PRICE RECOMMENDATION REPORT

COMPASS MARITIME SERVICES

GROUP 2:

GURPREET S. CHANDOK

NAITIK NAYAK

RAKSHIT MADAN

ASHIKUL KABIR

RISHIKA LALL

SANA PARAB

EXECUTIVE SUMMARY

The report delivers comprehensive statistical analysis to **forecast price** of Bet Performer, which is a capesize ship. To estimate the price we use variables like age at sale, year built, DWT, capesize Index.

In addition to these variables we also imply **Gross Weight Tonnage (GWT)** and **Fuel Price** to determine its effect while forecasting price for Bet Performer.

Our goal is to foster **collaboration** between the buyer and seller, navigating the complexities of this dynamic environment. We aim to facilitate a **mutually advantageous deal** following the identification, assessment, and acquisition of a capsize carrier vessel.

It also talks about **macroeconomic factors** fuel rate, geopolitics, technological factors, commodity prices, regulatory environment, credit availability and environment concern.

At initial stage to find out transactional ship, we used **market approach,** here we compare different market transactions taken place involving genuine knowledgeable buyer and sellers by considering all the hidden factors such as price of recent sales in the active market and ship inspection which ensure quality of structure, engine, pumping system, and other system like navigation etc.

We aim to facilitate a **mutually advantageous deal** following the identification, assessment, and acquisition of a capsize carrier vessel.

On the basis of that we discovered that **Sumihou ship** is best refer transaction ship.
We used multiple regression model to determine the price of Bet Performer.

Our goal is to foster **collaboration between the buyer and seller,** navigating the complexities of this dynamic environment. We aim to facilitate a mutually advantageous deal following the identification, assessment, and acquisition of a capsize carrier vessel.

METHODOLOGY

The aim of this project is to find a **reference transactional ship** for the Bet Performer and quote an equivalent price.

The dataset used for this analysis is sourced from a Harvard Business School case, and the research relies heavily on secondary data sources.

We use **statistical indicators** such as the mean and standard deviation in the initial stages to determine a representative transactional ship.

Further, to **estimate the price** of the ship and to derive single feature which is the best predictor of ship price we use regression.

Apart from the given data we all consider **additional variables** like fuel rate and gross weight ton (GWT) to detect how much these variables participate in deciding the price of Bet Performer. Fuel rate was derived from Sale Date ,which was month and year in this case. Gross Weight Ton was derived from sources available on the internet.

A **synthetic data** was generated which imitates our original data. Additional data were generated using Gretel.ai . This dataset was used as test data and multiple regression was tested to see performance of model on new data points,

VARIABLES

The data includes the following **variables**: "Sale Date, Vessel Name, Sale Price, Year Built, Age at Sale, Dead Weight Tons and the Average Monthly Baltic Dry Capesize Index."

The **Dead Weight Ton (DWT)** is a ship's weight measurement that takes into account the whole weight of cargo, fuel, passengers, and other things that the ship may be carrying on the waters. The Baltic Dry Capesize Index (BDI) is an index of average prices paid for dry bulk material transportation.

Besides that, we consider fuel rate and **Gross Weight Tonnage(GWT).** We calculate the fuel cost based on the rate in effect on the exact selling date in this case month and year, capturing the precise pricing associated with that transaction in the face of potential volatility caused by market conditions, supply and demand dynamics, and other relevant variables.

Gross Tonnage is a measurement of a ship's total interior capacity, including all enclosed compartments, and it is used to determine the ship's size.

External factors include **macroeconomic factors** such as the interest rates, ocean freight rates, oil prices, currency volatility, the dockyard where the ship was built or the original ship builders, the ship's location at the time of sale, and environmental factors such as air pollution from greenhouse gases and the potential threat of oil and chemical spills.

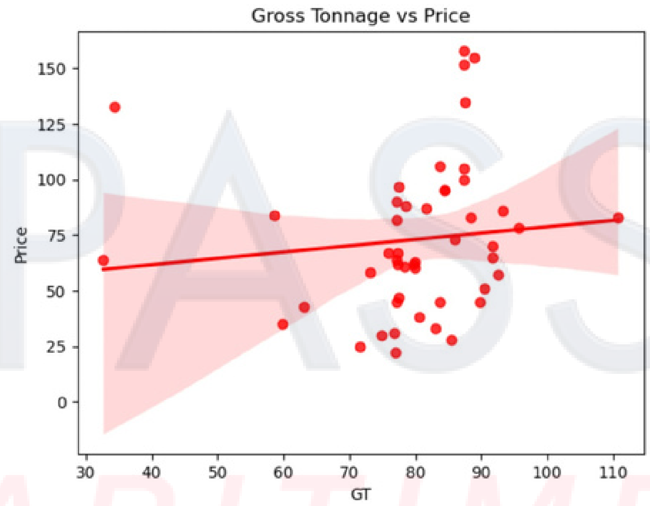
HYPOTHESIS

Apart from the given **variables** like sale date, sale price, year built, age at sale, dead-weight tons and capesize Index. We added fuel rate and Gross Weight Ton to show what correlation it has it towards price.

1) Gross Weight Ton (GWT)

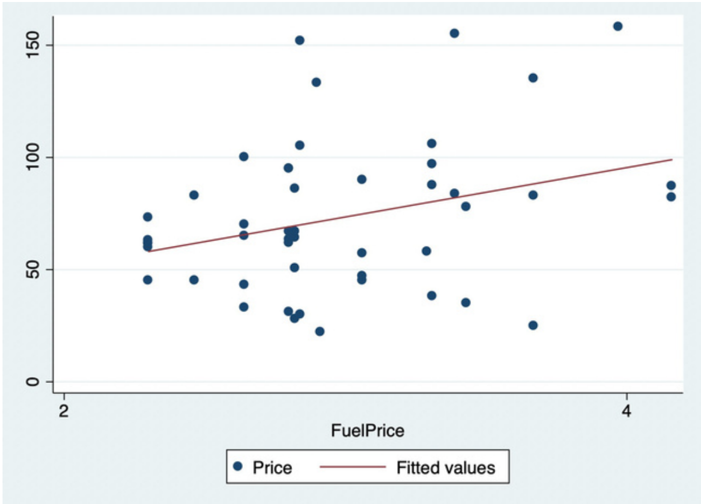
The GWT represents the ship's total internal volume, including all enclosed compartments. Larger ships with higher GWT require more materials for construction, which might raise the overall cost of the vessel but at the same time determine the capacity of carrying commodities.

When we find the relation of Sale Price with respect to GWT/GT is 11% which states that these two variables are weakly correlated and hence it can't be considered as a significant feature while evaluating price of Bet Performer



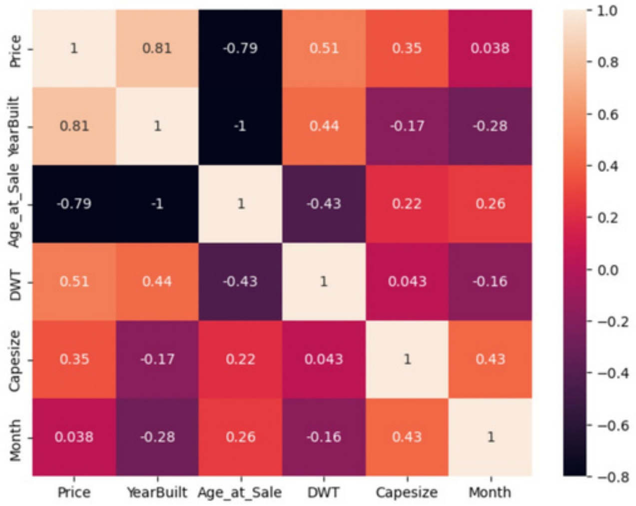
2) Fuel Rate

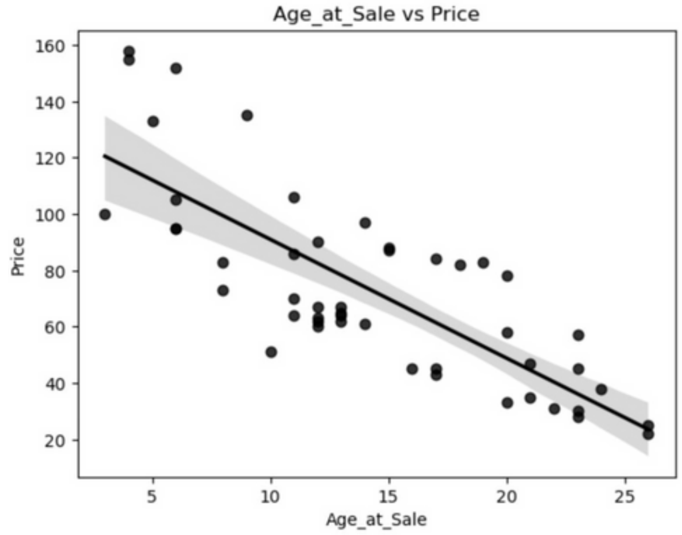
The price of fuel plays a vital role in determining the price of ship. Rising fuel prices may have an impact on market demand for more fuel-efficient vessels. In the given case the **correlation** between sale price of ship and fuel rate has a weak correlation of 0.31 and can't be considered while calculating the price for Bet Performer.

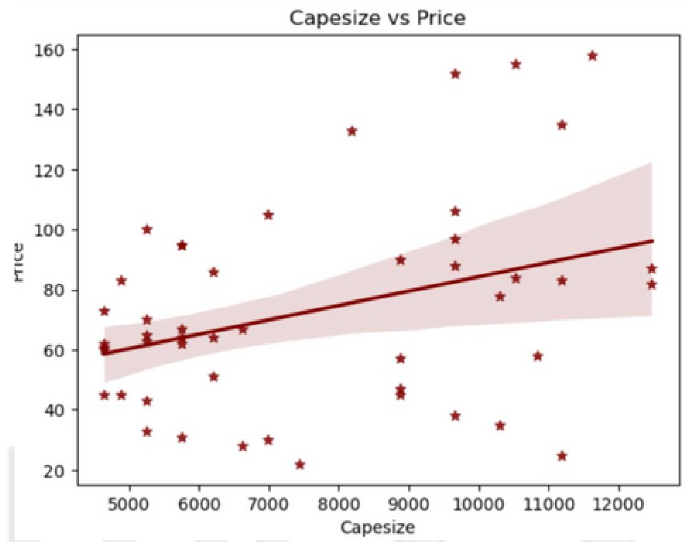


CORRELATION

In the given data, year built, age at sale , deadweight ton, capesize Index and month(sale date) are independent variables while sale price is dependent variable. Given **heat map** represents correlation of these variables with one other. We observed that **price and age at sale** have strong negative, indicating possibilities towards outdated technology, wear & tear, breakage, less efficient, etc. Hence, it is considered as a vital variable while determining price of ship.





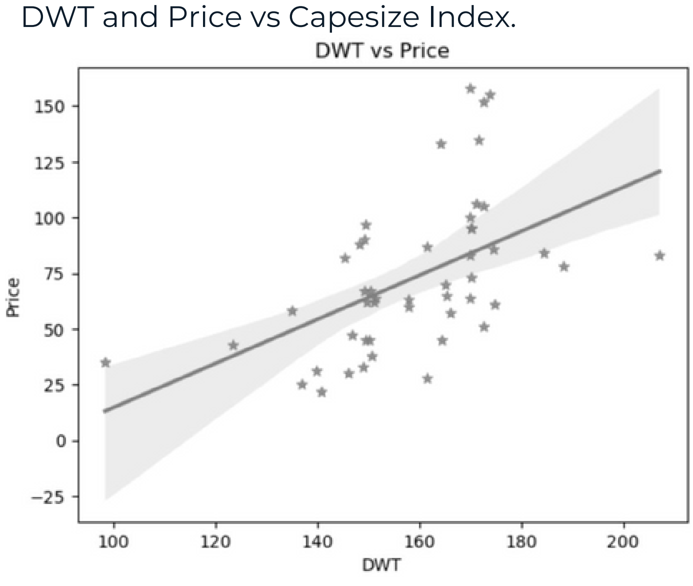


ASSUMPTIONS

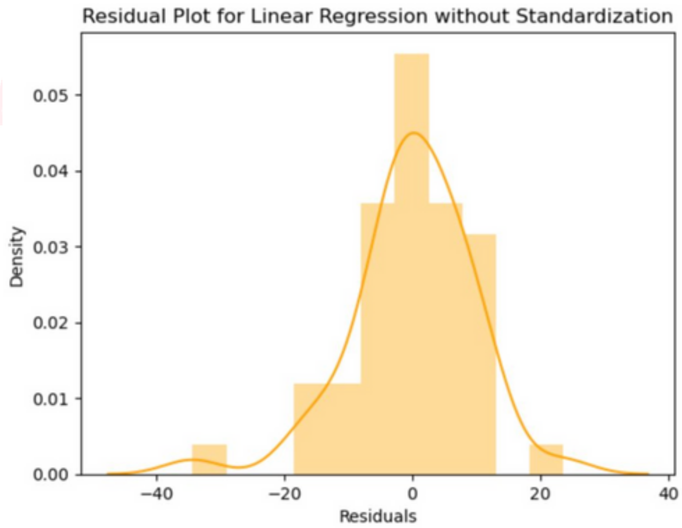
Multiple Regression Assumptions:

First, Linearity This states that the relationship between the dependent variable and the independent variables should be linear.

The following plots show that the **relationship between the variables is linear** and hence it satisfies the criteria. The below plots are for Price vs Age at Sale, Price vs DWT and Price vs Capesize Index.



Second, Residuals are normally distributed – From the below plot we can see that residuals are nearly normally distributed.

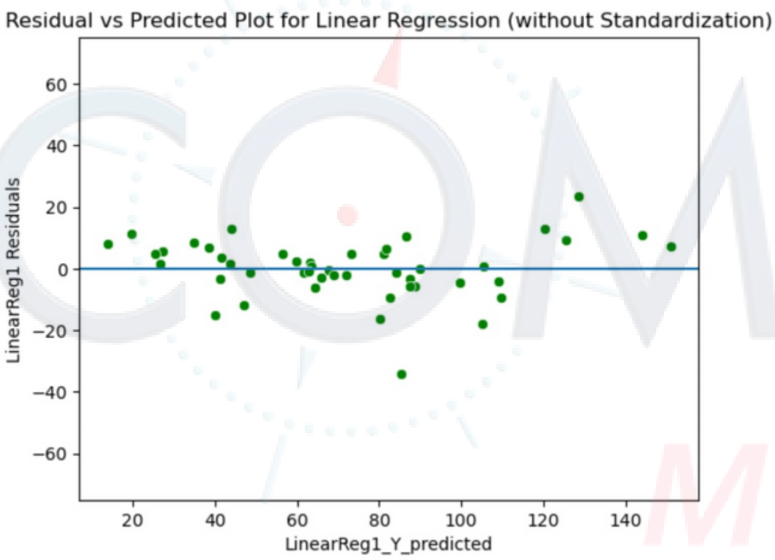


Third, Multicollinearity is used as Variance Inflation Factor (VIF) to conduct test for no multicollinearity.

There appears to be no multicollinearity if the VIF for each variable is between 1 to 5. Below table shows the above-mentioned criteria is satisfied.

	VIF	variable
0	131.944976	Intercept
1	1.258727	DWT
2	1.318729	Age_at_Sale
3	1.075427	Capesize

Fourth, Homoscedasticity is used to describe a situation in which the error term (that is, the "noise" or random disturbance in the relationship between the independent variables and the dependent variable) is the same across all values of the independent variables. The below plot shows that the spread of error term is constant.



LIMITATIONS

The **size of the dataset** is relatively small having only 48 data points which does not allow the model to be generalized.

Also the robustness of the interpretation might suffer from the limited availability of additional relevant variables, which might have helped test the model in different circumstances.

The model reports **lower accuracy** of results when run by creating synthetic dataset to train and test data validity.

It appears that there exists overfitting issues in the model. The external geopolitical or industry-specific factors are not considered since it is hard to capture comprehensive information in the relevant field of experimentation.

MULTIPLE REGRESSION USING SYNTHETIC DATA

We generated a **synthetic data** which mimics our original data. We created 200 additional data points using Gretel.ai which uses ACTGAN (Anyway Conditional Tabular Generative Adversarial Network).

We used our original dataset as train data and synthetic data as test data and ran multiple regression to see how our models performs on new data points. The R-squared dropped to 67% from 92%. This resulted in overfitting of the model.

CONCLUSION

The model reported a **high variance** level that explains its effectiveness in explaining the price of the ship considering the age at sale, DWT, and capesize index as independent variables.

The predicted price of **$125.7M** is reliable given the level of confidence in the model output. The magnitude of the individual impact of the variables is clearly explained by the coefficient values associated with each variable. The predicted ship price based on our model is approximately **$125.7M**, with a **95% confidence interval** ranging from **$116M to $135.4M.**

Finally, the study generates **valuable insights** for the ship industry stakeholders such as owners, investors, or policy makers. Given the **limitations,** the study can be useful to the researchers in extending the model validation process or exploring new variables that might help increase the explainability of the model.

SCENARIO ANALYSIS

Scenario 1:

- **Condition:** When the ship is 5 years younger.
- **Observation:** The price increases to $148.
- **Interpretation:** This suggests an inverse relationship between age and price, indicating that younger ships tend to have higher prices.

Scenario 2:

- **Condition:** When the ship is lighter by 20,000 DWT.
- **Observation:** The price decreases by 4.8 M to $120.69M.
- **Interpretation:** This implies a direct relationship between deadweight tonnage and price. A decrease in the ship's carrying capacity leads to a lower price.

Scenario 3:

- **Condition:** Lowering the capesize by 30%.
- **Observation:** The price decreases to $98.53 M
- **Interpretation:** This suggests a direct relationship between capesize and price. A reduction in the capesize results in a lower price.

REFERENCES

Fuel Price Distribution:
https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=EMD_EPD2DXL0_PTE_NUS_DPG&f=W

Info on Bet Performer & other ships-
https://www.shipspotting.com/photos/179754

Gross Tonnage distibution:
https://www.balticshipping.com/vessel/imo/9483243

Synthetic data generation:
https://gretel.ai

https://www.breakwaveadvisors.com/insights/2022/3/4/1aozoe8qjjo8zf5mcznujk9a5sunpo#:~:text=As%20the%20price%20of%20oil,(i.e.%20higher%20derived%20demand).

https://books.google.com/books?hl=en&lr=&id=HiVbBAAAQBAJ&oi=fnd&pg=PP1&ots=gTsQWnvksY&sig=UNoABwqgJR5uNxhqG9ys5t1b1HE#v=onepage&q&f=false.

https://link.springer.com/article/10.1007/s13437-022-00272-4#Sec2

For general references:
https://chat.openai.com/