

# Data Visualization of Student Performance in Exam

**Name:** Abdul Haq Khan

**Student ID:** 22093697

**Github:**

<https://github.com/abdulhaq555/StudentPerformanceVisualization>

**Dataset:**

<https://www.kaggle.com/datasets/spscientist/students-performance-in-exams>

## Report of Student Performance in Exams Dataset:

### Steps:

Import the necessary libraries first in order to create the visualization and load the dataset.

Use the panda library to import the dataset into a dataframe.

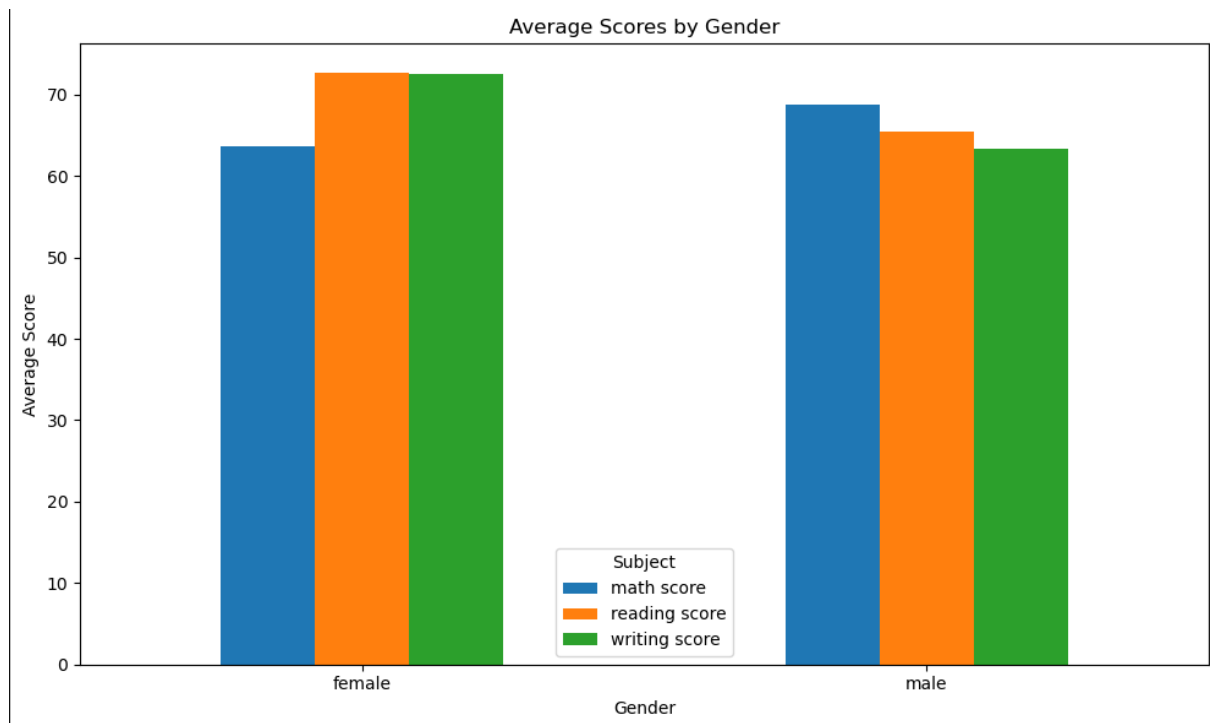
To deal with the missing numbers, we can either remove the null values or substitute the average values for the null ones. But in our case no missing value is founded over dataset is already cleaned

Print the dataset's summary description by utilising the data.describe()

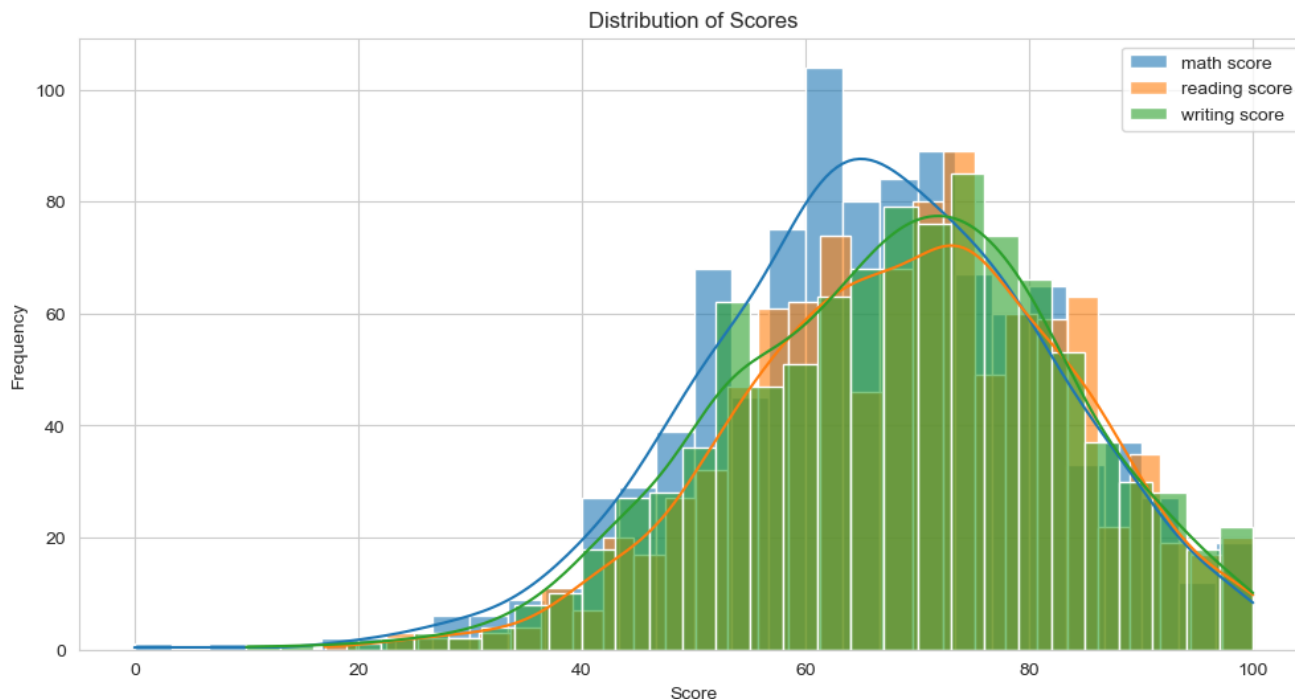
Extracting some meanings from data for better visualization and understanding by using other methods

Create several dataset visualisations to get various conclusions.

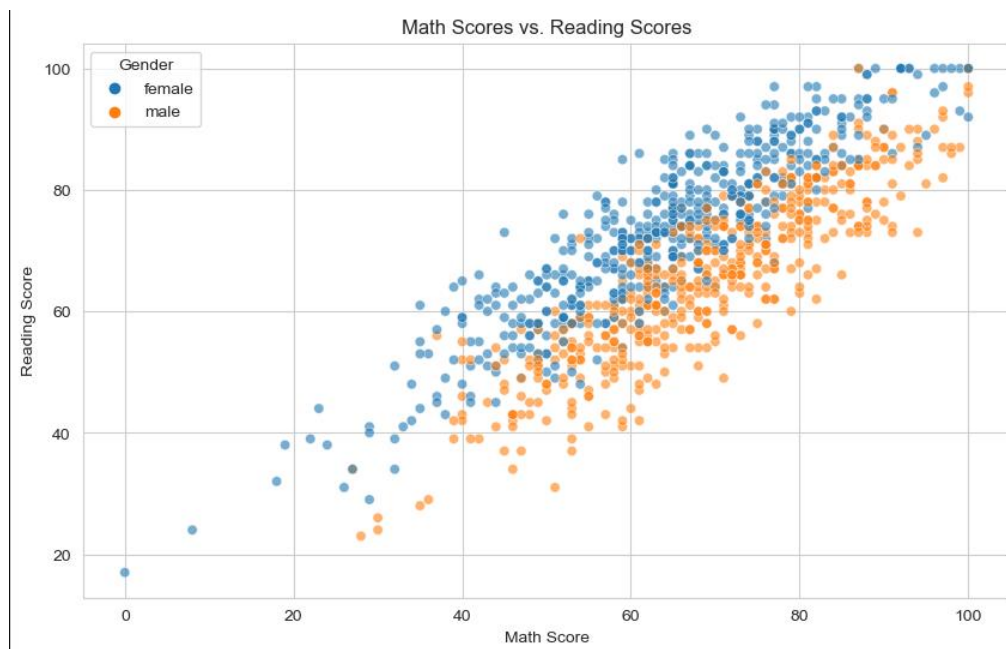
### Graphs:



This graph is showing the average scores by gender we can conclude from visualization that females got more average score in reading and writing subject than boys while males got higher score in math's then females and females have high score in average then male students.



From the shape of distributions we can conclude that most scores are clustered around middle range and with fewer scores at the lower and higher ends suggesting a normal-like distribution for each subject's scores and score in maths subject are relatively higher.



The scatter plot shows a positive correlation between maths and reading scores, meaning that generally students who score higher in maths also score higher in reading and vice versa. The distribution of scores for male and female students seems relatively similar and linearly distributed.