Project Requirements Document: Cyclistic

BI Analyst: Frida Ekner

Client/Sponsor: Jamal Harris, Director, Customer Data

Purpose:
The Customer Growth Team at Cyclistic is currently crafting a business strategy for the upcoming year. Their primary focus is to gain a comprehensive understanding of customer bike usage, particularly pinpointing demand across various station locations. With a dataset comprising millions of rides, the team seeks a dashboard that consolidates crucial insights. They understand that business plans informed by customer behaviors tend to outperform those solely reliant on internal observations. Therefore, it's crucial that summarized and aggregated key data points are included, enabling the leadership team to gain a clear perspective on Cyclistic's customer usage patterns.

Key dependencies:
To proceed with this project, a customer dataset has to be accessed, which necessitates approval from the Director of Customer Data. Additionally, it's essential to obtain consent from the teams responsible for specific product data, such as bike trip duration and bike identification numbers, to ensure accurate interpretation of the data. The key contacts for approval are Adhira Patel, Megan Pirato, Rick Andersson and Tessa Blackwell.

Stakeholder requirements:

To ensure ongoing enhancement and efficient product promotion, the dashboard should aid Cyclistic decision-makers in comprehending customer bike usage patterns and demand across various locations, considering factors that could impact this demand variably over time.

- Create a table or map visualization to explore the starting and ending station locations, grouped by location. (Required)
- Develop a visualization indicating the popularity of destination (ending) locations based on total trip minutes. (Required)
- Generate a visualization highlighting trends from the summer of 2020. (Desired)
- Design a visualization illustrating the percentage growth in the number of trips year over year. (Required)
- Collect insights regarding station congestion. (Nice to have)
- Gather insights on the volume of trips across all starting and ending locations. (Required)
- Obtain insights on peak usage patterns by time of day, season, and the influence of weather. (Required)

Success criteria:

- **Measurable:** Every trip should be assessed based on its starting and ending points, duration, and factors such as time of day, season, and weather. For instance, does Cyclistic usage decrease during rainy weather? Is there a consistent demand for bikeshare despite weather conditions? Is there variation across different locations and user categories (subscribers vs. non-subscribers)?
- **Action-oriented:** These results should either validate or invalidate the hypothesis that user demand is influenced by location, time, season, and weather. Subsequently, the Cyclistic team will utilize this information to enhance future product development strategies.
- **Relevant:** All metrics must align with the central inquiry: How can we enhance the Cyclistic experience?
- **Time-bound:** Analyze data covering a minimum of one year to assess the impact of seasonality on usage patterns. Examining data across multiple months will capture fluctuations in usage, including peak and off-peak periods.

User journeys:

Cyclistic's primary goal is to enhance the bike-share experience for its customers. A detailed analysis of trip trends will enable decision-makers to delve deeper into how customers presently utilize Cyclistic bikes and identify areas for enhancing this experience.

Assumptions:

- The dataset contains latitude and longitude coordinates of stations but lacks additional geographic aggregation details such as zip codes, neighborhood names, or boroughs. The team will furnish a separate database containing this information.
- The weather data provided does not specify the timing of precipitation events; it's conceivable that precipitation occurred during off-peak hours on certain days. Nonetheless, for the dashboard's purposes, it's advisable to consider any precipitation that transpired on the day of the trip as potentially impactful.
- Initiating bike trips from a location will be unfeasible if there are no bikes accessible at a station. Therefore, it may be necessary to contemplate alternative factors affecting demand.

Roll-out plan:

The stakeholders have set a deadline of six weeks for the completion of the BI tool:

- Week 1: Assignment of the dataset. Validation of initial design for fields and BikeIDs to ensure compliance with requirements.
- Weeks 2-3: Development of SQL queries and ETL processes.
- Weeks 3-4: Refinement of SQL queries. Design of the dashboard. Initial review of the first draft with peers.
- Weeks 5-6: Development and testing of the dashboard.