**DPA Assignment 2**

**Batch: 2018-20**

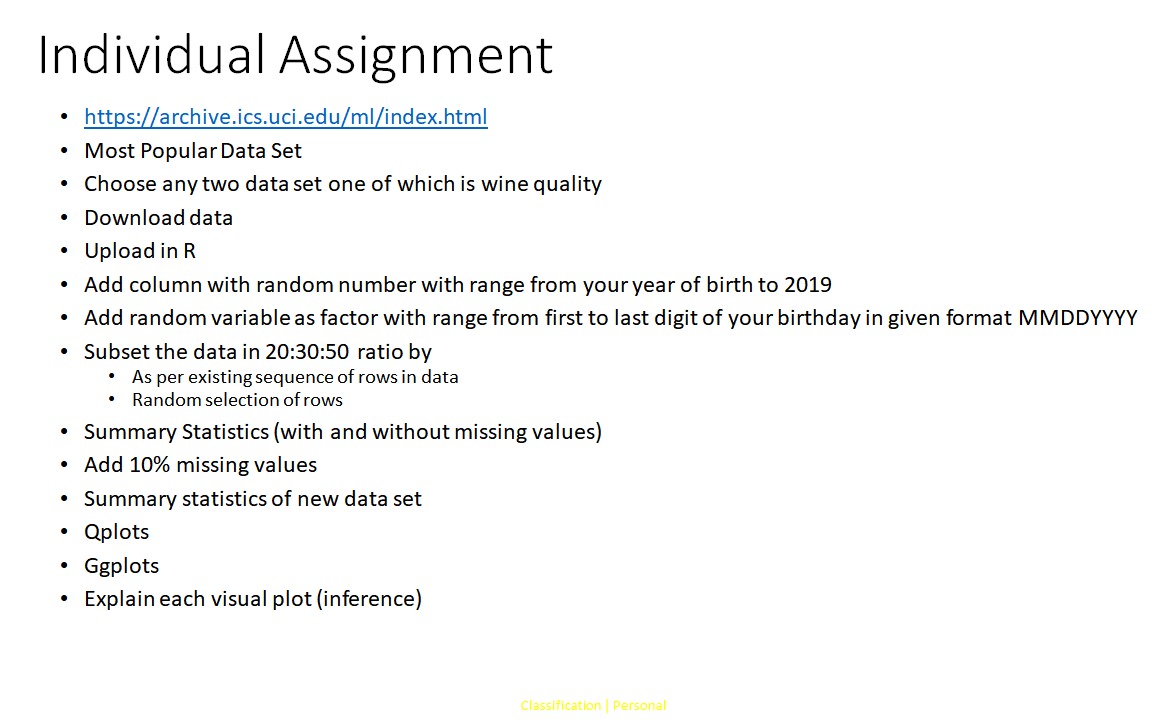
**Submitted to: Dr. Alok Yadav Submitted By:**

**Anmole Dewan-18PGDM014**

**Sanchit Taneja 18PGDM043**

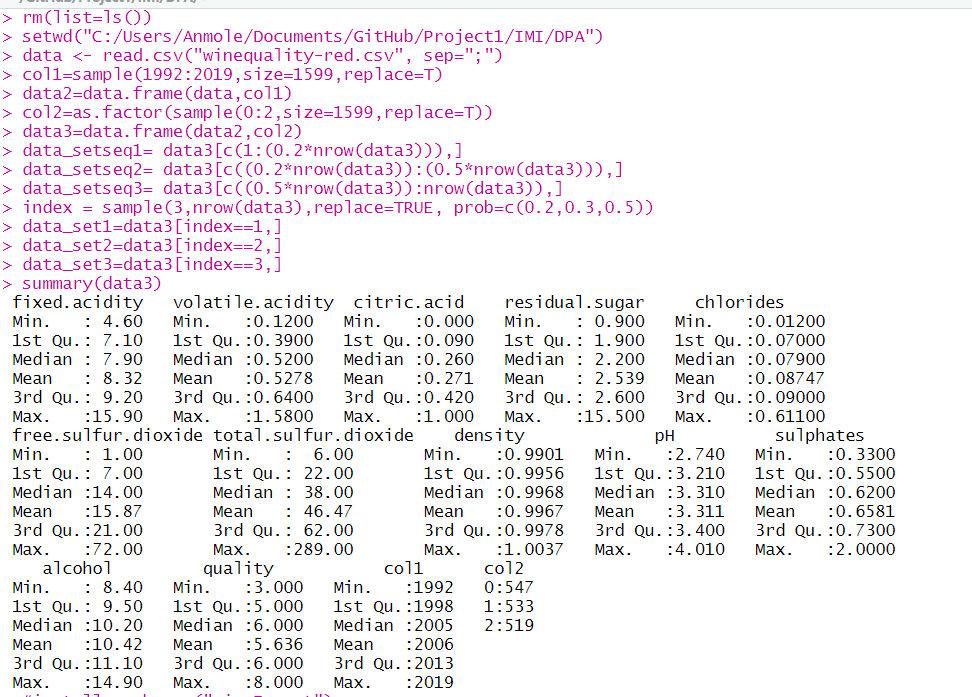
**Neha Singhal 18PGDM222**

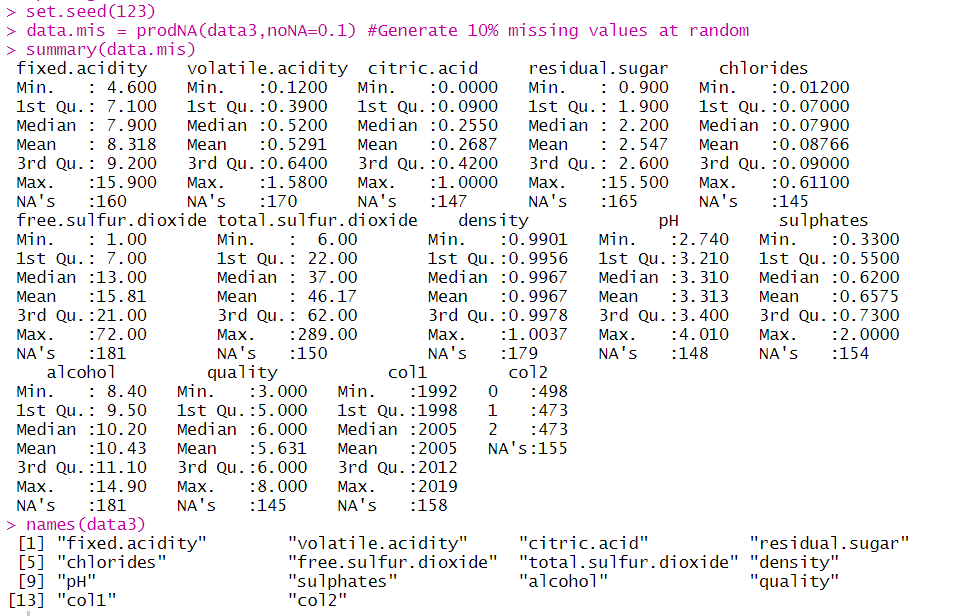
**Shivangi Sharma 18PGDM176**

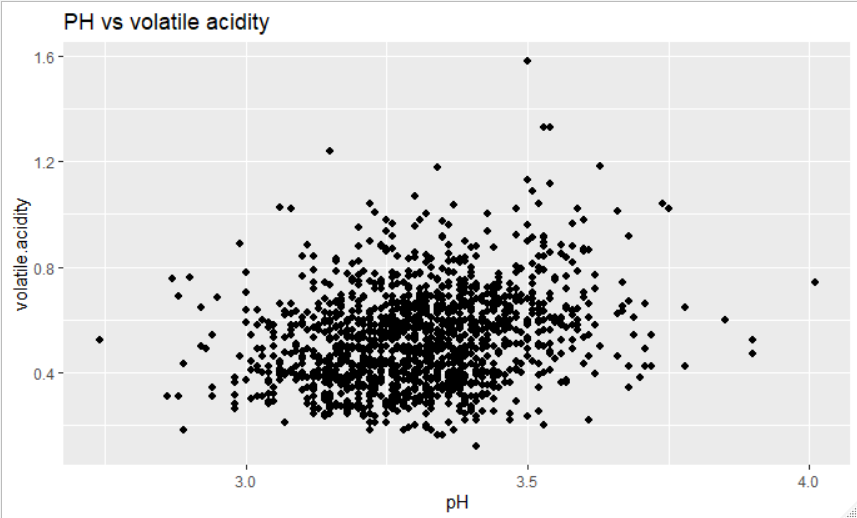


**Data Chosen:**

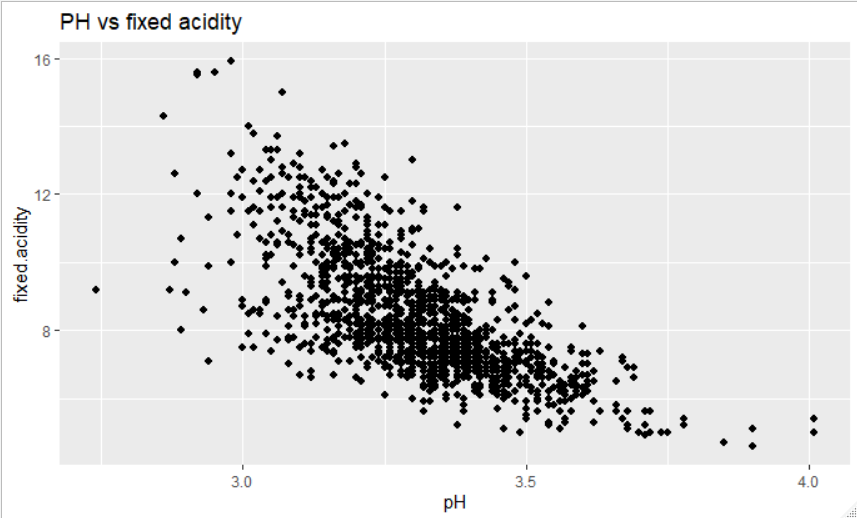
1. Wine Quality-Red
2. Bank-Full
3. **Wine Quality- Red**



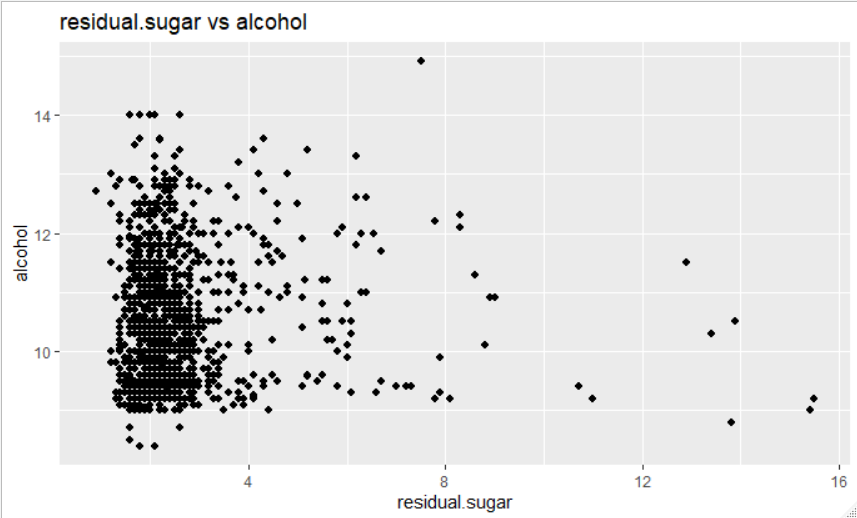




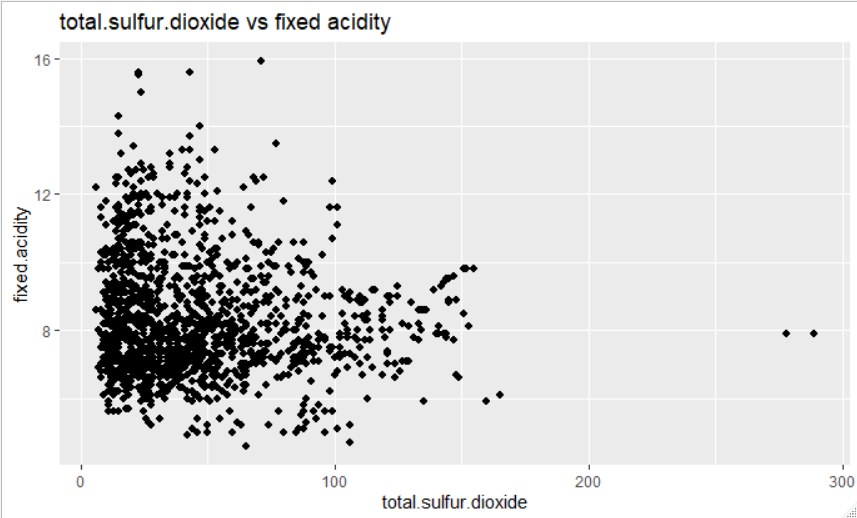
Volatile acidity is heterogeneously spread in the range 3.0 to 3.5 in most cases

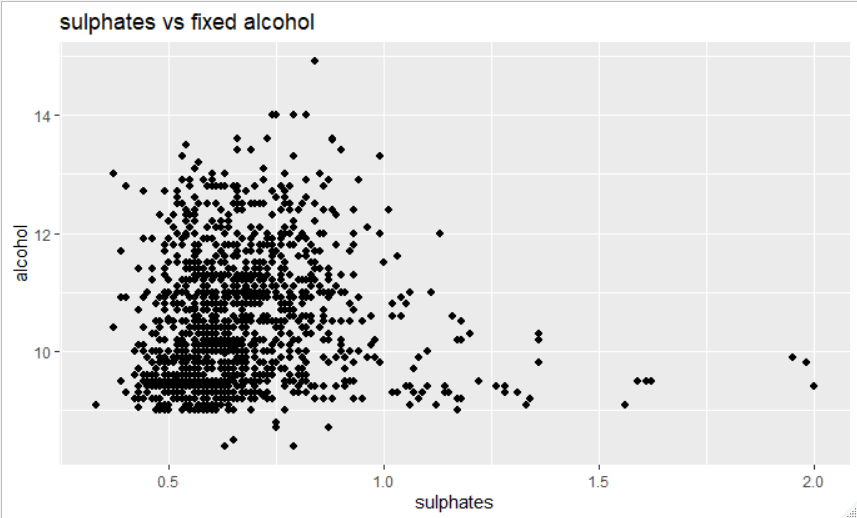


Fixed acidity is inversely proportional to PH

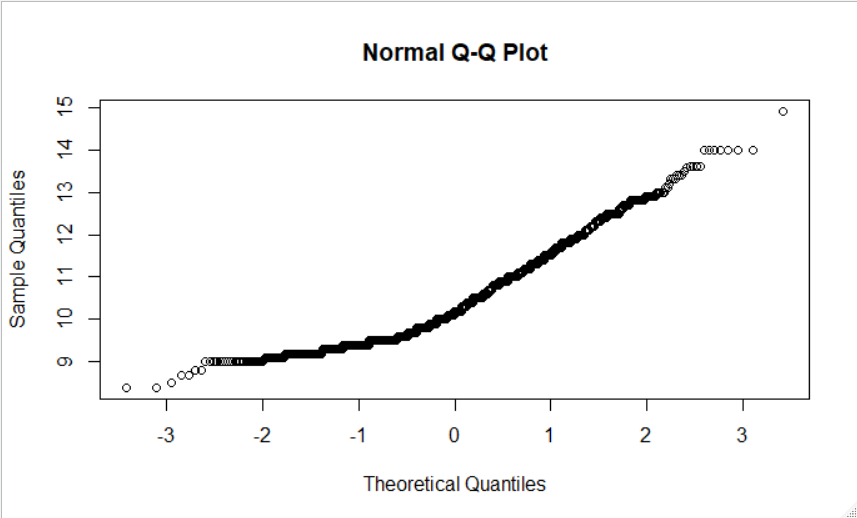


Residual sugar is spread majorly in the below 4 range for most alcohol ranges

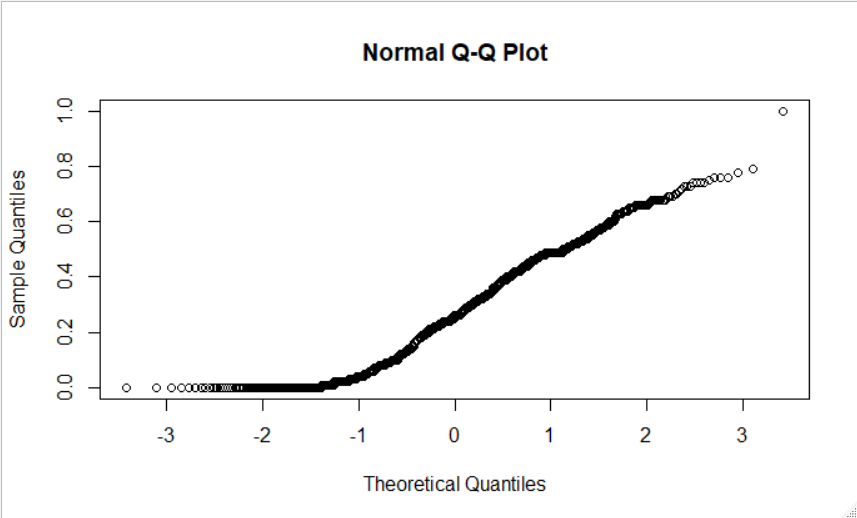




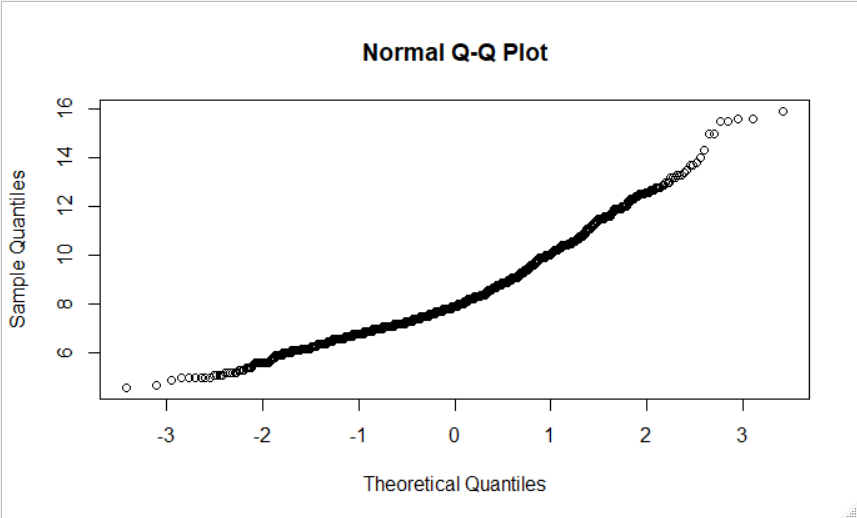
Alcohol- not that much normally distributed as curve can be seen



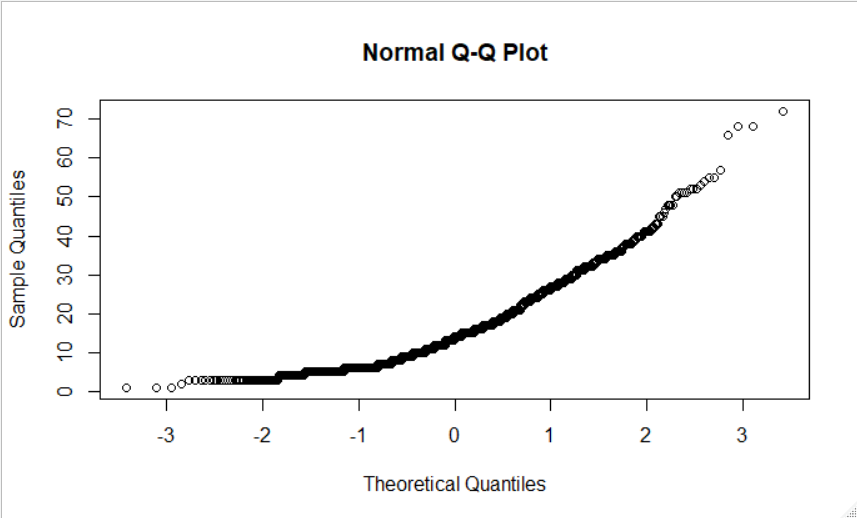
Citric Acid- Lower quartile is not normally distributed



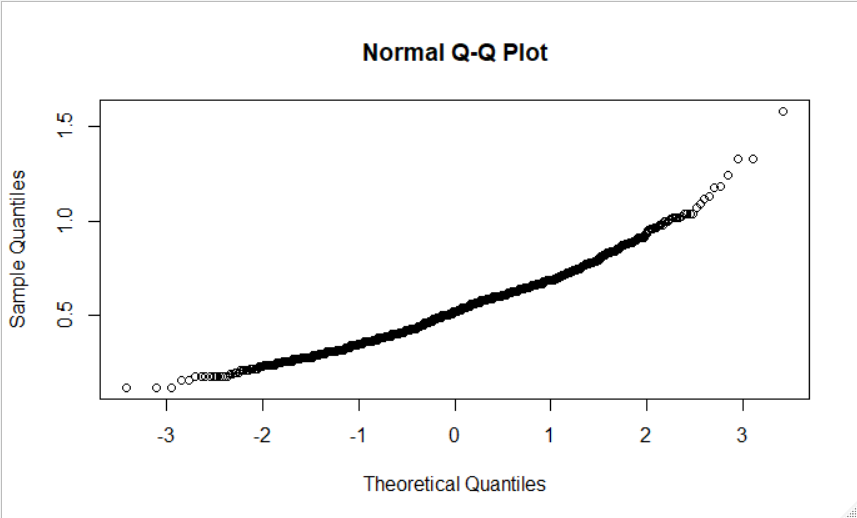
fixed.acidity



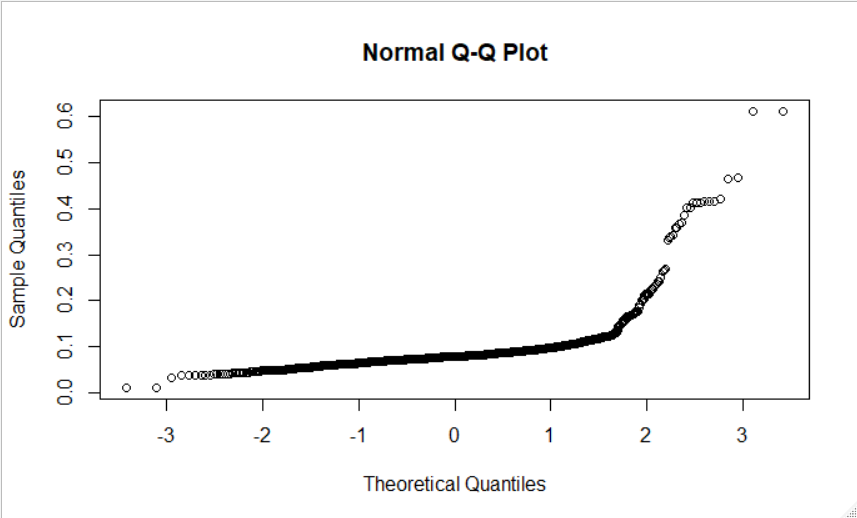
free.sulfur.dioxide



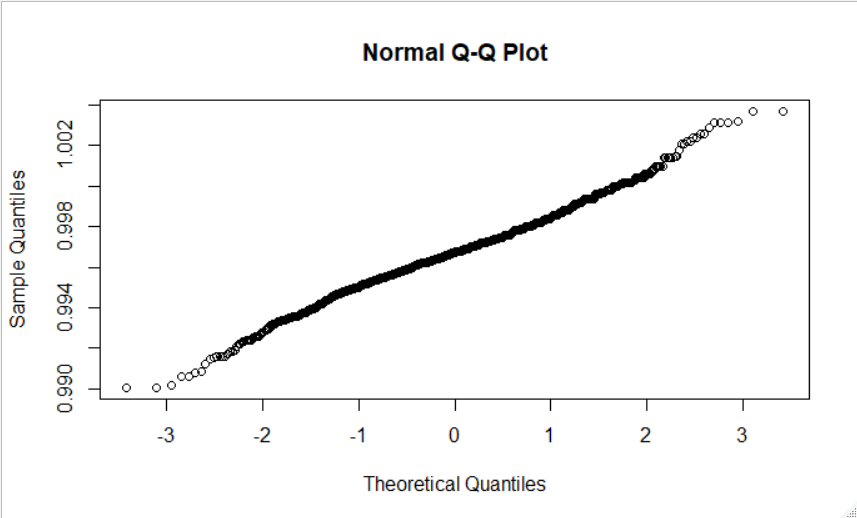
volatile.acidity



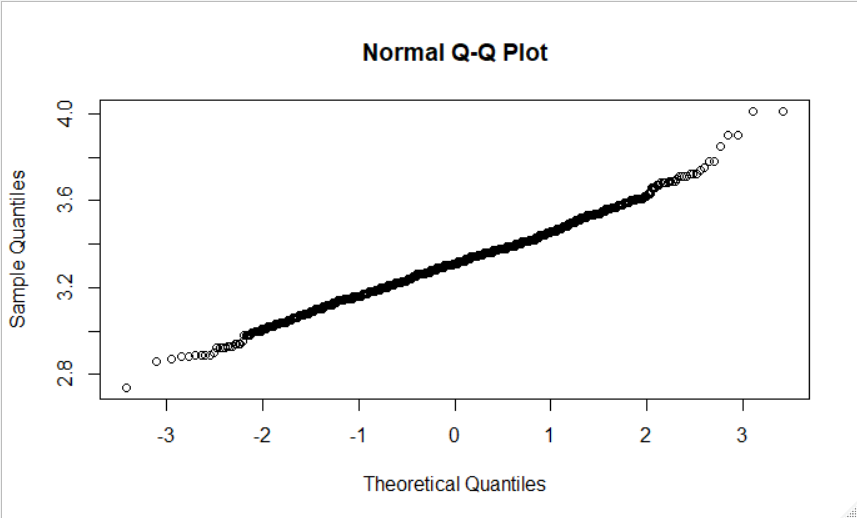
chlorides



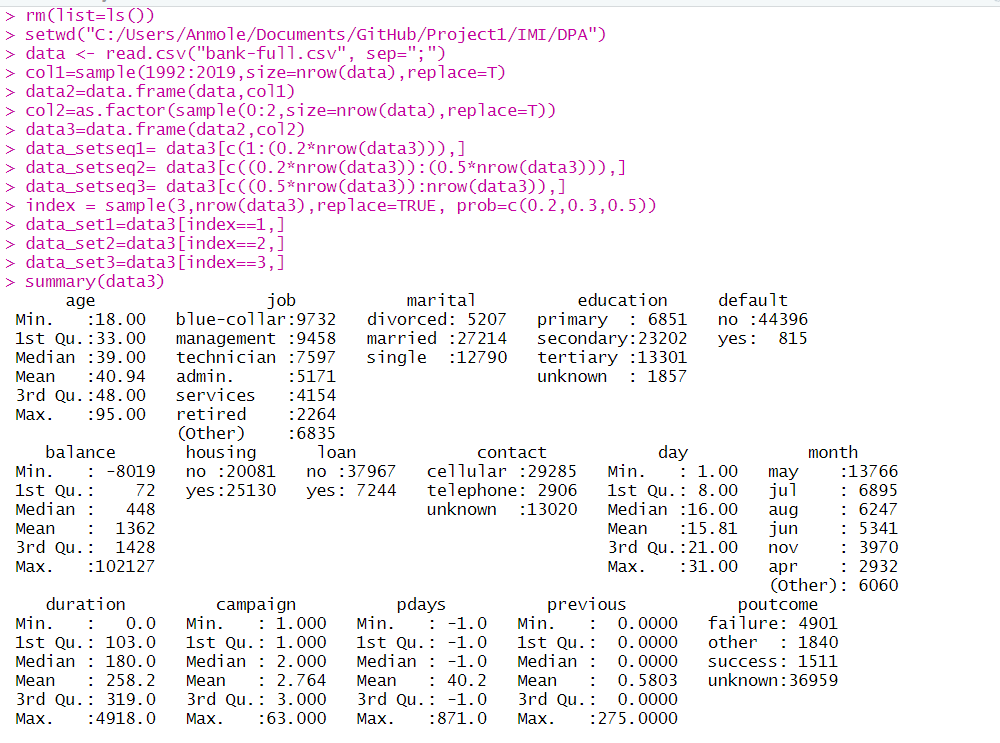
Density

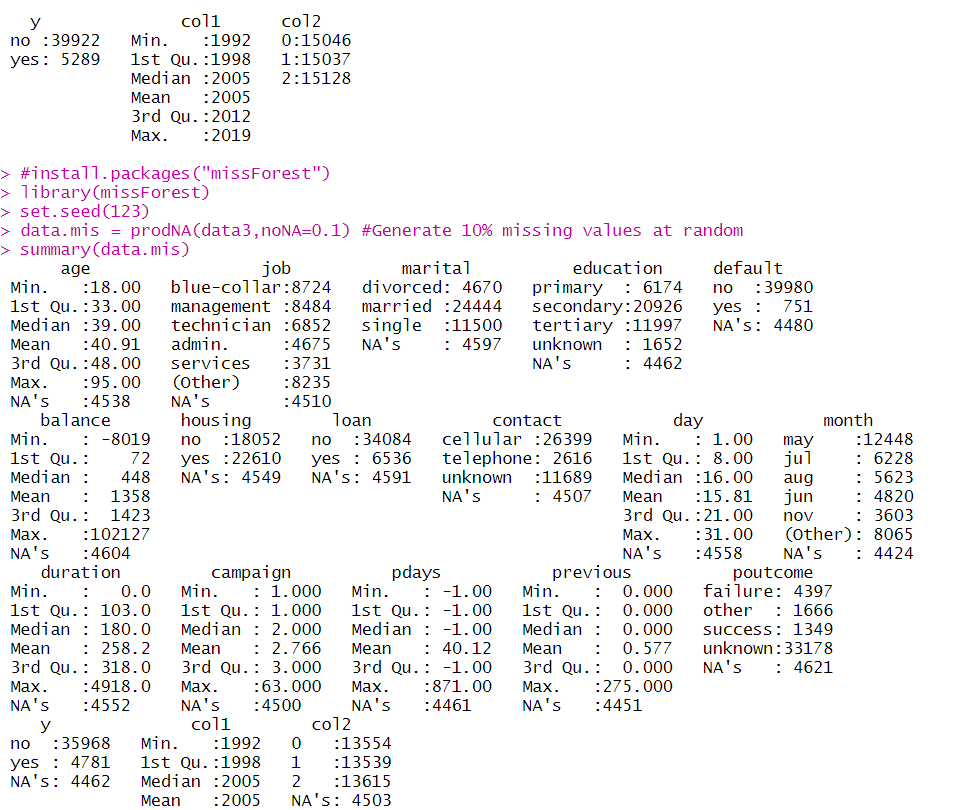


pH

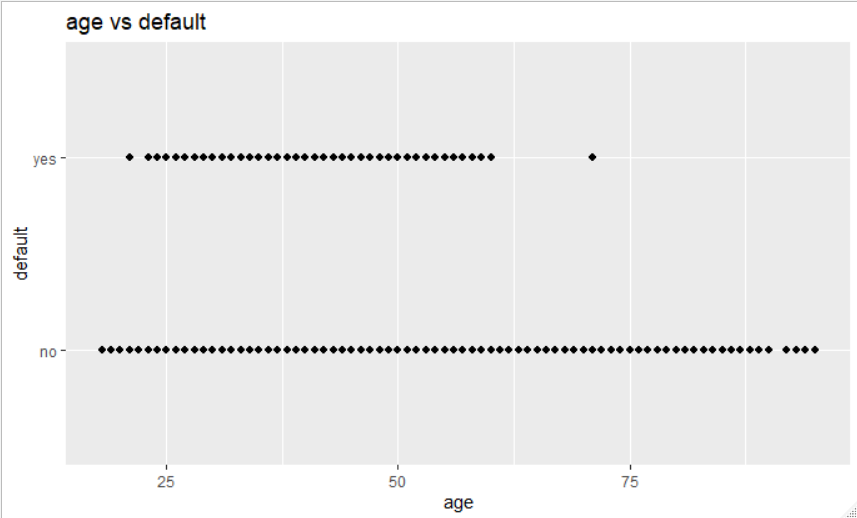


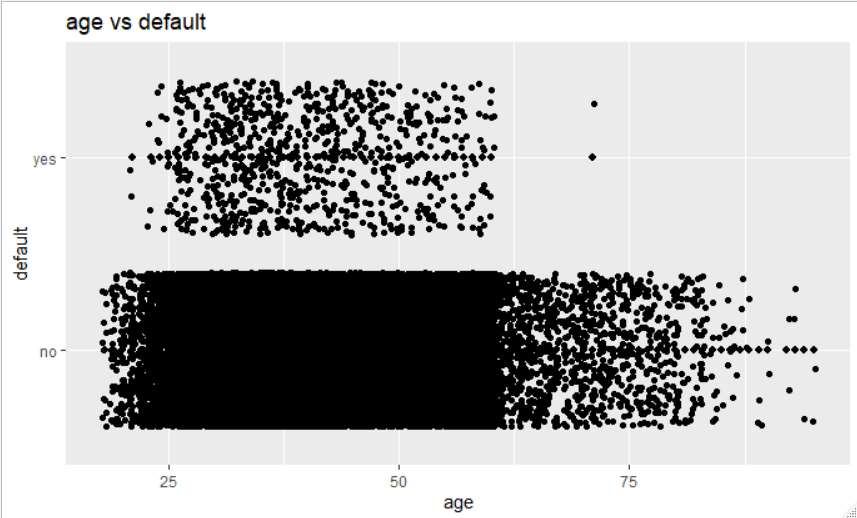
1. Bank Full

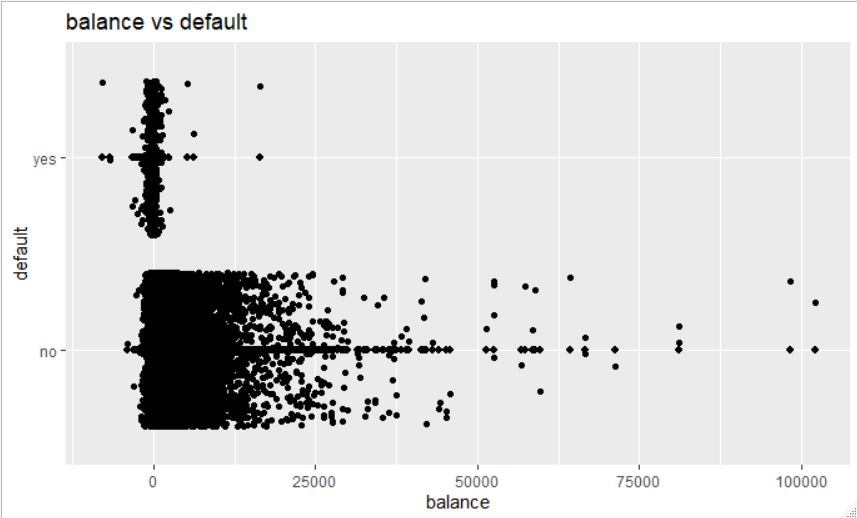


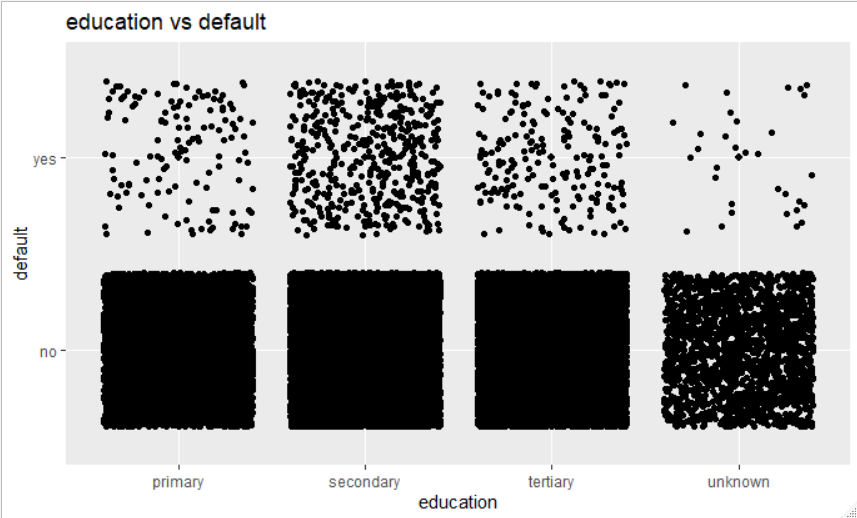


|  |
| --- |
| > names(data3)  [1] "age" "job" "marital" "education" "default" "balance" "housing"  [8] "loan" "contact" "day" "month" "duration" "campaign" "pdays"  [15] "previous" "poutcome" "y" "col1" "col2" |
|  |
| |  | | --- | | > | |

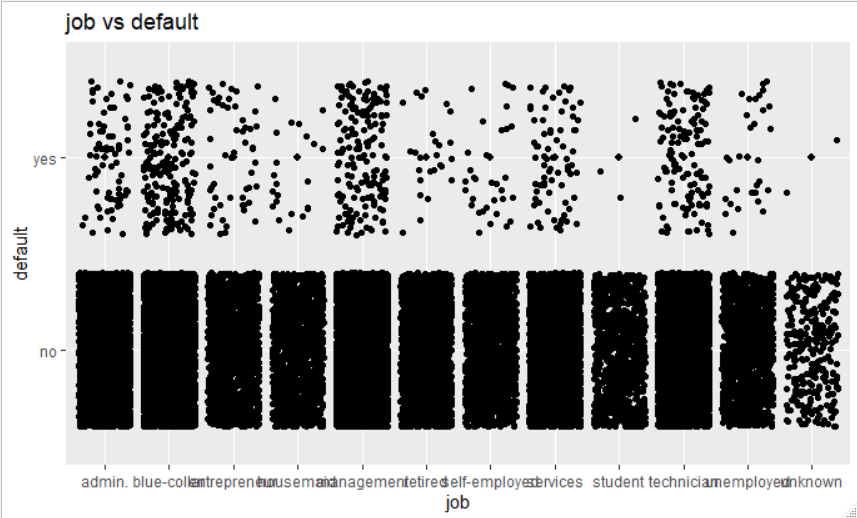


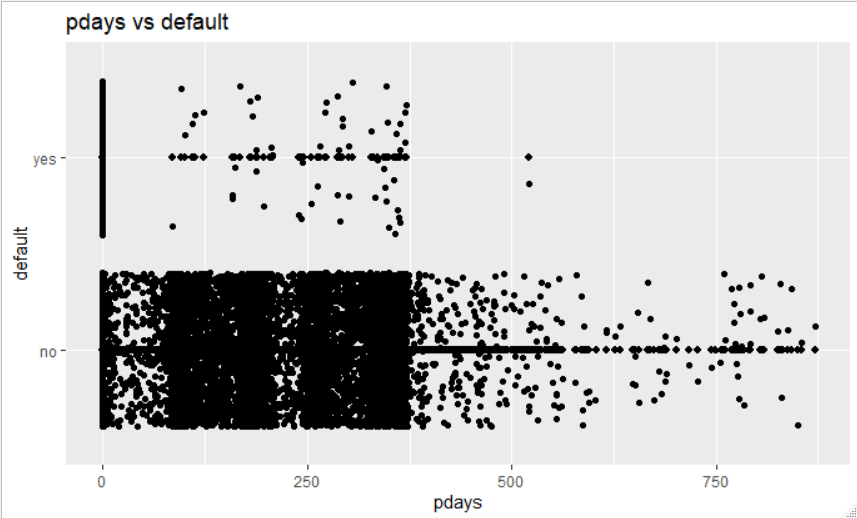


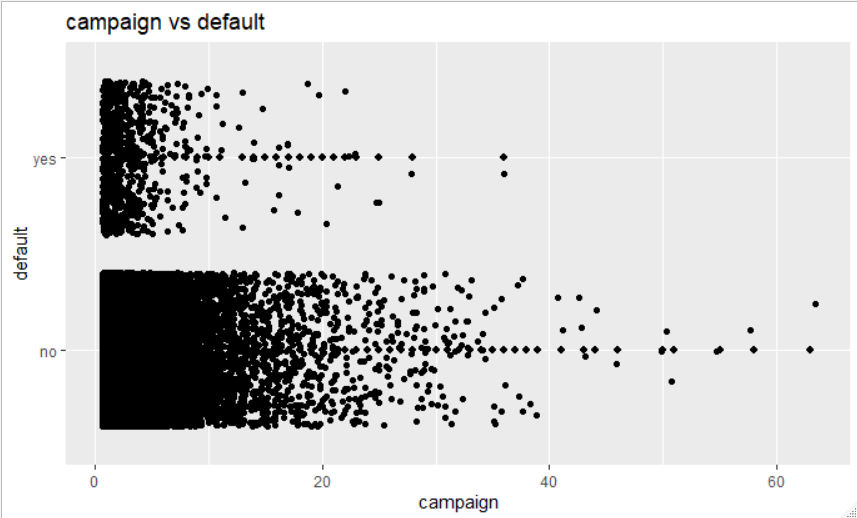


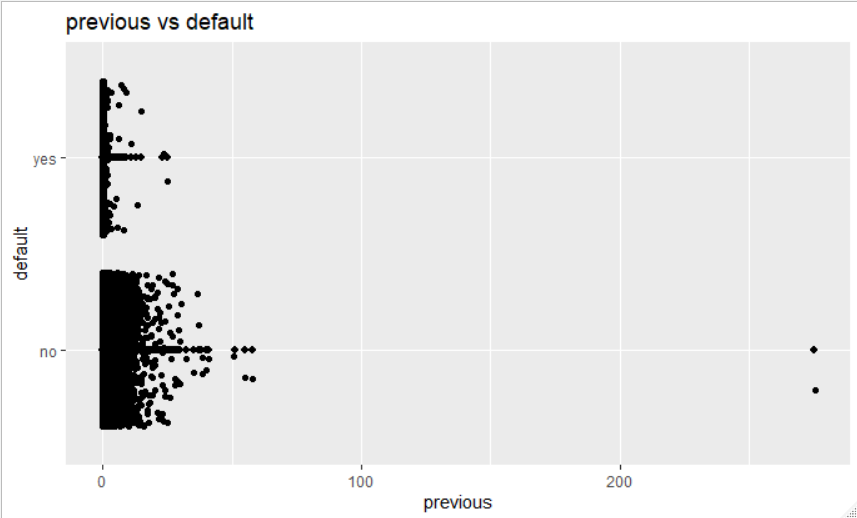


Acc to education, I can choose to merge Primary and secondary and keep others as Tertiary and unknown









**Age Quartile- 3 segments can be seen-**

1. In lower quartile
2. Above 50 to 75th percentile
3. Above 75th percentile

