

Nearly New Nautical



Views of Listings

Final Report (Summary)

Context

Look at data of the last 7 days and extract insights about the characteristics of the most viewed boat listings. The knowledge will be directed in preparing a weekly newsletter for boat owners.

Method

Use of common tools for data analysis, to create visualizations, aggregate and calculate results and for doing statistical analysis.

The used tool was the programming language R.

Data

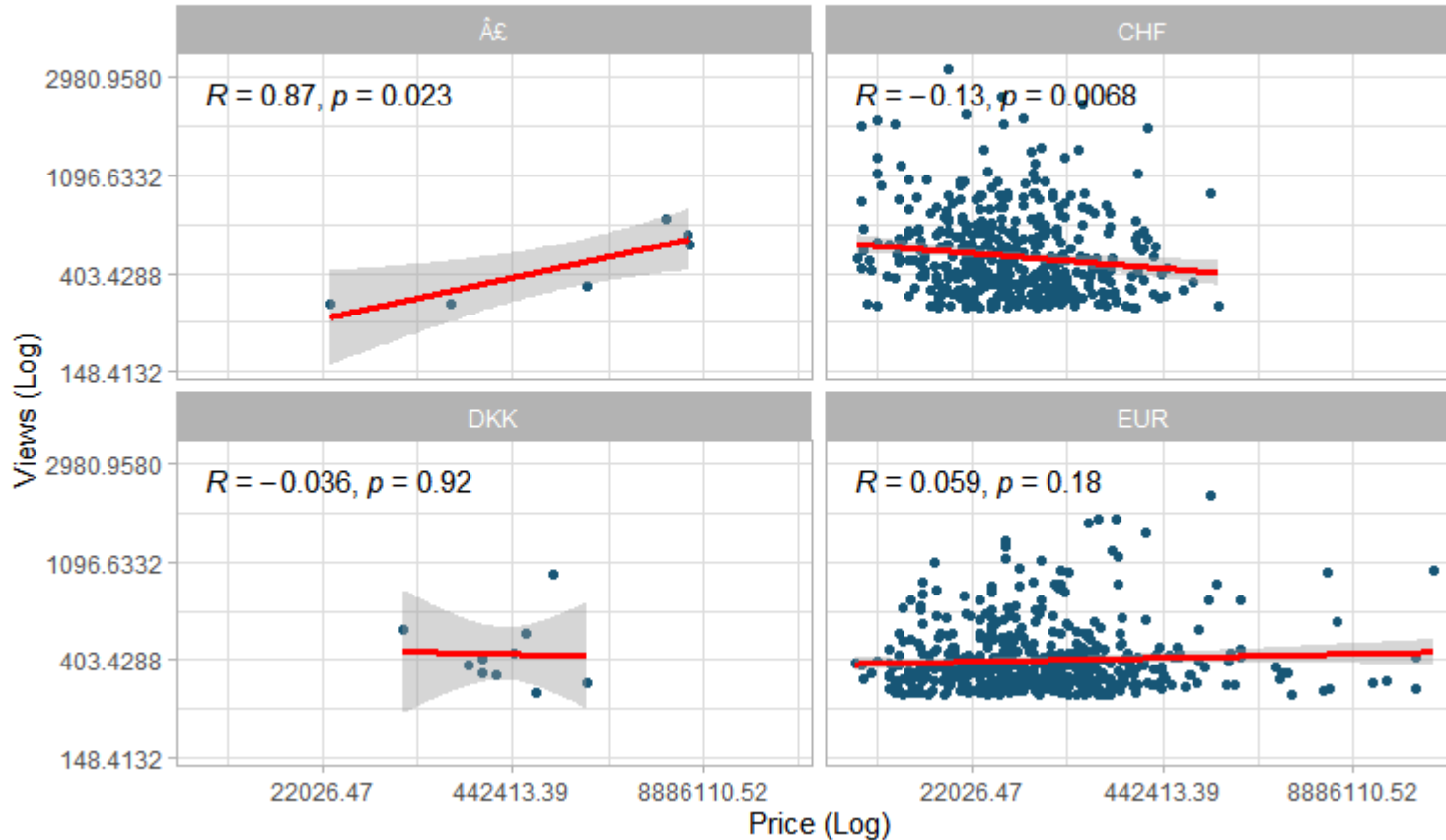
The data refers to posts made by users announcing their used boats for sales. It also contains the count of views in the last 7 days. $N = 990$ (10% of the original data)

- **Information contained in the data:** Number of views, price, boat type, manufacturer, type, year built, length, width, material, location

Price don't influence the number of Views

Number of Views compared to Boat Prices

Adjusted by the currency utilized



R is the Pearson correlation coefficient - it measures the linear relationship between two variables - and ranges between -1 and 1, in which these extremes mean strongly(negative and positive) association, values close to zero represent no association. The p assert the statistical significance of the measure, where values smaller than 0.05 indicate the existence of 'true' linear relationship.

For the CHF region, the most expensive boats get less views, with a very minimal impact

- Has a negative correlation coefficient, not so expressive (-0.13) but that indicates a negative impact of the price; we can guarantee a statistical support ($p < 0.05$)

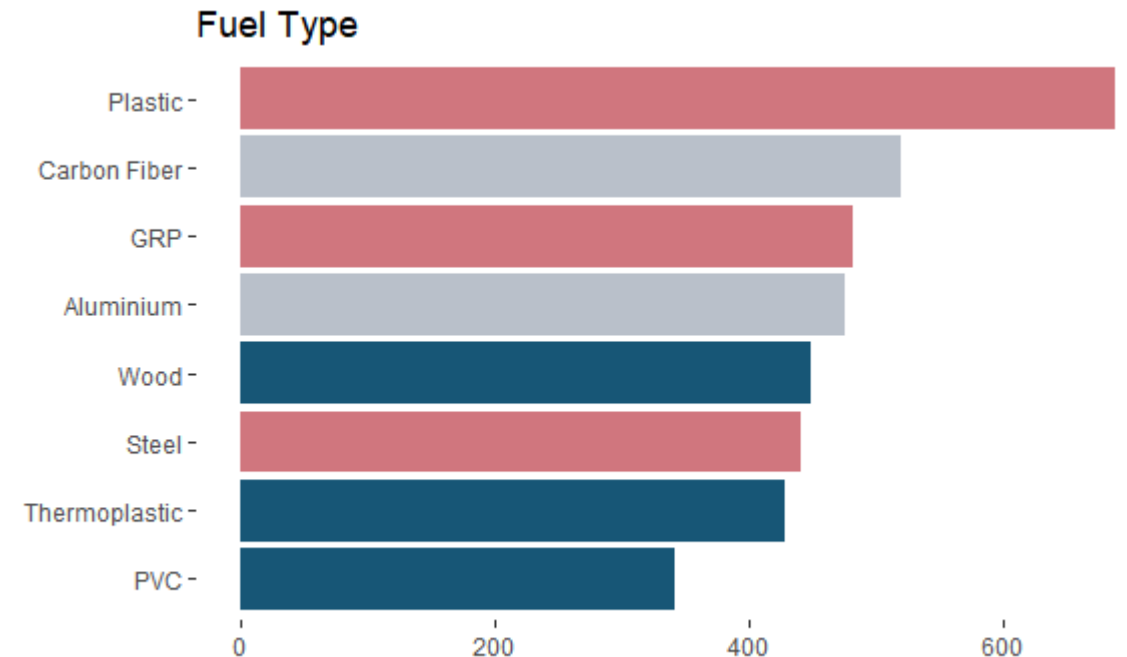
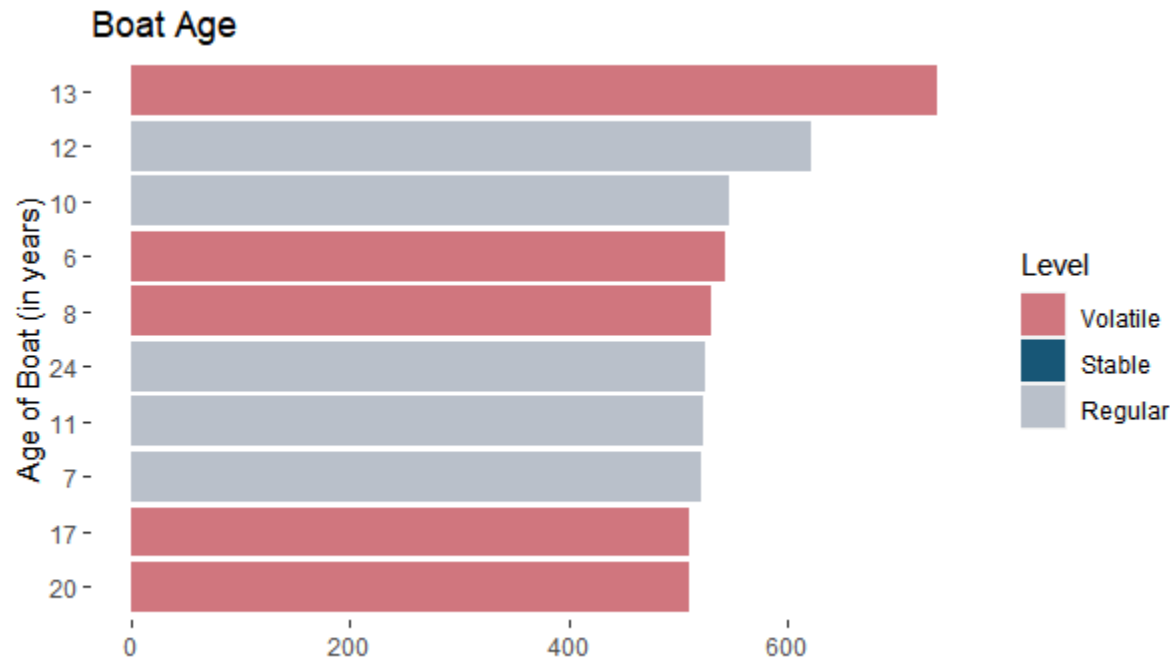
The others currencies dont have a pattern

- The others has inexpressive values - ranging from -0.1 to 0.1 - with different trends (negative and positive); theres no statistical support (except for the first, that has a very small sample)

Year built and Fuel type (substituir fuel type por material)

At the top we can see that boat with approximated age of ~10 years have more views. But the pattern is unclear, having a lot of unstable listings by the different ages.

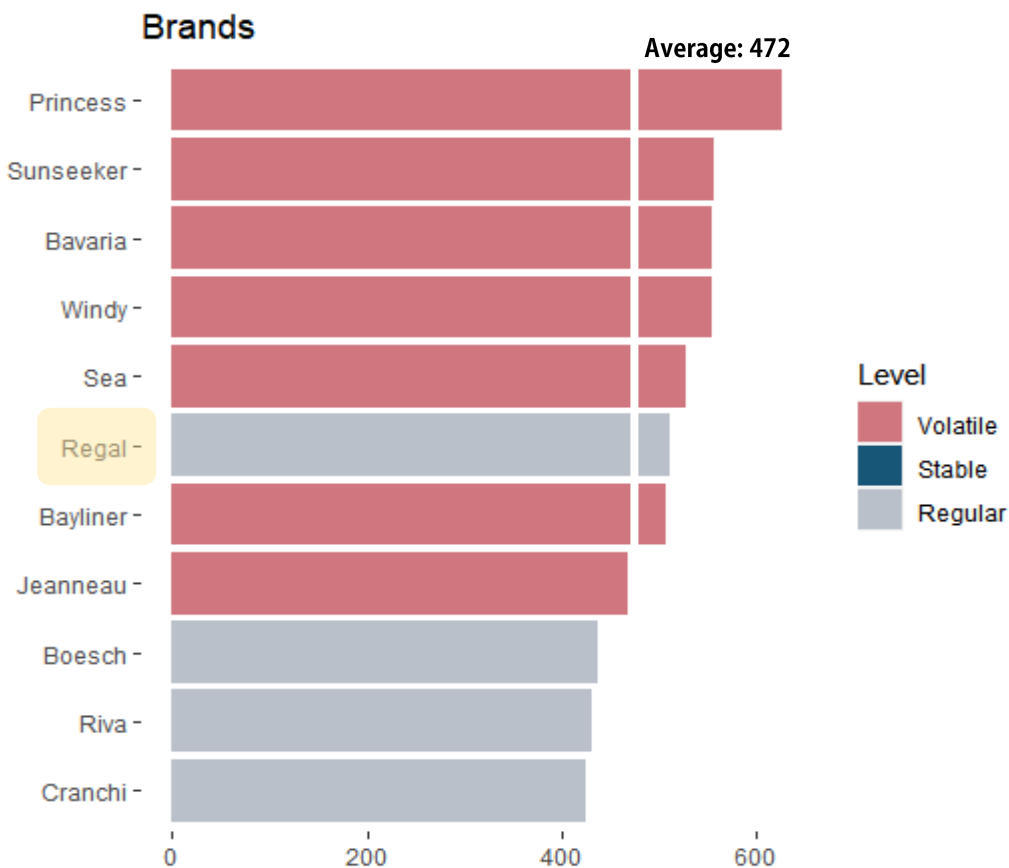
Boats made of plastic have 46% more engagement in comparison with the average. Carbon fiber and aluminium are stable.



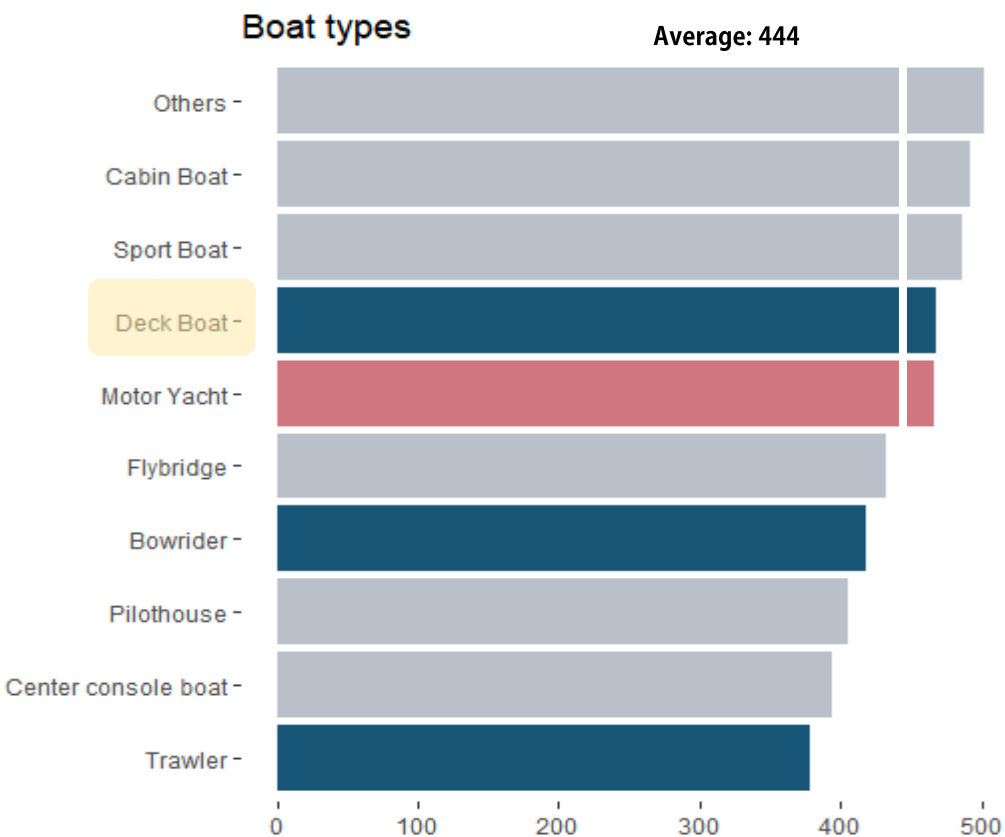
The label 'Level' is determined based on the variation of views of the posts. The variation measure utilized is coefficient of variation = Standard Deviation / Mean. To classify some record in a label, it was calculated the quintiles (divide the data into 5 pieces of 20%) and the values in the 1st and 5th quintiles are labeled as "Stable" and "Volatile" respectively; the others receive the "Regular" label.

Brand and Boat type

Princess is at the top. The average views of the top 5 are very similar and all of them are unstable, with a high variation of views. Regal is consistent and is above the average.



Four of the boat types have visualizations above the average. Deck boat is one of them and has had stable results.



The label 'Level' is determined based on the variation of views of the posts. The variation measure utilized is $\text{coefficient of variation} = \frac{\text{Standard Deviation}}{\text{Mean}}$. To classify some record in a label, it was calculated the quintiles (divide the data into 5 pieces of 20%) and the values in the 1st and 5th quintiles are labeled as "Stable" and "Volatile" respectively; the others receive the "Regular" label.

Regions

Focus on West and South Europe

- There two types of countires with good results: the ones with a high volume and the ones with a small volume. Theres a growing engagement in South Europe.

(avg.) and informations of well-performing regions



Switzerland (519)

Total: 257 thousand
Posts: 495 (1 of 11)



Croatia (486)

Total: 12 thousand
Posts: 25 (5 of 11)



Germany (422)

Total: 74 thousand
Posts: 177 (2 of 11)



Spain (443)

Total: 10 thousand
Posts: 24 (6 of 11)



Italy (399)

Total: 32 thousand
Posts: 81 (3 of 11)



France (443)

Total: 9 thousand
Posts: 22 (7 of 11)

Average views by Region



* Considering only regions with expressive number of posts (At least 75)

Conclusions

FEATURES



Some characteristics of the boats does not matter - price, fuel type, width and lenght.

Boats having **close to 10 years** have most views.

Plastic boats are the most sucessfull.

BRAND & BOAT TYPE



The manufacturers have a very unstable results. Even with some brands with more expressive engagement, the only one stable is **Regat**.

Cabin, Sport and specially **Deck boat** have more views than other types.

REGION



Countries with high volume like **Germany, Italy** and **Switzerland** have a good number of views.

Countries with little volume but good results like **Croatia, Spain** and **France** should increase the number of posts.