**Problem Statement:-**

**Predictive Modeling of Bike-Sharing Demand: Accelerating Revenue Post-Pandemic**

As the world anticipates a return to normalcy post-lockdown, US bike-sharing provider BoomBikes faces the challenge of reviving revenue amidst ongoing market disruptions. In collaboration with a consulting company, BoomBikes aims to decode the post-pandemic demand for shared bikes in the American market. Our mission is clear: to uncover the significant variables driving bike-sharing demand and develop a predictive model to guide strategic business decisions.

Through meticulous analysis of meteorological surveys and consumer behaviors, our project delves into the intricate dynamics of bike-sharing demand. By harnessing the power of data, we seek to unveil the nuanced relationships between various factors and bike demand patterns. Our ultimate goal? To equip BoomBikes with actionable insights that empower them to tailor their business strategy, meet customer expectations, and seize new market opportunities.

Join us on this journey as we unravel the mysteries of bike-sharing demand and pave the way for BoomBikes to thrive in a post-pandemic world. Let's ride towards a future of sustainable growth and innovation.

**Dataset Dictionary :-**

day.csv have the following fields:

- instant: record index

- dteday : date

- season : season (1:spring, 2:summer, 3:fall, 4:winter)

- yr : year (0: 2018, 1:2019)

- mnth : month ( 1 to 12)

- holiday : weather day is a holiday or not (extracted from http://dchr.dc.gov/page/holiday-schedule)

- weekday : day of the week

- workingday : if day is neither weekend nor holiday is 1, otherwise is 0.

+ weathersit :

- 1: Clear, Few clouds, Partly cloudy, Partly cloudy

- 2: Mist + Cloudy, Mist + Broken clouds, Mist + Few clouds, Mist

- 3: Light Snow, Light Rain + Thunderstorm + Scattered clouds, Light Rain + Scattered clouds

- 4: Heavy Rain + Ice Pallets + Thunderstorm + Mist, Snow + Fog

- temp : temperature in Celsius

- atemp: feeling temperature in Celsius

- hum: humidity

- windspeed: wind speed

- casual: count of casual users

- registered: count of registered users

- cnt: count of total rental bikes including both casual and registered

**Here is step by step approach :-**

1. **Data Preprocessing**: we've pre-processed our dataset by converting the date column to date time format, dropping redundant columns, and renaming columns for better understanding.
2. **Label Encoding and Visualization**: we've encoded categorical variables like season, weather situation, and working day, which must be text format were already encoded so we’ve changed them into categorical variables for analysis. Then, we've visualized the relationship between these variables and bike demand using bar plots.
3. **One-Hot Encoding**: Since we have categorical variables, we've performed one-hot encoding to convert them into a format suitable for regression analysis.
4. **Model Training**: we've split our data into training and testing sets, trained a linear regression model on the training data, and made predictions on the test data.
5. **Model Evaluation**: Finally, we’ve evaluated our model using the coefficient of determination (R-squared) to measure its performance in predicting bike demand.