WhiteWineAnalysis

October 24, 2018

```
In [9]: import pandas as pd
```

1 Load Data

```
In [8]: data = pd.read_csv('winequality-white.csv', sep=';')
        data.head()
```

Out[8]:

	fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates	alcohol	quality
0	7.0	0.27	0.36	20.7	0.045	45.0	170.0	1.0010	3.00	0.45	8.8	6
1	6.3	0.30	0.34	1.6	0.049	14.0	132.0	0.9940	3.30	0.49	9.5	6
2	8.1	0.28	0.40	6.9	0.050	30.0	97.0	0.9951	3.26	0.44	10.1	6
3	7.2	0.23	0.32	8.5	0.058	47.0	186.0	0.9956	3.19	0.40	9.9	6
4	7.2	0.23	0.32	8.5	0.058	47.0	186.0	0.9956	3.19	0.40	9.9	6

2 Explore Data

This section is for a short preliminary analysis of the data. It should help you get a feeling for the dataset and generate first insights or hypotheses on the data.

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

3 Select Features/Reduce Dimensionality

The dataset has 11 features so far. Regarding the small number of training examples, this might be too much. Also, not all of them might influence the target variable. Find a way to reduce the number of features or dimensionality with as little loss of information as possible.

In [ ]:

In [ ]:

In [ ]:

4 Train Model

In this section, we would like you to train a machine learning model. You can choose between a classification, regression or clustering technique.

In [ ]:

In [ ]:

In [ ]:

5 Results

Here, you can summarize, present and interpret your results.

In [ ]:

In [ ]:

In [ ]:

Looking forward to seeing your results!