

London Bike Sharing Data Analysis

1. Project Overview

This project analyzes the London Bike Sharing Dataset from Kaggle using a combination of Python (Jupyter Notebook) and an interactive Tableau dashboard. The goal is to understand how different factors—such as temperature, wind speed, weather conditions, seasons, and time—affect bike-sharing usage in London.

2. Dataset Summary

→ **Rows:** 17,414

→ **Columns:** 10

→ **Key Features:**

- ◆ **Timestamp:** Date and time of the bike rental
- ◆ **Count (cnt):** Total number of bikes rented at that timestamp
- ◆ **Temperature (t1, t2):** t1 – recorded temperature, t2 – adjusted

temperature

- ◆ **Humidity (hum):** Relative humidity at the time of rental
- ◆ **Wind Speed (wind_speed):** Wind speed in km/h
- ◆ **Season:** Season of the year (1 = spring, 2 = summer, 3 = fall, 4 = winter)

→ **Missing Data:** No missing values

3. Data Analysis Using Python

The data analysis process began with preparing, cleaning, and transforming the London bike-sharing dataset using Python.

Data Loading

- Imported the dataset using pandas.
- Loaded the file `london_merged.csv` into a DataFrame named `bikes` for exploration.

Initial Data Exploration

- Used `bikes.head()` to preview the first few rows and understand the structure.
- Executed `bikes.info()` to inspect column data types, ranges, and identify potential issues.
- Examined categorical distributions using:
 - `bikes.weather_code.value_counts()`
 - `bikes.season.value_counts()`

Column Standardization

To make the dataset more readable and analysis-friendly, column names were standardized:

- A dictionary (`new_cols_dict`) was created to map original column names to clear, descriptive labels, such as:

- cnt → count
- t1 → temp_real_C
- t2 → temp_feels_like_C
- hum → humidity_percent
- wind_speed → wind_speed_kph
- Renamed all these columns using `bikes.rename()`.

Data Cleaning

- Converted humidity values from 0–100 scale to a 0–1 scale by dividing the column by 100.
- Ensured consistent data types by converting:
 - season column to string before applying mappings.
 - weather column to string before applying mappings.

Feature Labeling and Mapping

Two dictionaries were created to replace numeric categorical codes with meaningful labels:

1. Season mapping (season_dict):
 - 0 → spring
 - 1 → summer
 - 2 → autumn
 - 3 → winter
2. Weather mapping (weather_dict):
 - 1 → Clear
 - 2 → Scattered clouds
 - 3 → Broken clouds
 - 4 → Cloudy
 - 7 → Rain
 - 10 → Rain with thunderstorm
 - 26 → Snowfall

These mappings improved interpretability for subsequent analysis and visualization.

Exporting the Cleaned Dataset

- After all transformations, the cleaned dataset was saved as `london_bikes_final.csv` for further analysis and visualization.

4. Dashboard in Tableau

Finally, we built an interactive dashboard in Tableau to present insights visually.

