Ahmedabad University

CSE523 Machine Learning

Weekly Report - 3

Section - 1

Submitted to faculty: Prof. Mehul Raval

Date of Submission: 25-02-2023

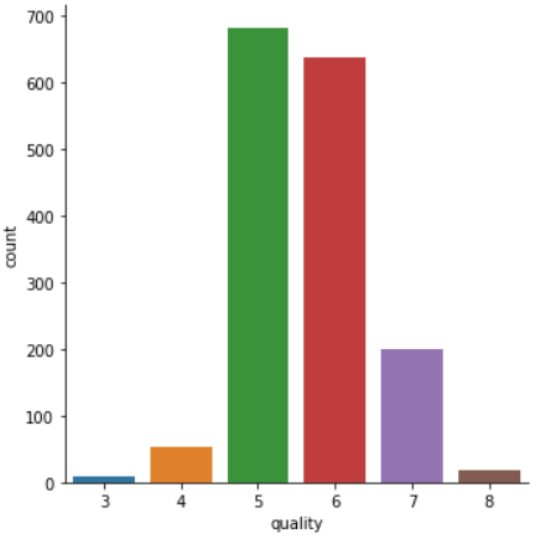
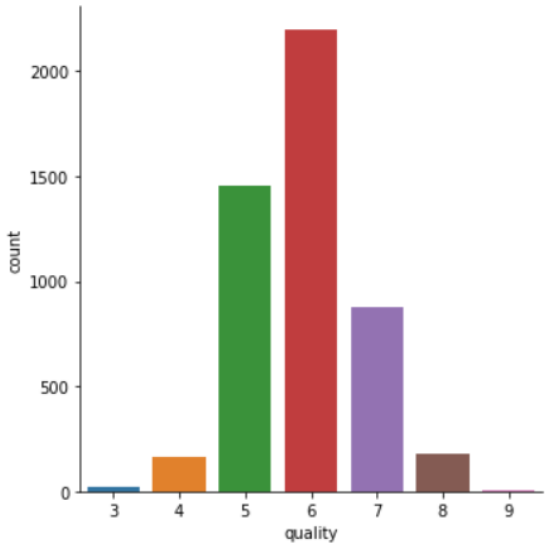
Roll No.	Name of the Student	Name of the Program
AU2040111	Kenil Shah	B. Tech CSE
AU2040215	Yesha Dhivar	B. Tech CSE
AU204087	Anshi Shah	B. Tech CSE
AU2040070	Rahi Shah	B. Tech CSE

2022-2023 (Winter Semester)

Tasks Performed: Data Analysis and Visualization

Outcomes:

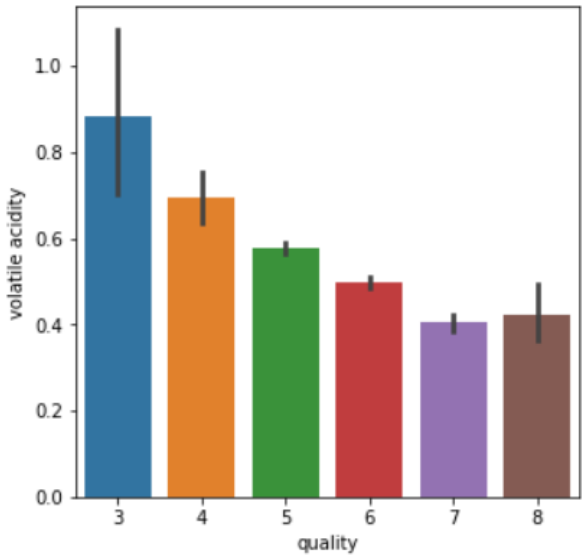
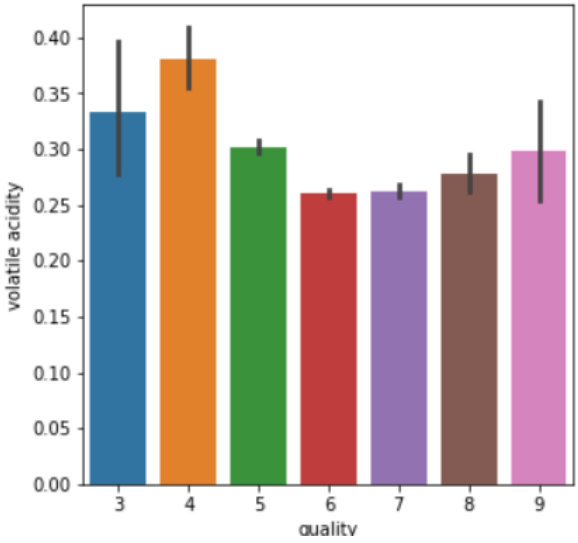
After completing the statistical measures, we got the total number of values for each quality and its spread. We got the following results for it.

Number of Values for each Quality	
Red wine data set	White wine data set
	

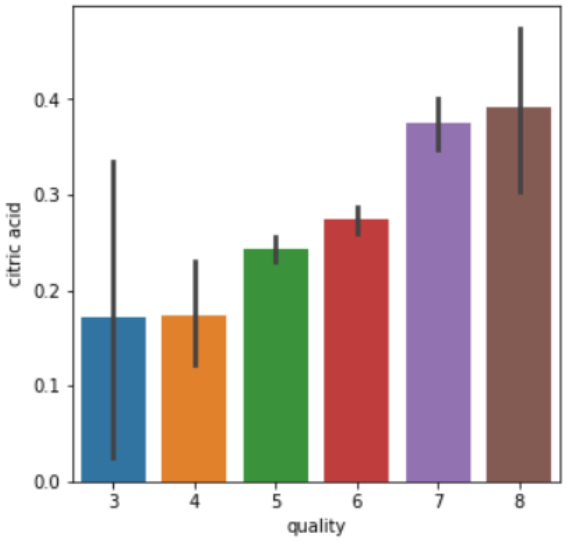
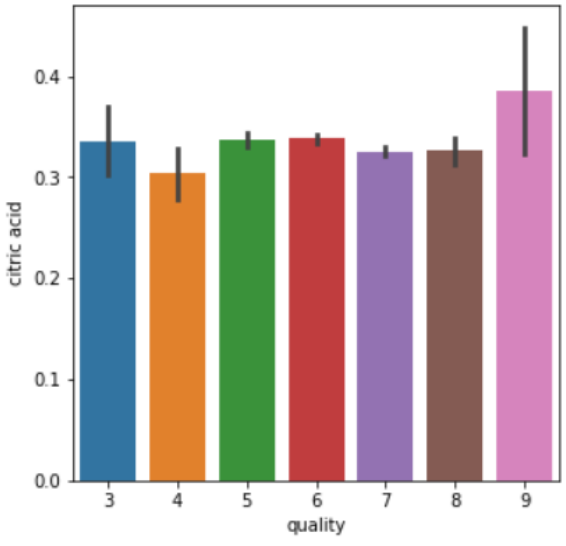
Next, we are comparing parameter values with the quality to see the relation of each parameter with the quality of the wine. We took first the ‘volatile acidity’ column (at random).
The comparison with red wine tells us that volatile acidity is almost inversely proportional to the quality of the wine. But, when comparing the white wine, we cannot exactly comment on how it is proportional to the data.

Here are the results–

…

Volatile acidity vs Quality (Red Wine)	
Red wine data set	White wine data set
	

Next we compared citric acid value to the quality and got the following plots:

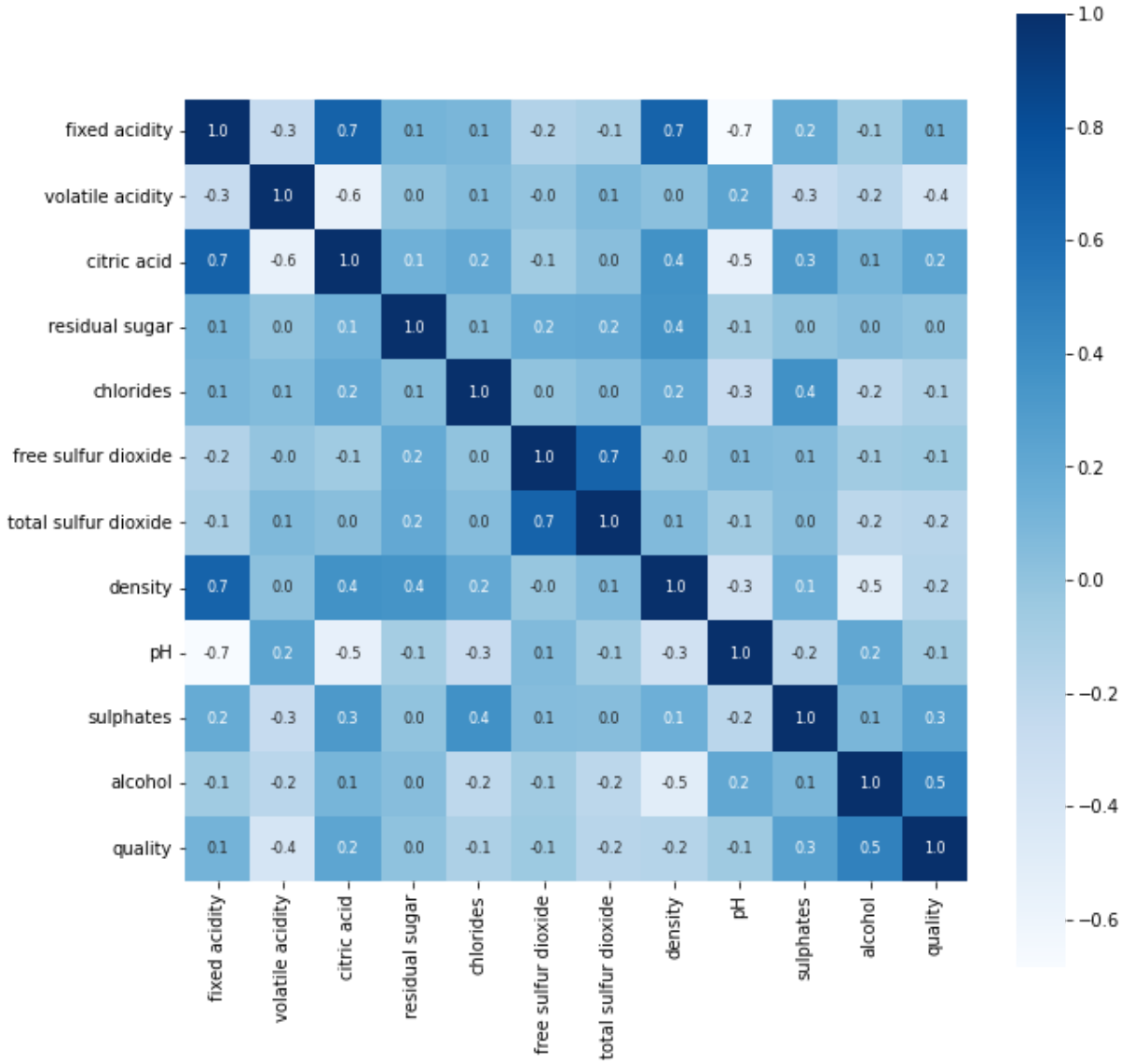
Citric Acid vs Quality (Red Wine)	
Red wine data set	White wine data set
	

Here we see a proportionality between the citric acid value and the quality of red wine.

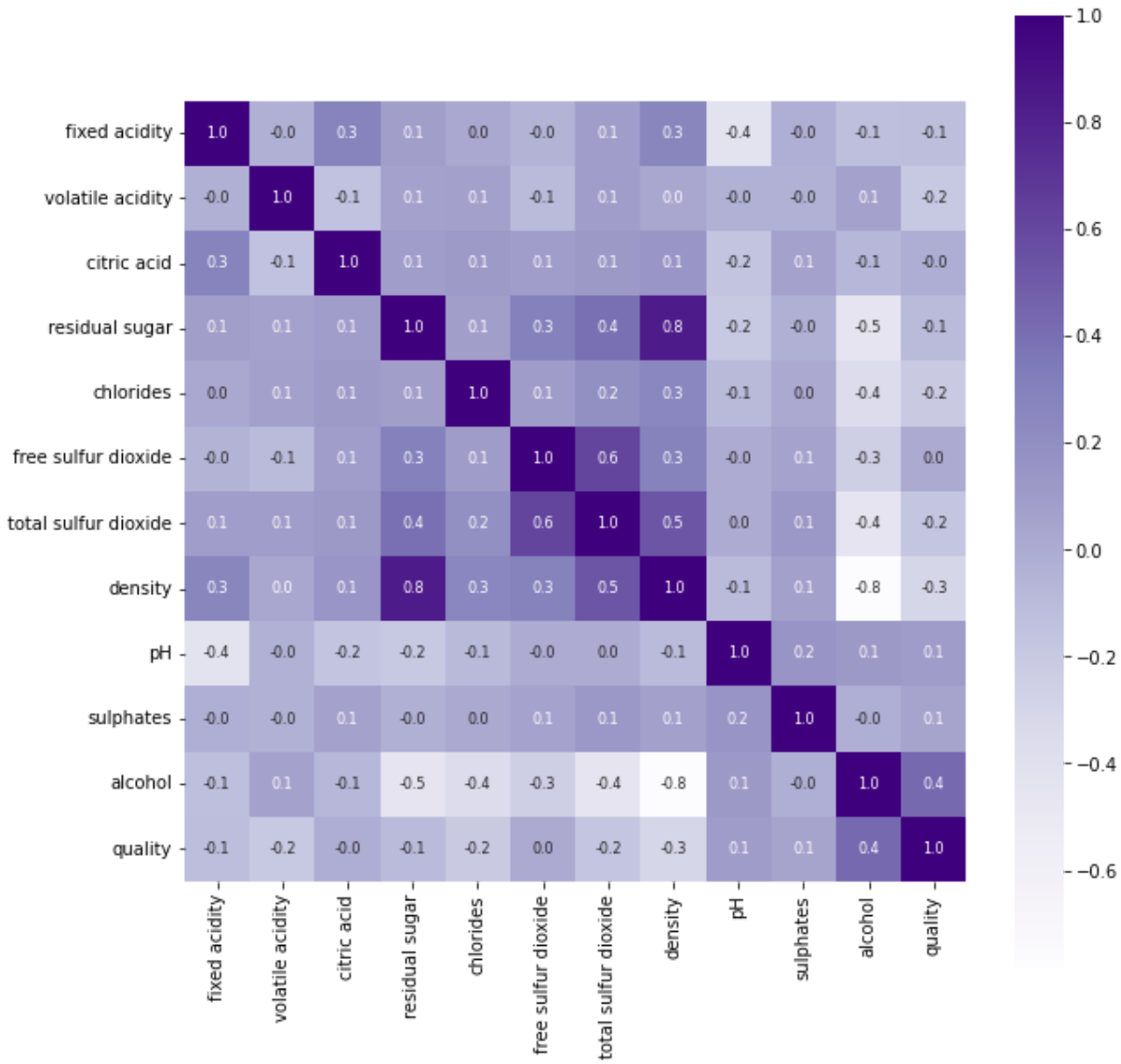
Similarly, we tried to know the relations between all the parameters and the quality of the wine to get an idea about the parameters and their impact on the quality of the wine.

Correlation Matrix

Correlation Matrix for Red Wine



Correlation Matrix for White Wine



By these correlation matrices, we got to know which parameters play an important role in deciding the wine quality of the dataset.