Megha Gupta: mg9428@rit.edu

Project Proposal for Statistics 614.01

Dataset: Student exam performance

StudentPerformace.csv

Description of the Dataset: This dataset contains the score of the students about math, reading and writing. Along with this, the data includes different categorical variables of the students such as ('gender', 'race/ethnicity', 'parental level of education', 'lunch', 'test preparation course').

Types of variables

Predictor Variables: test preparation of the course, parental level of education, gender, race, lunch.

Response Variables: math score, reading score, writing score.

Total Number of observations: 100

Objective of the project: Whether the categorical data in the dataset is correlated to the reading, writing and math scores Also, we can identify the correlation between the performance of students based on gender.

Questions to be answered based on the analysis

1. Which gender performed better in each math, reading and writing sections?
2. Does parental education have any effect on student's performance?
3. What was the type of meal consumed by the students who performed better in respective sections?
4. Does practice help to improve the scores?
5. The best performed exam.
6. Whether a student who performs well on a particular exam is likely to perform good in any other exam too? Is there any correlation between them?

Goal of the project: To identify the features which effects a student's performance on a particular course. Analyzing this can be helpful, to design a successful plan for the student and thereby it can help to improve the performance.

Statistical Analysis:

1. Linear regression to identify if the math score, reading score and writing score are statistically significant to one another.
2. 2 – sampled t test using 95% confidence interval.
3. 1-way ANOVA.