# # (Students Performance Exploration)

## ## by (Basant Hussein)

## Dataset

> The Inspiration is to understand the influence of the parents background, test preparation etc on students performance. the dataset comprises of 1,000 rows and 8 columns, with which we continue to determine what all the features which plays a vital role in affecting the student's performance.

The dataset consists of the marks secured by the students in various subjects, which accessible from Kaggle Student Performance in Exams. there are 1000 occurance and 8 columns:

1. gender
2. race / ethnicity
3. parental level of education
4. lunch
5. test preparation course
6. math score
7. reading score
8. writing score

> I start explore the data, cleaning the data, remove outliers in the subject’s scores, replace the space in the columns names with underscore.

## Summary of Findings

### In univariate exploration

> 1. The distribution is normal in the subject’s scores.

> 2. The median in the reading score is the highest.

> 3. The female's count seems to be more than the male's count, but not with a large difference.

> 4. The counts of some college is the largest, and master's degree is the lowest.

> 5. The count of none in test preparation course is larger than completed.

# # # in bivariate exploration

> 1. The masters degree's median is the largest, and high school is the lowest, with math score

> 2. The median also in master's degree is the largest in master's degree with reading score

> 3. The median also in master's degree is the largest in writing score

> 4. The median of male in math\_score is larger than the female

> 5. The median of female in reading\_score is larger than the male

> 6. The median of female in writing\_score is larger than the male

> 7. The median of completed in test preparation course is larger than none in math score

> 8. The median of completed in test preparation course is larger than none in reading score

> 9. The median of completed in test preparation course is larger than none in writing score

> 10. Their is a positive correlation between math score and reading score

> 11. Their is a positive correlation between math score and writing score

> 12. Their is a positive correlation between reading score and writing score

# # # in multivariate exploration

> 1. The multivariate exploration of subjects\_scores with the gender shows that the scores of females is higher than male's score as a combination of subjects.

> 2. The concentration of subjects scores in test preparation course =none is more than test preparation course =completed, the concentration of scores also is between 60 and 80.

> 3. The concentration of parental level education of associate degree,some collage,high school in math score with reading score is the highest.

> 4. The concentration of parental level education of associate degree,some collage in math score with writing score is the highest.

> 5. The concentration of parental level education of high school, associate degree,some collage in reading score with writing score is the highest.

#### > 6. the test preparation course has an effect on the student's scores in the subject, after doing hypothesis testing.

> I will bring in the presentation :

> in bivariate : the relation between the subjects scores with each others, parental level of education with the subjects scores, test preparation course with the subjects scores.

> in multivariate the relation between the gender and the subjects.

## Key Insights for Presentation

> the polishing will be in color, title, legend, axes, subtitle.