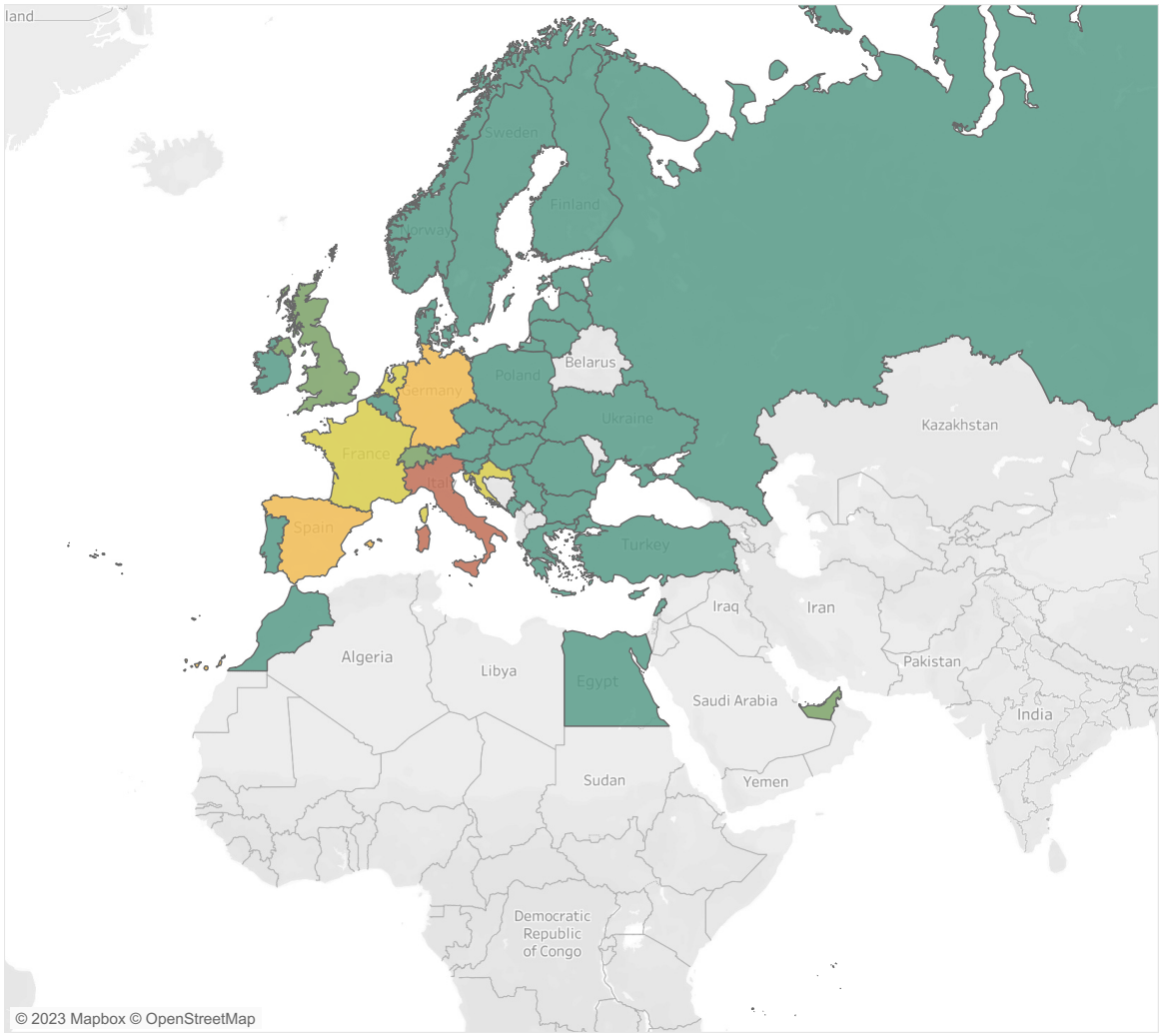
Final Story

Boat Pricing Analysis



USD Total

12,750 592M

The sale of boats is a prevelant business in all corners of the Earth. This notion especially true in countries that border the sea like multiple European nations.

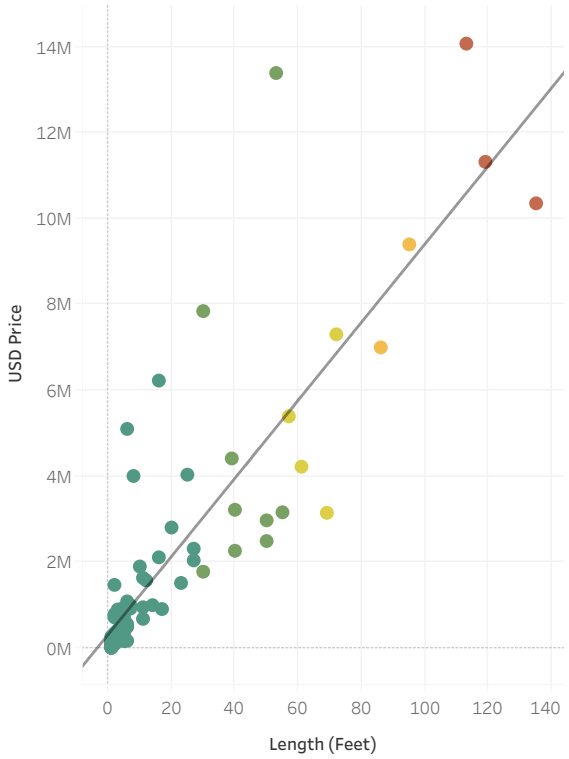
In this case study, we will be exploring the amount of profit made in boat sales per country using data collected over multiple decades.

We will also be exploring potential correlations between the pricing of boats based off of boat size to explain what characteristics, if any, dictate price.

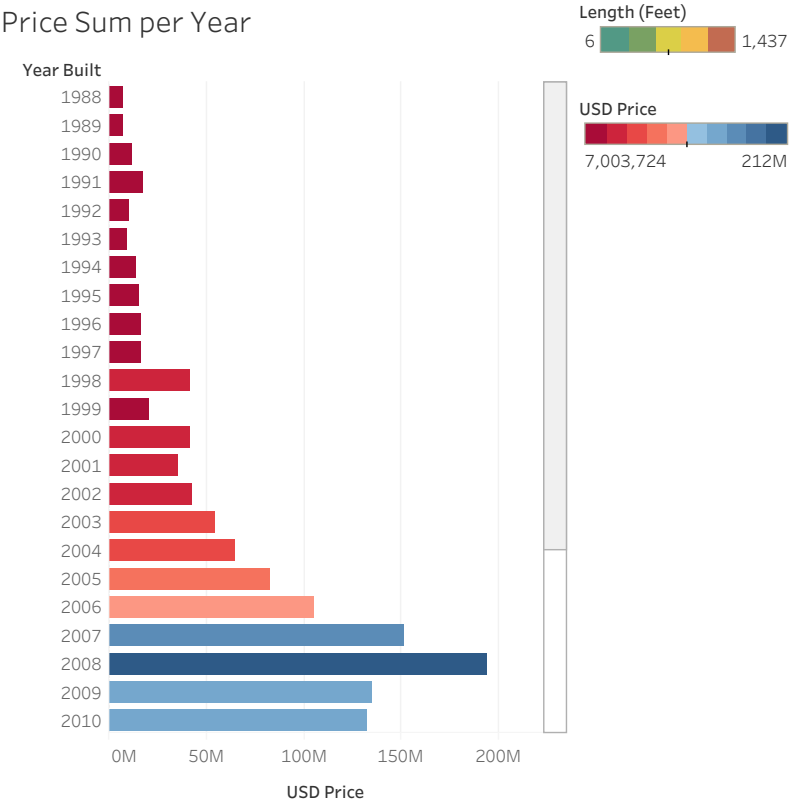
Final Story

Introduction	Exploratory Analysis	Linear Regression	Cluster Analysis	Cluster Analysis Results	Final Results and Recommendations

Boat Price by Length



Price Sum per Year



The first step was to look at linear relationships.

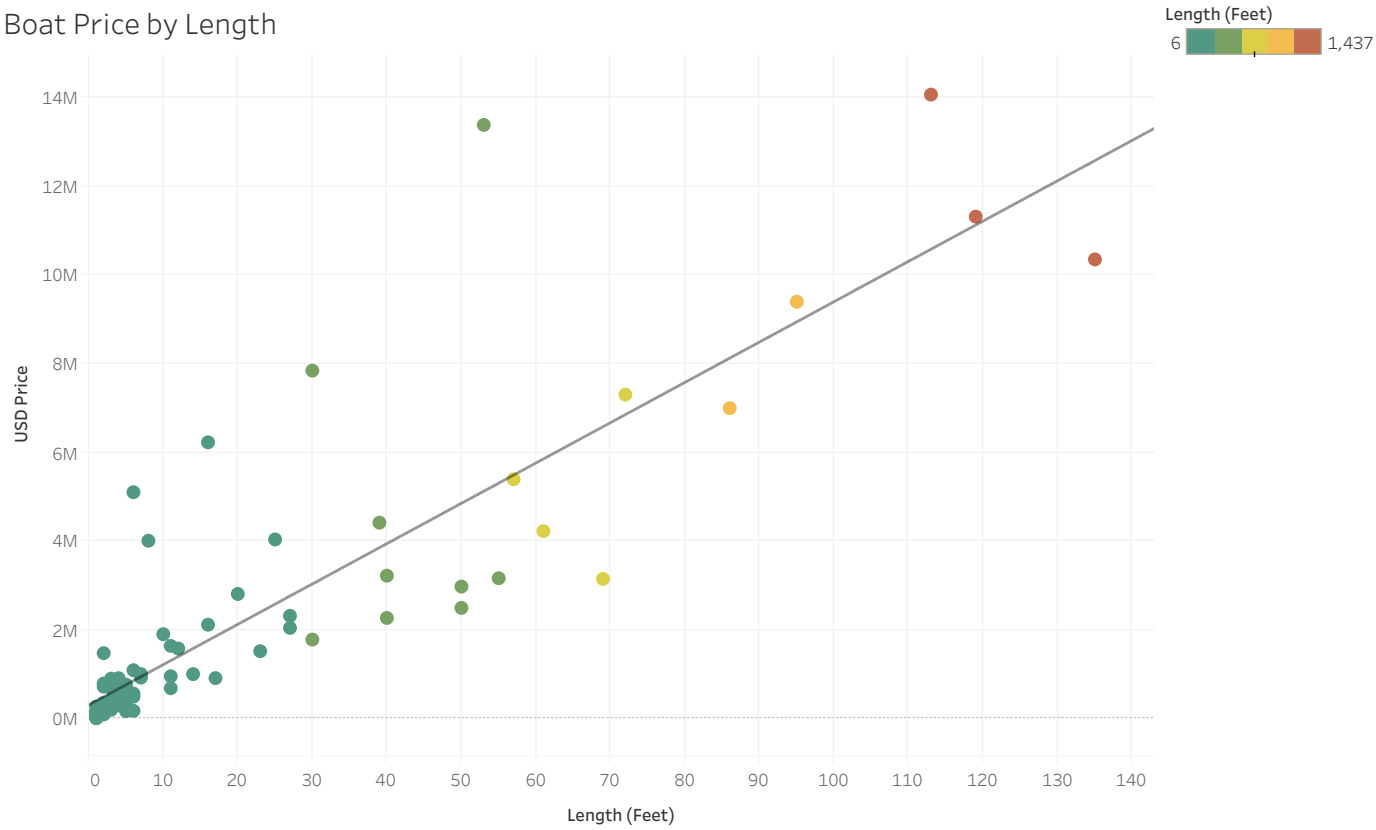
As expected, boat length tends to trend upward with price as shown in the first graph. However, there are outliers that exist in the data that are currently impossible. For consistency, those values were removed.

The bar chart shows an estimate of the dollar amount spent per year on boats. I excluded the values prior to 1988 as they were very similar in values. After 1988 is where the largest flunctuation of dollar amount is.

With this data, our hypothesis is: in majority of cases, as boat length increases, boat price also increases...

Final Story

Boat Price by Length



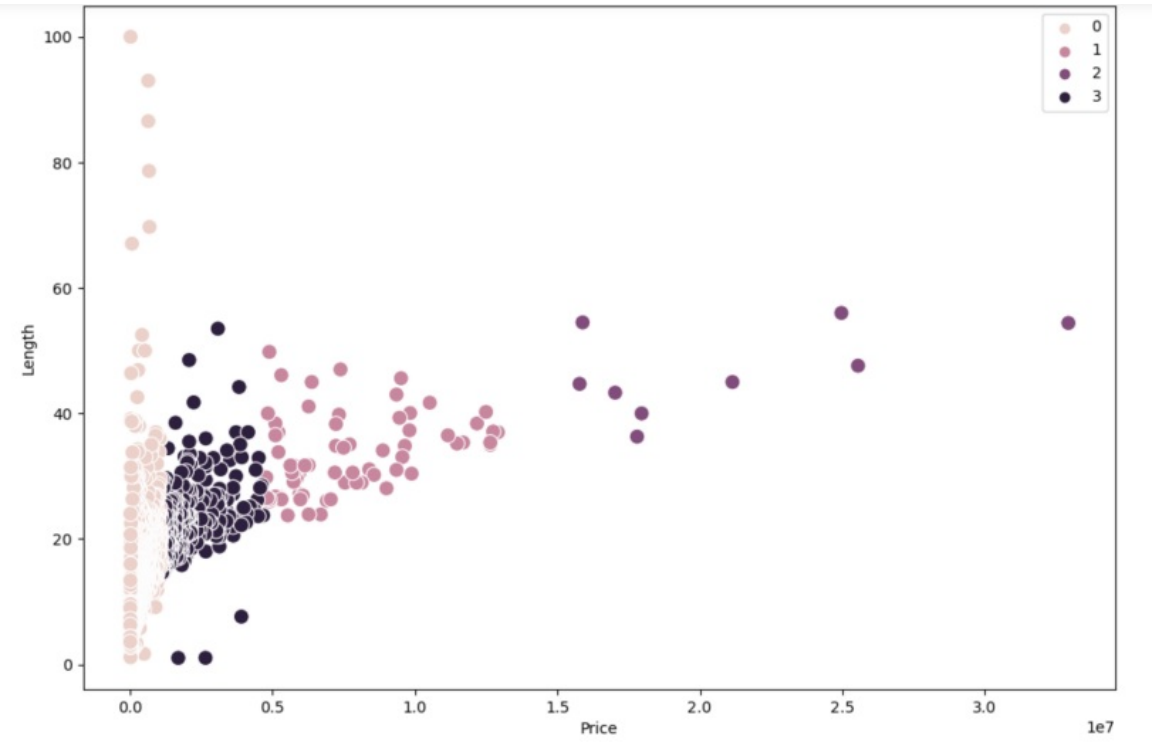
To test the hypothesis, a linear regression was conducted.

The relationship between the data is very linear with the lower values but as length increases, price tends to flunctuate more inconsistently. There are multiple points that are nowehre near the regression line, even with the impossible outliers removed.

The linear regression method is unsuitable to properly interpret this data.

Final Story

Cluster Analysis



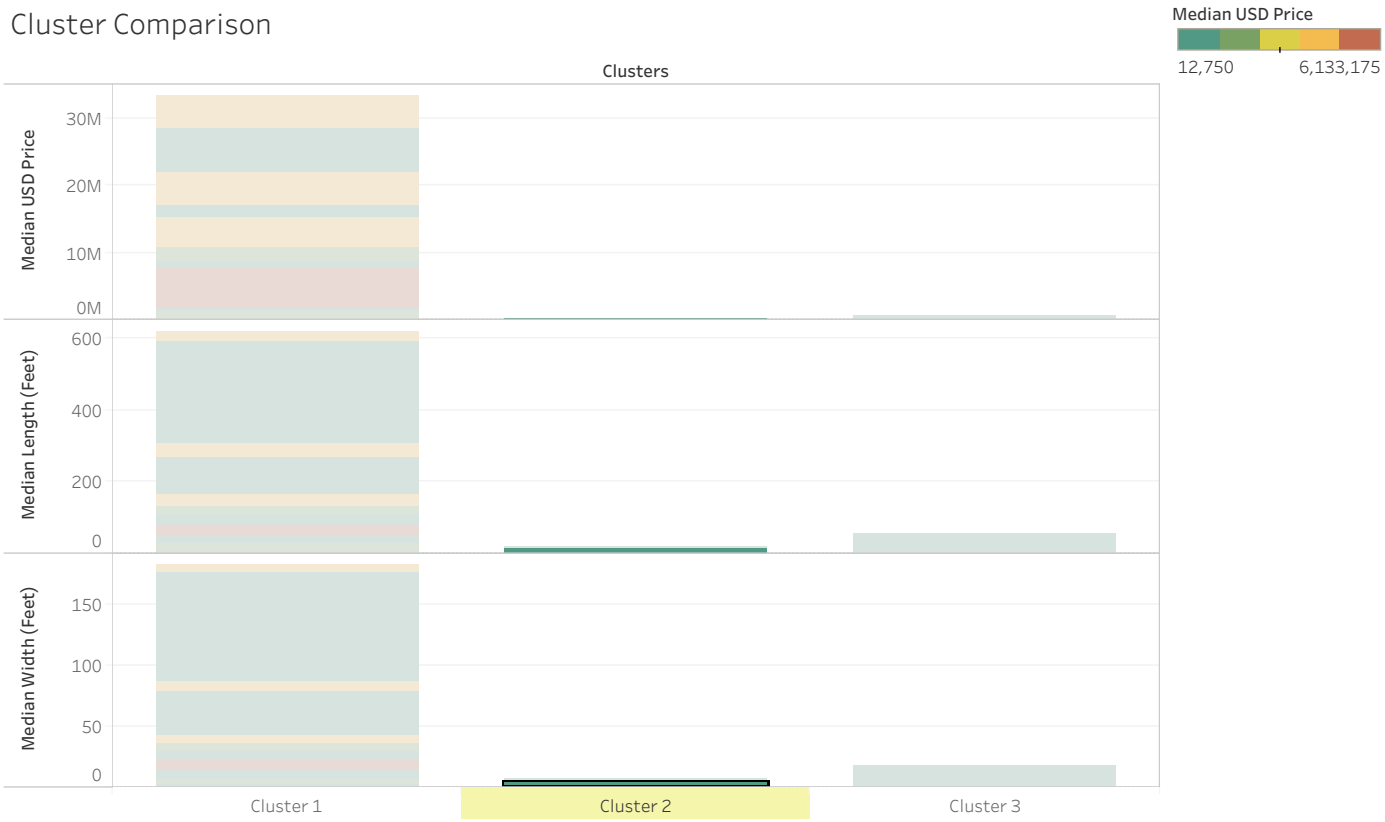
I decided to conduct a cluster analysis to determine if any new patterns exist within the data. This graph shows that despite the very long length of certain boats, they still fall into the lower price range. Additionally, smaller boats still have much higher prices the outliers with longer length.

The least expensive boats are represented by the beige points while the most expensive are represented by the medium purple points. This presents an interesting point of view as that even the longest ship documented in the graph is not the most expensive. It shows that the most expensive boats are between roughly 35 to 60 feet in length.

A conclusion could be drawn that there are multiple factors that play into the cost of a boat other than length. Some factors may include engine size, horsepower, model of boat, demand for that particular make or model, luxery amenities, the pur..

Final Story

Cluster Comparison

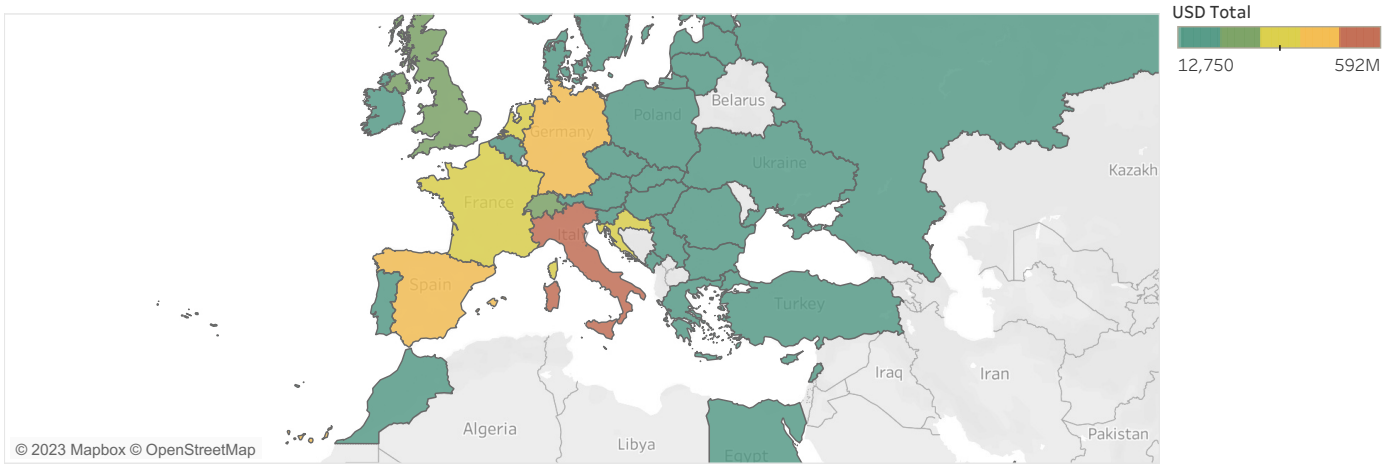


These clusters represent the median prices boats have sold for in specific countries when compared to the median lengths and widths of the boats being sold in those nations. These clusters again demonstrate that the amount a boat is being paid for within that nation is not caused solely by length or width of the boat itself; pointing to other factors that dictate the overall price.

A prime example of this analysis is the United Arab Emirates. The median cost spent in that nation is 61,000,000 while the median length and width of the boats sold is about 31 feet and 7 feet, respectively.

Final Story

Introduction	Exploratory Analysis	Linear Regression	Cluster Analysis	Cluster Analysis Results	Final Results and Recommendations



Conclusion:

While boat size (length and width) play key roles in the pricing of a boat, they are not the only factors to take into consideration. Other variables worth exploring would include:

- Engine size/horsepower
- Model
- Years of usage if any
- Purpose of the boat (fishing, shipping, luxery, etc.)
- Demand
- Amenities

Limitations:

While conducting my analysis, many of the variables listed above were missing which could have provided valuable insight to the cost of a boat. Additionally, many data values were missing from the raw data set resulting in data that was inconsistent or columns that were not useful. Of the columns that included data, the data was not very granular and did not provide much insight on their own.

Next Steps:

- Gather a more complete data set for each boat.
- Ensure the data collection method is collecting the listed variables.
- Rerun analytical tests to see if new conclusions can be drawn.