QMBU 450 Burak EROZ

Advanced Data Analysis in Python 54390

Wine Quality, Linear Regression and 95% Prediction Intervals

I analysed a multivariate dataset which is about white wine and predicted regressed wine quality based on other dimensions of white wine which are;

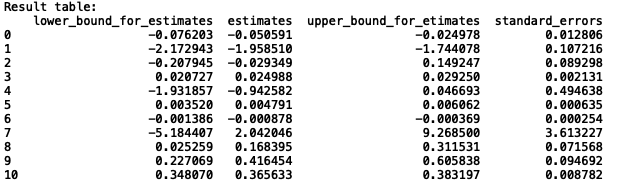
fixed acidity , volatile acidity, citric acid, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, pH, sulphates and alcohol.

I took the data from <https://archive.ics.uci.edu/ml/machine-learning-databases/wine-quality/>

A screenshot of a cell phone

Description automatically generated

﻿



Basically, when we interpret the data,

* **Fixed acid** is correlated with wine quality negatively and the beta is -0.05.
* **Volatile acidity** is negatively correlated with wine quality which has the beta of [-1.95851023]
* **Citric acid** is negatively correlated with wine quality which has the beta of [-0.02934924]
* **Residual sugar** is positively correlated with wine quality which has the beta of [ 0.0249884 ]
* **Chlorides** is negatively correlated with wine quality which has the beta of [-0.94258237]
* **Free sulfur dioxide** is positively correlated with wine quality which has the beta of [ 0.00479079]
* **Total sulfur dioxide** is negatively correlated with wine quality which has the beta of [-0.00087763]
* **Density** is positively correlated with wine quality which has the beta of [ 2.04204607]
* **pH** is positively correlated with wine quality which has the beta of [ 0.16839514]
* **Sulphates** is positively correlated with wine quality which has the beta of [ 0.41645356]
* **Alcohol** is positively correlated with wine quality which has the beta of [ 0.36563338]