Statistical foundation of Data Sciences

Practical- 03

Roll number: GF202344767

Name: Ishita Mehta

Workflow summary:

1. Open Jupyter Notebook or Python IDE.
2. Import required libraries: pandas, numpy, matplotlib, seaborn.
3. Load the updated dataset teacher_ratings_updated.csv.
4. Display the first few records using head().
5. Identify duplicate entries using the prof column.
6. Calculate mean and standard deviation for age (all observations).
7. Remove duplicate professors and create a filtered dataset.
8. Recalculate mean and standard deviation for the filtered dataset.
9. Compare evaluation scores by course division using a bar chart.
10. Plot scatter plot of age vs eval to analyze relationships.
11. Create gender-based scatter plot to visualize evaluation trends.
12. Plot combined scatter plot differentiated by both gender and tenure.
13. Interpret results and write final observations and conclusions.

Github Repository link:

https://github.com/pineapplesdontbelongonpizza/CSU1658_practical1_Testing_Pandas_and_Numpy.git