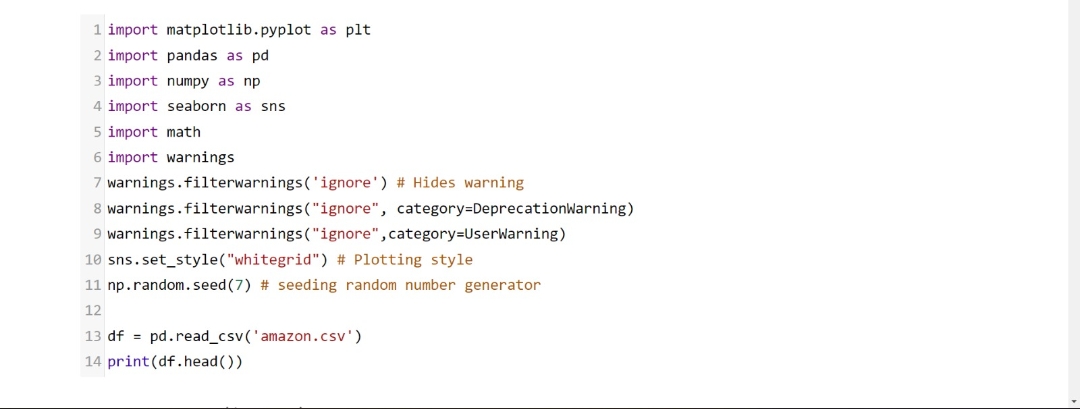
Sentimental Analysis for Marketing

Phase-4

***We further building our project by loading the Data Set and Describing that and Cleaning the data and Visualising the distributions and Evaluation in Google colab Notebook.***

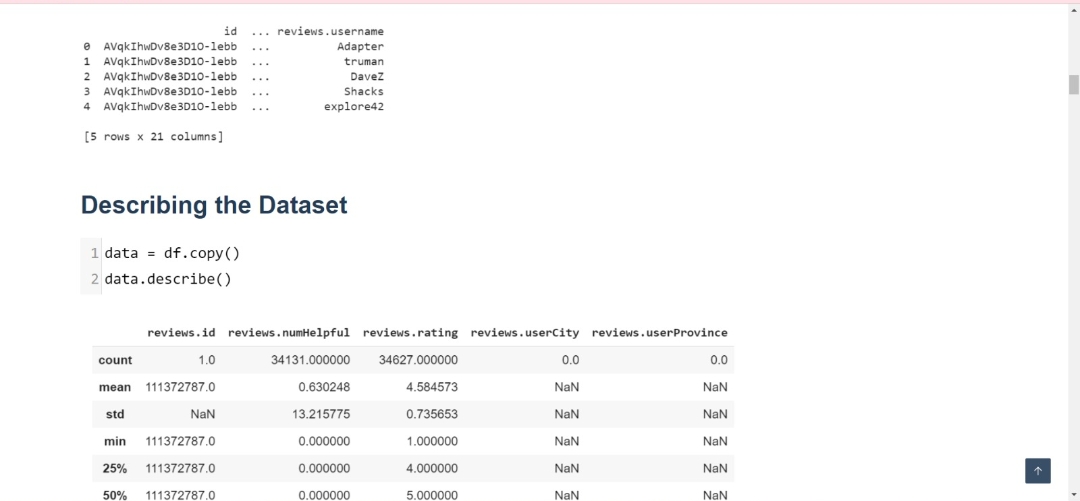
### **Let’s Import the necessary Modules and take a look at the data:**



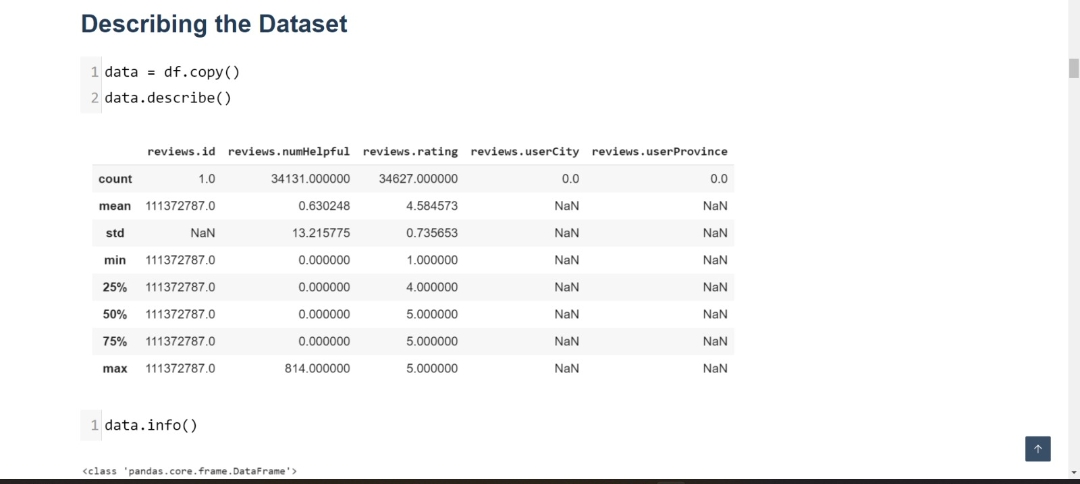
**Describing the Dataset:**

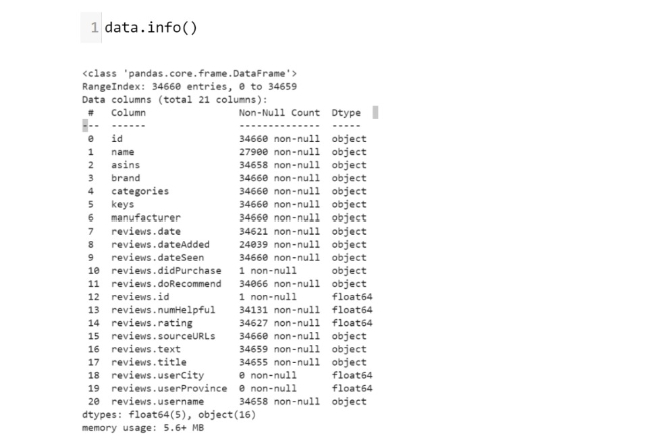
*Overall description about the dataset should be contain in this*

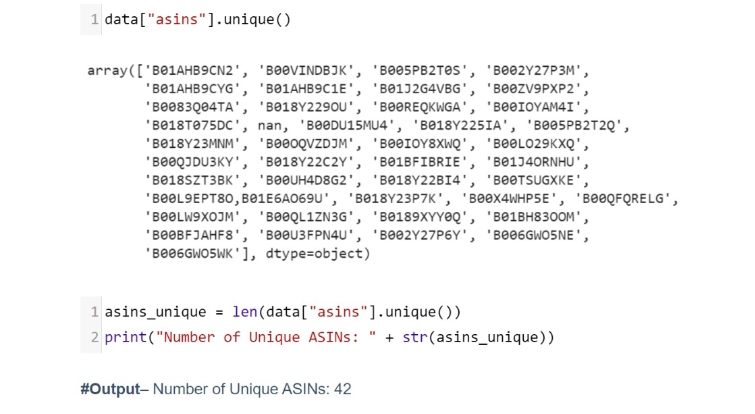
The purpose of this to done a overall relationship between the data and the future predictions based upon that.

****

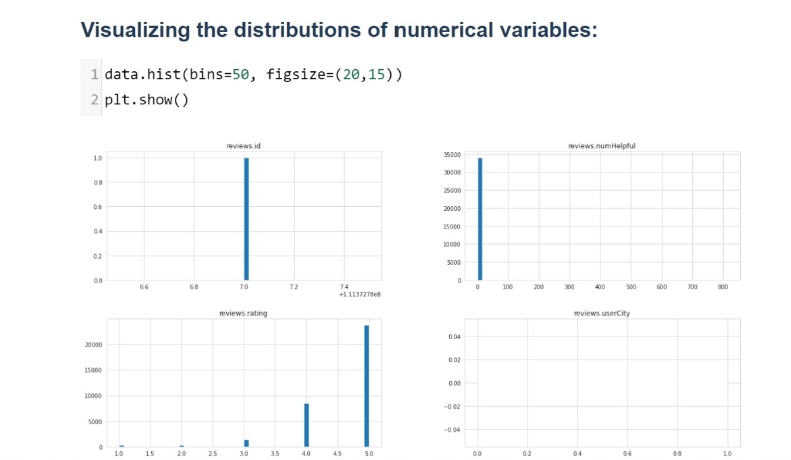
**We need to clean up the name column by referencing asins (unique products) since we have 7000 missing values:**

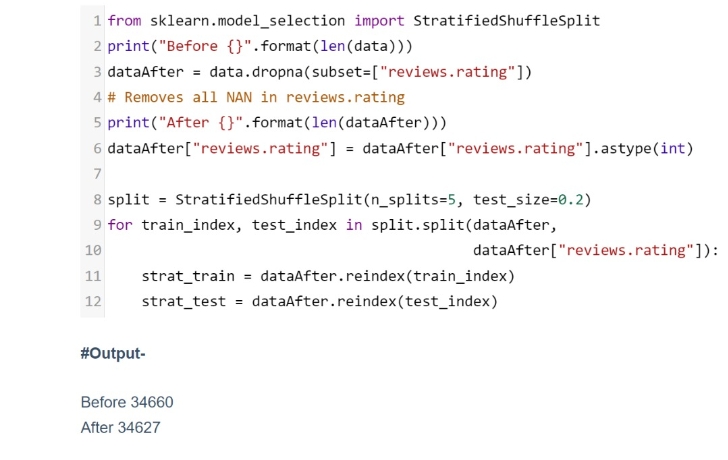
****

****

****

**Visualizing the distributions of numerical variables:**

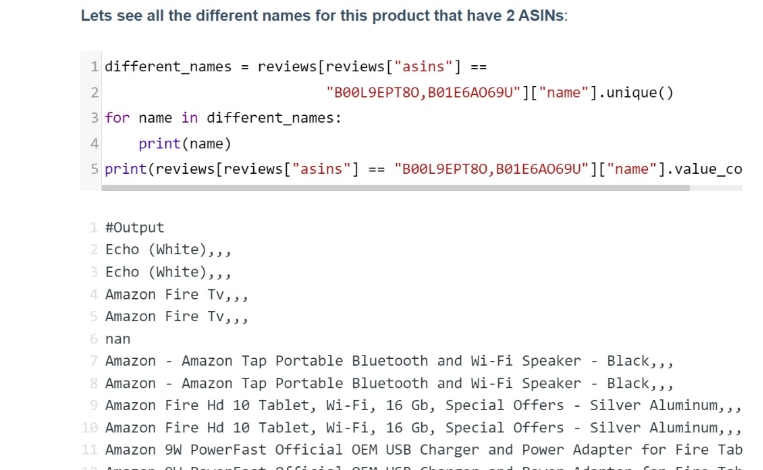
****

****

**Outliers in this case are valuable, so we may want to weight reviews that had more than 50+ people who find them helpful.**

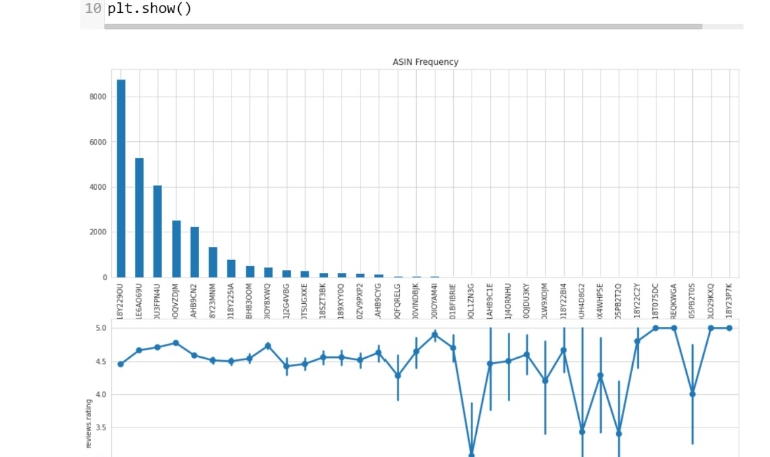
**Majority of examples were rated highly (looking at rating distribution). There is twice amount of 5 star ratings than the others ratings combined.**

**Split the data into Train and Test:**

****

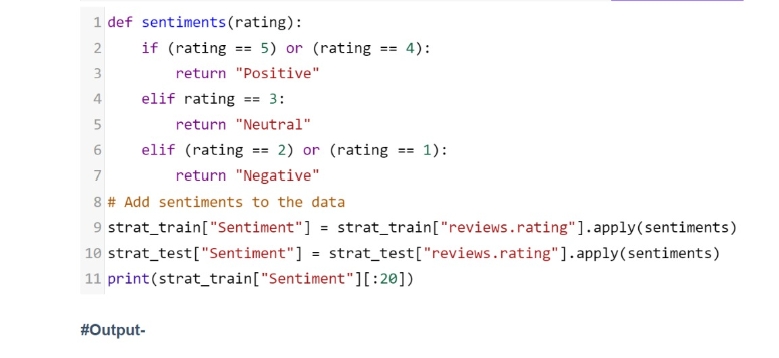
**Data Exploration (Training Set):**

***We will use regular expressions to clean out any unfavorable characters in the dataset, and then preview what the data looks like after cleaning.***

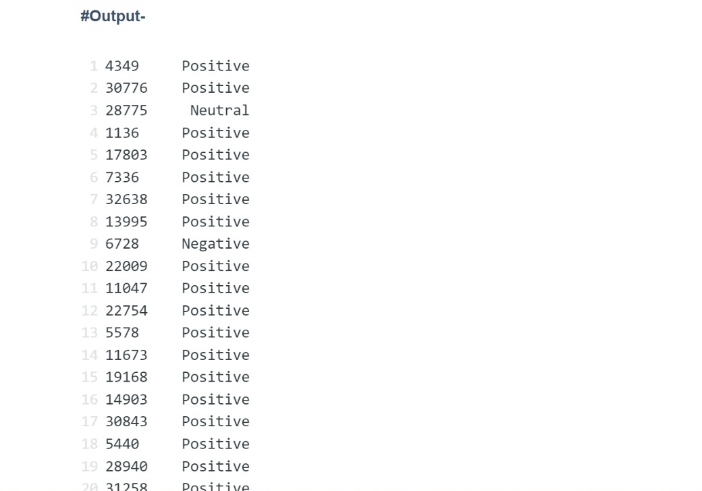
******

***Sentimental Analysis:***

Using the features in place, we will build a classifier that can determine a review’s sentiment.



**Output:**

****

**Team Members**

**I.JaharaAsmi**

**M.Muthulakshmi**

**K.Sandhiya**

**M.Shajitha Barveen**

**M.SathiyaJothi**

**JP College Of Engineering**