# Project: Excel Filters

Purpose: Advance Python Class

Author: Muzaffar Ali

Version: 1.0

Created: 31/07/2024

Copyright: (c) Muzaffar Ali

License: Public

## Requirements:

* pandas: Used for data manipulation and analysis, providing powerful data structures like DataFrames and Series.
* os: Provides a way to interact with the operating system, handling file and directory operations.
* matplotlib.pyplot: Used for creating visualizations, including various types of plots and charts.

## Description:

This project provides a comprehensive tool for filtering, searching, analyzing, sorting, grouping, handling missing values, and visualizing data from an Excel file. The tool is menu-driven, allowing users to perform various operations on the data interactively.

## Functions:

### load\_excel(file\_path):

Purpose: Load an Excel file.  
Args: file\_path (str): The path to the Excel file.  
Returns: pd.DataFrame: The loaded DataFrame.

### display\_columns(df):

Purpose: Display the columns of a DataFrame.  
Args: df (pd.DataFrame): The DataFrame whose columns are to be displayed.

### filter\_data(df):

Purpose: Filter the DataFrame based on user input.  
Args: df (pd.DataFrame): The DataFrame to filter.  
Returns: pd.DataFrame: The filtered DataFrame.

### filter\_even\_odd(df):

Purpose: Filter the DataFrame to show even or odd numbers in a selected column.  
Args: df (pd.DataFrame): The DataFrame to filter.  
Returns: pd.DataFrame: The filtered DataFrame with even or odd numbers.

### search\_data(df):

Purpose: Search for a keyword in the DataFrame.  
Args: df (pd.DataFrame): The DataFrame to search.  
Returns: pd.DataFrame: The search results DataFrame.

### analyze\_data(df):

Purpose: Perform basic data analysis on the DataFrame.  
Args: df (pd.DataFrame): The DataFrame to analyze.

### sort\_data(df):

Purpose: Sort the DataFrame based on user input.  
Args: df (pd.DataFrame): The DataFrame to sort.  
Returns: pd.DataFrame: The sorted DataFrame.

### group\_data(df):

Purpose: Group the DataFrame based on user input and perform aggregation.  
Args: df (pd.DataFrame): The DataFrame to group.  
Returns: pd.DataFrame: The grouped DataFrame.

### handle\_missing\_values(df):

Purpose: Handle missing values in the DataFrame.  
Args: df (pd.DataFrame): The DataFrame with missing values.  
Returns: pd.DataFrame: The DataFrame with missing values handled.

### visualize\_data(df):

Purpose: Visualize data from the DataFrame using plots.  
Args: df (pd.DataFrame): The DataFrame to visualize.

### save\_results(df):

Purpose: Save the DataFrame results to a file.  
Args: df (pd.DataFrame): The DataFrame to save.

### main():

Purpose: Main function to handle the workflow.  
Description: Interactively guides the user through various operations on the Excel data.

## Usage:

1. Load Excel File:  
- User provides the path to the Excel file.  
- The program checks if the file exists and is accessible.  
  
2. Menu Options:  
- Filter Data: Filter data based on user-defined criteria.  
- Search Data: Search for a specific keyword within the data.  
- Analyze Data: Perform basic statistical analysis on the data.  
- Sort Data: Sort data based on a selected column.  
- Group Data: Group data and apply aggregation functions.  
- Handle Missing Values: Fill or drop missing values in the data.  
- Visualize Data: Create visual representations of the data.  
- Filter Even/Odd Numbers: Filter numeric data to show even or odd numbers.  
  
3. Save Results:  
- The user can choose to save the results of any operation to an Excel file.  
  
4. Exit:  
- The user can exit the program.

## Example:

- Filtering Data:  
 - The user selects a column and provides a value to filter by. The filtered data is displayed and can be saved.  
  
- Searching Data:  
 - The user enters a keyword to search within the data. The search results are displayed and can be saved.  
  
- Analyzing Data:  
 - Summary statistics, data types, and missing values are displayed to help the user understand the dataset better.  
  
- Sorting Data:  
 - The user selects a column to sort by, and the data is displayed in sorted order.  
  
- Grouping Data:  
 - The user selects a column to group by and an aggregation function to apply. The grouped data is displayed.  
  
- Handling Missing Values:  
 - The user can choose to fill missing values with a specified value or drop rows with missing values.  
  
- Visualizing Data:  
 - The user selects columns for the x and y axes, and a scatter plot is displayed.