

Statistical foundation of Data Sciences

Practical- 01

Roll number: GF202344767

Name: Ishita Mehta

Workflow summary:

1. Set random seed and created a synthetic dataset with NaN values.
2. Calculated mean, median, and weighted mean, explaining why weighted mean is important.
3. Standardized income using z-score, identified outliers using $|z| > 3$, and handled NaNs correctly.
4. Created age bins and computed count, mean, and median for each bin, showing tidy DataFrame.
5. Created a 2D NumPy array, performed shape/resize operations, showcased negative indexing and slicing errors, applied broadcasting, dot product, and basic linear algebra (determinant, inverse).

Github Repository link:

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