**Dataset Description:**

The Cyclistic bike-share dataset contains information about bike rides, including details such as ride ID, rideable type, start and end timestamps, start and end station names, coordinates, and user type (member or casual).

**Dataset Features:**

ride\_id: Unique identifier for each ride.

rideable\_type: Type of bike used for the ride.

started\_at: Start timestamp of the ride.

ended\_at: End timestamp of the ride.

start\_station\_name: Name of the station where the ride started.

end\_station\_name: Name of the station where the ride ended.

start\_lat, start\_lng: Latitude and longitude of the starting station.

end\_lat, end\_lng: Latitude and longitude of the ending station.

member\_casual: Type of user (member or casual).

**Insights:**

The dataset provides insights into how Cyclistic bikes are used by different user groups.

Analysis of the dataset can help in optimizing bike availability and improving user experience.

**Problem Definition:**

The dataset had the following issues:

* Missing Values: Approximately 4.4% of the data contained missing values, which needed to be addressed to ensure accuracy.
* Data Type Definition: The data types of columns needed to be defined.
* Inconsistent Data: Some rows contained incomplete or irrelevant data, which needed to be removed.

These issues could affect the accuracy and reliability of future analysis and the functionality of the data.

**Methodology:**

**Data Cleaning:**

* Missing values were removed.
* Rows with ride length less than or equal to zero were removed.
* Incomplete rows were removed.

**Data Transformation:**

* Datetime columns were converted to the correct data type.
* Insightful columns such as day of the week, month, year, etc., were added.
* Ride distance and ride speed columns were added.

**Column Removal:**

Unnecessary columns like ride\_id, start\_station\_id, end\_station\_id, and bike\_id were dropped.

**Results:**

* The dataset was cleaned and processed to ensure accuracy and relevance for future analysis.
* Columns were defined with appropriate data types.
* Additional insightful columns such as day of the week, month, year, etc., were added for further analysis.