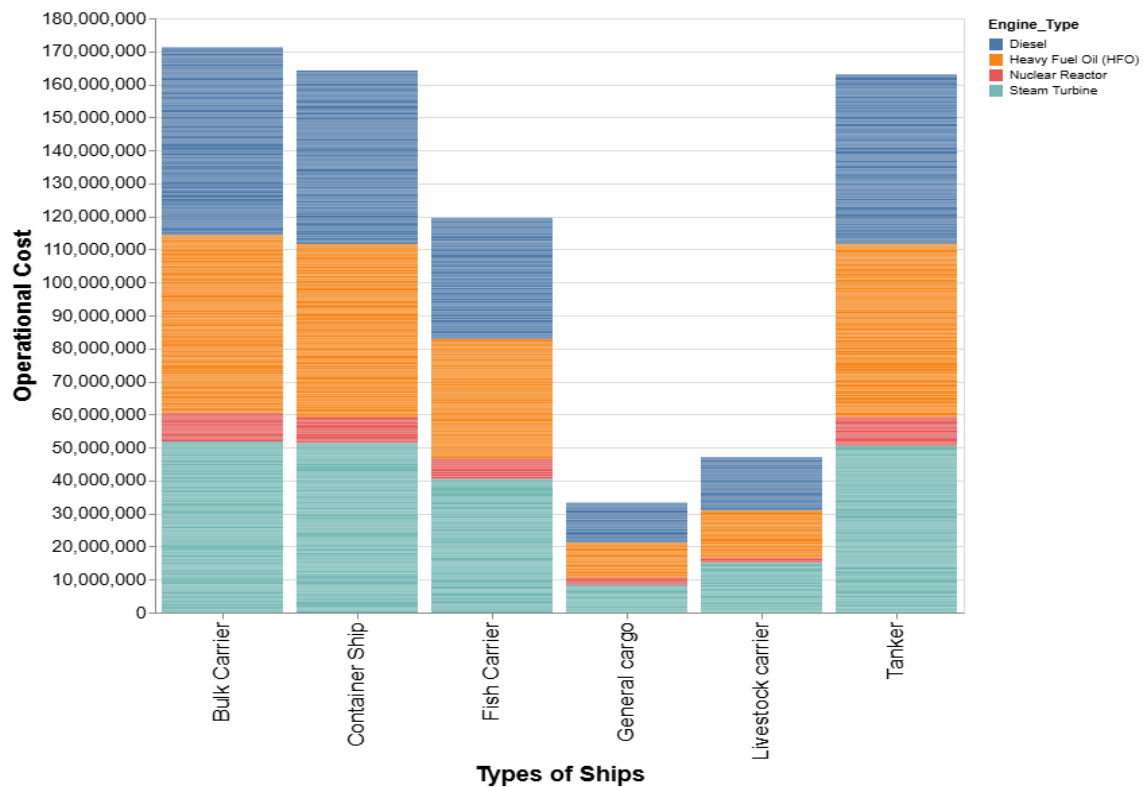Goal: -

My dataset is primarily focused on the performance of the ships and the factors affecting it. These ships are carrier ships which are used for transporting goods. In this following data visualization assignment, I have tried to visualize the operational cost of different type of ships with respective to their engine type.

Image of visualization: -

Insights: -

- Chart Type: Stacked Bar Chart.
- The bar in the above chart represents the operational cost (in USD) of individual carrier ships.
- Multi colored line in the bar represents the cost recorded with different engine types.
- X-axis represents the cost and Y-axis represents the different types of ships present in data sets.
- Legend: Displays engine types with distinct colours for easy identification.

Data Abstraction: -

This dataset focuses on analyzing the performance of carrier ships and the factors affecting their efficiency. It includes various operational and environmental factors affecting the ship performance.

- ❖ Key Attributes: -
    - Ship_Types: - Contains data on different ships mainly carrier ships used for transporting goods.
    - Engine_Types: - Stores information on different engine types.
    - Operational_Cost_USD: - It represent the operational cost to operate each ship which is measured in USD.

Task Abstraction: -

By analyzing the dataset, I have used Ship_types and Operational_Cost_USD to map out the average cost of individual ships using bar chart, and then used Engine_Types to find out detailed cost of same ships having different engines. This allows me to understand which engine types are more cost-effective, providing me with valuable insights into whether specific engine type helps to reduces operational costs for certain ship types.

- ❖ Channels: -
    - I. Hue.
    - II. Height.
    - III. Vertical Position.
- ❖ Mark: -
    - I. Line.
- ❖ Data type: -
    - I. Categorical Type.
    - II. Numeric Type.
- ❖ Numerical data
    - I. Operational_Cost_USD
- ❖ Categorical data
    - I. Ship_Types
    - II. Engine_Types

Data Source: -

https://www.kaggle.com/datasets/jeleeladekunlefijabi/ship-performance-clustering-dataset

Code: -

```python
import pandas as pd
import altair as alt
import warnings

# Suppress Future Warnings
warnings.simplefilter(action='ignore', category=FutureWarning)

# Load the dataset
file_path = "Ship_Performance_Dataset.csv"
ship = pd.read_csv(file_path)
#drops the cell with null values
ship.dropna(axis=1)

#Shows first five rows
ship.head()

#It is used to convert ship into a Pandas DataFrame.
df= pd.DataFrame(ship)

#To plot the bar diagram
chart = alt.Chart(ship).mark_bar().encode(
    x=alt.X('Ship_Type', title="Types of Ships",
        axis=alt.Axis(labelFontSize=16, titleFontSize=18)),
    y=alt.Y('Operational_Cost_USD', title="Operational Cost",
        axis=alt.Axis(labelFontSize=14, titleFontSize=18)),
    color=('Engine_Type'),
).properties(
    width=600,
    height=500,
)
chart.display()
```