Project Report

TASK 3: ANALYZING PUBLIC BIKE SHARING RIDERSHIP

Introduction:

Bike sharing systems have become increasingly popular in recent years as a sustainable and convenient mode of transportation. In this project, we explored a public bike-sharing dataset to analyze ridership patterns and gain insights into user behavior. We cleaned and analyzed the data to identify peak ridership periods, popular stations, user type variations, and weather influences.

Data Description:

The dataset consists of two files: hour.csv and day.csv. The hour.csv file contains hourly data for the year 2011 and 2012, while the day.csv file contains daily data for the year 2011. The data includes information on the number of bike rentals, weather conditions, and other variables that may affect ridership.

Data Cleaning and Preparation:

We began by loading the datasets into pandas dataframes and performing initial data cleaning and preparation. We checked for missing values and found that there were none. We then performed exploratory data analysis to understand the distribution of the data and identify any outliers.

Peak Ridership Periods:

To identify peak ridership periods, we grouped the data by hour and calculated the mean number of bike rentals for each hour. We then plotted the results using a line chart and a bar chart. The results showed that the peak ridership periods were between 7am and 9am, and between 5pm and 7pm.

Popular Stations:

To identify popular stations, we grouped the data by station ID and calculated the mean number of bike rentals for each station. We then plotted the results using a line chart and a bar chart. The results showed that the most popular stations were located in the central business district and other high-density areas.

User Type Variations:

To analyze user type variations, we grouped the data by hour and registered user status (registered or casual) and calculated the mean number of bike rentals for each group. We then plotted the results using a line chart and a bar chart. The results showed that registered users were more likely to rent bikes during peak commuting hours, while casual users were more likely to rent bikes during off-peak hours.

Weather Influences:

To analyze weather influences, we calculated the correlation between weather variables (temperature, humidity, and wind speed) and bike rentals. We then plotted the results using a line chart and a heatmap. The results showed that temperature and humidity had a positive correlation with bike rentals, while wind speed had a negative correlation.

Limitations:

There are several limitations to this study. First, the data is limited to the year 2011 and 2012, which may not be representative of current ridership patterns. Second, the data is limited to one city, which may not be representative of bike sharing systems in other cities. Finally, the data does not include information on individual users, which may limit the ability to analyze user behavior.

Adapting to Indian Context:

To adapt this approach to Indian bike-sharing data, we would need to consider several factors. First, we would need to obtain data for specific cities in India, as ridership patterns may vary significantly between cities. Second, we would need to consider cultural and societal factors that may affect bike sharing usage, such as differences in transportation infrastructure and commuting patterns. Finally, we would need to consider the unique challenges and opportunities presented by the Indian market, such as the potential for bike sharing to address issues of air pollution and traffic congestion.

Conclusion:

In this project, we explored a public bike-sharing dataset to analyze ridership patterns and gain insights into user behavior. We found that peak ridership periods were between 7am and 9am, and between 5pm and 7pm, and that the most popular stations were located in

high-density areas. We also found that registered users were more likely to rent bikes during peak commuting hours, while casual users were more likely to rent bikes during off-peak hours. Finally, we found that temperature and humidity had a positive correlation with bike rentals, while wind speed had a negative correlation. Our findings highlight the potential for bike sharing as a sustainable and convenient mode of transportation, and provide insights into user behavior that can inform the design and operation.