Ahmedabad University

# CSE523 Machine Learning

## Weekly Report - 7

**Section - 1**

**Submitted to faculty: Prof. Mehul Raval**

**Date of Submission: 01-04-2023**


| Roll No. | Name of the Student | Name of the Program |
|---|---|---|
| AU2040087 | Anshi Shah | B. Tech CSE |
| AU2040215 | Yesha Dhivar | B. Tech CSE |
| AU2040111 | Kenil Shah | B. Tech CSE |
| AU2040070 | Rahi Shah | B. Tech CSE |


**2022-2023 (Winter Semester)**

**Tasks Performed:** Hypothesis Testing and ANOVA

## Outcomes:

```
Red Wine

                            OLS Regression Results
==============================================================================
Dep. Variable:     total_sulfur_dioxide   R-squared:                       0.054
Model:                              OLS   Adj. R-squared:                  0.053
Method:                   Least Squares   F-statistic:                     45.71
Date:                  Sat, 01 Apr 2023   Prob (F-statistic):           4.97e-20
Time:                          14:53:53   Log-Likelihood:                -7809.7
No. Observations:                  1599   AIC:                          1.563e+04
Df Residuals:                      1596   BIC:                          1.564e+04
Df Model:                             2
Covariance Type:              nonrobust
================================================================================================
                                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------------------------
Intercept                     33.4444      7.545      4.433      0.000      18.645      48.244
C(quality_mark)[T.low]        21.2007      7.636      2.776      0.006       6.223      36.178
C(quality_mark)[T.medium]      6.0346      7.626      0.791      0.429      -8.923      20.993
==============================================================================
Omnibus:                      483.807   Durbin-Watson:                   1.705
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1849.727
Skew:                           1.432   Prob(JB):                         0.00
Kurtosis:                       7.422   Cond. No.                         20.0
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Means for total sulfur dioxide by quality marks of wine
```

```
              total_sulfur_dioxide
quality_mark
high                     33.444444
low                      54.645161
medium                   39.479092
```

Standard deviation for total sulfur dioxide by quality marks of wine

```
              total_sulfur_dioxide
quality_mark
high                     25.433240
low                      36.720468
medium                   27.291245
```

```
 Multiple Comparison of Means - Tukey HSD, FWER=0.05
====================================================
group1 group2 meandiff p-adj   lower    upper   reject
----------------------------------------------------
  high    low  21.2007 0.0153   3.2875  39.1139   True
  high medium   6.0346 0.7084 -11.8551  23.9244  False
   low medium -15.1661    0.0 -18.9499 -11.3822   True
----------------------------------------------------
```

White Wine

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     total_sulfur_dioxide   R-squared:                       0.030
Model:                              OLS   Adj. R-squared:                  0.030
Method:                   Least Squares   F-statistic:                     76.66
Date:                  Sat, 01 Apr 2023   Prob (F-statistic):           1.65e-33
Time:                          14:53:53   Log-Likelihood:                -25239.
No. Observations:                  4898   AIC:                         5.048e+04
```

```
Df Residuals:                     4895   BIC:                         5.050e+04
Df Model:                            2
Covariance Type:              nonrobust
==================================================================================================
                                 coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------------------------
Intercept                    125.8833      3.120     40.350      0.000     119.767     132.000
C(quality_mark)[T.low]        22.7145      3.287      6.911      0.000      16.271      29.158
C(quality_mark)[T.medium]      7.7525      3.210      2.415      0.016       1.460      14.045
==============================================================================
Omnibus:                      124.307   Durbin-Watson:                   1.598
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              162.002
Skew:                           0.306   Prob(JB):                     6.63e-36
Kurtosis:                       3.648   Cond. No.                         11.3
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Means for total sulfur dioxide by quality marks of wine

              total_sulfur_dioxide
quality_mark
high                    125.883333
low                     148.597866
medium                  133.635802

Standard deviation for total sulfur dioxide by quality marks of wine

              total_sulfur_dioxide
quality_mark
high                     32.719653
low                      46.914579
medium                   39.400592
```

```
 Multiple Comparison of Means - Tukey HSD, FWER=0.05
====================================================
group1 group2 meandiff p-adj   lower   upper   reject
----------------------------------------------------
  high    low  22.7145    0.0 15.0095  30.4195   True
  high medium   7.7525 0.0417  0.2276  15.2774   True
   low medium -14.9621    0.0 -17.962 -11.9621   True
----------------------------------------------------
```

**To be performed next week:**

Model Building

(K-NN)