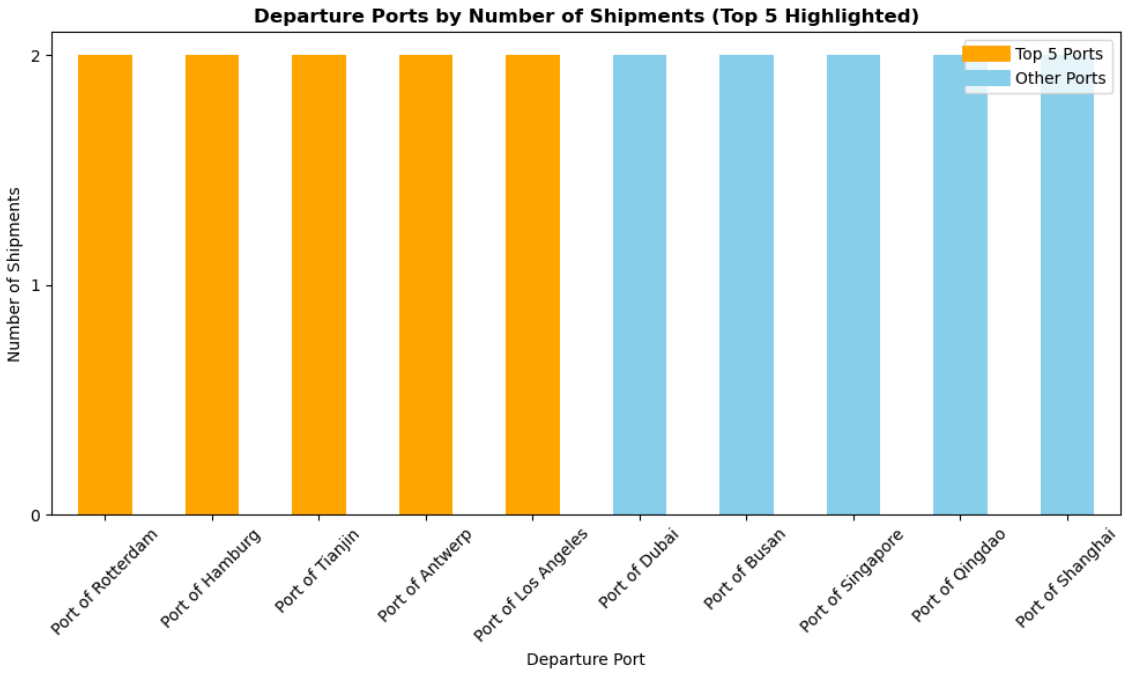
Primanta Holand Bangun – phb2123

Shipping Analysis Assignment

Please find the assignment in this github link below
https://github.com/primantah/Shipping-Analysis

Tasks and Result

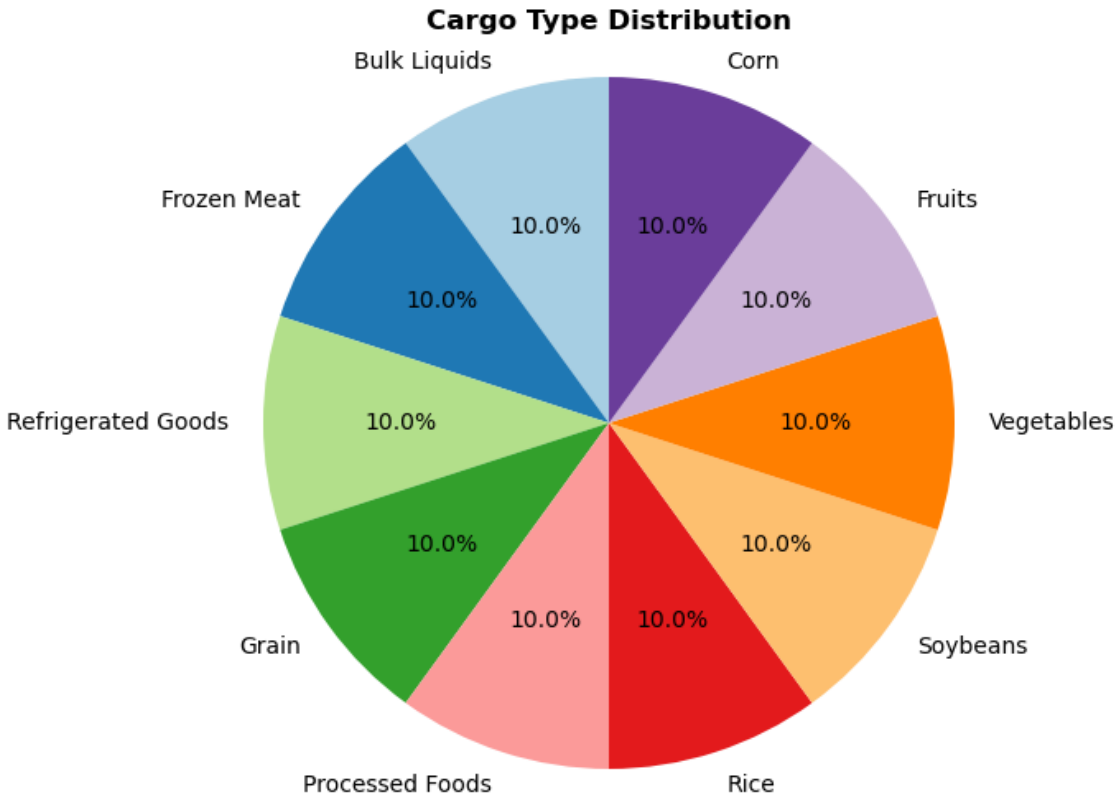
1. Top 5 Departure Ports:
   Create a bar chart showing the number of shipments departing from each port. Highlight the top 5 ports with the highest number of departures.



Insights:

- Each port has exactly 2 shipments, indicating an equal distribution of the shipping load across all ports in the dataset.
- There is no dominant port in terms of departures, and the top 5 ports are highlighted based on their detection order rather than volume.
- The chart also shows a global distribution of ports across Europe, Asia, the Middle East, and North America, reflecting a well-balanced global shipping network.

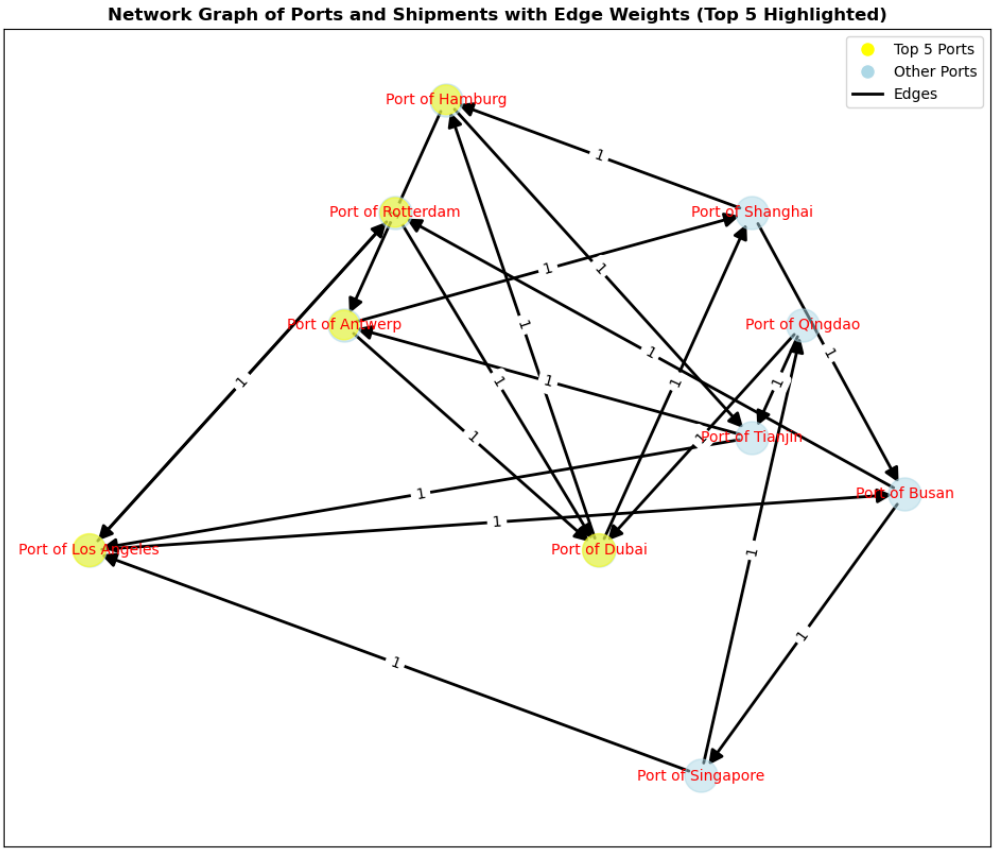
2. Cargo Type Distribution:
   Generate a pie chart illustrating the distribution of different cargo types across all shipments.



Insights:

- All cargo types—Bulk Liquids, Corn, Fruits, Vegetables, Soybeans, Rice, Processed Foods, Grain, Refrigerated Goods, and Frozen Meat—make up an equal share (10%) of the total shipments.
- This indicates a balanced variety of goods, with no emphasis on any single type, serving industries like agriculture, food processing, and energy
- The inclusion of temperature-controlled goods (Frozen Meat, Refrigerated Goods) suggests the network supports such shipments, highlighting the importance of speed and efficiency.
- This equal distribution may also reflect a strategy to reduce risk by diversifying the cargo base, making operations more resilient to market disruptions.

3. Network:
    - Construct a Network Graph:
        - Nodes: Represent ports.
        - Directed Edges: Represent shipments from a departure port to an arrival port.
        - Edge Weight: Number of shipments between two ports.
    - Visualize the Network:
        - Create a network graph highlighting the top 5 most connected ports based on the number of shipments.
        - Use node size or color to indicate the degree of connectivity.



Insights:

- Dubai and Los Angeles are the most active ports, with 5 connections to other key ports, highlighting their role as central hubs in the shipping network.
- Hamburg, Rotterdam, Antwerp, Tianjin, and Shanghai also play significant roles with 4 connections, reinforcing their importance in global shipping.
- Each edge has a weight of 1, indicating only one shipment between each pair of connected ports, suggesting that no single route is heavily trafficked.

- Each port has exactly 2 shipments, indicating an equal distribution of the shipping load across all ports in the dataset.
- The chart also shows a global distribution of ports across Europe, Asia, the Middle East, and North America, reflecting a well-balanced global shipping network.

Additional Insights

4. Displaying the Shipment Data in Table

Shipment ID	Departure Port	Arrival Port	Cargo Type	Departure Time	Arrival Time	Duration
SHIP_0001	Port of Rotterdam	Port of Dubai	Bulk Liquids	2024-01-10 08:00:00	2024-01-15 14:00:00	5 days 6 hours 0 minutes
SHIP_0002	Port of Hamburg	Port of Antwerp	Frozen Meat	2024-02-15 12:30:00	2024-02-20 18:30:00	5 days 6 hours 0 minutes
SHIP_0003	Port of Tianjin	Port of Los Angeles	Refrigerated Goods	2024-03-20 16:45:00	2024-03-25 22:45:00	5 days 6 hours 0 minutes
SHIP_0004	Port of Antwerp	Port of Shanghai	Grain	2024-04-25 09:15:00	2024-04-30 15:15:00	5 days 6 hours 0 minutes
SHIP_0005	Port of Los Angeles	Port of Rotterdam	Processed Foods	2024-05-30 14:20:00	2024-06-04 20:20:00	5 days 6 hours 0 minutes
SHIP_0006	Port of Dubai	Port of Hamburg	Rice	2024-06-05 07:50:00	2024-06-10 13:50:00	5 days 6 hours 0 minutes
SHIP_0007	Port of Busan	Port of Singapore	Soybeans	2024-07-10 11:25:00	2024-07-15 16:25:00	5 days 5 hours 0 minutes
SHIP_0008	Port of Singapore	Port of Qingdao	Vegetables	2024-08-15 18:40:00	2024-08-20 23:40:00	5 days 5 hours 0 minutes
SHIP_0009	Port of Qingdao	Port of Dubai	Fruits	2024-09-20 05:10:00	2024-09-25 11:10:00	5 days 6 hours 0 minutes
SHIP_0010	Port of Shanghai	Port of Busan	Corn	2024-10-25 22:55:00	2024-10-30 04:55:00	4 days 6 hours 0 minutes
SHIP_0011	Port of Rotterdam	Port of Los Angeles	Bulk Liquids	2024-11-30 13:35:00	2024-12-05 19:35:00	5 days 6 hours 0 minutes
SHIP_0012	Port of Hamburg	Port of Tianjin	Frozen Meat	2024-12-05 09:00:00	2024-12-10 15:00:00	5 days 6 hours 0 minutes
SHIP_0013	Port of Tianjin	Port of Antwerp	Refrigerated Goods	2025-01-10 16:45:00	2025-01-15 21:45:00	5 days 5 hours 0 minutes
SHIP_0014	Port of Antwerp	Port of Dubai	Grain	2025-02-14 20:30:00	2025-02-19 02:30:00	4 days 6 hours 0 minutes
SHIP_0015	Port of Los Angeles	Port of Busan	Processed Foods	2025-03-19 11:10:00	2025-03-24 17:10:00	5 days 6 hours 0 minutes
SHIP_0016	Port of Dubai	Port of Shanghai	Rice	2025-04-24 03:55:00	2025-04-29 09:55:00	5 days 6 hours 0 minutes
SHIP_0017	Port of Busan	Port of Rotterdam	Soybeans	2025-05-29 10:20:00	2025-05-04 14:20:00	-25 days 4 hours 0 minutes
SHIP_0018	Port of Singapore	Port of Los Angeles	Vegetables	2025-06-30 19:40:00	2025-06-09 23:40:00	-21 days 4 hours 0 minutes
SHIP_0019	Port of Qingdao	Port of Tianjin	Fruits	2025-07-15 08:25:00	2025-07-14 13:25:00	-1 days 5 hours 0 minutes
SHIP_0020	Port of Shanghai	Port of Hamburg	Corn	2025-08-20 14:50:00	2025-08-25 19:50:00	5 days 5 hours 0 minutes

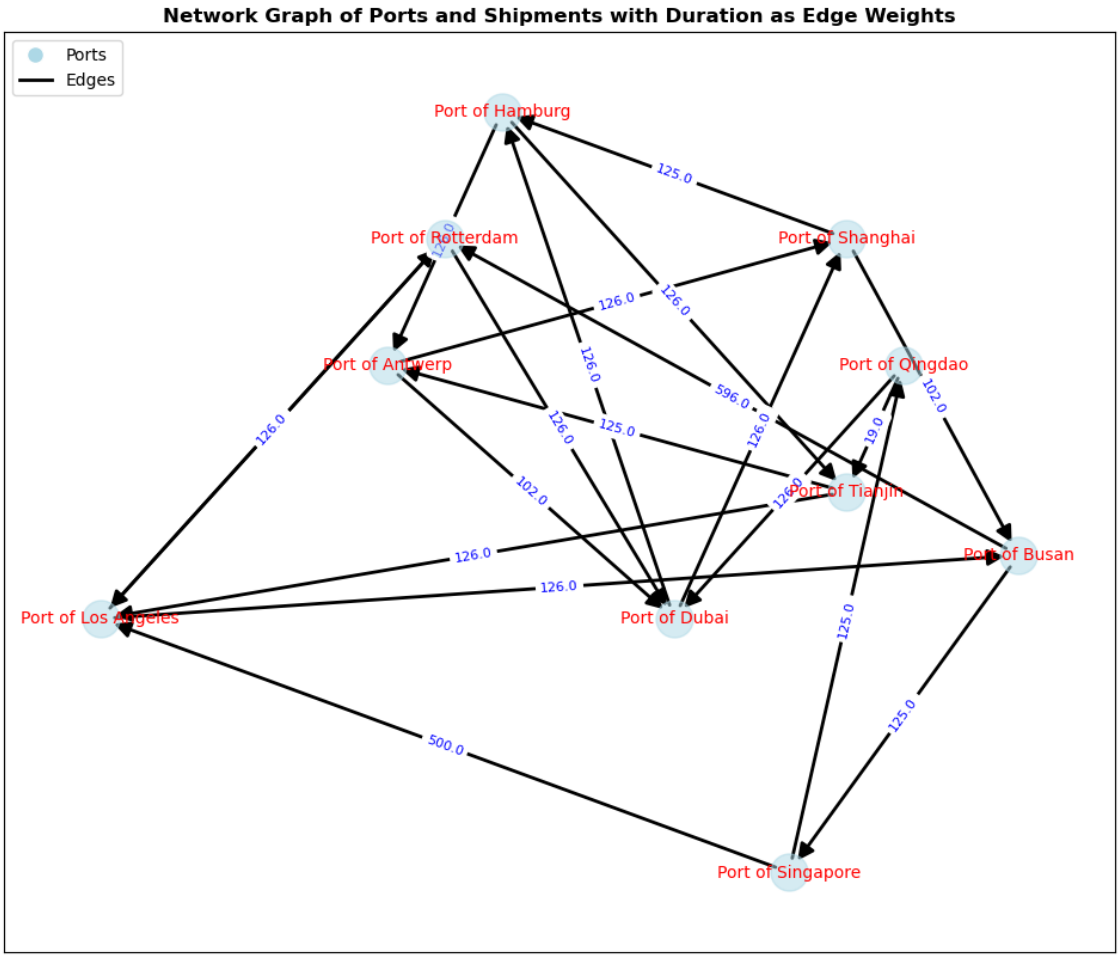
Insight:

- There are several entries where the duration is **negative**, such as:
  - SHIP_0017: -25 days 4 hours 0 minutes
  - SHIP_0018: -21 days 4 hours 0 minutes

    - SHIP_0019: -1 days 5 hours 0 minutes
- These negative durations suggest that the 'Arrival Time' is earlier than the 'Departure Time,' which is a clear data anomaly. This likely indicates an issue with how the data was recorded or how the times were swapped during analysis.
- **Action Needed:** I should investigate these entries further. For now, I assume I could **swap** the Departure Time and Arrival time.

5. Displaying the Revised Shipment Data in Table and Network

Shipment ID	Departure Port	Arrival Port	Cargo Type	Departure Time	Arrival Time	Duration
SHIP_0001	Port of Rotterdam	Port of Dubai	Bulk Liquids	2024-01-10 08:00:00	2024-01-15 14:00:00	5 days 6 hours 0 minutes
SHIP_0002	Port of Hamburg	Port of Antwerp	Frozen Meat	2024-02-15 12:30:00	2024-02-20 18:30:00	5 days 6 hours 0 minutes
SHIP_0003	Port of Tianjin	Port of Los Angeles	Refrigerated Goods	2024-03-20 16:45:00	2024-03-25 22:45:00	5 days 6 hours 0 minutes
SHIP_0004	Port of Antwerp	Port of Shanghai	Grain	2024-04-25 09:15:00	2024-04-30 15:15:00	5 days 6 hours 0 minutes
SHIP_0005	Port of Los Angeles	Port of Rotterdam	Processed Foods	2024-05-30 14:20:00	2024-06-04 20:20:00	5 days 6 hours 0 minutes
SHIP_0006	Port of Dubai	Port of Hamburg	Rice	2024-06-05 07:50:00	2024-06-10 13:50:00	5 days 6 hours 0 minutes
SHIP_0007	Port of Busan	Port of Singapore	Soybeans	2024-07-10 11:25:00	2024-07-15 16:25:00	5 days 5 hours 0 minutes
SHIP_0008	Port of Singapore	Port of Qingdao	Vegetables	2024-08-15 18:40:00	2024-08-20 23:40:00	5 days 5 hours 0 minutes
SHIP_0009	Port of Qingdao	Port of Dubai	Fruits	2024-09-20 05:10:00	2024-09-25 11:10:00	5 days 6 hours 0 minutes
SHIP_0010	Port of Shanghai	Port of Busan	Corn	2024-10-25 22:55:00	2024-10-30 04:55:00	4 days 6 hours 0 minutes
SHIP_0011	Port of Rotterdam	Port of Los Angeles	Bulk Liquids	2024-11-30 13:35:00	2024-12-05 19:35:00	5 days 6 hours 0 minutes
SHIP_0012	Port of Hamburg	Port of Tianjin	Frozen Meat	2024-12-05 09:00:00	2024-12-10 15:00:00	5 days 6 hours 0 minutes
SHIP_0013	Port of Tianjin	Port of Antwerp	Refrigerated Goods	2025-01-10 16:45:00	2025-01-15 21:45:00	5 days 5 hours 0 minutes
SHIP_0014	Port of Antwerp	Port of Dubai	Grain	2025-02-14 20:30:00	2025-02-19 02:30:00	4 days 6 hours 0 minutes
SHIP_0015	Port of Los Angeles	Port of Busan	Processed Foods	2025-03-19 11:10:00	2025-03-24 17:10:00	5 days 6 hours 0 minutes
SHIP_0016	Port of Dubai	Port of Shanghai	Rice	2025-04-24 03:55:00	2025-04-29 09:55:00	5 days 6 hours 0 minutes
SHIP_0017	Port of Busan	Port of Rotterdam	Soybeans	2025-05-04 14:20:00	2025-05-29 10:20:00	24 days 20 hours 0 minutes
SHIP_0018	Port of Singapore	Port of Los Angeles	Vegetables	2025-06-09 23:40:00	2025-06-30 19:40:00	20 days 20 hours 0 minutes
SHIP_0019	Port of Qingdao	Port of Tianjin	Fruits	2025-07-14 13:25:00	2025-07-15 08:25:00	0 days 19 hours 0 minutes
SHIP_0020	Port of Shanghai	Port of Hamburg	Corn	2025-08-20 14:50:00	2025-08-25 19:50:00	5 days 5 hours 0 minutes

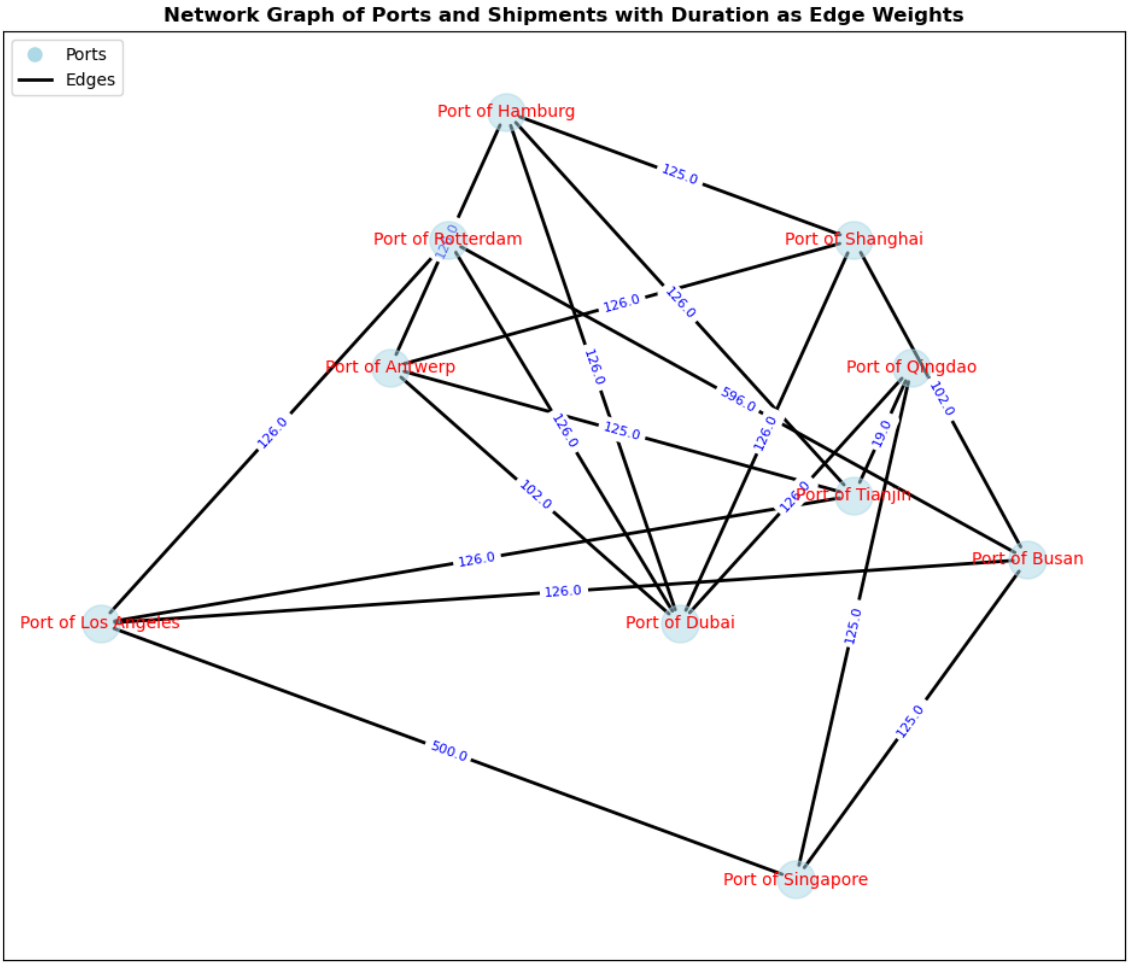


From this representation, we can observe how the network between ports is constructed, with the edge being the shipment duration between each port.

Insights:

- There are a few entries with unusually long durations:
  - SHIP_0017 has a duration of 24 days 20 hours.
  - SHIP_0018 has a duration of 20 days 20 hours.
  - These durations are far longer than the typical 5–6-day range seen in most other entries. This could indicate an issue with the shipment, a possible delay in transit, or because the distance between the ports is far.
- SHIP_0019 shows a duration of 0 days 19 hours, which is significantly shorter than other entries. This could mean it is a very short trip.
- There is a possibility to analyze are the shipment path in the raw data is the most optimum path or not, by assuming that the graph is not have direction.

6. Analyzing the Shortest Path



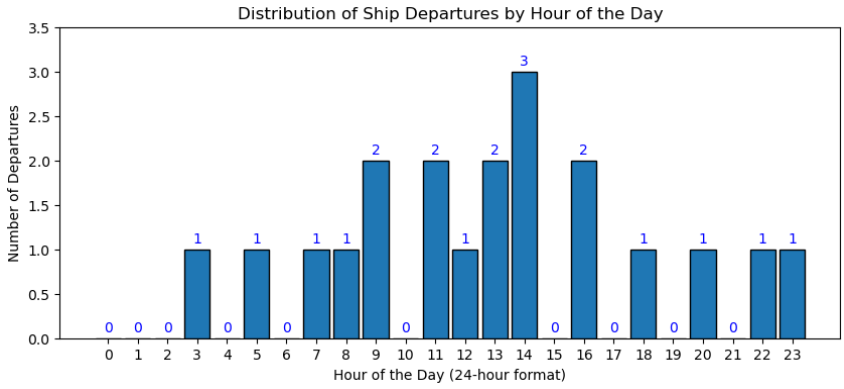
	Shipment ID	Cargo Type	Departure	Arrival	Direct Duration (hours)	Shortest Duration (hours)	Shortest Path	Transit Needed
0	SHIP_0001	Bulk Liquids	Port of Rotterdam	Port of Dubai	126.0	126.0	Port of Rotterdam -> Port of Dubai	False
1	SHIP_0002	Frozen Meat	Port of Hamburg	Port of Antwerp	126.0	126.0	Port of Hamburg -> Port of Antwerp	False
2	SHIP_0003	Refrigerated Goods	Port of Tianjin	Port of Los Angeles	126.0	126.0	Port of Tianjin -> Port of Los Angeles	False
3	SHIP_0004	Grain	Port of Antwerp	Port of Shanghai	126.0	126.0	Port of Antwerp -> Port of Shanghai	False
4	SHIP_0005	Processed Foods	Port of Los Angeles	Port of Rotterdam	126.0	126.0	Port of Los Angeles -> Port of Rotterdam	False
5	SHIP_0006	Rice	Port of Dubai	Port of Hamburg	126.0	126.0	Port of Dubai -> Port of Hamburg	False
6	SHIP_0007	Soybeans	Port of Busan	Port of Singapore	125.0	125.0	Port of Busan -> Port of Singapore	False
7	SHIP_0008	Vegetables	Port of Singapore	Port of Qingdao	125.0	125.0	Port of Singapore -> Port of Qingdao	False

8	SHIP_0009	Fruits	Port of Qingdao	Port of Dubai	126.0	126.0	Port of Qingdao -> Port of Dubai	False
9	SHIP_0010	Corn	Port of Shanghai	Port of Busan	102.0	102.0	Port of Shanghai -> Port of Busan	False
10	SHIP_0011	Bulk Liquids	Port of Rotterdam	Port of Los Angeles	126.0	126.0	Port of Rotterdam -> Port of Los Angeles	False
11	SHIP_0012	Frozen Meat	Port of Hamburg	Port of Tianjin	126.0	126.0	Port of Hamburg -> Port of Tianjin	False
12	SHIP_0013	Refrigerated Goods	Port of Tianjin	Port of Antwerp	125.0	125.0	Port of Tianjin -> Port of Antwerp	False
13	SHIP_0014	Grain	Port of Antwerp	Port of Dubai	102.0	102.0	Port of Antwerp -> Port of Dubai	False
14	SHIP_0015	Processed Foods	Port of Los Angeles	Port of Busan	126.0	126.0	Port of Los Angeles -> Port of Busan	False
15	SHIP_0016	Rice	Port of Dubai	Port of Shanghai	126.0	126.0	Port of Dubai -> Port of Shanghai	False
16	SHIP_0017	Soybeans	Port of Busan	Port of Rotterdam	596.0	252.0	Port of Busan -> Port of Los Angeles -> Port of Rotterdam	True
17	SHIP_0018	Vegetables	Port of Singapore	Port of Los Angeles	500.0	251.0	Port of Singapore -> Port of Busan -> Port of Los Angeles	True
18	SHIP_0019	Fruits	Port of Qingdao	Port of Tianjin	19.0	19.0	Port of Qingdao -> Port of Tianjin	False
19	SHIP_0020	Corn	Port of Shanghai	Port of Hamburg	125.0	125.0	Port of Shanghai -> Port of Hamburg	False

Insights:

- For most shipments, the direct duration matches the shortest duration, implying that no transit or intermediary steps are needed between the departure and arrival ports.
- Exceptions are seen in the shipments for SHIP_0017 (Soybeans) and SHIP_0018 (Vegetables) where the shortest duration is significantly less than the direct duration, indicating that transits reduce overall shipping time.

7. Distribution of Ship Departure by Hours



Insights:

- The highest number of ship departures occurs around 14:00 (2:00 PM), with 3 departures. This could indicate that this is a preferred time for starting shipments, perhaps due to optimal port operations or other logistical reasons.