

Development phase - 2

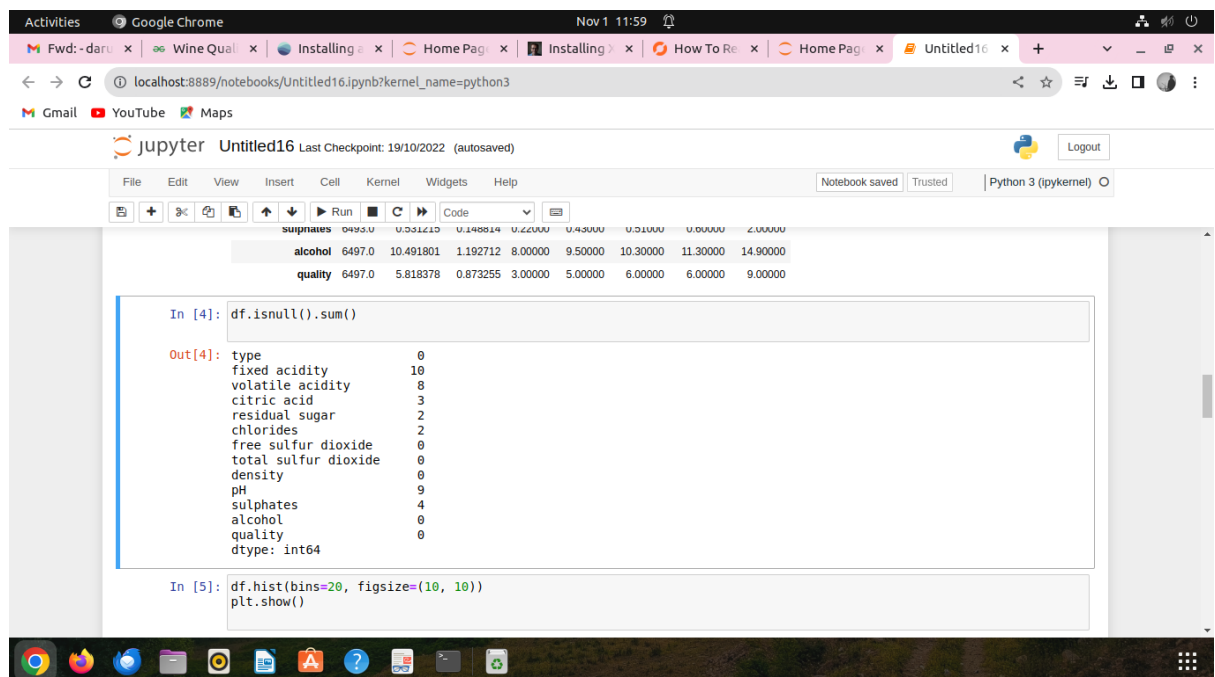
Machine learning model deployment using IBM Watson Studio

Wine quality prediction

Exploratory Data Analysis

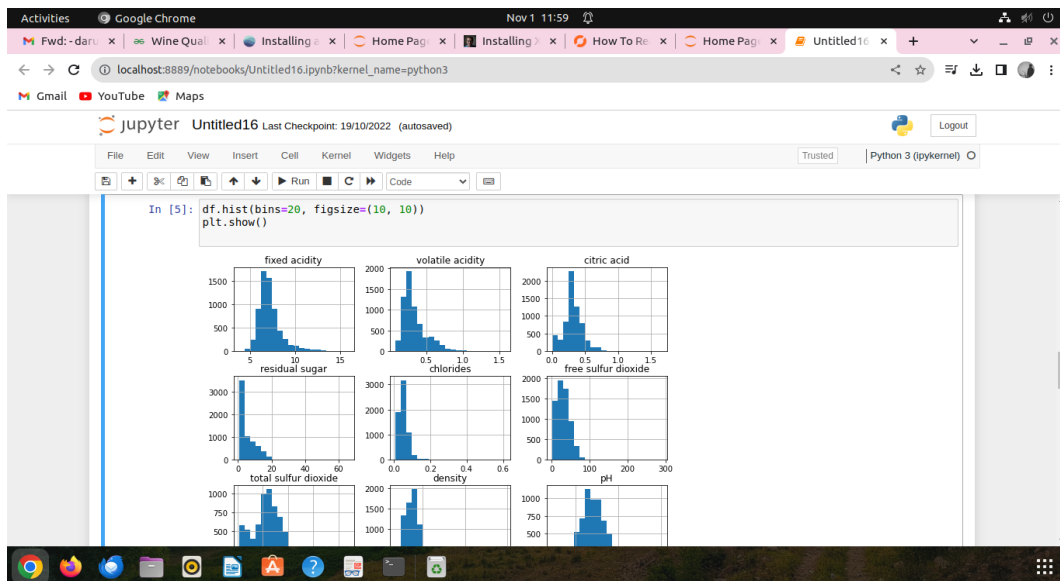
[EDA](#) is an approach to analysing the data using visual techniques. It is used to discover trends, and patterns, or to check assumptions with the help of statistical summaries and graphical representations. Now let's check the number of null values in the dataset columns.

```
df.isnull().sum()
```



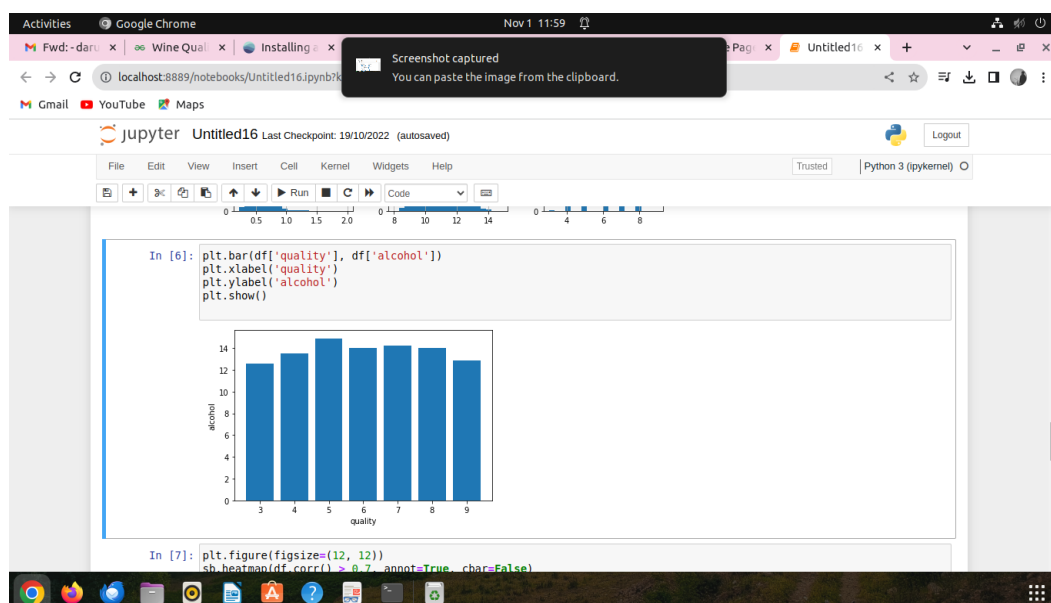
Let's draw the [histogram](#) to visualise the distribution of the data with continuous values in the columns of the dataset.

```
df.hist(bins=20, figsize=(10, 10))  
plt.show()
```



Now let's draw the count plot to visualise the number data for each quality of wine.

```
plt.bar(df['quality'], df['alcohol'])  
plt.xlabel('quality')  
plt.ylabel('alcohol')  
plt.show()
```

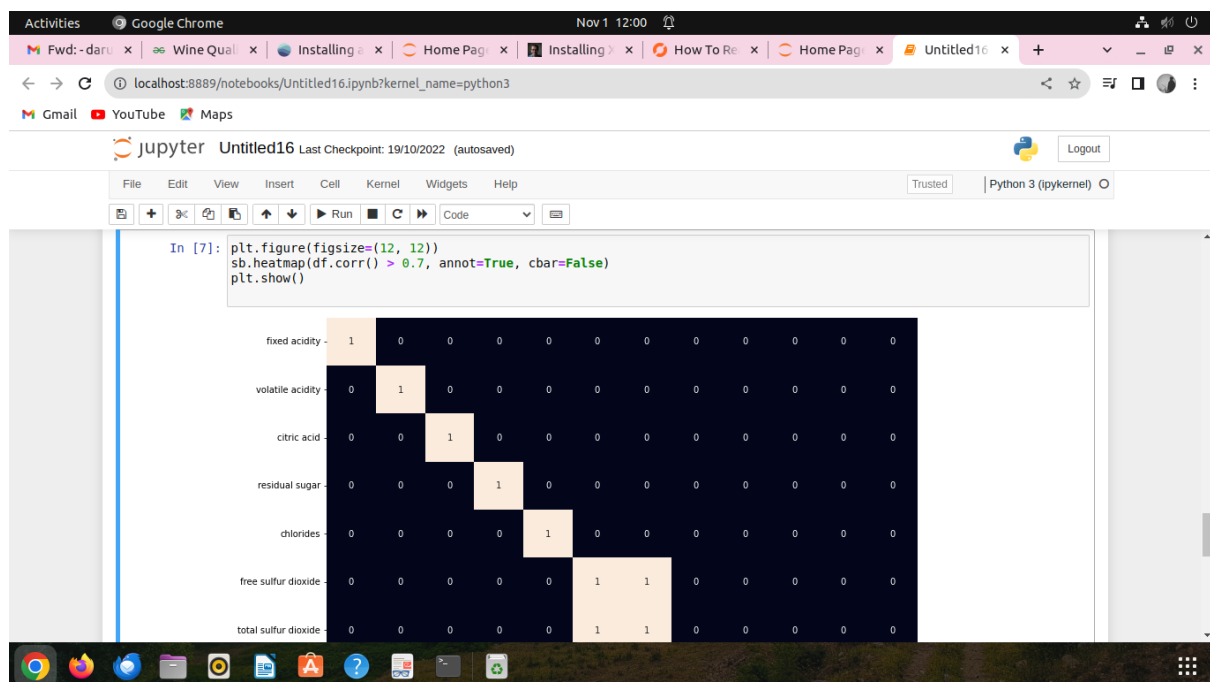


There are times the data provided to us contains redundant features they do not help with increasing the model's performance. That is why we remove them before using them to train our model.

```
plt.figure(figsize=(12, 12))
```

```
sb.heatmap(df.corr() > 0.7, annot=True, cbar=False)
```

```
plt.show()
```



Summary

Wine quality prediction refers to the use of machine learning algorithms and statistical models to assess and predict the quality of wine based on various input features. These features can include factors such as acidity, sweetness, alcohol content, and pH levels. By analysing historical data of different wines along with their quality ratings, machine learning models can learn patterns and relationships within the data to predict the quality of new or unseen wines. This predictive analysis aids winemakers, retailers, and enthusiasts in making informed decisions about production, pricing, and selection, ultimately enhancing the overall wine industry.