**Project Title: Prediction of London Bike sharing based on Past data**

**Dataset Link:** <https://www.kaggle.com/hmavrodiev/london-bike-sharing-dataset>**.** The dataset has 10 columns and 5000 rows which is suitable for big-data analysis

**Problem Statement:** With the cost-of-living increasing day by day, and congestion charges levied in many parts of London, the choice of using public transport makes sense to people. But what if you want to travel shorter distances, still have your independence, and maintain a healthy lifestyle – that’s where bike sharing started (is my assumption). Many organizations like Santander offer public bike-sharing schemes with several docking stations across London to help and encourage this lifestyle. However, it is also a challenge to maintain the requisite number of bikes. The goal of this project is to try and predict the bike share numbers using Machine Learning

**Aim & Objectives:**

The aim of this project is to try to predict future bike shares in London based on past information on bike sharing in London.

* To identify the best prediction model for the data collected
* Analysis of the previous data and visualize the hidden patterns using Python and Tableau
* Check the suitable machine learning classification type (Classification or Regression)
* Compare the models generated and evaluate the prediction performance

**Tools & Techniques:**

* Python
* Py-spark
* [Exploratory data analysis (EDA)](https://github.com/nindate/mlzoomcamp-capstone-project/blob/b4b32cd046becac7e7693a360bda7bbcd1bf07e9/#eda)
* Feature Engineering
* [Machine Learning model](https://github.com/nindate/mlzoomcamp-capstone-project/blob/b4b32cd046becac7e7693a360bda7bbcd1bf07e9/#baseline-model)s
  + [Linear Regression](https://github.com/nindate/mlzoomcamp-capstone-project/blob/b4b32cd046becac7e7693a360bda7bbcd1bf07e9/#linear-regression)
  + [Decision Tree](https://github.com/nindate/mlzoomcamp-capstone-project/blob/b4b32cd046becac7e7693a360bda7bbcd1bf07e9/#decision-tree)
  + [Random Forest](https://github.com/nindate/mlzoomcamp-capstone-project/blob/b4b32cd046becac7e7693a360bda7bbcd1bf07e9/#random-forest)
  + [XGBoost](https://github.com/nindate/mlzoomcamp-capstone-project/blob/b4b32cd046becac7e7693a360bda7bbcd1bf07e9/#xgb)