

practice problems using the Pandas module in Python:

1. Load Data and Basic Operations:

- Load a CSV file into a DataFrame using Pandas.
- Display the first 5 rows of the DataFrame.
- Display the column names of the DataFrame.
- Check the shape of the DataFrame (number of rows and columns).
- Find the data types of each column in the DataFrame.

2. Data Selection and Filtering:

- Select a single column from the DataFrame.
- Select multiple columns from the DataFrame.
- Filter rows based on a condition (e.g., select rows where a certain column value is greater than 100).
- Apply multiple filters on the DataFrame and display the result.

3. Data Cleaning:

- Check for missing values in the DataFrame.
- Replace missing values with appropriate values (e.g., mean, median, mode).
- Drop rows with missing values.
- Drop columns that are not required for analysis.

4. Data Manipulation:

- Create a new column in the DataFrame based on existing columns (e.g., calculate the total of two columns).
- Apply a function to a column (e.g., convert string values to lowercase).
- Sort the DataFrame based on a column.
- Group data based on a column and perform aggregation (e.g., find the sum, mean, count for each group).

5. Data Visualization:

- Plot a histogram of a numerical column.
- Plot a bar chart of a categorical column.
- Plot a line chart of a numerical column over time (if applicable).
- Use seaborn or matplotlib along with Pandas to create more complex visualizations (e.g., scatter plot, box plot).

6. Data Analysis:

- Calculate descriptive statistics (mean, median, standard deviation) for numerical columns.
- Find the most frequent value in a categorical column.
- Calculate the correlation between numerical columns.
- Perform cross-tabulation between two categorical columns.

7. Data Transformation:

- Convert a categorical column into dummy variables.
- Apply one-hot encoding to a categorical column.
- Binning numerical data into categories.
- Normalize numerical data using Min-Max scaling or Standardization.

8. Combining DataFrames:

- Merge two DataFrames based on a common column.
- Concatenate two DataFrames along rows or columns.
- Join two DataFrames based on a common index.

[Solution](#)

https://github.com/amirkhan1092/Batch2023-24/blob/main/module2/pandas%20basics/solution_practiceSheet.md