**Some of the information from from the data are :**

1) U.S. is the top user of wine with 38573 reviews and France is the second highest with 15948 reviews and Italy is the third highest user with review counts 9112.

The reviews counts shows that America and European countries are active user of wine.

2) Most of the reviews are positive with 76594 as positive and 6063 as negative reviews. It show that users like using wine a lot.

3) The cost of the price largely lies between 4 units to 400 units. The lowest price of wine is 4 units and the highest price of wine is 3300 units.

4) Pinot Noir is the most popular variety of wine with 10587 counts, Chardonnay is the second most popular with counts as 9403 and Cabernet Sauvignon is the third most popular variety with counts as 7552. Other popular varieties are Red Blend, Bordeaux-style Red Blend, Riesling and Sauvignon Blanc.

5) Their are a total of 28 varieties of wine and 38 countries are active user of it.

Short Description on Model and Features used :

I have used ‘review\_description’ as the feature for the model. I first applied Count Vectorizer on the selected feature and directly applied the Machine Learning models.

I tried three Machine Learning Models which are Logistic Regression, Naive Bayes and Random Forest Classifiers.

Logistic Regression comes out to be the better model with almost 70% accuracy. Naive Bayes gave 60% accuracy and Random Forest gave 54% accuracy.