Business Intelligence Project: Cyclistic

End-of-course project for Google Business Intelligence Professional Certificate

PART 1

Welcome to Cyclistic!

Congrats on your new job with the business intelligence team at Cyclistic, a fictional bike-share company in New York City. In order to provide your team with both BI business value and organizational data maturity, you will use your knowledge of the BI stages: capture, analyze, and monitor.

Your meeting notes

You recently attended a meeting with key stakeholders to gather details about this BI project. The following details are your notes from the meeting. Use the information they contain to complete the Stakeholder Requirements Document, Project Requirements Document, and Planning Document.

Project background:

Primary dataset: NYC Citi Bike Trips
Secondary dataset: Census Bureau US Boundaries

Cyclistic has partnered with the city of New York to provide shared bikes. Currently, there are bike stations located throughout Manhattan and neighboring boroughs. Customers are able to rent bikes for easy travel between stations at these locations. Cyclistic's Customer Growth Team is creating a business plan for next year. The team wants to understand how their customers are using their bikes; their top priority is identifying customer demand at different station locations.

Cyclistic has captured data points for every trip taken by their customers, including:

- Trip start time and location (station number, and its latitude/longitude)
- Trip end time and location (station number, and its latitude/longitude)
- The rented bike's identification number
- The type of customer (either a one-time customer, or a subscriber)

The dataset includes millions of rides, so the team wants a dashboard that summarizes key insights. Business plans that are driven by customer insights are more successful than plans driven by just internal staff observations. The executive summary must include key data points that are summarized and aggregated in order for the leadership team to get a clear vision of how customers are using Cyclistic.

Stakeholders:

- Sara Romero, VP, Marketing
- Ernest Cox, VP, Product Development
- Jamal Harris, Director, Customer Data
- Nina Locklear, Director, Procurement

Team members:

- 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
  *Primary contacts are Adhira, Megan, Rick, and Tessa.

Per Sara: Dashboard needs to be accessible, with large print and text-to-speech alternatives.

Project approvals and dependencies:

The datasets will include customer (user) data, which Jamal will need to approve. Also the project might need approval by the teams that own specific product data, including bike trip duration and bike identification numbers. So I need to make sure that stakeholders have data access to all datasets.

Project goal: Grow Cyclistic's Customer Base

Details from Ms. Romero:

- Understand what customers want, what makes a successful product, and how new stations might alleviate demand in different geographical areas.
- Understand how the current line of bikes are used.
- How can we apply customer usage insights to inform new station growth?
- The customer growth team wants to understand how different users (subscribers and non-subscribers) use our bikes. We'll want to investigate a large group of users to get a fair representation of users across locations and with low- to high-activity levels.
- Keep in mind users might use Cyclistic less when the weather is inclement. This should be visible in the dashboard.

The deliverables and metrics:

- A table or map visualization exploring starting and ending station locations, aggregated by location. I can use any location identifier, such as station, zip code, neighborhood, and/or borough. This should show the number of trips at starting locations.
- A visualization showing which destination (ending) locations are popular based on the total trip minutes.
- A visualization that focuses on trends from the summer of 2024.
- A visualization showing the percent growth in the number of trips year over year.
- Gather insights about congestion at stations.
- Gather insights about the number of trips across all starting and ending locations.
- Gather insights about peak usage by time of day, season, and the impact of weather.
  *Dashboard must be created in 6 weeks!

Measure success:

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 usage. Evaluate each trip on the number of rides per starting location and per day/month/year to understand trends. For example, do customers use Cyclistic less when it rains? Or does bikeshare demand stay consistent? Does this vary by location and user types (subscribers vs. nonsubscribers)? Use these outcomes to find out more about what impacts customer demand.

Other considerations:

The dataset includes 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 provided does not include what time precipitation occurred; it's possible that on some days, it precipitated during off-peak hours. However, for the purpose of this dashboard, I should assume any amount of precipitation that occurred on the day of the trip could have an impact. Starting bike trips at a location will be impossible if there are no bikes available at a station, so we might need to consider other factors for demand. Finally, the data must not include any personal info (name, email, phone, address). Personal info is not necessary for this project. Anonymize users to avoid bias and protect their privacy.

People with dashboard-viewing privileges:

Adhira, Brianne, Ernest, Jamal, Megan, Nina, Rick, Shareefah, Sara, Tessa

Roll-out:

- 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

Questions:

- How were bikes used by our customers?
- How can we apply insights from the data generated by trip data?