|  |  |
| --- | --- |
| **Topic** | **Subtopic** |
| **NUMPY** | |
| Introduction to NumPy | (Importing NumPy, Benefits of NumPy over Python Lists, Homogeneity of data, Speed and efficiency with NumPy (written in C)) |
| Creating and Understanding Arrays | (Creating arrays using np.array, Type conversion (astype, dtype), Array properties: Dimensions (ndim), Shape (shape), Size (size)) |
| Indexing and Slicing | (Basic indexing, Negative indexing, Slicing arrays, Fancy indexing using lists and masks, Boolean indexing and masking) |
| Working with 2D Arrays (Matrices) | (Creating 2D arrays, Indexing and slicing in 2D arrays, Transpose of matrices, Extracting rows, columns, and diagonals) |
| Array Operations | (Element-wise operations, Aggregate functions (np.sum, np.mean, np.min, np.max), Axis-based operations, Logical operations (np.any, np.all), Conditional selection using np.where) |
| Array Reshaping | (Reshaping arrays (reshape), Flattening arrays (ravel, flatten), Expanding dimensions (expand\_dims, squeeze)) |
| Broadcasting | (Understanding broadcasting rules, Using broadcasting for efficient computation) |
| Sorting and Searching | (Sorting arrays using np.sort, Sorting along specific axes, Searching with conditions using np.where) |
| Copying and Views | (Shallow copy vs Deep copy, Using .view() for shallow copies, Using .copy() for deep copies, Memory sharing with np.shares\_memory) |
| Random Number Generation | (Generating random numbers using np.random, Setting seeds for reproducibility using np.random.seed) |
| Linear Algebra Operations | (Matrix multiplication (np.dot, @ operator, np.matmul), Element-wise multiplication, Matrix transpose, Determinants, Eigenvalues, Solving linear systems) |
| Advanced Array Manipulations | (Splitting arrays (split, hsplit, vsplit), Stacking arrays (hstack, vstack, concatenate), Tiling and repeating arrays (np.tile, np.repeat)) |
| File I/O Operations | (Loading data from text files using np.loadtxt, Saving arrays to files using np.save, Loading arrays using np.load) |
| Statistical Functions | (Mean, Median, Variance, Standard deviation, Correlation (np.corrcoef), Covariance (np.cov), Percentiles and quantiles) |
| Use Cases and Applications | (Fitness data analysis, Net Promoter Score (NPS) analysis) |
| Miscellaneous Topics with Case Study | (Handling errors and debugging, Using NumPy for data preprocessing, Exploratory Data Analysis (EDA)) |
| **PANDAS** | |
| Introduction to Pandas | (Installation of Pandas, Importing Pandas, Why use Pandas, Difference from NumPy, Dataframes and Series) |
| Dataframe Operations | (Basic operations on Dataframes, df.info(), df.head(), df.tail(), df.shape, Creating Dataframe from scratch) |
| Column Operations | (Accessing columns, Checking unique values, Renaming columns, Deleting columns, Creating new columns) |
| Row Operations | (Implicit and explicit indexing, df.index, Indexing in Series, Slicing in Series, loc, iloc) |
| Working with Rows and Columns Together | (Appending rows, Sorting data, Concatenation, Merge and Join operations) |
| Data Selection and Filtering | (Using conditions, Boolean indexing, Filtering data with masks) |
| Grouping and Aggregation | (groupby(), Group-based aggregates, Group-based filtering, Group-based apply) |
| Data Restructuring | (Multi-indexing, Restructuring data with pd.melt, pd.pivot, pd.pivot\_table, pd.cut) |
| Handling Missing Values | (None and NaN values, isna(), isnull(), Dropping missing values, Filling missing values) |
| String Methods in Pandas | (Handling text data, String operations) |
| Datetime Operations | (Handling datetime objects, Extracting date and time components, Time series analysis) |
| File I/O Operations | (Reading CSV files, Writing to CSV files, Exporting data) |
| Apply Function | (Using apply() on DataFrames, Row-wise and column-wise operations, Using custom functions with apply()) |
| Advanced Grouping Techniques | (Multi-level aggregation, aggregate() with multiple columns, Calculating active years, Productivity analysis) |
| Data Visualization and Insights | (Visualizing grouped data, Insights from aggregated data) |
| Case Study: McKinsey Dataset Analysis | (Analyzing GDP per capita vs Life Expectancy, Filtering and grouping data, Drawing insights for business decisions) |
| Case Study: IMDB Movie Business Use-case | (Analyzing movie data, Merging datasets, Exploring movie budgets, revenues, and ratings) |
| Case Study: Director Productivity Analysis | (Evaluating directors based on active years and movie release rate, Multi-indexing and grouping techniques) |
| Case Study: Pfizer Drug Stability Monitoring | (Monitoring drug stability using temperature and pressure data, Handling time-series data, Restructuring data using pd.melt and pd.pivot) |
| **MATPLOTLIB AND SEABORN** | |
| Introduction to Data Visualization | (Uses of Matplotlib, Introduction to Seaborn, Difference between Matplotlib and Seaborn, Exploratory vs Explanatory Data Visualization) |
| Anatomy of a Plot | (Figure, Axes, Axis, Labels, Ticks, Legends, Titles) |
| Types of Data Visualization | (Univariate Analysis, Bivariate Analysis, Multivariate Analysis) |
| Univariate Data Visualization | (Categorical Data: Bar Chart, Countplot, Pie Chart; Continuous Data: Histogram, KDE, Box Plot) |
| Bivariate Data Visualization | (Continuous-Continuous: Line Plot, Scatter Plot; Categorical-Categorical: Dodged Countplot, Stacked Countplot; Continuous-Categorical: Box Plots) |
| Multivariate Data Visualization | (CCN (Continuous-Continuous-Categorical), CNN (Continuous-Numerical-Numerical), NNN (Numerical-Numerical-Numerical)) |
| Advanced Plotting Techniques | (Joint Plot, Pair Plot, Heatmap, Correlation Analysis) |
| Styling and Customization | (Plot Titles, Labels, Ticks, Legends, Color Customization, Adding Grids, Highlighting Points) |
| Subplots and Dashboards | (Using plt.subplots(), Creating dashboards, Comparing multiple plots) |
| Case Study: Tencent Video Games Analysis | (Analyzing sales data for video games, Identifying trends in North America, Europe, and Japan, Exploring the correlation between game ranks and sales) |
| Case Study: Genre and Publisher Analysis | (Analyzing top genres and publishers, Visualizing distribution of games across platforms, Sales analysis by genre and publisher using box plots) |
| Case Study: Global Sales Analysis | (Using Bubble Charts to visualize sales impact based on game rankings, Using heatmaps to identify correlations among sales regions) |
| **SQL** | |
| Introduction to SQL | (Understanding SQL, Importance in Data Analysis, SQL Syntax Basics, Introduction to Databases) |
| Data Types in SQL | (Numeric, String, Date & Time, Boolean, Handling NULL values) |
| Filtering and Subqueries | (Filtering data with WHERE clause, Inline calculations, Subqueries, Using IN, NOT IN, and BETWEEN, Using LIKE for pattern matching) |
| Joins and Relationships | (INNER JOIN, LEFT JOIN, RIGHT JOIN, FULL OUTER JOIN, Self Joins, Joining multiple tables, Handling duplicates in joins) |
| Group By and Aggregation | (GROUP BY clause, Using aggregate functions: SUM, COUNT, AVG, MIN, MAX, HAVING clause for conditional filtering after grouping) |
| Window Functions | (ROW\_NUMBER, RANK, DENSE\_RANK, NTILE, Aggregated Window Functions, Using PARTITION BY and ORDER BY, Calculating running totals) |
| Advanced Constructs | (Common Table Expressions (CTEs), Creating Views, Ad-hoc reporting, Designing reusable analytical datasets) |
| Date & Time Functions | (Extracting components using EXTRACT(), Date and time formatting, DATE\_ADD, DATE\_SUB, DATEDIFF, Handling timestamps, Time series analysis) |
| String Manipulation in SQL | (Using CONCAT, UPPER, LOWER, String functions like LEFT, RIGHT, SUBSTRING, Trimming whitespace) |
| Data Transformation and Cleaning | (Using CASE statements for conditional data transformations, Handling duplicates, Data type conversions) |
| Advanced Filtering Techniques | (AND, OR, NOT conditions, Filtering on multiple columns, Filtering based on string patterns using wildcards, Filtering on ranges using BETWEEN) |
| Use Cases and Applications | (Case Study 1: Sales Analysis for Farmerâ€™s Market |
| Use Cases and Applications | Case Study 2: Target Customer Satisfaction |