

# **Case Study: Yacht and Boat Sales Website Analysis Report**

**Author:** Kundai Shoniwa

**Date:** 3 September 2023

*This project was completed in partial fulfillment of the Google Data Analysis  
Professional Certificate.*

Introduction	3
Objectives	3
Limitations	3
Phase 1: Stakeholder Identification	4
Key Stakeholders	4
Phase 3: Data Cleaning and Manipulation	5
Identifying Data Issues	5
Phase 4: Exploratory Data Analysis	6
Customer Behavior Insights	6
Market Trends and Preferences	6
Website Performance	6
Phase 5: Key Findings	7
Customer Behaviour and Insights	7
Boat Type and Popularity	7
Average Price of Yachts and Boats	7
Geographical Trends and Boat Interest	8
Boat Listings and Views	8
Market Trends and Preferences	9
Manufacturer Trends	9
Material Preferences	10
Boat Size and Views	12
Boat Year of Build and Popularity	12
Boat Type and Condition	14
Website Performance	15
Phase 6: Recommendations	16
Conclusion	16

## Introduction

The primary objective of this project was to gain insights into the performance and trends of a Yacht and Boat sales website. Specifically, the project aimed to analyze data related to boat prices, manufacturer trends, visitor locations, view counts, boat types, boat conditions, dimensions, build years, and materials used.

## Objectives

The client represents a yacht and boat sales website. They have tasked me with analyzing their data to achieve the following objectives:

- Understand customer preferences and behavior on their website.
- Identify key factors influencing boat and yacht sales.
- Provide insights and recommendations to optimize the website's user experience and sales strategies.

## Limitations

- Data Availability: The analysis relied on the availability and quality of data from the website's database. Incomplete or inaccurate data impacted the analysis.
- External Factors: The analysis did not account for external factors such as economic conditions, industry regulations, or global events that could influence boat sales.
- Data Privacy: Care was taken to ensure user privacy and compliance with data protection regulations when handling user demographic data.

# Phase 1: Stakeholder Identification

## Key Stakeholders

The primary stakeholders for this project were the customers and the personnel within the business as follows:

1. **Customers:** These are potential boat buyers and yacht enthusiasts. Previous customers who may provide testimonials or referrals. Customer service representatives who interact with customers.
2. **Boat Manufacturers and Dealers:** Suppliers of the boats and yachts listed on the website. Dealerships or manufacturers who use the platform for sales.
3. **Website Development Team:** Developers, designers, and IT professionals who are responsible for maintaining and updating the website.
4. **Marketing and Sales Team:** Sales and marketing professionals who are responsible for promoting listings and driving sales. Digital marketing experts and SEO specialists.
5. **Management and Executives:** Senior management and executives who oversee the strategic direction of the business. Decision-makers who allocate resources to the website.

## Phase 2: Preparing the data

The data used in this project was from the [Boat Sales](#) dataset (CC0: Public Domain, dataset made available through Kaggle). This Kaggle dataset contains 9889 rows and 10 columns of data, and is saved in CSV format. It has details of the boats regarding the year it was built, the type of boat, the price, the number of views in the last 7 days, and much more. It was noted that there were no citations mentioned and the dataset was last updated 2 years ago. Whilst the dataset came with its business task, for the purpose of my course, I decided to define new objectives and ask further questions.

## Phase 3: Data Cleaning and Manipulation

### Identifying Data Issues

The first step in the data cleaning process was to identify issues with the dataset. The following issues were identified:

1. Missing Values
2. Duplicate Records and Whitespace
3. Incorrect Data Types
4. Inconsistent formatting
5. Deleting Irrelevant Data

### Data Cleaning and Validation

1. **Handling Missing Values:** Because of the nature of the data, I chose to replace missing values with a default value (i.e., NULL for string/varchar datatypes, and 0 for integers).
2. **Removing Duplicates and Trimming Whitespaces:** To effectively remove duplicates and trim whitespaces, I used Google Sheets' inbuilt data cleanup function. No duplicate entries were found and 52 whitespaces were trimmed from the data.
3. **Correcting Data Types:** The price column was set to 'text' datatype and this was changed to currency.
4. **Standardizing Data:** The location column consisted of a string comprised of the country, city/town, area, and in some cases the area code. To ensure consistent formatting, such as uppercase or lowercase, and variations, I split the contents of this column by country and town/city. I also replaced cities/towns with their respective countries instead by using the filter and find and replace functions. The price column had 4 different currencies. I went on to standardize the different currencies to USD for consistency in comparison. I assumed that the exchange rate was based on the foreign exchange rate of the day of 2 September 2023.
5. **Deleting Irrelevant Data:** I proceeded to delete irrelevant data by deleting the location, city/town, and postcode/Area code columns created in the previous step.

I then proceed to validate the data, for quality and consistency, by checking for errors that may have been introduced during preparation and cleaning.

## **Phase 4: Exploratory Data Analysis**

To understand the distribution of key variables and identify patterns, I used Google Sheets' pivot tables as the primary tool for analysis. The choice to use spreadsheets was based on the size of the dataset, and the ease of preparation and processing for analysis. For the analysis, the following questions were asked:

### **Customer Behavior Insights**

1. Which boat types receive the most views? Are there certain boat types that attract more interest than others?
2. What is the average price range for boats on the website? Are there price trends based on boat type or material used?
3. Which countries are the website's visitors primarily from? Are there geographic trends in boat interest?
4. Is there a correlation between the number of views a boat listing receives and the country the boat is listed from?

### **Market Trends and Preferences**

1. Which manufacturers' boats are most frequently viewed on the website? Are there specific manufacturers that are more popular among buyers?
2. Are there material preferences among buyers, such as fiberglass, wood, or aluminum? Does material affect pricing or popularity?
3. Do larger boats receive more views or have higher price tags? Are there size preferences based on boat type?
4. Is there a correlation between the year a boat was built and its price or popularity? Are older or newer boats more sought after?

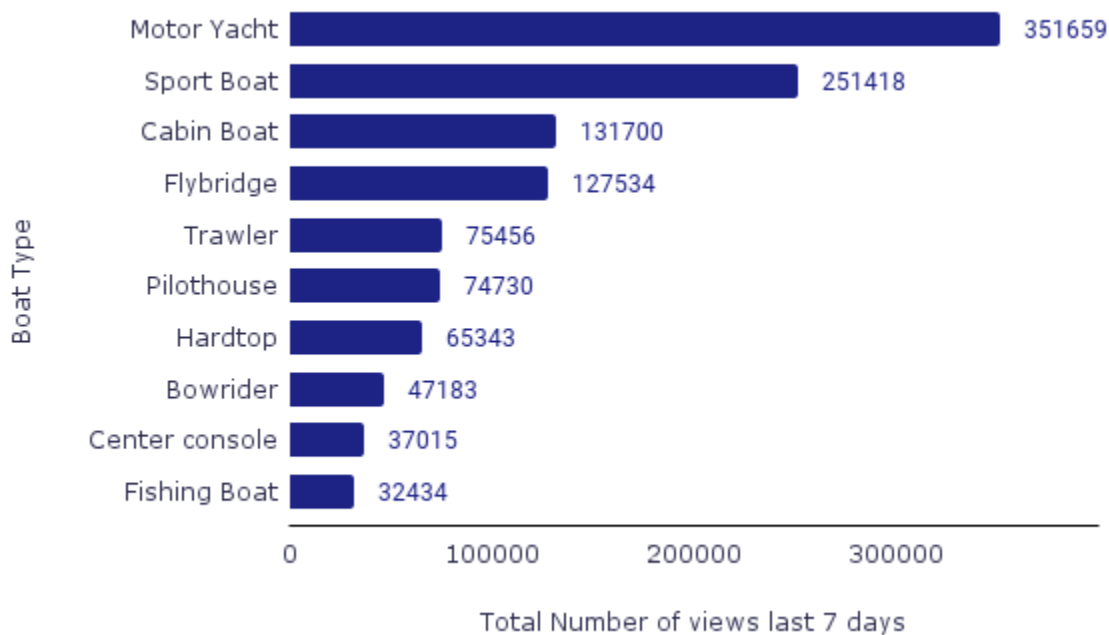
### **Website Performance**

1. Do visitors from specific countries tend to engage more with the website?
2. Are there factors (e.g., boat type, price, location) that significantly influence the number of views a listing receives?

# Customer Behavior and Insights

## Boat Type and Popularity

Motor yachts and sports boats receive the highest number of views on the website, indicating their popularity among visitors as shown in Figure 1.

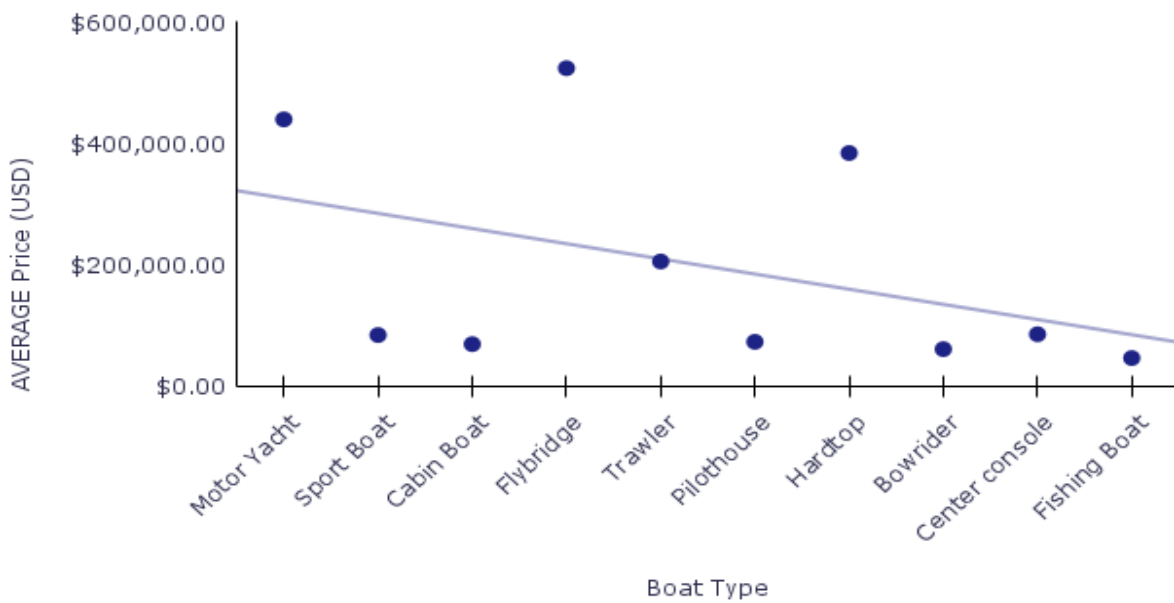


**Figure 1:** Showing the relationship between boat type and number of views

## Average Price of Yachts and Boats

The average price range for boats on the website varies widely, spanning from approximately \$5,000 to \$6,000,000. Prices depend on factors such as boat type, condition, size, age and the material used in construction.

Notably, there are distinct price trends based on boat type and the material used. Larger boats tend to be more expensive, as do those made from materials like carbon fiber, aluminum and GRP (Glass-Reinforced Plastic). The popularity of boat types tends to inversely affect their price trends, often resulting in a downward trajectory.



**Figure 2:** Relationship between average price and boat popularity

### Geographical Trends and Boat Interest

The dataset lacks information regarding the source of website traffic. Consequently, it is challenging to determine the specific origins of the website's visitors. To determine if there are geographic trends in boat interest, additional data about the origin of the website's visitors would be necessary.

### Boat Listings and Views

The dataset reveals that listings in Europe, particularly in Switzerland, Germany, and Italy, garner the highest number of views. While this observation provides an initial insight, the absence of additional data related to user behavior and location presents challenges in making definitive correlations between viewing statistics and the countries of the website's visitors. To establish a robust relationship between these variables and their significance for marketing and targeting strategies, further statistical analysis and comprehensive data collection are imperative.



Country	Total Number of views last 7 days
Switzerland	385006
Germany	306033
Italy	214310
Netherlands	122568
France	118645
Spain	81986
Croatia	80237
Denmark	23334
United Kingdom	19601
Austria	16990

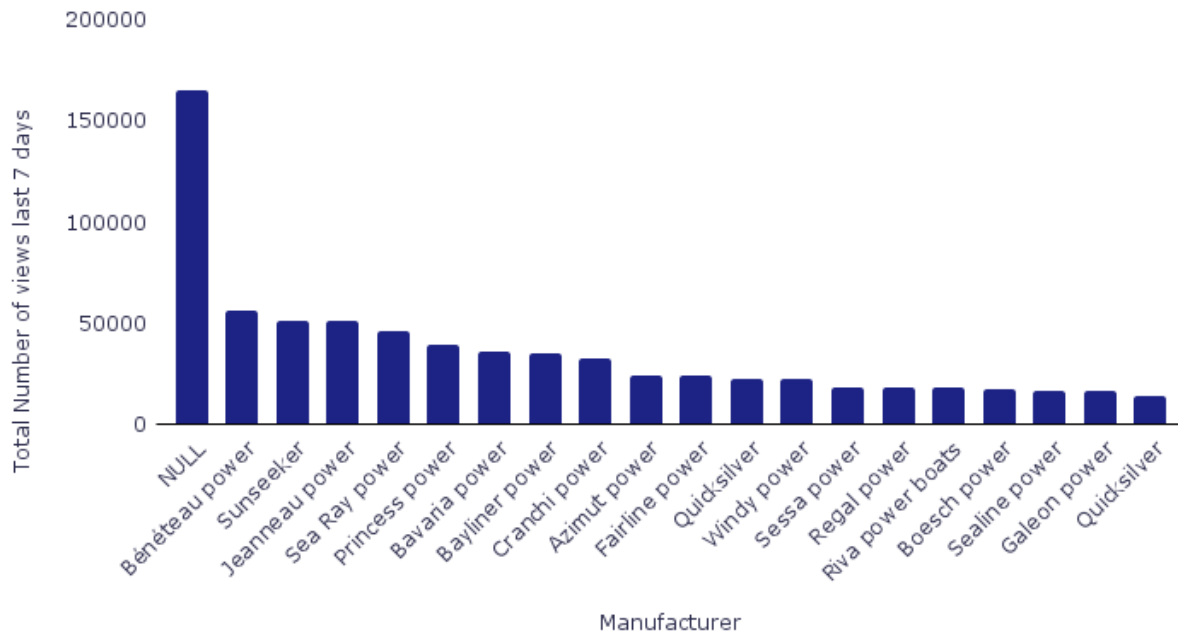
**Table 1:** Relationship between country of boat listing and views received

## Market Trends and Preferences

### Manufacturer Trends

The analysis of the dataset reveals that Bénéteau power boats, Sunseeker power boats, and Jeanneau power boats emerge as the most popular manufacturers in terms of the number of views. These manufacturers appear to attract a higher level of interest from website visitors compared to others.

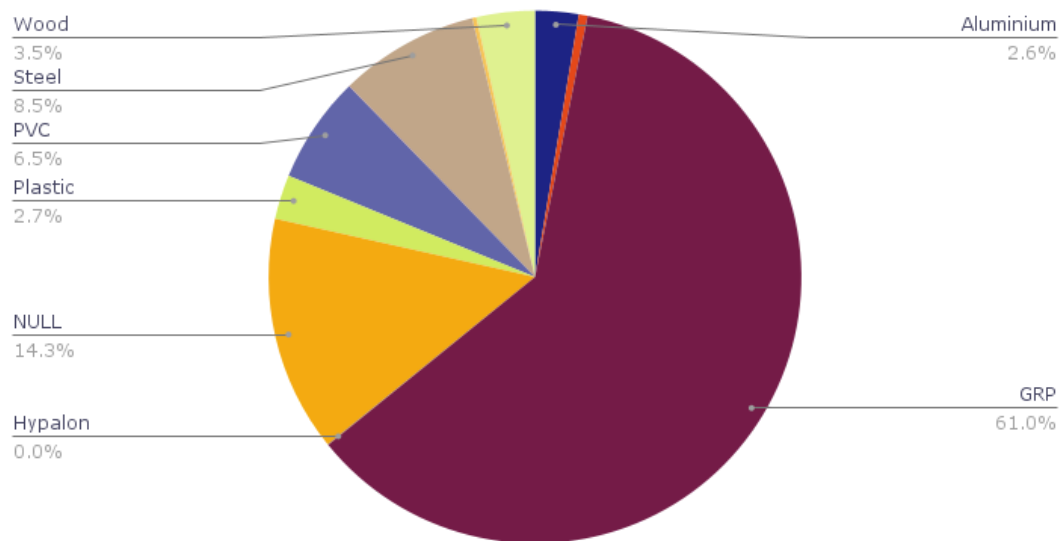
Additionally, it's worth noting a particularly intriguing observation: where the boat manufacturer is omitted or not mentioned, there appears to be the highest number of views. This raises questions about whether boats were intentionally listed without a manufacturer name, or if this omission was due to errors during data collation. Further investigation into the reasons behind this phenomenon could provide valuable insights into user behavior and listing practices on the website.



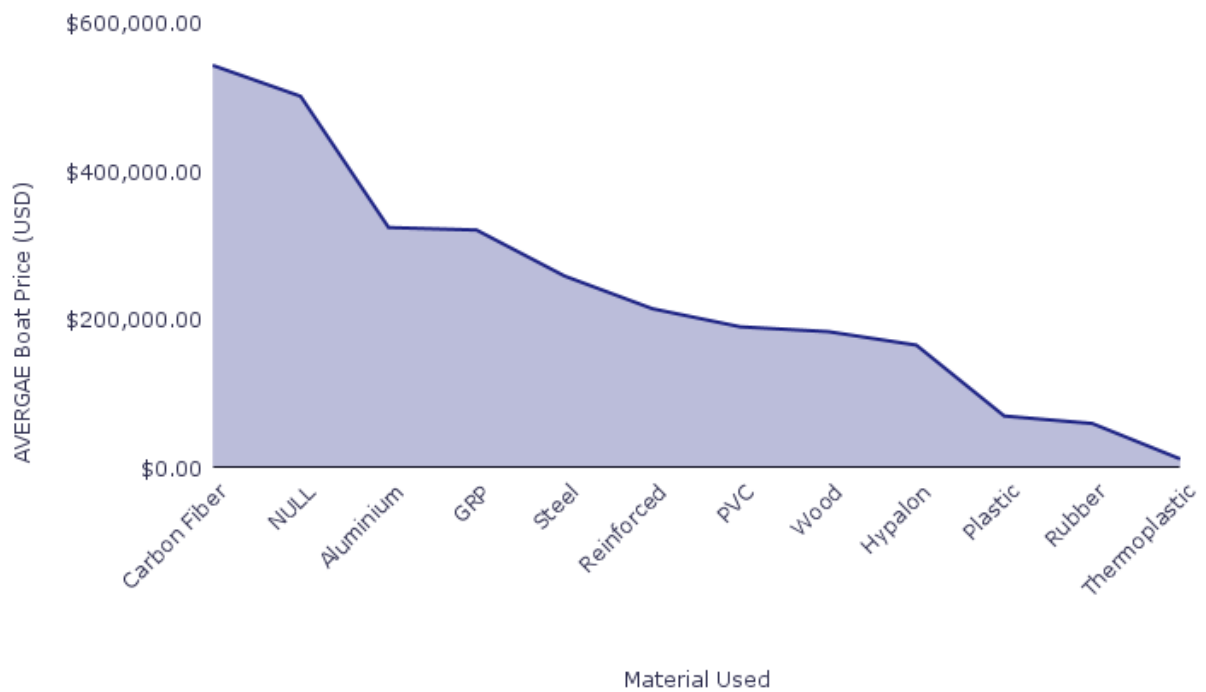
**Figure 3:** Relationship between the total number of views and manufacturer

### Material Preferences

Visitors to the website exhibit a strong preference for boats made from Glass Reinforced Plastic (GRP), commonly known as fiberglass. This material appears to be more popular among buyers, potentially due to its durability, versatility, and lower maintenance requirements. While other materials like wood or aluminum are available, they do not seem to attract the same level of popularity as GRP. Remarkably, despite the fact that boats made from GRP tend to be more expensive, this price difference does not appear to deter visitors' interest in these models.



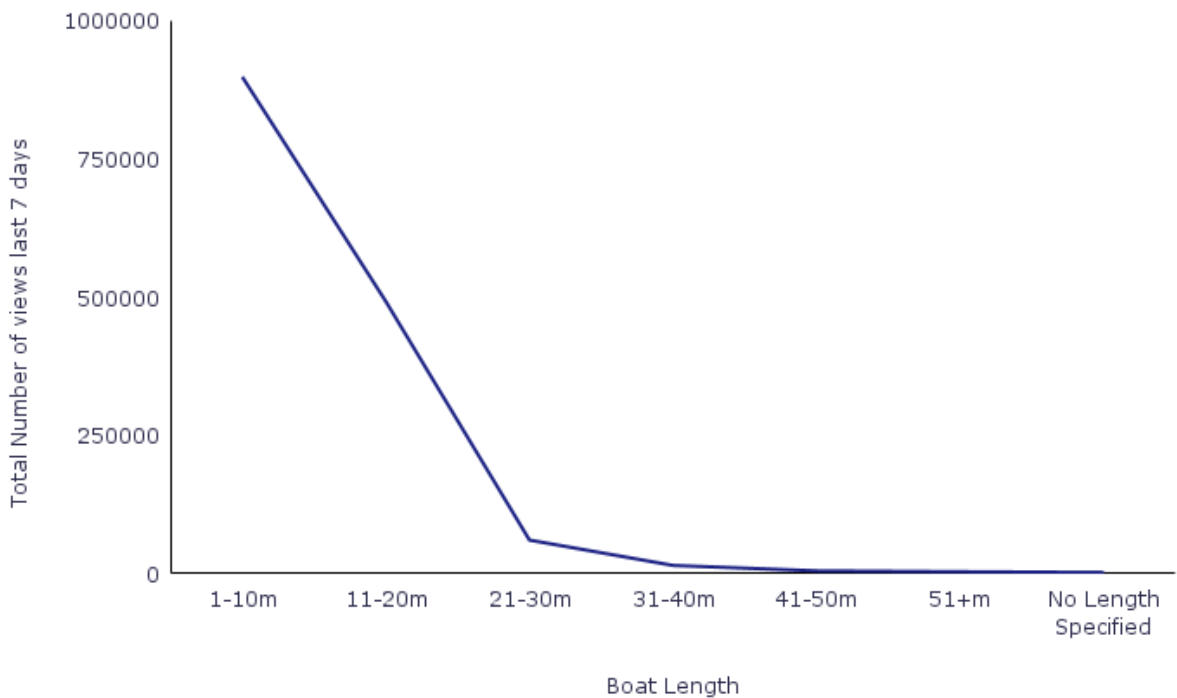
**Figure 4:** Relationship between the material used and boat popularity



**Figure 5:** Relationship between average boat price and material used

### Boat Size and Views

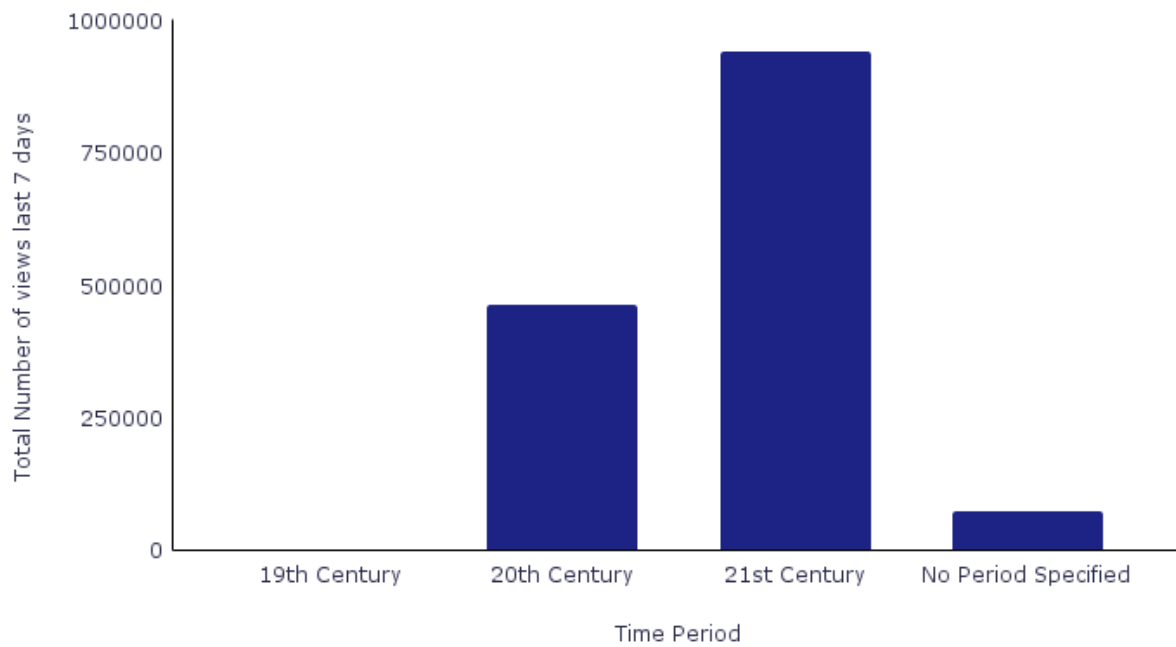
The analysis indicates that larger boats do not necessarily receive more views. There is a notable trend where smaller boats, particularly those in the 1-10 meter length range, tend to attract more views. Beyond the 10-meter mark, there is a significant drop in the number of views, suggesting that visitors may prefer smaller, more accessible boats over larger, potentially more expensive ones.



**Figure 6:** Relationship between the boat size and boat popularity

### Boat Year of Build and Popularity

There is a clear correlation between the year a boat was built and its popularity. Newer model boats receive more attention from visitors to the website. This trend is further supported by the observation that newer model boats also tend to command higher average prices as shown in Table 2. Buyers appear to be more interested in modern, recently built boats, potentially due to advancements in design, technology, and overall condition.



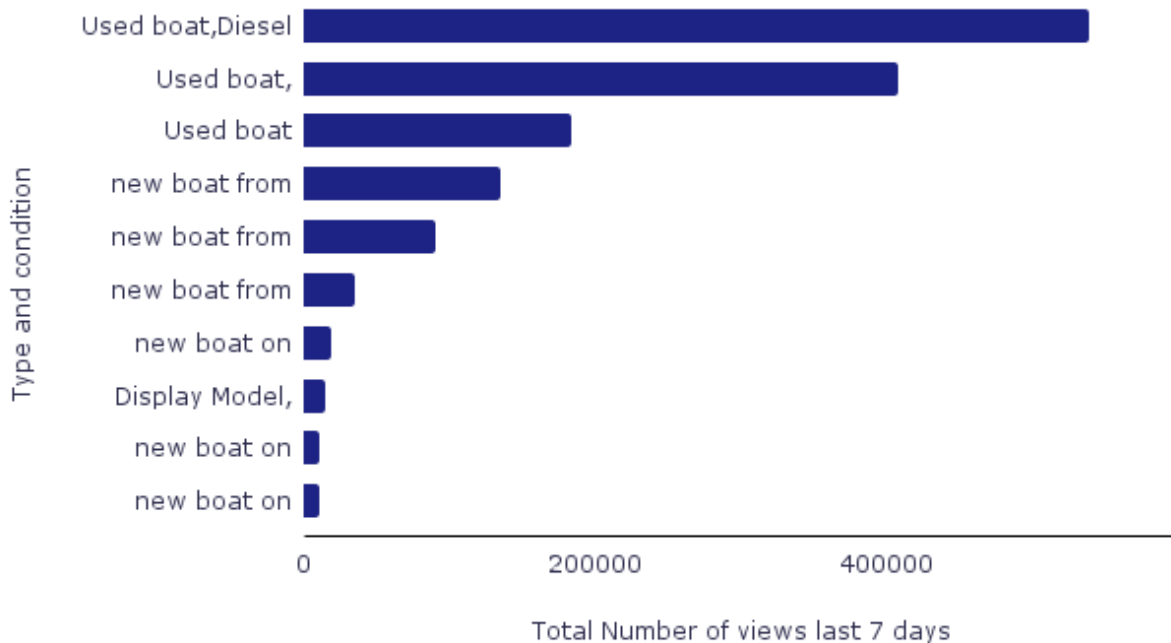
**Figure 7:** Relationship between boat age and number of views

Time Period	AVERAGE of Price (USD)
19th Century	\$124,946.24
20th Century	\$122,753.44
21st Century	\$398,190.82
No Period Specified	\$298,890.34

**Table 2:** Relationship between boat age and the average price

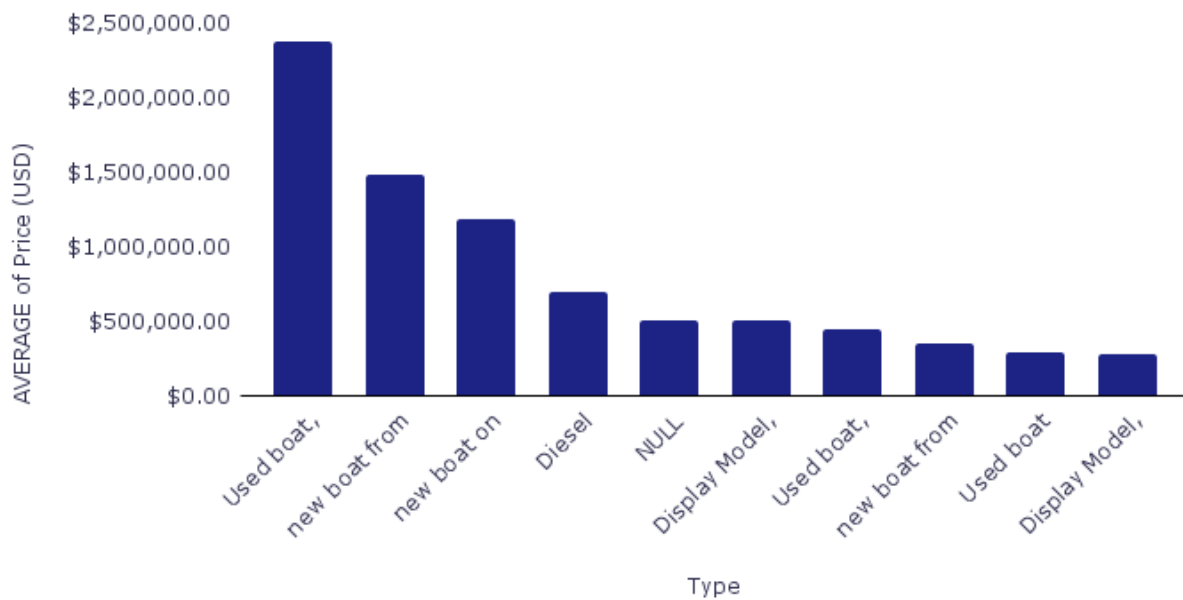
## Boat Type and Condition

There is evidence that suggests that the condition of boats significantly influences the number of views they receive on the website. Notably, used boats tend to be more popular among visitors, despite the variable pricing observed for both new and old boats.



**Figure 8:** Relationship between average price and boat popularity

The pricing of boats does appear to be influenced by their condition, but interestingly, there is often little differentiation in pricing between used boats and older boats. Despite the variable condition of these vessels, pricing tends to be quite similar. This observation suggests that buyers in the boat market may prioritize factors beyond just the age or condition of the boat when making purchasing decisions. Other factors, such as boat type, size, material, and included features, likely play a significant role in determining the price of a boat.



**Figure 9:** Relationship between average price and boat type and condition

## Website Performance

The dataset, while providing valuable insights, does not include information about the user experience and behavior whilst navigating the website, and source locations of the website's visitors. To gain a deeper understanding of user engagement patterns and potential geographic trends, further investigation is warranted.

It's worth noting that the website receives a substantial 1,474,902 views per week, demonstrating its significance and popularity within the boat and yacht market. However, without location data, drawing concrete conclusions about the geographic origins of user engagement remains challenging.

Nevertheless, the analysis uncovers several factors that significantly influence the number of views a boat listing receives. These include the type and condition of the boat, boat size, boat type, and the location of the listing. These findings suggest that potential buyers and enthusiasts exhibit particular interest in specific boat characteristics and the geographical locations of the listed boats. Further exploration of how these factors interact and drive user engagement could yield valuable insights for optimizing the website's performance and marketing strategies.

## Phase 6: Recommendations

1. Give customers access to longer-term market trend data to aid in making informed purchasing decisions. Encourage satisfied customers to submit testimonials or referrals to build trust and attract more potential buyers. Train customer service representatives to use the dataset to understand customer preferences better.
2. Manufacturers and dealers should consider the popularity of GRP boats and focus on offering a variety of boat sizes, particularly in the 1-10 meter range. Manufacturers could also emphasize the features of newer models to capitalize on their popularity.
3. Enhance website features to collect and analyze visitor location data. This will enable a deeper understanding of geographic user engagement patterns. Implement analytics tools to track user interactions and behavior for more robust insights.
4. Tailor marketing strategies to emphasize the boat type, condition, size, age and location factors, which significantly influence user engagement. Explore further statistical analysis to refine targeting systems based on visitor geography and preferences.
5. Consider allocating resources for the collection and analysis of data over longer time frames to identify and act upon long-term market trends effectively. Make data-driven decisions in alignment with the observed trends and customer preferences to drive the business's strategic direction.

## Conclusion

The analysis is limited to a 7-day timeframe, restricting the understanding of longer-term market dynamics and sales trends. Moreover, the absence of sales data hinders insights into key factors influencing boat and yacht sales, such as inquiries and negotiations. The analysis identifies correlations but not causality. External factors like economic conditions and holidays are not considered. Additionally, the dataset lacks geographic details, limiting the analysis of regional variations and user engagement. Despite these limitations, the analysis highlights the popularity of motor yachts and sports boats, price trends, and strong visitor engagement for boats and yachts listed in Europe. Further statistical analysis is recommended for insights into user behavior and marketing strategies.