Project Requirements Document: Cyclistic

Client/Sponsor: Jamal Harris, Director, Customer Data

Purpose: The purpose of this project is to create a dashboard that summarizes key insights based on a dataset including millions of rides. Since business plans driven by customer insights are more successful than plans driven by internal staff observations alone, Cyclistic's Customer Growth Team wants to use the generated dashboard to understand customer bike usage and identify customer demand at different station locations. The dashboard should include summarized and aggregated key data points that give the leadership team a clear picture of customer bike usage.

Key dependencies: This project requires a dataset of customer data, so the Director of Customer Data will need to approve the request. Approval should also be given by the teams that own specific product data including bike trip duration and bike identification numbers to validate that the data is being interpreted correctly. Primary contacts are Adhira Patel, Megan Pirato, Rick Andersson, and Tessa Blackwell.

Stakeholder requirements: Stakeholders want insight into how customers use the bikes and their demand in different locations, including factors that might influence demand at different times. Such insight is critical to stakeholders' decision making as it helps to continuously improve and effectively market products. (Requirements are prioritized as: R - Required, D - Desired, or N - Nice to have)

- A table or map visualization exploring starting and ending station locations, aggregated by location. R
- A visualization showing which destination (ending) locations are popular based on the total trip minutes. R
- A visualization that focuses on trends from the summer of 2015. D
- A visualization showing the percent growth in the number of trips year over year. R
- Gather insights about congestion at stations. N
- Gather insights about the number of trips across all starting and ending locations. R
- Gather insights about peak usage by time of day, season, and the impact of weather. R

Success criteria: SMART criteria are used to define clear objectives.

Specific: BI insights clearly identifies the specific characteristics of a successful product. They demonstrate how customers are currently using bikes and what impacts demand at station locations. **Measurable**: Each trip is evaluated based on start and end location, duration, and variables such as time of day, season, and weather. Insights evaluate locations by the number of start and end trips. They can also answer questions such as: Do customers use Cyclistic less when it rains? Or does the demand for bikeshare remain constant? Does it vary by location and user type (subscribers vs. non-subscribers)? **Action-oriented**: Results prove or disprove the theory that location, time, season, and weather affect user demand. The team will use this knowledge to refine future product development. **Relevant**: All metrics must support the primary question: How can we build a better Cyclistic experience? **Time-bound**: Analysis spans at least one year to see how seasonality affects usage, and captures peaks and valleys in usage.

User journeys: Cyclistic's main goal is to provide a better bike-share experience for its customers. A deeper dive into trip trends will help decision makers explore how customers are currently using Cyclistic bikes and how that experience can be improved.

Assumptions: The dataset includes latitude and longitude of the stations, but does not identify more geographic aggregation details such as zip code, neighborhood name, or borough. The team will provide a separate database with this data.

The weather data provided does not include the time of day the precipitation occurred; it's possible that some days the precipitation occurred during off-peak hours. However, for the purposes of this dashboard, it is assumed that any amount of precipitation that occurred on the day of the trip could have an impact.

If there are no bikes available at a station, it will be impossible to start a bike trip at that location. For such stations, other demand factors may need to be considered.

Compliance and privacy: The data must not include any personal data such as name, email address, phone number, or physical address. It is important that users are anonymous to prevent biases.

Accessibility: The dashboards should offer text alternatives including large print and text-to-speech.

Roll-out plan: The dashboard is expected to be completed in 6 weeks.

Week 1: Dataset assigned. Initial design for fields and BikeIDs validated to fit the requirements.

Weeks 2–3: SQL and ETL development

Weeks 3–4: Finalize SQL. Dashboard design. 1st draft review with peers.

Weeks 5–6: Dashboard development and testing