Chapter 4 - CABA

**Analysis Report**

Students Performance in an Admission test

By Dilkasha Tarannum

horizontal line

# 

# Introduction:

Several students take admission tests and they face different types of challenges. It is very crucial for schools to understand and analyze the performance of students in those tests and how they can play their part in enhancing it. Student performance can vary among different groups and can be affected by various factors. A better understanding of these factors can help the school to devise strategies to improve student performance. Based on the dataset our problem statement is:

**The school wants to analyze how parental education level and test preparation courses would affect a student's academic performance?**

This dataset has been taken from an open source platform and it represents the “*Student performance in an Admission test*”. As the name suggests; Data set is gathered from different students undertaking an admission test. The dataset represented the scores of students in an admission test which had 3 sections ; math, reading and writing. Along with this, the dataset provides some additional information about the students such as their gender, test preparation courses, etc.

In this report we will briefly discuss the data description, methodology of performing data analysis followed by the insights generated from the dataset.

**Data Description:**

This dataset has 9 columns and 1000 rows. This dataset contained no Null or Na values hence we have not used any Data imputation techniques for data preprocessing.

Column types are both categorical and numerical and they provide us the information about the students such as their Gender, Test preparation courses, parental level of education, lunch, race/ethnicity groups and their scores in different sections of the admission test such as writing, reading and math.

Along with these columns we have one feature engineered column; Obtained Score, which is a sum of scores in different sections.Numerical columns are providing the marks of students in different sections and in the overall test. However, categorical columns are providing some extra details about their test preparation, or their background which can help us in our analysis.

## Methodology for Data Analysis:

From different data analysis strategies such as Data aggregation and Data summary techniques we have performed different strategies such as filtering our dataset into subsets and aggregating dataset based on either mean,mode or median. We have conducted analysis using the Minimum, maximum and range values of the numerical columns helping us to perform our data aggregation based on quartiles. These quartiles have also helped us in performing Inter quartile range analysis to find the outliers in a dataset.

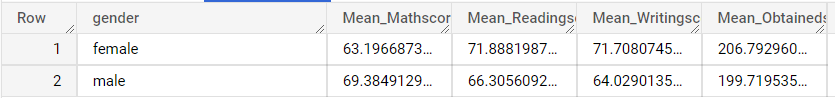
Along with implementing Measures of Location and Measures of Spread we have also conducted data visualization analysis such as plotting the scatter plots, bar plots and correlation graphs of our dataset.

**Insights generated:**

Following are some insights that we have been able to generate form our aforementioned analysis:

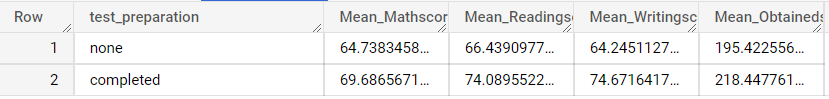
**Based on Gender:**

1. If we consider the overall performance of students in the admission test based on Obtained Score, Female students were able to perform better than male students.
2. However, If we consider the individual scores we come to a conclusion that female students were not able to perform better in Maths as compared to male students. But their performance in writing and reading was better than male students involved in the study.



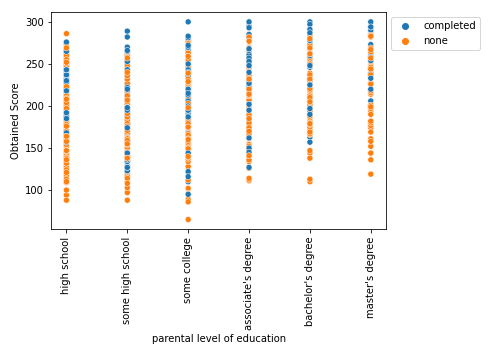
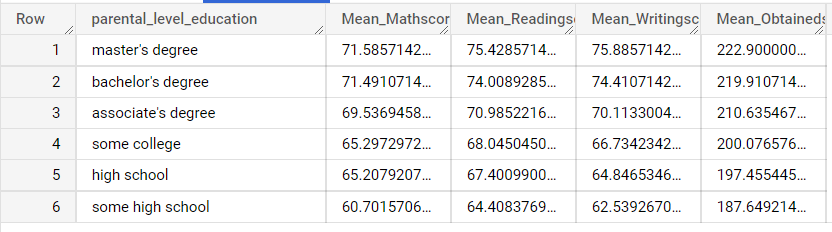
**Based on Test Preparation:**

1. All the students who were able to complete the test preparation courses were relatively performing better than other students in overall scores as well as the section wise scores.



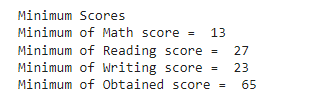
**Based on Parental Level of Education:**

In order to understand the importance of your parent’s level of education we summarized our analysis at parental level of education.

1. Students whose parents had relatively higher levels of education such as Masters, bachelors were able to secure good marks as compared to the students whose parents had some college or some high school as their parental level of education. 

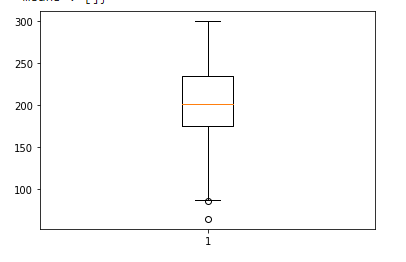
**Section wise performance:**

Based on the minimum, mean and median scores in the Math section we have identified that students have struggled in the Math section the most as compared to other sections.



### Outlier Identification:

Taking help from the Interquartile Range, we have identified two students who secured lower than the minimum of IQR or the lower bound threshold of data distribution.

****

**Conclusion:**

1. We have seen that parents hold a significant role in the improvement of a student's performance. Such as if parents are educated, having Bachelor's or Master’s degree they can better motivate and train their children for the admission test. Likewise, schools can arrange extra support and training sessions for the parents who do not have a high level of education. So that such parents can also motivate and play a better role in the success of their children.
2. From the dataset analysis we have seen that students who have taken their test preparation courses can perform better in all sections of their admission test and score better marks. Therefore, schools should encourage the students to complete preparation courses enabling them to have a conceptual and hands-on practice of all sections so that students can perform better in their test scores.

### Thank You