Project: Nearly New Nautical

Table of Contents

Introduction

> **Descriptions:** This dataset has information about Nearly New Nautical which is a website that allows users to advertise their used boats for sale. The marketing team is preparing a weekly newsletter for boat owners. The newsletter is designed to help sellers to get more views of their boat, as well as stay on top of market trends. The main target is to take a look at the recent data and help the company to learn more about the characteristics of the most viewed boat listings in the last 7 days.

dataset columns description:

- **Price**: Character, boat price listed in different currencies (e.g. EUR, Â£, CHF etc) on the website.
- **Boat Type**: Character, type of the boat.
- **Manufacturer**: Character, manufacturer of the boat.
- **Type**: Character, condition of the boat and engine type(e.g. Diesel, Unleaded, etc).
- **Year Built**: Numeric, year of the boat built.
- **Length**: Numeric, length in meter of the boat.
- **Width**: Numeric, width in meter of the boat.
- **Material**: Character, material of the boat (e.g. GRP, PVC, etc).
- **Location**: Character, location of the boat is listed.
- **Number of views last 7 days**: Numeric, number of the views of the list last 7 days.

Questions:

- Do the most expensive boats get the most views?
- Which boat type has the most views on average?
- Which boat and engine condition has the most views on average?
- Which boat material has the most views on average?
- Are there common features among the most viewed boats?

Data Wrangling

This is a three step process:

- Gathering the data from Dataset and investegate it trying to understand more details about it.
- Assessing data to identify any issues with data types, structure, or quality.
- Cleaning data by changing data types, replacing values, removing unnecessary data and modifying Dataset for easy and fast analysis.

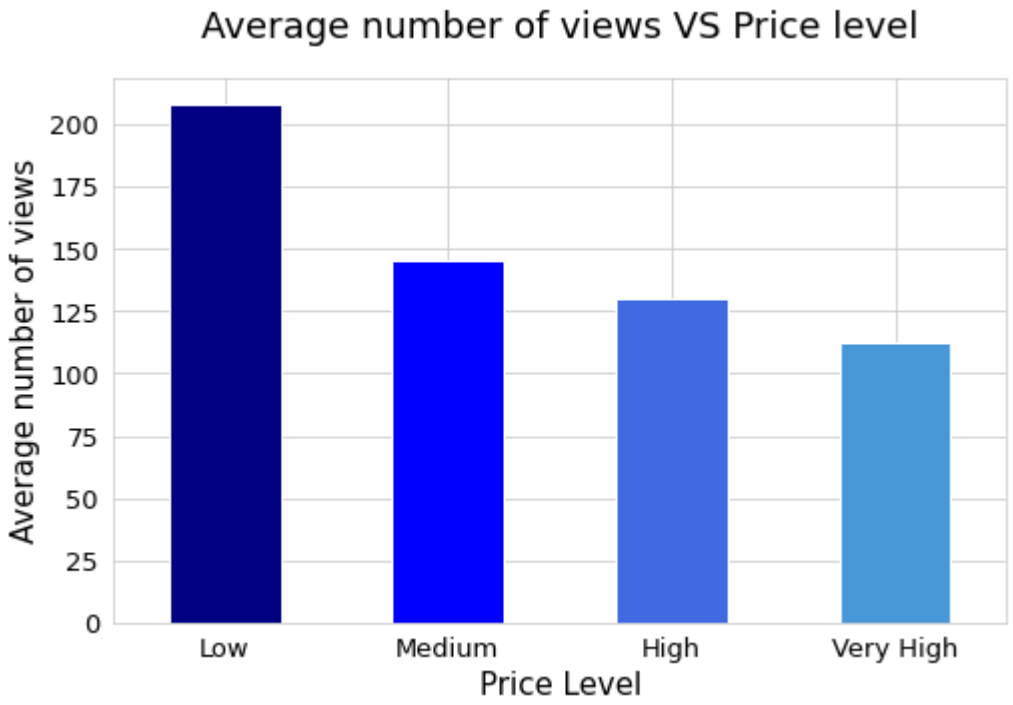
Exploratory Data Analysis

> Now I'm going to explore this data and try to find patterns in it, compute statistics and visualize the relationships to answer the questions and detect the characteristics that affect patients attending to their hospital appointment

Research Question 1

> Do the most expensive boats get the most views?

In [56]:
```
# make bar plot of Average number of views VS Prcie levels
```



The cheapest boats get the most views and The most expensive boats get the least views on average.

Answer 1

- From above visualization and information we can conclude that:
  - The cheapest boats get the most views with average 208 views.

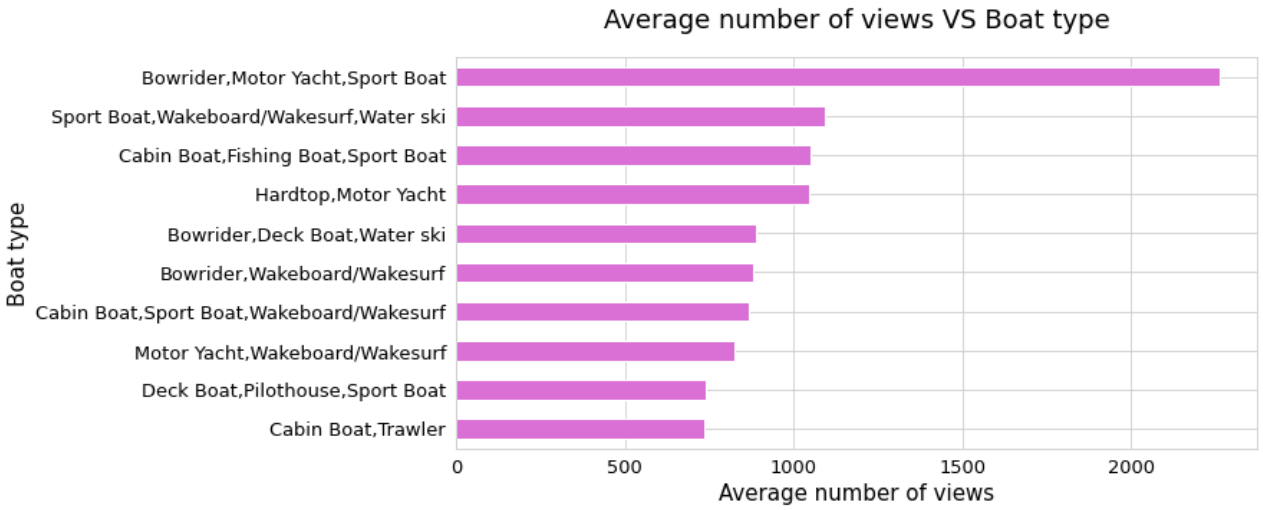
- Boats in medium price level get 145 average views.
- Boats in high price level get 130 average views.
- The most expensive boats get the least views with average 112 views.

(The most expensive boats get the least views and the cheapest boats get the most views on average.)

Research Question 2

> Which boat type has the most views on average?

In [59]:
```
# make bar plot for top 10 boat types vs number of views
```



"Bowrider,Motor Yacht,Sport" Boats get the most views on average

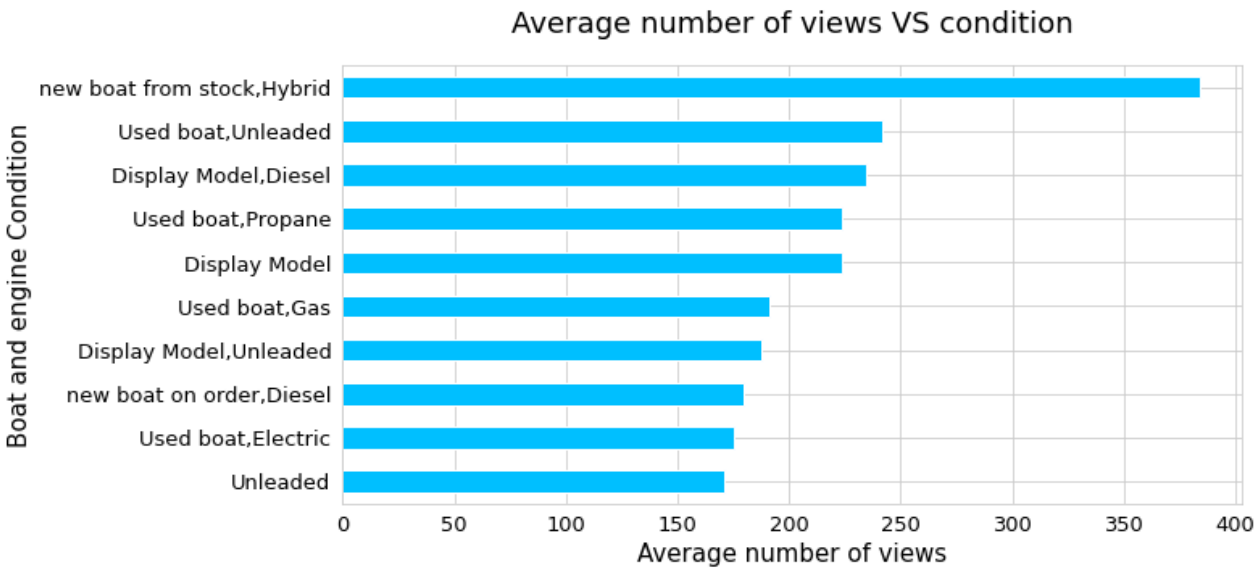
Answer 2

- From above visualization we can conclude that:
  - "Bowrider,Motor Yacht,Sport" Boats get the most views with average 2261 views.
  - "Sport Boat,Wakeboard/Wakesurf,Water ski" Boats come in the second place with 1095 views on average.

Research Question 3

> Which boat and engine condition has the most views on average?

In [61]:
```
# make bar plot for top 10 types vs number of views
```



Hybrid new boats get the most views on average.

Answer 3

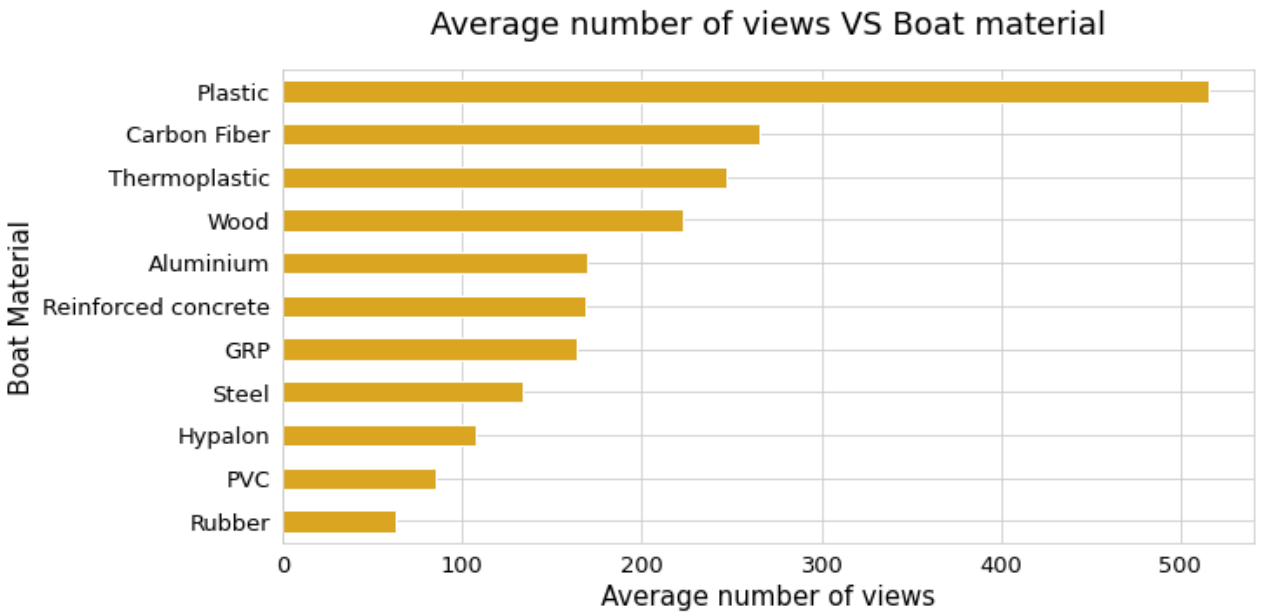
- From above visualization we can conclude that:
  - Hybrid new boats get the most views with average 384 views.
  - Unleaded used boats come in the second place with 242 views on average.

Research Question 4

> Which boat material has the most views on average?

In [488…

```
# make bar plot for boats material vs average number of views
```



Answer 4

- From above visualization we can conclude that:

- Plastic boats get the most views with average 515 views.
- Carbon Fiber boats come in the second place with 265 views on average.
- Rubber and PVC boats get the least views with average 63 and 85 views respectively.

Research Question 5

> Are there common features among the most viewed boats?

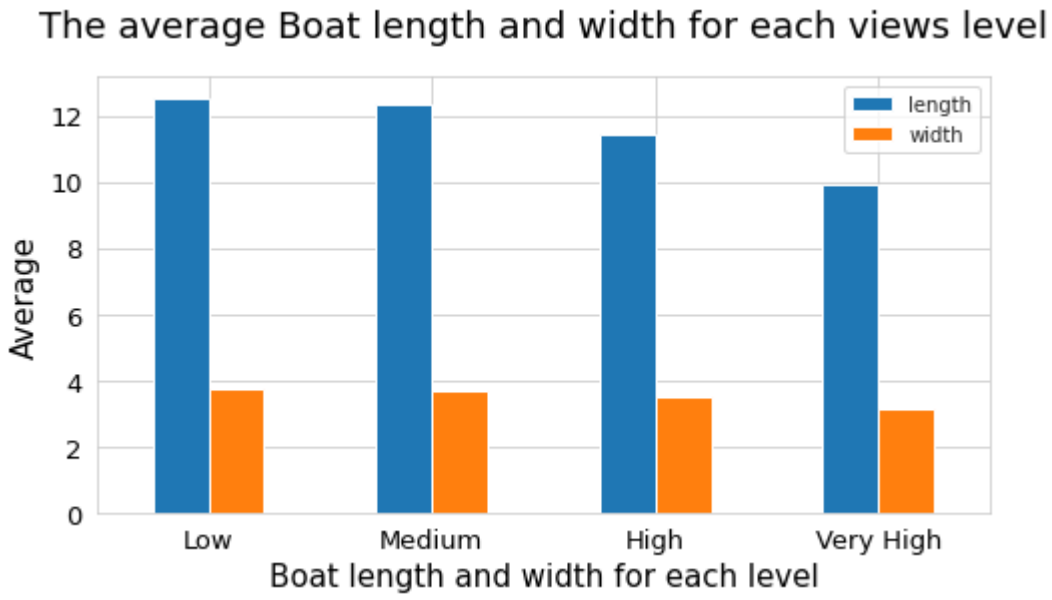
In [480…
```
# check the common characteristics for quantitative variables on each views leve
```

Out[480…

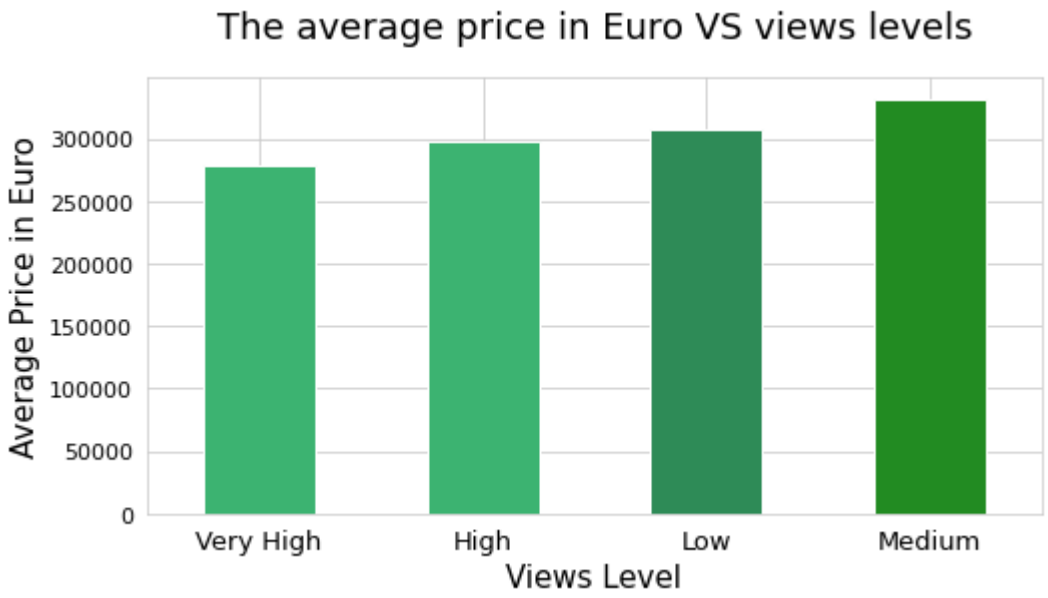
views_level	price_in_euro	length	width
Low	306688.058552	12.533124	3.760860
Medium	331775.429545	12.358767	3.696139
High	297021.508799	11.422328	3.501307
Very High	277990.211799	9.924699	3.111706

Very high views level has the smalest boat length and width average. Low views level has the largest boat length and width average.

In [487…
```
# bar chart for The average Boat length and width for each views level
```



In [501…
```
# The average Price for each views level
```



Very high views level has the least expensive boat price on average. Medium views level has the most expensive boat price on average.

In [512…
```
# check the common characteristics for categorical variables on each views level
```

Out[512…

	boat_type				type				mater		
	count	unique	top	freq	count	unique	top	freq	count	unique	top	fr
views_level												
Low	2514	33	Motor Yacht	747	2511	16	Used boat,Diesel	1241	1989	9	GRP	12
Medium	2481	33	Motor Yacht	755	2481	20	Used boat,Diesel	1165	2006	9	GRP	12
High	2447	41	Motor Yacht	717	2445	19	Used boat,Diesel	981	2000	10	GRP	13
Very High	2446	116	Motor Yacht	501	2445	21	Used boat,Unleaded	846	2144	8	GRP	16

Answer 5

- From above visualisation we can conclude that:
  - Very high views level has the smalest boat length and width average.
  - Low views level has the largest boat length and width average.
  - Very high views level has the least expensive boat price on average.
  - Medium views level has the most expensive boat price on average.
  - The most frequant boat type is motor yacht for all views levels.
  - The most frequant boat type is "motor yacht" for all view levels.
  - The most frequant engine condition is "Used boat,Diesel" for low, medium and high views levels and "Used boat,Unleaded" for very high views level.
  - The most frequant boat material is "GRP" for all views levels.

Conclusions

> **1. The cheapest boats get the most views and the most expensive boats get the least views on average.**
>
> **2. "Bowrider,Motor Yacht,Sport" Boats get the most views on average.**
>
> **3. Plastic,Carbon Fiber and Thermo Plastic boats get the most views on average respectively.**
>
> **4. Rubber,PVC boats get the least views on average respectively.**
>
> **5. The smaller the boat, the more views it gets:**
>
> - Boats with the smalest length and width average get the most views.
> - Boats with the largest length and width average get the least views.
>
> **6. Very high views level has the least expensive boat price on average.**
>
> **7. The most frequant boat type is motor yacht for all views levels.**
>
> **8. The most frequant boat material is "GRP" for all views levels.**
>
> **9. The most frequant engine condition for very high views level is "Used boat,Unleaded".**

Limitation and recommendations

> **1. There are a lot of empty values on this data set (Nans) especially on Manufacturing and Material columns.**
>
> **2. The minimum boat width is 0.01 Meter which is obviously a typo issue.**
>
> **3. There are 551 observations with value 0 on Year Built column.**

It's very important to give these points more attention in the future because they may affect our results accuracy and we need avoid such mistakes.

In [ ]: