## **Description of Bike Count Prediction Data**

Bike-sharing systems are a means of renting bicycles where the process of obtaining membership, rental, and bike return is automated via a network of kiosk locations throughout a city. Using these systems, people are able to rent a bike from one location and return it to a different place on an as-needed basis. Currently, there are over 500 bike-sharing programs around the world.

The data generated by these systems make them attractive for researchers because of the duration of travel, departure location, arrival  
location, and time elapsed is explicitly recorded. Bike-sharing systems therefore function as a sensor network, which can be used for studying mobility in a city.

**Problem Statement**

In this project, you are asked to combine historical usage patterns with weather data in order to forecast hourly bike rental demand.

**Data**

You are provided with the following files:

1. train.csv : Use this dataset to train the model. This file contains all the weather-related features as well as the target variable “count”. The training dataset is comprised of the first 18 months.
2. test.csv : Use the trained model to predict the count of total rentals for each hour during the next 6 months.

<https://www.kaggle.com/datasets/brajeshmohapatra/bike-count-prediction-data-set>

**Data Dictionary**

Here is the description of all the variables :

|  |  |
| --- | --- |
| **Variable** | **Definition** |
| datetime | hourly date + timestamp |
| season | Type of season (1 = spring, 2 = summer, 3 = fall, 4 = winter) |
| holiday | whether the day is considered a holiday |
| workingday | whether the day is neither a weekend nor holiday |
| weather | weather |
| temp | temperature in Celsius |
| atemp | "feels like" temperature in Celsius |
| humidity | relative humidity |
| windspeed | wind speed |
| casual | number of non-registered user rentals initiated |
| registered | number of registered user rentals initiated |
| count | number of total rentals |