

6.1: Sourcing Open Data

Boat Sales Analysis:

Data Source: The goal of the analysis is to prepare 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 data set provided by Kartik Bhandary. And it can be downloaded from [Kaggle](#).

Data Profiling:

Variable	Structured/ Unstructured	Quntitative/Qualitative	Discrete/Continuous	Nominal/Ordinal/Binary
Price	Structured	Quntitative	Continuous	-
Boat Type	Structured	Qualitative	-	Nominal
Manufacturer	Structured	Qualitative	-	Nominal
Type	Structured	Qualitative	-	Nominal
Year Built	Structured	Quantitative	Descrete	-
Length	Structured	Quntitative	continuous	-
Width	Structured	Quntitative	continuous	-
Material	Structured	Qualitative	-	Nominal
Location	Structured	Qualitative	-	Nominal
Number of views last 7 days	Structured	Quntitative	descrete	-

ethics:

The dataset is open data and displayed on website. It appears that there are no ethical concerns with this data.

Limitations:

There are so many missing values in columns Material and Manufacturer which are important for the analysis. Also, the Price values in the price column are in different currencies.

Key Questions:

1. What are the characteristics of the most viewed boat listings?

2. is it the most expensive boats that get the most views?
3. Are there common features among the most viewed boats?