**Cyclistic Case Study NYC**

Data Analytics, BigQuery, and Tableau

**Links**

Github:

Tableau:

**Scenario**

Cyclistic, a fictitious bicycle rental company in partnership with NYC, operates numerous bike stations throughout the city. With a strategic focus on expanding its customer base, the Customer Growth Team is actively developing a comprehensive business plan for the upcoming year. The team’s primary objective is to gain a nuanced understanding of customer behavior and bike usage, precisely pinpointing high-demand stations.

In crafting the executive summary, the Customer Growth Team aims to present a concise yet insightful view of critical data points and aggregations. This strategic summary will provide leadership with a rapid and clear understanding of prevailing customer trends. By identifying the stations experiencing high demand, Cyclistic intends to make informed decisions that contribute to the overall growth and success of the business in the coming year.

**Problem Identification**

The central challenge for the Customer Growth Team is to identify and leverage high-demand locations for Cyclistic’s bike stations. This project will focus on addressing the following questions:

1. What patterns and distinctions exist in using Cyclistic bikes at high-demand locations compared to other stations?
2. How is Cyclistic Bike usage affected by the summer months?
3. What are the most popular starting neighborhoods?

By meticulously analyzing the available data, we aim to uncover key trends associated with high-demand locations. This analysis will provide valuable insights into customer behaviors and preferences specific to these locations, laying the groundwork for a nuanced understanding of the factors contributing to their heightened demand.

These insights will be pivotal in informing strategic decisions for the Customer Growth Team and Cyclistic’s leadership. The company can refine its operational and marketing strategies by identifying and capitalizing on the characteristics that drive demand at specific locations, ultimately fostering enhanced customer satisfaction and bolstering Cyclistic’s overall growth trajectory.

**ETL Processing**

**Data Sources**

Primary dataset: [NYC Citi Bike Trips](https://console.cloud.google.com/marketplace/details/city-of-new-york/nyc-citi-bike?project=coffee-test-379520)

Secondary dataset: [Census Bureau US Boundaries](https://console.cloud.google.com/marketplace/product/united-states-census-bureau/us-geographic-boundaries?project=coffee-test-379520)

The primary dataset consists of NYC Citi Bike Trips, supplemented by the secondary dataset from the Census Bureau delineating US Boundaries. The approach involves utilizing Citi Bike’s historical data spanning the last 12 months and merging it with the Census Bureau’s US Boundaries dataset to visualize bike-related information effectively and map it within the confines of NYC boundaries.

We executed the integration process through BigQuery, where we extracted data from Cyclistic’s datasets and unified the data into a single target table. This consolidated table is the foundation for developing a comprehensive dashboard in Tableau, facilitating streamlined data visualization and analysis.

SQL Query for Target Table:

A screenshot of a computer code

Description automatically generated

For the NYC Citi Bike Trips dataset (bigquery-public-data.new\_york\_citibike.citibike\_trips), we extracted the following field: user-type, start-time renamed as start\_day, stop-time renamed as stop\_day, trip-duration renamed as trip\_minutes, and COUNT(bikeid) renamed as trip count.

Simultaneously, we gathered information on zip\_code, borough, and neighborhood from the Census Bureau’s US Boundaries dataset (bigquery-public-data.geo\_us\_boundaries.zip\_codes), we gathered information on zip\_code, borough, and neighborhood.

Additionally, we incorporated data from the National Oceanic and Atmospheric Administration’s (NOAA) Global Surface Summary of the Day (GSOD) dataset (bigquery-public-data.noaa\_gsod.gsod20\*). Specifically, the fields extracted include temperature (temp) as day\_mean\_temperature, wind speed (wdsp) as day\_mean\_wind\_speed, and precipitation (prcp) as day\_total\_precipitation.

The resulting fields in the target table:

A screenshot of a computer

Description automatically generated

**Analysis**

As a BI developer focusing on user engagement and strategic decision-making for the Customer Growth Team, the first dashboard centers around seasonality trends and starting locations. It incorporates the Trip Totals chart and the Trip Counts by Starting Neighborhood table.

The second dashboard maps out bike usage for the Customer Growth Team. The primary objective of the maps is to provide visual insights into the seasonal trends of bike trips across New York City boroughs.

**Dashboard 1: Cyclistic Seasonality**

A screenshot of a graph

Description automatically generatedCyclistic Seasonality

**Dashboard 1: Cyclistic Seasonality**

this dashboard comprises several components:

**Trip Totals Chart:**

* Visualizes the total number of bike trips throughout 2014, distinguishing between customers and subscribers.
* Highlights that subscribers constitute a significantly more significant portion of Cyclistic’s user base than regular customers.
* Demonstrates a clear seasonal trend, indicating a surge in users during warmer months (May–October) and a decrease in colder months, aligning with the expected behavior of reduced bicycle usage in colder weather.

**Trip Counts by Starting Neighborhood Table:**

* Lists the total number of bike trips started in each neighborhood monthly for 2014, organized by borough and neighborhood.
* Utilizes a color gradient to emphasize monthly trip counts, with darker values representing higher trip counts.
* Ensures readability with light text on darker values, enhancing accessibility.
* The dashboard emphasized the significance of starting locations with key active stations identified in the Lower East Side, Chelsea, and Clinton neighborhoods.

This comprehensive dashboard empowers the Customer Growth Team to make informed decisions on station placement by leveraging insights into yearly usage trends and the importance of starting locations for bike trips.

**Dashboard 2: Cyclistic’s Summer trend**

A screenshot of a map

Description automatically generatedCyclistic’s summer trend

**Dashboard 2: Cyclistic Summer Trend**

 The dashboard comprises several components:

1. **Overview Map:**

* A large, comprehensive map showcasing each of the New York boroughs.
* Utilizes color coding to represent the intensity of bike trips, allowing quick identification of high-traffic areas.
* It enables the team to visualize the distribution of bike usage across the city at a glance.

1. **Neighborhood Comparison Table:**

* A tabular presentation detailing the number of trips and average trip durations for customers and subscribers in each neighborhood.
* Facilitates easy comparison and identification of neighborhoods with high customer engagement and potential growth opportunities.

1. **Monthly Focus Maps**:

* We dedicated three smaller maps to July, August, and September, historically exhibiting the highest bike traffic.
* Enables more granularity into seasonal effects on bike usage.
* It helps identify specific neighborhoods that experience increased activity during peak months.

This dashboard empowers the Customer Growth Team to discern patterns in bike usage throughout the year, enabling strategic decision-making regarding resource allocation during peak seasons.

**Conclusion:**

These dashboards provide a holistic understanding of user behavior, enabling the team to anticipate and capitalize on seasonal fluctuations. The insights gained can be transformative for the following reasons:

1. Optimized Resource Allocation: Understanding yearly trends allows for resource allocation that aligns with periods of heightened user activity, ensuring efficient use of resources during peak months.
2. Station Placement Precision: The emphasis on starting locations, as highlighted in both dashboards, provides a clear guide for where to place new bike stations, strategically catering to user demand.
3. Subscriber-Centric Approach: Recognizing the dominance of subscribers in user numbers informs a subscriber-centric approach, enabling tailored strategies to retain and expand this key user segment.

In conclusion, these dashboards empower the Customer Growth Team to navigate the dynamic landscape of bike usage in New York City precisely. The comprehensive insights into seasonal and yearly trends, coupled with the emphasis on starting locations, position the team to implement targeted growth strategies, fostering increased user engagement and ensuring the success of Cyclistic’s expansion initiatives.