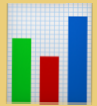


PYTHON FOR STATISTICS



Notes

Libraries

Pandas → Data manipulation and analysis & handling csv files
Matplotlib.pyplot → used to create visuals & graphs
Random → used to generate random numbers

Reading Data From CSV File

- used `pd.read_csv()` to read the specific column containing marks into a variable

Performing Basic Statistics

- Pandas provides relevant functions to perform basic statistics i.e. Mode, mean, median, Range, standard Deviation, variance & covariance.
- Random provides function to create a random integer

Plotting Graphs

→ matplotlib provides relevant functions for each kind of graph

Subplots: `plt.subplot` is used to create multiple subplots to enter multiple graphs

→ relevant graphing functions are provided data accordingly to produce accurate representations of data.

→ Data is then adjusted and shown through `plt.show()`

`marks — frequency = marks.value_counts().sort_index()`

• `value_counts()` → returns a new set of data that consists of the frequencies of values found in variable 'marks' and stores in `marks — frequency.values`.

• `sort_index()` → returns a new set of data that consists of the unique values from the values in variable 'marks' and stores them in `marks — frequency.index`