Project Requirements Document: Cyclistic

BI Analyst: Nachiket Dixit

Client/Sponsor: Jamal Harris, Director, Customer Data

Purpose:

Cyclistic's Customer Growth team wants to understand how customers use their bikes; their top priority is identifying customer demand at different station locations. The dataset includes millions of rides, so the team wants a dashboard that summarizes key insights. Business plans driven by customer insights are more successful than internal staff observations. The executive summary must include key data points outlined and aggregated for the leadership team to understand how customers use Cyclistic.

Key dependencies:

The project requires a dataset of customer data, so the Director of Customer Data will need to approve the request. The project also needs approval 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. The following team will work on this project.

- Adhira Patel, API Strategist
- Megan Pirato, Data Warehousing Specialist
- Rick Andersson, Manager, Data Governance
- Tessa Blackwell, Data Analyst
- Brianne Sand, Director, IT
- Shareefah Hakimi, Project Manager

The primary contacts are Adhira, Megan, Rick, and Tessa.

Stakeholder requirements:

The dashboard with the below requirements will be helpful for the Cyclistic decision-makers to understand how their customers are using the bikes and the demand at different locations.

- R - A table or map visualization exploring starting and ending station locations, aggregated by location.
- R - A visualization showing the popular destination locations based on the total tip

minutes.

- D - A visualization that focuses on trends from the summer of 2015.
- R - A visualization shows the percent growth in the number of trips yearly.
- N - Gather insights about congestion at stations.
- R - 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.

Success criteria:

Specific: How customers use bikes and what impacts demand at station locations.

Measurable: To understand trends, each trip should be evaluated using starting and ending location, duration, and variables such as time, season, and weather. For example, do customers use Cyclistic less when it rains? Or does bike-share demand stay consistent? Does this vary by location and user types (subscribers vs. non-subscribers)?

Action-Oriented: The outcomes must prove or disprove the theory that location, duration, season, and weather impact user demand. The Cyclist team will use this knowledge to refine future product development.

Relevant: All metrics must support the primary goal of improving the Cyclist experience.

Time-bound: Analyze data that spans at least one year to see how seasonality affects usage. Exploring data that spans multiple months will capture peaks and valleys in use.

User journeys:

The primary purpose of the Cyclist is to provide customers with a better bike experience. The insights into the trip trends will help decision-makers explore how customers use bikes and how that experience can be improved.

Assumptions:

The dataset includes the latitude and longitude of 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 does not include what time precipitation occurred; it’s possible that it precipitated during off-peak hours on some days. However, for this dashboard, I should assume any precipitation that happened on the day of the trip could have an impact.

Starting bike trips at a location will be impossible if no bikes are available at a station, so we might need to consider other factors for demand.

Compliance and privacy:

The data must not include personal info such as name, email, phone, or address. The user provides this data as part of their device activation, which is unnecessary for this project. It is paramount that the users be anonymized to avoid any bias.

Accessibility:

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

Roll-out plan:

The stakeholders have requested to complete the dashboard in 6 weeks.

- Week 1: Dataset assigned. The initial design for fields and BikeIDs was 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