Analyzing & Identifying Factors That Affects Student's Academic Performance.

- Group Members –

Rohit Buddiga

Ashwin Chudasama

Shivani Uday Jahagirdar

Needhi Devan Senthil Kumar

INTRODUCTION

COMPREHENDING DATA

DATA PREPARATION

RESULTS

CONCLUSION

OBJECTIVE

INTRODUCTION

For higher education institutions whose goal is to contribute to the improvement of quality of education.

Therefore, the prediction of students' success is crucial for education institutions.

Benefit of predicting student performance : Detecting at risk students, assurance of student retention, course and resource allocations, and many others.

Participants and datasets - In this study, the records of students were collected from the administration department.

Tools - To utilize from the provisioned dataset, multiple modifications have been created to prepare the dataset for analysis.

R studio (version 1.1.456) was used to visualize the dataset and select the key attributes.
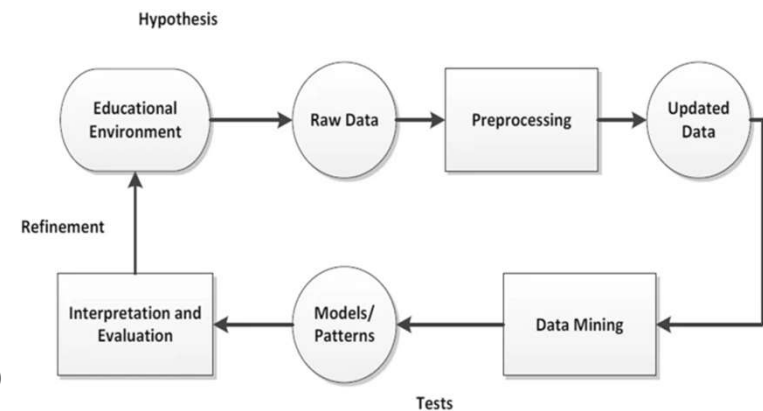
COMPREHENDING DATA

1) Dataset pre-processing phase –

- Initially, the datasets contained valueless attributes, missing instances, inadequate attributes .

- Therefore, the datasets were passed through the following preparation stages:

2) Dataset cleaning –

- Students with incomplete records, like those who had no grades' details in most of their courses or those who didn't have any course records were excluded from the list.

DATA PREPARATION

Data consolidation	→	Data Cleaning	→	Data transformation

- Data selection :

  ➢ Academic variables of interest identified.

- Missing values report :
  ➢ No reported values missing.
- Reported outliers :
  ➢ No reported outliers detected.

- Variable creations :
  ➢ Age
  ➢ Gender
  ➢ Free time
  ➢ Health
  ➢ Absences

Main dataset attributes

```
## 'data.frame':    100 obs. of  13 variables:
## $ school   : chr  "GP" "GP" "GP" "GP" ...
## $ sex      : chr  "M" "F" "F" "M" ...
## $ age      : int  16 16 16 16 16 16 16 15 15 16 ...
## $ Medu     : int  2 4 2 4 3 4 4 4 3 4 ...
## $ Fedu     : int  3 1 2 2 3 4 3 3 4 3 ...
## $ studytime: int  1 5 5 4 3 2 5 5 3 4 ...
## $ walc     : int  5 3 3 2 2 1 1 1 1 1 ...
## $ health   : int  5 5 5 5 5 5 5 5 5 5 ...
## $ absences : int  18 4 4 0 2 10 2 2 0 0 ...
## $ G1       : int  8 11 6 13 12 15 15 15 12 12 ...
## $ G2       : int  6 12 10 13 12 15 15 15 15 14 ...
## $ G3       : int  7 11 10 12 14 15 16 14 15 15 ...
## $ mean.G   : int  7 11 9 13 13 15 15 15 14 14 ...
```

• The Student performance is evaluated based on variables such as Grades, Health, Alcohol consumption, Parents education, Attendance, Study time.

DATA CONSOLIDATION

- DATA ATTRIBUTES USED :

 Grades, Health ,Alcohol consumption, Parents' education, Attendance, Study time , etc.

- Academic performance identified is dependent variable.

- Understanding Factors or variables affecting students' academic performance , which are interconnected and interdepended.

- Student Data summary

```
summary(student_data)
```

```
##     school              sex                 age            Medu
## Length:100         Length:100         Min.   :15.00   Min.   :1.00
## Class :character   Class :character   1st Qu.:15.00   1st Qu.:1.00
## Mode  :character   Mode  :character   Median :15.00   Median :2.00
##                                       Mean   :15.53   Mean   :2.45
##                                       3rd Qu.:16.00   3rd Qu.:4.00
##                                       Max.   :18.00   Max.   :4.00
##      Fedu         studytime          walc           health        absences
## Min.   :1.00   Min.   :1.00   Min.   :0.00   Min.   :1.00   Min.   : 0.00
## 1st Qu.:1.75   1st Qu.:2.00   1st Qu.:1.00   1st Qu.:2.00   1st Qu.: 4.00
## Median :3.00   Median :3.00   Median :3.00   Median :3.00   Median :22.00
## Mean   :2.55   Mean   :3.08   Mean   :3.21   Mean   :3.02   Mean   :20.18
## 3rd Qu.:4.00   3rd Qu.:4.00   3rd Qu.:5.00   3rd Qu.:4.00   3rd Qu.:34.00
## Max.   :4.00   Max.   :5.00   Max.   :7.00   Max.   :5.00   Max.   :46.00
##       G1             G2             G3            mean.G
## Min.   : 5.00   Min.   : 5.00   Min.   : 5.00   Min.   : 5.00
## 1st Qu.: 8.00   1st Qu.: 9.00   1st Qu.: 9.75   1st Qu.: 9.00
## Median :11.00   Median :11.00   Median :11.00   Median :11.00
## Mean   :10.81   Mean   :11.38   Mean   :11.49   Mean   :11.23
## 3rd Qu.:13.00   3rd Qu.:14.00   3rd Qu.:14.25   3rd Qu.:14.00
## Max.   :19.00   Max.   :19.00   Max.   :20.00   Max.   :19.00
```
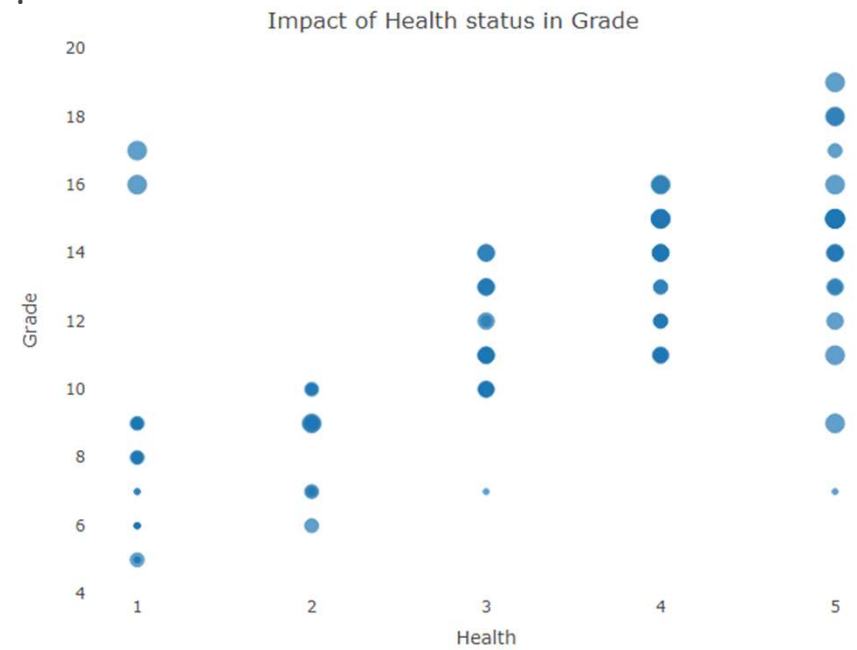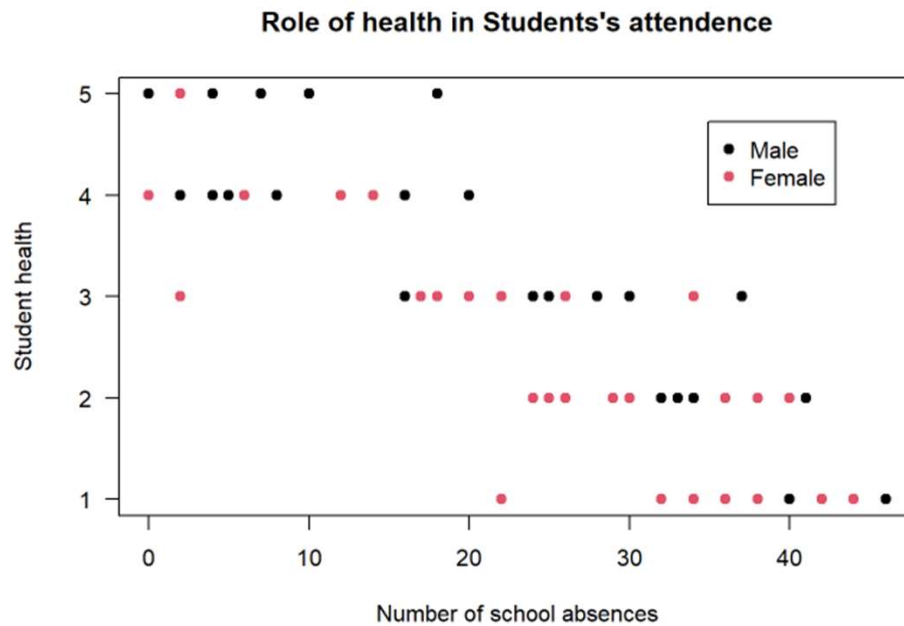
```
class(student_data)
```

```
## [1] "data.frame"
```
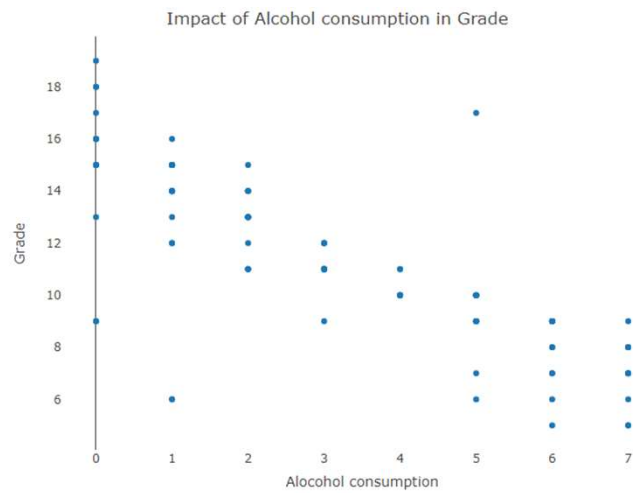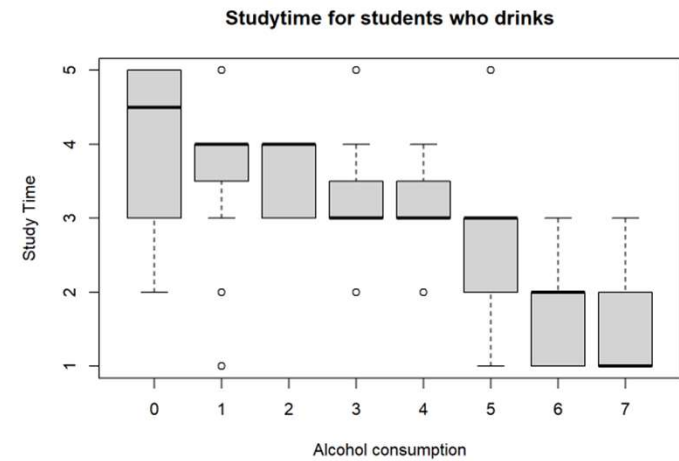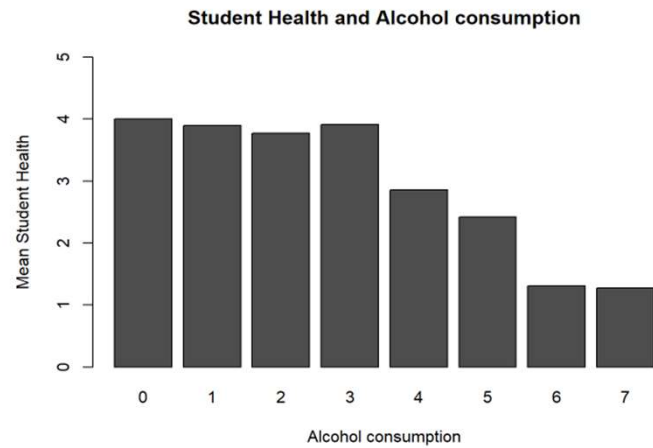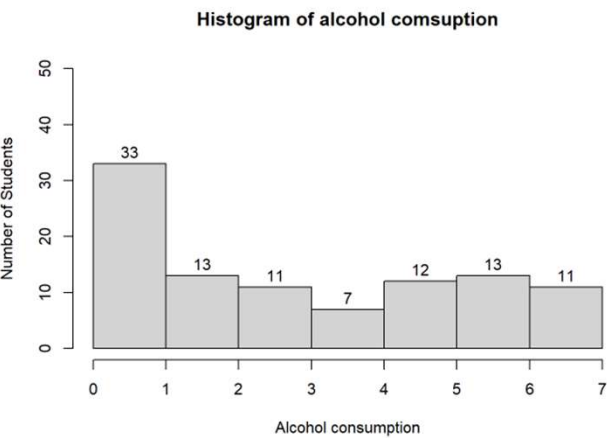
```
dim(student_data)
```
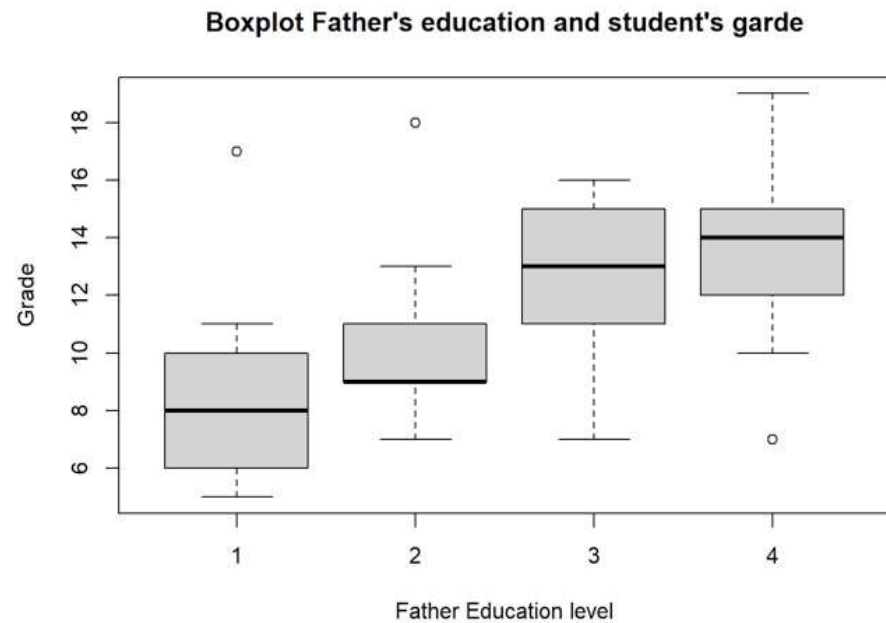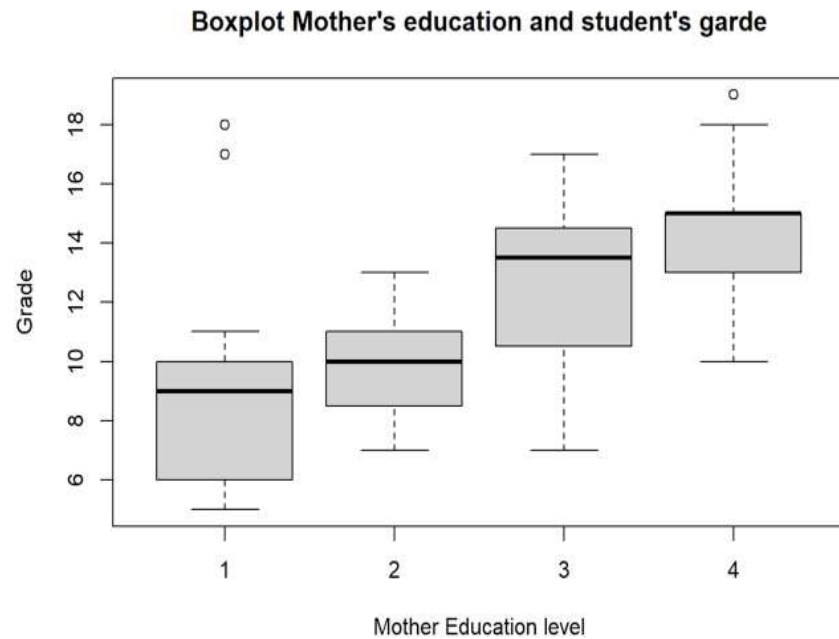
```
## [1] 100  13
```

RESULTS

A) Prognosis and cumulative effects of health status and free time on educational attainment.
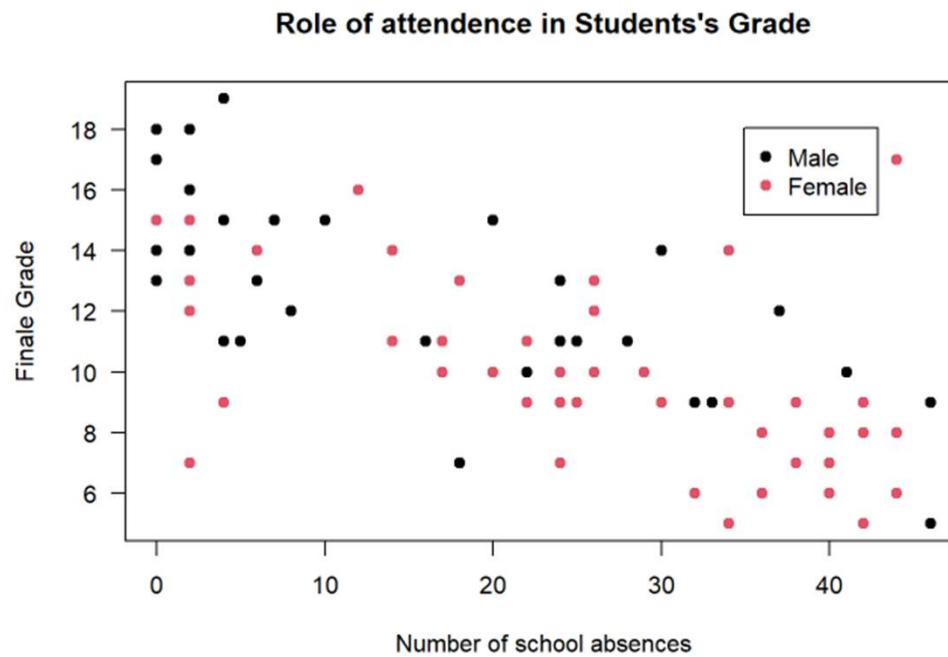
B) Finding out residual effects of weekly alcohol consumption on academic performance.

C) How parental education affects students' academic performance.

D) Evaluating students' academic performance based upon the consistency in attendance.

Role of attendence in Students's Grade

RESULTS :

1) Students with better health condition had better grades .

2) Students with lower alcohol consumption and who utilized their study time efficiently obtained good grades.

3) Students with a high ratio of parent education performed better.

4) Students having consistent attendance performed well in academics.

5) There is no gender-specific difference in the result of the grades.

CONCLUSION

- Encourage students to attend classes consistently.

- Encourage students to attend office hours.

- Promote mentorship programs.

- Promote Alcohol and drug prevention programs for students.

- Obtain new perspectives by interviewing students and staff, to further investigate on lower academic performance.

THANKYOU