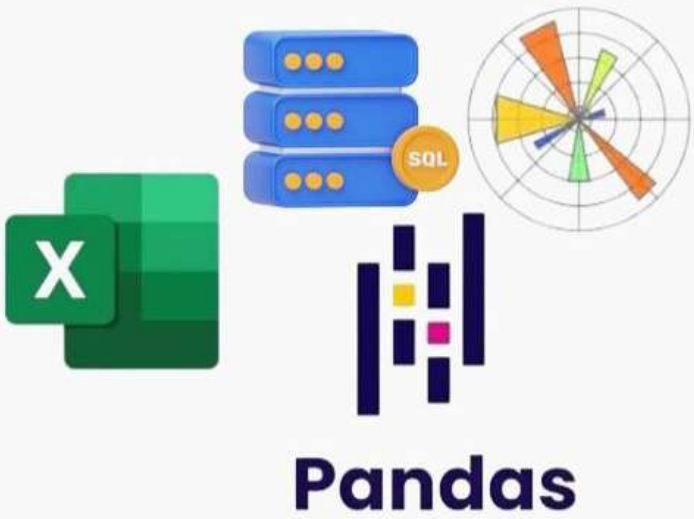
Functions Used By Data Analysts



EXCEL FUNCTIONS

1. **VLOOKUP:** Searches for a value in the first column of a table array and returns a value in the same row from another column.
2. **HLOOKUP:** Similar to VLOOKUP but searches for a value in the first row of a table array and returns a value in the same column.
3. **SUMIF/SUMIFS:** Adds the cells specified by a given condition or criteria.
4. **COUNTIF/COUNTIFS:** Counts the number of cells specified by a given condition or criteria.
5. **AVERAGEIF/AVERAGEIFS:** Calculates the average of cells specified by a given condition or criteria.
6. **INDEX/MATCH:** Returns the value in a cell at the intersection of a particular row and column, based on matching a criteria.
7. **PivotTables:** Summarizes, sorts, and filters data in Excel.
8. **IF/IFERROR:** Executes a specific action based on a condition or returns a value if an error occurs.
9. **CONCATENATE/CONCAT:** Joins two or more strings together.
10. **TEXT/DATEVALUE:** Converts text to date values.
11. **INDEX/MATCH Combination:** Provides more flexibility than VLOOKUP and HLOOKUP for searching values in a table.
12. **Conditional Formatting:** Allows formatting cells based on certain conditions, making data visualization more intuitive.
13. **Data Validation:** Restricts the type of data that users can enter into a cell, ensuring data integrity.
14. **Array Formulas:** Perform multiple calculations on one or more items in an array.
15. **Solver:** An Excel add-in used for optimization and what-if analysis.
16. **Pivot Charts:** Visual representations of PivotTable data.
17. **Goal Seek:** Finds the input needed to achieve a desired result in a formula.
18. **Advanced Filter:** Allows filtering data by multiple criteria and copying the filtered results to another location.
19. **Text Functions (e.g., LEFT(), RIGHT(), MID(), etc.):** Extract or manipulate text data in cells.
20. **Data Tables:** Allows performing sensitivity analysis by calculating multiple versions of a formula with different input values.

SQL FUNCTIONS

1. **SELECT:** Used to retrieve data from a database.
2. **WHERE:** Filters data based on specified conditions.
3. **GROUP BY:** Groups rows that have the same values into summary rows.
4. **HAVING:** Filters records returned by a GROUP BY clause.
5. **ORDER BY:** Sorts the result set in ascending or descending order.
6. **JOIN:** Combines rows from two or more tables based on a related column.
7. **DISTINCT:** Returns unique values in a specified column or expression.
8. **COUNT():** Returns the number of rows in a specified table or view.
9. **SUM():** Calculates the sum of a set of values.
10. **AVG():** Calculates the average of a set of values.
11. **CASE Statement:** Allows conditional logic within SQL queries.
12. **UNION:** Combines the result sets of two or more SELECT statements.
13. **CTE (Common Table Expressions):** Temporary result sets that can be referenced within a SELECT, INSERT, UPDATE, or DELETE statement.
14. **Window Functions (e.g., ROW_NUMBER(), RANK(), etc.):** Perform calculations across a set of rows that are related to the current row.
15. **Stored Procedures:** Precompiled SQL code that can be executed by calling the procedure name.
16. **INDEX:** Improves the speed of data retrieval operations on a database table at the cost of additional space and decreased performance for insert, update, and delete operations.
17. **TRIGGER:** A database object that automatically performs an action in response to certain events on a particular table or view.
18. **EXISTS:** Tests for the existence of any rows in a subquery and returns true if the subquery returns one or more rows.
19. **ROLLUP:** Generates subtotal values for the data, based on one or more columns.
20. **EXPLAIN:** Analyzes the execution plan of a SELECT statement to help optimize query performance.

PANDAS FUNCTIONS

1. **read_csv():** Reads a CSV file into a DataFrame.
2. **head():** Returns the first n rows of a DataFrame.
3. **info():** Provides a concise summary of a DataFrame, including data types and non-null values.
4. **describe():** Generates descriptive statistics of the DataFrame.
5. **loc[]:** Accesses a group of rows and columns by label(s) or a boolean array.
6. **iloc[]:** Accesses a group of rows and columns by integer position(s).
7. **merge():** Combines two DataFrames by a common column.
8. **groupby():** Groups DataFrame using a mapper or by a Series of columns.
9. **pivot_table():** Creates a spreadsheet-style pivot table as a DataFrame.
10. **to_csv():** Writes DataFrame to a CSV file.
11. **pd.concat():** Concatenates pandas objects along a particular axis with optional set logic along the other axes.
12. **pd.melt():** Unpivots DataFrame from wide to long format.
13. **pd.pivot_table():** Creates a spreadsheet-style pivot table as a DataFrame.
14. **pd.cut():** Bin values into discrete intervals.
15. **pd.qcut():** Quantile-based discretization function.
16. **pd.merge():** Combines DataFrame objects by performing a database-style join operation.
17. **pd.DataFrame.apply():** Applies a function along an axis of the DataFrame.
18. **pd.DataFrame.groupby():** Groups DataFrame using a mapper or by a Series of columns.
19. **pd.DataFrame.drop_duplicates():** Removes duplicate rows from the DataFrame.
20. **pd.DataFrame.to_excel():** Writes DataFrame to an Excel file.

MATPLOTLIB FUNCTIONS

1. **plt.plot():** Creates a line plot.
2. **plt.scatter():** Creates a scatter plot.
3. **plt.bar():** Creates a bar plot.
4. **plt.hist():** Creates a histogram.
5. **plt.boxplot():** Creates a boxplot.
6. **plt.xlabel():** Sets the label for the x-axis.
7. **plt.ylabel():** Sets the label for the y-axis.
8. **plt.title():** Sets the title of the plot.
9. **plt.legend():** Adds a legend to the plot.
10. **plt.show():** Displays the plot.
11. **plt.savefig():** Saves the plot to a file.
12. **plt.subplots():** Creates a figure and a set of subplots.
13. **plt.figure():** Creates a new figure.
14. **plt.xticks():** Sets the tick labels on the x-axis.
15. **plt.yticks():** Sets the tick labels on the y-axis.
16. **plt.grid():** Adds grid lines to the plot.
17. **plt.xlim():** Sets the limits for the x-axis.
18. **plt.ylim():** Sets the limits for the y-axis.
19. **plt.annotate():** Adds annotations to the plot.
20. **plt.subplots_adjust():** Adjusts the spacing between subplots.