**New York Citi Bikes Analysis**

Analyzing New York Citi Bike Usage Patterns for Data-Driven Insights

**Project Idea**

* Python
* Excel
* Power BI

**Tools Used**

* Data cleaning
* Exploratory Data Analysis (EDA)
* Data Visualization
* Trend Identification

**Tasks**

**Objective**

***The main objective of this project is to analyze NY Citi Bike usage patterns and extract meaningful insights to help stakeholders make data-driven decisions. This includes understanding user demographics, trip duration trends, popular pick-up/drop-off locations, peak usage times, and subscriber behavior. The analysis will assist in optimizing station placements, improving bike availability, and enhancing customer satisfaction.***

**Skills Used**

* Data Cleaning & Preprocessing (Handling missing values, filtering, and formatting data)
* Exploratory Data Analysis (EDA) (Statistical insights, summary statistics, trend analysis)
* Data Visualization (Creating charts, maps, and interactive dashboards)
* Business Intelligence (BI) Reporting (Developing insights for stakeholders)
* SQL Queries (Extracting and manipulating data from databases)
* Statistical Analysis (Understanding patterns and trends in the data)

**Tools & Technologies Used**

* Python (Pandas, NumPy, Matplotlib, Seaborn for data analysis and visualization)
* Power BI (For dashboard creation and interactive reports)
* Excel (Basic EDA, pivot tables, and data summaries)
* SQL (To query large datasets efficiently)

**Key Analysis Questions**

* What are the most popular pick-up locations across the city for Citi Bike rental?
* How does the average trip duration vary across different age groups?
* Which age group rents the most bikes?
* How does bike rental vary across the two user groups (one-time users vs long-term subscribers) on different days of the week?
* Does the factor of user age impact the average bike trip duration?

**Summary of Findings**

* The most popular pick-up locations are in business districts and tourist hotspots.
* The highest bike rental activity is observed during morning and evening rush hours.
* Annual subscribers tend to take shorter trips compared to casual users.
* Younger users (aged 20-35) dominate bike rentals, while older users take longer trips.
* Weekdays have higher usage among subscribers, while weekends see more casual users.
* Seasonality affects ridership, with peaks in summer and drops in winter.
* Certain stations experience congestion, requiring an increase in bike availability.

**Dashboard Overview**

The Power BI Dashboard includes:

* Heatmaps of pick-up and drop-off locations
* Line charts showing daily and hourly trends
* Bar charts for age group and gender distribution
* Pie charts comparing subscriber vs. one-time users
* Trend analysis for monthly usage patterns

**Conclusion & Recommendations**

* Optimize bike distribution by increasing bikes at high-demand stations.
* Expand Citi Bike stations in areas with high trip demand but fewer bikes.
* Targeted marketing for casual users to convert them into long-term subscribers.
* Improve bike availability during peak hours to prevent shortages.
* Seasonal promotions and discounts to encourage winter usage**.**

**Future Scope**

* Predictive Analytics: Forecasting demand using machine learning.
* Real-time Data Analysis: Monitoring live data for dynamic decision-making.
* Integration with Traffic Data: Understanding the impact of traffic patterns on bike rentals.