|  |
| --- |
| Practice exercises with PANDAS |
| Exercise #01 |
| Tips Dataset Exercises  Exercise 1: Load and Display the Dataset  Task: Load the tips dataset from the seaborn library and display the first 5 rows.  Exercise 2: Basic Statistics  Task: Calculate and display the basic statistics (mean, median, mode, standard deviation) for the total\_bill column.  Exercise 3: Group By and Aggregate  Task: Group the data by day and calculate the average total\_bill and tip for each day.  Exercise 4: Filtering Data  Task: Filter the dataset to include only dinners (time is 'Dinner') and display the first 5 rows.  Exercise 5: Adding a New Column  Task: Add a new column tip\_percentage which is the tip amount as a percentage of the total bill.  Exercise 6: Filtering Based on Conditions  Task: Filter the dataset to include only rows where the tip\_percentage is greater than 20%.  Exercise 7: Sorting Data  Task: Sort the dataset by total\_bill in descending order and display the first 5 rows.  Exercise 8: Visualization  Task: Create a scatter plot of total\_bill vs tip using matplotlib.  Exercise 9: Counting Values  Task: Count the number of times each day appears in the dataset.  Exercise 10: Pivot Table  Task: Create a pivot table that shows the average total\_bill and tip for each combination of day and time. |
| Exercise #02 |
| Wine Dataset Tasks   1. Load and Display the Dataset    * Task: Load the wine dataset and display the first 5 rows to understand the structure of the data. 2. Basic Statistics    * Task: Calculate and display basic statistics (mean, median, mode, standard deviation) for the 'alcohol' column. 3. Group By and Aggregate    * Task: Group the data by 'target' and calculate the average 'alcohol' and 'malic\_acid' for each class. 4. Filtering Data    * Task: Filter the dataset to include only rows where 'alcohol' is greater than 13. Display the first 5 rows of the filtered dataset. 5. Adding a New Column    * Task: Add a new column 'alcohol\_to\_malic\_acid\_ratio' which is the ratio of 'alcohol' to 'malic\_acid'. 6. Filtering Based on Conditions    * Task: Filter the dataset to include only rows where the 'alcohol\_to\_malic\_acid\_ratio' is greater than 10. 7. Sorting Data    * Task: Sort the dataset by 'alcohol' in descending order and display the first 5 rows. 8. Visualization - Scatter Plot    * Task: Create a scatter plot of 'alcohol' vs 'malic\_acid' to visualize the relationship between these two features. 9. Visualization - Box Plot    * Task: Create a box plot of 'alcohol' for each 'target' class to compare the distributions. 10. Correlation Matrix     * Task: Compute and display the correlation matrix for the dataset to understand the relationships between different features. |
|  |
|  |
|  |
|  |