Brief of the Project

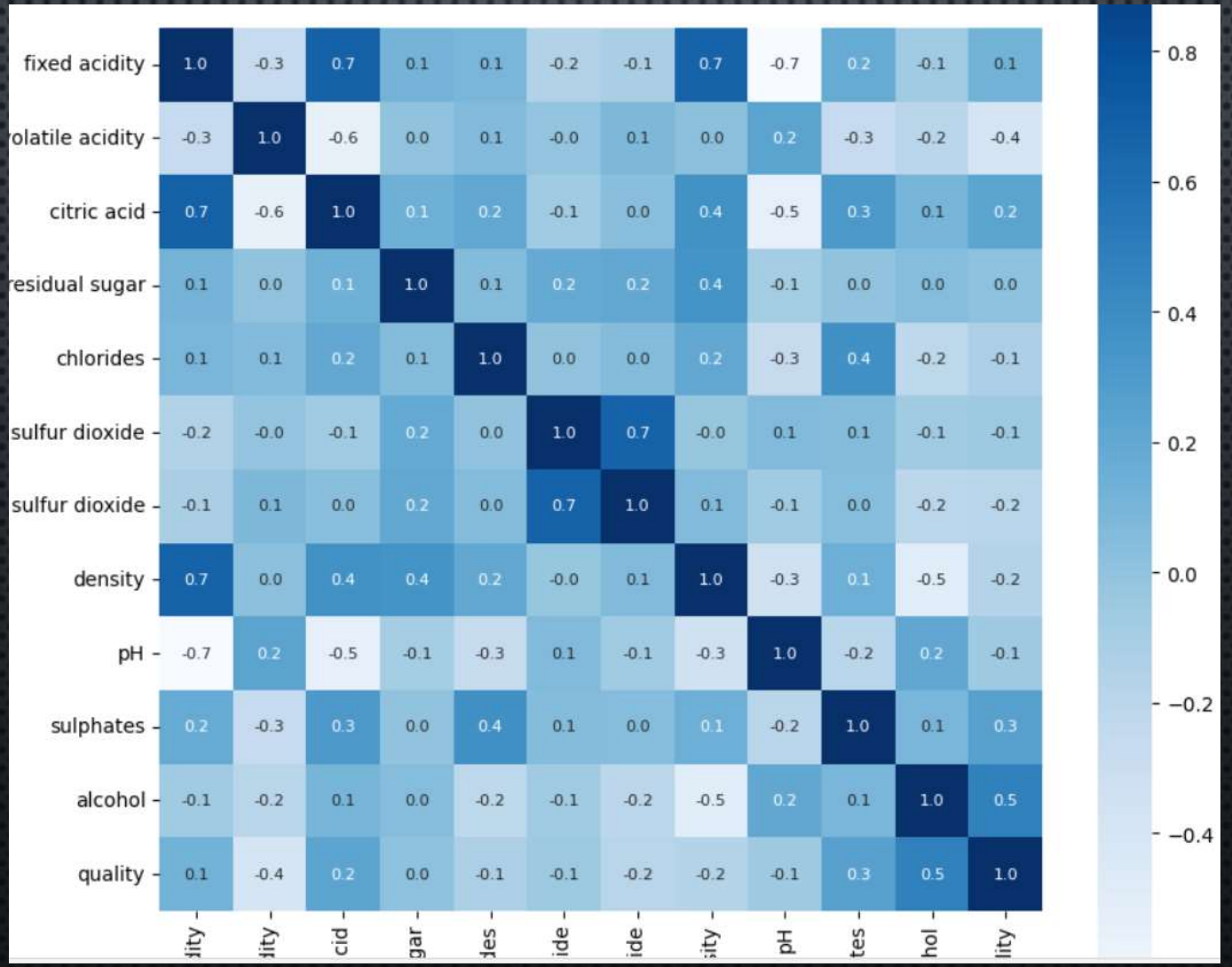
Good Quality

Bad Quality

Importing the dependencies

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score
```

Correlation between Parameters



Splitting the dataset into training and test
In which 80 percent is training and 20 percent is test dataset

Train and test split

```
6]: X_train, X_test, Y_train, Y_test = train_test_split(X,Y,test_size=0.2,random_state=2)

7]: print(Y.shape,Y_train.shape,Y_test.shape)
```

```
(1599,) (1279,) (320,)
```

Training the model with RFC

Model Training: Random Forest Classifier

```
model = RandomForestClassifier()
```

```
model.fit(X_train,Y_train)
```

```
RandomForestClassifier()
```

Accuracy of the Model

Model Evaluation

```
# accuracy on test data
X_test_prediction  = model.predict(X_test)
test_data_accuracy = accuracy_score(X_test_prediction,Y_test)
```

```
print('Accuracy : ',test_data_accuracy)
```

```
Accuracy :  0.9875
```