Pandas Topics for Data Science

Basics of Pandas

1. Installing and Importing Pandas
   import pandas as pd

2. Data Structures
   - Series
   - DataFrame

Data Input and Output

3. Reading Data
   - pd.read_csv()
   - pd.read_excel()
   - pd.read_json()
   - pd.read_sql() (with SQLAlchemy)

4. Writing Data
   - df.to_csv()
   - df.to_excel()

Data Exploration

5. Basic Functions
   - df.head(), df.tail()
   - df.shape, df.info(), df.describe()
   - df.columns, df.index, df.dtypes

6. Summary Statistics
   - df.mean(), df.median(), df.std()

Pandas Topics for Data Science

- df.value_counts(), df.unique()

Data Cleaning

7. Handling Missing Values
   - df.isnull(), df.notnull()
   - df.fillna(), df.dropna()

8. Replacing Values
   - df.replace()

9. Renaming Columns
   - df.rename()

10. Changing Data Types
   - df.astype()

Data Manipulation

11. Filtering / Slicing / Subsetting
   - df.loc[], df.iloc[], df[df['col'] > 10]

12. Sorting
   - df.sort_values(), df.sort_index()

13. Adding / Removing Columns
   - df['new_col'] = ...
   - df.drop(columns=['col1'])

14. Merging / Joining

Pandas Topics for Data Science

- pd.merge()
- pd.concat()
- df.join()

15. Grouping and Aggregation
    - df.groupby('col').sum()
    - df.agg(), df.transform()

Data Visualization with Pandas

16. Basic Plotting (via Matplotlib)
    - df.plot(), df.hist(), df.boxplot(), df.plot(kind='bar')

Advanced Topics (For Real Projects)

17. Pivot Tables
    - df.pivot_table()

18. MultiIndex / Hierarchical Indexing

19. Time Series Analysis
    - pd.to_datetime(), df.resample()

20. Window Functions
    - df.rolling(), df.expanding()