

Step-by-Step Guide for Data Tools

1. CSV Data with Excel

Goal: Clean and explore data.

1. Open the CSV File:

- Download a dataset from Kaggle (e.g., sales_data.csv).
- Open it in Excel.

2. Explore the Data:

- Check the columns and rows for missing or inconsistent data.
- Example columns: CustomerID, PurchaseAmount, Date.

3. Clean the Data:

- Remove duplicate rows: Go to Data > Remove Duplicates.
- Fill or remove missing values: Use Find & Select > Go To Special > Blanks.
- Format the Date column: Go to Home > Number > Date to standardize.

4. Save the Cleaned File:

- Save as cleaned_sales_data.csv for later use.

2. URL Data with Python

Goal: Fetch online data and save it as CSV.

1. Set Up Python Environment:

- Install Python and required libraries: pip install pandas requests.

2. Fetch Data from URL:

- Example script:

```
import pandas as pd
```

```
# URL for the data
```

```
url = 'https://example.com/data.csv'
```

```
# Fetch and save as CSV
```

```
data = pd.read_csv(url)
```

```
data.to_csv('url_data.csv', index=False)
```

```
print('Data saved as url_data.csv')
```

3. Run the Script:

- Save the code as fetch_data.py.

- Run it: python fetch_data.py.

4. Check the CSV File:

- Open url_data.csv in Excel to ensure data is correct.

3. SQL Database with MySQL

Goal: Store a Kaggle dataset in a database and query it.

1. Install MySQL:

- Download and install MySQL Community Edition from the official website.

2. Create a Database:

- Open MySQL Workbench and run:

```
CREATE DATABASE customer_data;
```

```
USE customer_data;
```

3. Create a Table:

- Design the table structure to match your CSV columns:

```
CREATE TABLE customer_info (
```

```
    CustomerID INT,
```

```
    Name VARCHAR(50),
```

```
    Age INT,
```

```
    Gender VARCHAR(10),
```

```
    Income FLOAT
```

);

4. Import the CSV:

- Go to Server > Data Import in MySQL Workbench.
- Choose your CSV file and map the columns.

5. Query the Data:

- Fetch all data to check it:

```
SELECT * FROM customer_info;
```

4. Visualization with Power BI

Goal: Analyze and visualize data.

1. Import Datasets:

- Open Power BI.
- Go to Home > Get Data and choose:
 - * Excel for CSV files.
 - * MySQL for SQL data (install MySQL connector if needed).

2. Clean and Transform Data:

- Use Transform Data to:
 - * Rename columns for consistency.
 - * Merge tables based on CustomerID.

3. Create Visuals:

- Add charts and visuals for each dataset:
 - * Bar chart: PurchaseAmount by ProductCategory.
 - * Line chart: SessionDuration over Time.
 - * Pie chart: Gender Distribution.

4. Save and Share:

- Save the report as .pbix.
- Export to PDF or publish to Power BI Service if needed.

Recap: Tools and Steps

This guide covered the following tools and steps:

- Cleaning CSV data using Excel.
- Fetching online data using Python.
- Storing and querying data in MySQL.
- Visualizing data with Power BI.