marine transport, this section uses worldwide port operations data from 480 international ports and subdivides it into only the top 20 major ports belonging to five port states: China, Singapore, Hong Kong SAR, The Netherlands, and Belgium, as depicted by the scatter plot titled “Port Operations of the Top Five Countries.” As demonstrated by the top twenty list, China’s port is the largest. Because of this, it may be inferred that the major ports with the greatest number of ships coming and departing are crucial for global commerce.

```
#Create data frame and specify the range of data set
portofcall <- Column[c(1:20),]
```

```
#Summarize the whole data set of "portofcall" (20 obs. of 5 variables)
```

```
plot(portofcall, main = "Port Operations of the Top Five Countries in 24 Hours (20 Ports obs.)")
```

### Port Operations of the Top Five Countries in 24 Hours (20 Ports obs.)



As mentioned in the 480 port data, the clustering of the data demonstrates that several ports have fewer vessels having operation than other sectors that are distinct from the majority. Also, the scatter plot displays the association between each activity individually. Vessels in port and departure illustrate the relationship between these activities that affect ships

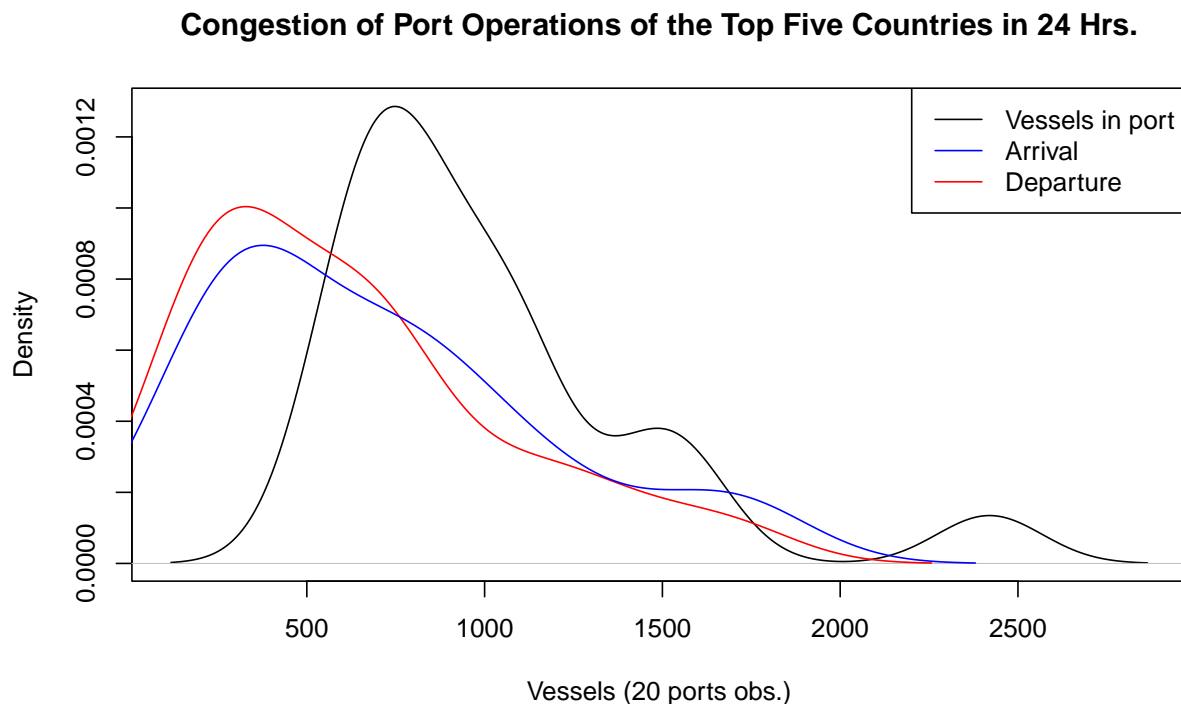
awaiting cargo operations at the port. In contrast, the relationship between vessels leaving the harbor and those waiting to come alongside is almost identical, allowing port management to effectively regulate the timing of their activities to fulfill their objectives. In addition, the anticipated arrivals of other vessels can be coordinated to ensure optimal fuel consumption and avoidance of standby fees in anchorage areas.

```
topfive_portofcall1 <- (portofcall$Vessels.in.Port)
topfive_portofcall2 <- (portofcall$Departures.Last.24.Hours.)
topfive_portofcall3 <- (portofcall$Arrivals.Last.24.Hours.)

plot(density(topfive_portofcall1),
     main = "Congestion of Port Operations of the Top Five Countries in 24 Hrs.", xlab = "Vessels (20 ports obs.)",
     lines(density(topfive_portofcall2), col = "red")+
     lines(density(topfive_portofcall3), col = "blue"))
```

```
## integer(0)
```

```
legend("topright",
      legend = c("Vessels in port", "Arrival", "Departure"),
      col = c("black", "blue", "red"), lty = 1)
```



The scatter plot can be redrawn as line graphs using only three activities (in port, arrival, and departure) to illustrate the relative performance of a cargo port. In addition, the graph

displays the top 20 ports where fewer than 2,000 ships are berthed. Only a single port in China has more than 2,000 ships alongside here, and Port of Singapore has more than 1,600 ship movements. Whereas the majority of port states have operations fewer than such mentioned ports.

```
# Assigning Operators by Grouping Specific Data (to prepare for further calculations)
portofcall_china <- portofcall[c(1:5, 7:11, 14:15, 17:20),]
china_inport <- sum(portofcall_china$Vessels.in.Port)
china_dep <- sum(portofcall_china$Departures.Last.24.Hours.)
china_arr <- sum(portofcall_china$Arrivals.Last.24.Hours.)
china_exp <- sum(portofcall_china$Expected.Arrivals)

portofcall_singapore <- portofcall[c(6),]
sing_inport <- portofcall_singapore$Vessels.in.Port
sing_dep <- portofcall_singapore$Departures.Last.24.Hours.
sing_arr <- portofcall_singapore$Arrivals.Last.24.Hours.
sing_exp <- portofcall_singapore$Expected.Arrivals

portofcall_hongkong <- portofcall[c(12),]
HK_inport <- portofcall_hongkong$Vessels.in.Port
HK_dep <- portofcall_hongkong$Departures.Last.24.Hours.
HK_arr <- portofcall_hongkong$Arrivals.Last.24.Hours.
HK_exp <- portofcall_hongkong$Expected.Arrivals

portofcall_netherland <- portofcall[c(13),]
ND_inport <- portofcall_netherland$Vessels.in.Port
ND_dep <- portofcall_netherland$Departures.Last.24.Hours.
ND_arr <- portofcall_netherland$Arrivals.Last.24.Hours.
ND_exp <- portofcall_netherland$Expected.Arrivals

portofcall_belgium <- portofcall[c(16),]
belgium_inport <- portofcall_belgium$Vessels.in.Port
belgium_dep <- portofcall_belgium$Departures.Last.24.Hours.
belgium_arr <- portofcall_belgium$Arrivals.Last.24.Hours.
belgium_exp <- portofcall_belgium$Expected.Arrivals

# Assigning Operators & Doing Calculations Using Variables

total_inport <- sum(china_inport, portofcall_singapore$Vessels.in.Port,
                    portofcall_hongkong$Vessels.in.Port,
                    portofcall_netherland$Vessels.in.Port,
                    portofcall_belgium$Vessels.in.Port)

# Calculation to Find Percentage of a Specific Port Operation
```

```

china_inport_in_percent <- china_inport*100/total_inport
sing_inport_in_percent <- sing_inport*100/total_inport
HK_inport_in_percent <- HK_inport*100/total_inport
ND_inport_in_percent <- ND_inport*100/total_inport
belgium_inport_in_percent <- belgium_inport*100/total_inport

#Create Data Frame, Prepare for Further Process
In_Port_in_Percent <- c(china_inport_in_percent, sing_inport_in_percent,
                        HK_inport_in_percent, ND_inport_in_percent,
                        belgium_inport_in_percent)
Country<- c("China", "Singapore", "Hong Kong SAR", "Netherlands", "Belgium")
totaldata_inport <- data.frame(Country, In_Port_in_Percent)

# Assigning Operators & Doing Calculations Using Variables
total_dep <- sum(china_dep, sing_dep, HK_dep, ND_dep, belgium_dep)

# Calculation to Find Percentage of a Specific Port Operation
china_dep_in_percent <- china_dep*100/total_dep
sing_dep_in_percent <- sing_dep*100/total_dep
HK_dep_in_percent <- HK_dep*100/total_dep
ND_dep_in_percent <- ND_dep*100/total_dep
belgium_dep_in_percent <- belgium_dep*100/total_dep

#Create Data Frame, Prepare for Further Process
Departure_in_Percent <- c(china_dep_in_percent, sing_dep_in_percent,
                          HK_dep_in_percent, ND_dep_in_percent,
                          belgium_dep_in_percent)
totaldata1 <- data.frame(Country, Departure_in_Percent)

# Assigning Operators & Doing Calculations Using Variables
total_arr <- sum(china_arr, sing_arr, HK_arr, ND_arr, belgium_arr)

# Calculation to Find Percentage of a Specific Port Operation
china_arr_in_percent <- china_arr*100/total_arr
sing_arr_in_percent <- sing_arr*100/total_arr
HK_arr_in_percent <- HK_arr*100/total_arr
ND_arr_in_percent <- ND_arr*100/total_arr
belgium_arr_in_percent <- belgium_arr*100/total_arr

#Create Data Frame, Prepare for Further Process
Arrival_in_Percent <- c(china_arr_in_percent, sing_arr_in_percent,
                        HK_arr_in_percent, ND_arr_in_percent,

```

```

        belgium_arr_in_percent)
totaldata2 <- data.frame(Country, Arrival_in_Percent)

# Assigning Operators & Doing Calculations Using Variables
total_exp <- sum(china_exp, sing_exp, HK_exp, ND_exp, belgium_exp)

# Calculation to Find Percentage of a Specific Port Operation
china_exp_in_percent <- china_exp*100/total_exp
sing_exp_in_percent <- sing_exp*100/total_exp
HK_exp_in_percent <- HK_exp*100/total_exp
ND_exp_in_percent <- ND_exp*100/total_exp
belgium_exp_in_percent <- belgium_exp*100/total_exp

#Create Data Frame, Prepare for Further Process
Expectation_of_Arrival_in_Percent <- c(china_exp_in_percent, sing_exp_in_percent,
                                       HK_exp_in_percent, ND_exp_in_percent,
                                       belgium_exp_in_percent)
totaldata3 <- data.frame(Country, Expectation_of_Arrival_in_Percent)

newdataframe <- data.frame(totaldata_inport, totaldata1, totaldata2, totaldata3)

newdataframe_1 <- newdataframe[c(1, 2, 4, 6, 8)]

topport <- portofcall %>%
  select(Country) %>%
  group_by(Country) %>%
  summarize(count = sum(n()))%>%
  arrange(desc(count))%>%
  head(5)

```

The 20 ports of call can be classified into five geographic categories, with 16 ports in China accounting for 80% of the total, followed by Singapore (6th out of 20), Hong Kong (12th out of 20), The Netherlands (13th out of 20), and Belgium (16th out of 20), which each contributed for 5%.

As previously indicated, I would like to know the ratio of port activity from the 20 ports of call, determined based on the nations in which these port numbers are located. It can be divided into five countries based on the proportions provided in the following table.

```

knitr::kable(newdataframe_1, "pipe", col.names = c("Country", "A: In Port %",
                                                  "B: Departure %", "C: Arrival %", "D:
                                                  align = c("l", "c"))

```

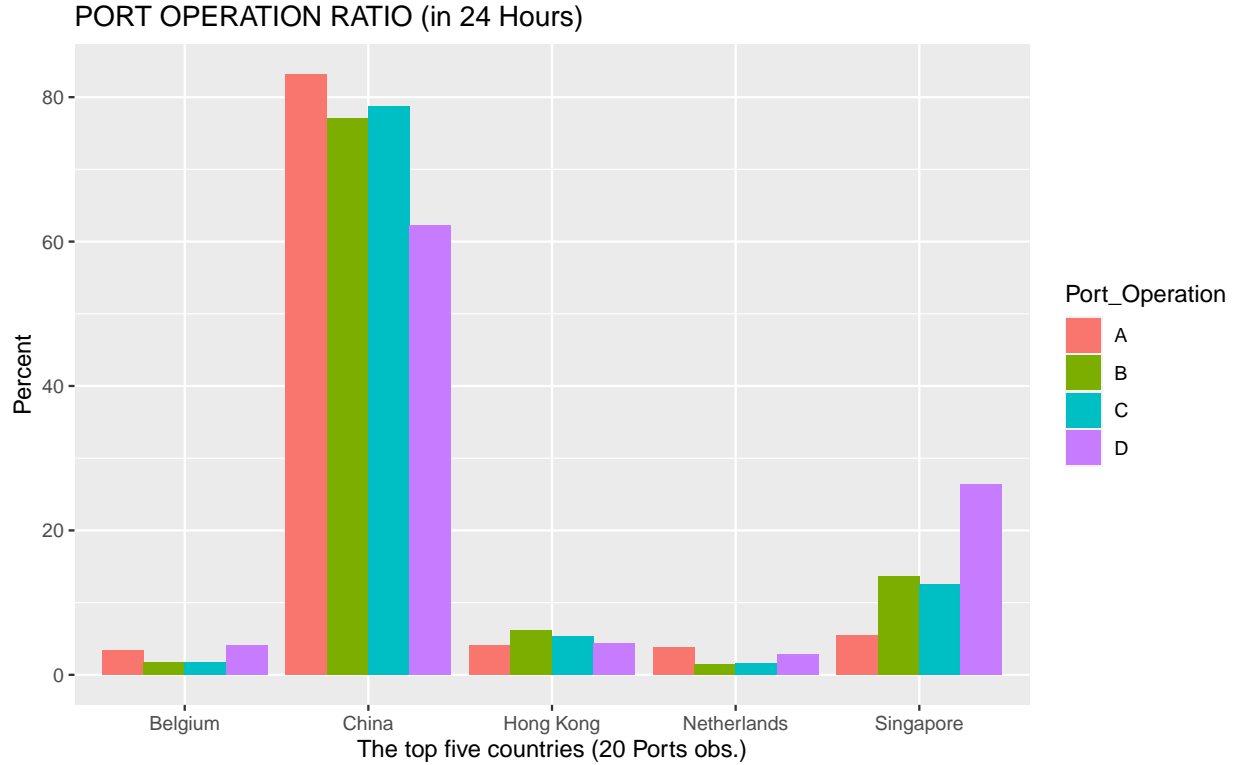
Country	A: In Port %	B: Departure %	C: Arrival %	D: ETA %
China	83.182246	77.046263	78.721579	62.245568
Singapore	5.493634	13.604012	12.498212	26.373386
Hong Kong SAR	4.126418	6.195406	5.383955	4.421099
Netherlands	3.799475	1.447752	1.630202	2.867148
Belgium	3.398227	1.706568	1.766052	4.092799

Table 3: Port operation ratios

The information in the above table can be used to create a group bar graph to compare the port operations of each country. From the table, the port operations symbols are **A** for ships docking in port, **B** for ships departing, **C** for arrival ships waiting to berth, and **D** for ships expected to arrive in the future which is implied as an ETA. It is depicted as follows on the graph:

```
dataset_1 <- data.frame(values = c(83.182246, 77.046263, 78.721579, 62.245568,
                                   5.493635, 13.604012, 12.498212, 26.373386,
                                   4.126418, 6.195406, 5.383955, 4.421099,
                                   3.799475, 1.447752, 1.630202, 2.867148,
                                   3.398227, 1.706567, 1.766052, 4.092799),
                        Group = rep(c("China",
                                      "Singapore",
                                      "Hong Kong",
                                      "Netherlands",
                                      "Belgium"),
                                   each = 4),
                        Port_Operation = LETTERS [1:4])

ggplot(dataset_1,
       aes(x = Group,
           y = values,
           fill = Port_Operation)) +
  geom_bar(stat = "identity",
           position = "dodge")+
  labs(title = "PORT OPERATION RATIO (in 24 Hours)")+
  xlab(label = 'The top five countries (20 Ports obs.)')+
  ylab(label = 'Percent')
```



China, which has the most ports, has the greatest port operations rate, as indicated by converting the data from the table displaying port operation ratios into the bar graph depicted above. This leads us to the conclusion that any nation with a significant number of ports and ships entering and leaving ports has a great potential for maritime cargo handling.

In comparison to the past, modern seaports serve a range of objectives. Recent tectonic developments, such as China's emergence as a manufacturing powerhouse, have ushered in dramatic transformations in the corporate world. However, the most significant development is the growing significance of seaports to the global economy as a whole.

## 8 Conclusion

Seaports play a vital part in the supply chain of modern industry and are evaluated using a range of factors. The sophistication of their automation technologies, the knowledge of their staff, the closeness to key markets, and the size of their docking facilities are examples of such characteristics. The traditional port environment is evolving as demographic, technological, and sustainability drivers have an impact on daily operations and form a variety of considerable changes as mentioned in the maritime market analysis and segments.

Diverse types of commercial ships and the merchant fleet are projected to increase in the future, particularly in countries that are leaders in fleet and maritime transportation, such as China and other Asian nations where they have the largest fleet, totaling over fifty percent of the world's maritime fleets. Whereas Europe, North America, Latin America, Africa, and Oceania are the subsequent continents, in that order. Moreover, in terms of a large number



of ports according to a list of 480 global ports reveals that the country with the most port operations is among the top five, beginning with China, which has the most ports in the world, followed by Germany, the Netherlands, the United Kingdom, and the United States, respectively. What represents the maritime transport industry, however, is not only the number of ports, but also the degree of congestion in port operations, which consists mostly of vessel movement in such port states. As seen by the top twenty list of the world's shipping ports, China's port is the largest and most congested, with the highest number of vital ship movements for international trade.

## 9 A Journey of R-Studio Project

When problems arise while coding each component, a great deal of time is spent searching for the correct solution. Perhaps students are scared that something will go wrong with any aspect of their Rstudio project. Consequently, we waste time searching for concrete answers. Occasionally, we fail to consider the big picture of the project, its actual objectives, and identify difficulties that we cannot handle directly. I, therefore, bring up the issues during the project because it will serve as a study guide to find solutions for the next cohort of students enrolling in the upcoming Data Science course. In short, if there is a problem that cannot be solved, the student should consult the subject teacher for advice on solving the problem.

However, even after receiving advice, the problem remains unsolved. Students are encouraged to identify such issues in their Data Science course project reports so that the course instructor is aware of them and can guide them in exploring new solutions. This will lead to developments in education for the next generation of students.

### Computer related issues

In this part, many functions, and problems are compiled. It is believed that the employment of R Studio will serve as a model for improving the work of people who employ such tools. Before detailing the issues that arose during the project utilizing the R Studio program, the computer's operating system is illustrated as follows:

- macOS Monterey (Version 12.6)
- MacBook Air (13-inch, Early 2015)
- Processor 1,6 GHz Dual-Core Intel Core i5
- Memory 8 GB 1600 MHz DDR3
- Startup Disk Macintosh HD
- Graphics Macintosh HD

The reasons given in the computer's operating system are problems caused by using an old computer that cannot update its software, which makes it impossible to update R Studio applications to a certain extent; having less memory in the computer's storage; or using too many programs and not knowing how to keep the computer in working order, including the capabilities of different computer models that affect the writing of code in R Studio.

## Inserting picture in R markdown

Importing images is problematic owing to downloading “.webp” image files (such as “Offshore Vessel.webp”) and being unable to retrieve images placed in the folder. Therefore, whenever you download an image file from a website, you must ensure that it is compatible with R Studio (e.g. jpeg, png, etc.).

A second issue I noticed while importing image files was that normal coding might not function on a Macbook (produced in early 2015). If such a problem does exist, it can be resolved by writing code similar to the example shown below.

```
# ```{r echo=FALSE, out.width="100%"}  
# knitr::include_graphics("Global_Maritime_Growth_Rate_2020.jpeg", error = FALSE)  
# ```
```

Also ensure the below code has been written prior to importing the image file.

```
# ```{r include=TRUE, echo=TRUE}  
# getwd()  
# ```
```

The `getwd()` function is an integral R function that returns an absolute file path indicating the current working directory of an R process. If the code is not adhered to and then evaluated for data validation issues (such as image file data), a problem such as the following will occur.

```
# !Package pdftex.def Error: File `shipatport.jpg' not found: using draft setting.`  
# Error: LaTeX failed to compile MarineLogisticsProject.tex. `  
# See https://yihui.org/tinytex/r/#debugging for debugging tips. `  
# See MarineLogisticsProject.log for more info. `  
# Execution halted`
```

## Bibliography

When coding to display a bibliography, the problem is that the code is entered, but the obtained date does not appear in the reference section, or the year the source was created on appears again when the retrieved date does appear. Here are some examples of such issues and their respective solutions.

As you can see in the outcome explained in the box, year “2022” is duplicated:

```
#@online{maritimetransport,  
#   author = {UNCTAD},  
#   title = {Merchant Fleet},  
#   year = {2022},  
#   url = {https://hbs.unctad.org/merchant-fleet/},
```

```
#          howpublished = {Retrieved on 2022-10-26},}

# The outcome of this coding methods is:
# UNCTAD. 2022. "Merchant Fleet." Retrieved on 2022-10-26. 2022.
# https://hbs.unctad.org/merchant-fleet/.
```

The use of this code, all essential information for bibliography are included but the problem is the alignment:

```
#@misc{maritimetransport,
#      author = {UNCTAD},
#      title = {Merchant Fleet},
#      year = {2022},
#      url = {https://hbs.unctad.org/merchant-fleet/},
#      howpublished = {Retrieved on 2022-10-26},}

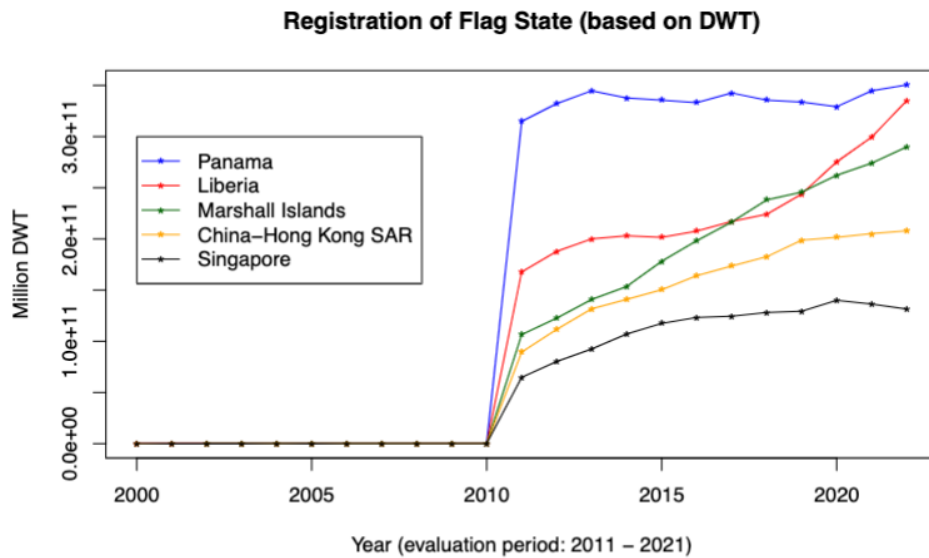
# The outcome of this coding methods is:
# UNCTAD. 2022. "Merchant Fleet." Retrieved on 2022-10-26.
# https://hbs.unctad.org/merchant-fleet/.
```

Nevertheless, even if a 'Retrieved date' can be identified and the information is complete, the results will still display the incorrect data alignment. This issue might necessitate additional study into the ideal coding approach.

### **Converting CVS to Excel format**

Converting a CSV file to an Excel file can result in data that is inconsistent with the originals and distorted graph values. Different programs might define the format of integers with decimals or commas differently. For example, in CSV files, decimals are used to define thousands, but when converted to Excel files where decimals are displayed as the value behind the digits. The below picture is an example of converting the CVS file to the Excel format and has such issue.

Even if it has no effect on the prediction period (2011-2021) but what if the data needs to be assessed further backward? Besides, the graph will look better if the historical data adheres to the original format.

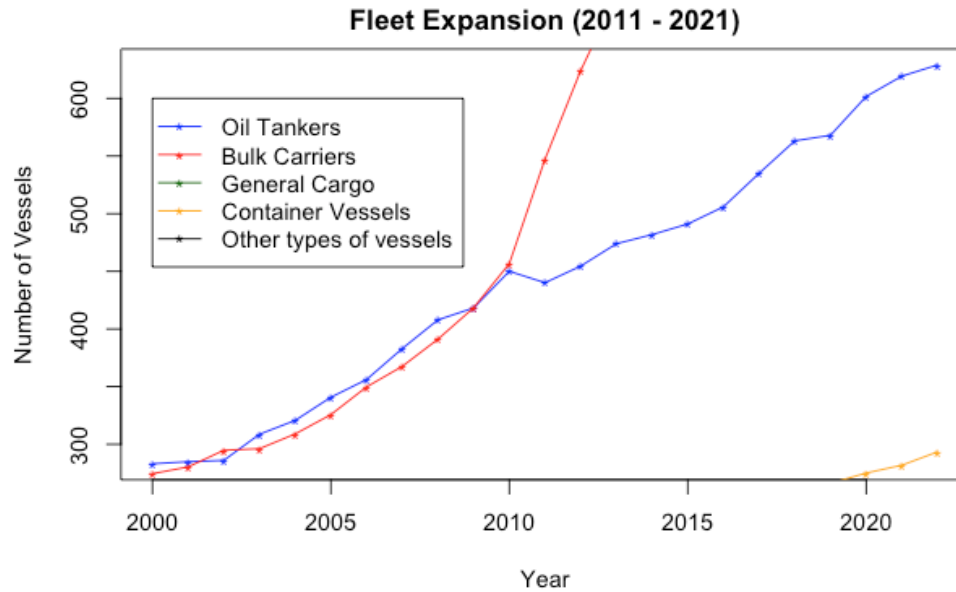


## Data set

The problem with the downloaded data set is that the number in the table is formatted by default as a text, even though it appears as a number. In this instance, the data in the data set cannot be correctly transformed for graph display. Therefore, it is always a good idea to examine the data set's quality to ensure that it is assigned to the relevant category.

From the sample provided below, I am still uncertain as to whether the reason is the downloaded data set or improper code. I spent a great deal of effort writing the code for the graph below, but I was unable to configure it.

```
# typeofvessels <- read.csv("~/Downloads/R-Project/HBS2022_5_2fig1 (1).csv",
#                             dec=",", nrow = 42)
#
# a <- typeofvessels[c(20:42),]
# plot(a$Category, a$`Oil.tankers`, type="o", col="blue", pch="*",
#      main = "Fleet Expansion (2011 - 2021)",
#      xlab="Year", ylab="Number of Vessels", lty=1)
# points(a$Category, a$`Bulk.carriers`, col="red", pch="*")
# lines(a$Category, a$`Bulk.carriers`, col="red", lty=1)
# points(a$Category, a$`General.cargo`, col="dark green", pch="*")
# lines(a$Category, a$`General.cargo`, col="dark green", lty=1)
# points(a$Category, a$`Container.ships`, col="orange", pch="*")
# lines(a$Category, a$`Container.ships`, col="orange", lty=1)
# points(a$Category, a$`Other.types.of.ships`, col="black", pch="*")
# lines(a$Category, a$`Other.types.of.ships`, col="black", lty=1)
# legend(2000, 600, legend=c("Oil Tankers", "Bulk Carriers", "General Cargo",
#                            "Container Vessels", "Other types of vessels"),
#       col=c("blue", "red", "dark green", "orange", "black"),
#       pch=c("*", "*", "*", "*", "*"), lty=c(1, 1, 1, 1, 1), ncol=1)
```



One last issue is `## integer (0)`

It appears that this last issue is a visual problem when plotting a density line graph. In the following example of written code, when knit “Rmd”, it contains ‘`## integer (0)`’ in the outcome. Even though it was a small point, I still need to find ways to find flaws that may be caused by writing code or by missing something in the syntax.

```
# plot(density(topfive_portofcall1),
#       main = "Congestion of Port Operations in 24 Hrs.",
#       xlab = "Vessels (20 ports obs.)")
# lines(density(topfive_portofcall2), col = "red")+
# lines(density(topfive_portofcall3), col = "blue")
# legend("topright",
#       legend = c("Vessels in port", "Arrival", "Departure"),
#       col = c("black", "blue", "red"), lty = 1)
```

## 10 Affidavit

I hereby affirm that this submitted paper was authored unaided and solely by me. Additionally, no other sources than those in the reference list were used. Parts of this paper, including tables and figures, that have been taken either verbatim or analogously from other works have in each case been properly cited with regard to their origin and authorship.

This paper either in parts or in its entirety, be it in the same or similar form, has not been submitted to any other examination board and has not been published.

Cologne, 13th February 2023

*James*

## 11 References