Project Planning

Course 1 workplace scenario overview: Cyclistic

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

My meeting notes: I recently attended a meeting with key stakeholders to gather details about this BI project. My task is to use the information 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.

Project goal: Grow Cyclistic's Customer Base

- 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 2015.
- 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!

Activity: Complete the business intelligence project documents for Cyclistic

In this activity, I will complete the three key business intelligence project documents: the Stakeholder Requirements Document, the Project Requirements Document, and the Strategy Document.

Stakeholder Requirements Document: Cyclistic

BI Professional: Ishan Perera

Client/Sponsor: Director, Customer Data

Business problem: *(What is the primary question to be answered or problem to be solved?)*

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 Primary question: How can we apply customer usage insights to inform new station growth?

Stakeholders: *(Who are the major stakeholders of this project, and what are their job titles?)*

- VP, Marketing
- VP, Product Development
- Director, Customer Data
- Director, Procurement

Stakeholder usage details: *(How will the stakeholders use the BI tool?)*

To effectively develop new station locations, the team wants to understand how customers use the current line of bikes. They will use this BI tool in order to gain insights related to data generated by the bikes when being used by customers. Then, this information will be used to understand what customers want, what makes a successful product, and how new stations might alleviate demand in different geographical areas.

Primary requirements: *(What requirements must be met by this BI tool in order for this project to be successful?)*

- A table or map visualization exploring starting and ending station locations, aggregated by location.
- 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 2015.
- 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.

Project Requirements Document: Cyclistic

BI Analyst: Ishan Perera

Client/Sponsor: Director, Customer Data

Purpose: *(Briefly describe why the project is happening and why the company should invest resources in it.)*
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. 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 view 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.

Key dependencies: *(Detail the major elements of this project. Include the team, primary contacts, and expected deliverables.)*

This project will require 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. The primary contacts are Adhira Patel, Megan Pirato, Rick Andersson, and Tessa Blackwell.

Stakeholder requirements: *(List the established stakeholder requirements, based on the Stakeholder Requirements Document. Prioritize the requirements as: R - required, D - desired, or N - nice to have.)*

In order to continuously improve and effectively market products, the dashboard must help Cyclistic decision-makers understand how their customers are using the bikes and the demand at different locations, including factors that might influence that demand at different times.

- 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: *(Clarify what success looks like for this project. Include explicit statements about how to measure success. Use SMART criteria.)*

- **Specific:** BI insights must clearly identify the specific characteristics of a successful product. They must demonstrate how customers are currently using bikes and what impacts demand at station locations.
- **Measurable:** Each trip should be evaluated using starting and ending location, duration, variables such as time of day, season, and weather. 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. non-subscribers)?
- **Action-oriented:** These outcomes must prove or disprove the theory that location, time, season, and weather impact user demand. Then, the Cyclistic 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:** 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.

User journeys: (Document the current user experience and the ideal future experience.)

The main purpose of Cyclistic is to provide customers with a better bike-share experience. 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: (Explicitly and clearly state any assumptions you are making.)

The dataset includes latitude and longitude of stations but does not identify more geographic aggregation details like 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, you 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.

Compliance and privacy: (Include compliance, privacy, or legal dimensions to consider.)

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

Accessibility: (List key considerations for creating accessible reports for all users.)

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

Roll-out plan: (Detail the expected scope, priorities and timeline.)

The stakeholders have requested a completed BI tool in six weeks:

- Week 1: Dataset assigned. Initial design for fields and BikeIDs validated to fit the requirements.
- Weeks 2-3: SQL & ETL development
- Weeks 3-4: Finalize SQL. Dashboard design. 1st draft review with peers.
- Weeks 5-6: Dashboard development and testing

Strategy Document: Cyclistic

Proposer: Director, Customer Data

Status: Draft

Primary dataset: NYC Citi Bike Trips

Secondary dataset: Census Bureau US Boundaries

User Profiles *[Who is the intended audience for this dashboard? How do you expect them to use this dashboard?]*

VP, Marketing

VP, Product Development

Director, Customer Data

Director, Procurement

API Strategist

Data Warehousing Specialist

Manager, Data Governance

Data Analyst

Director, IT

Project Manager

Dashboard Functionality

Dashboard Feature	Your Request
Reference dashboard *[Should this dashboard be modeled on an existing dashboard? If so, provide a link and describe the similarity.]*	Build a new dashboard to display the starting and ending locations, aggregated by location. This should show the number of trips at starting locations.
Access *[How should access to the dashboard be limited? Who needs to have access?]*	Access will be provided as read-only to the user profiles listed in this document.
Scope *[What data should be included (or excluded) in this dashboard?]*	Fields include: station, zip code, neighborhood, and/or borough, year, month, trip count, weather

Date filters and granularity *[Should the dashboard include date filters? If so, what time frame should be displayed by default? Should the dashboard include a "Granularity" drop-down? If so, what granularity should be selected by default?]*	Data filters can be applied for the following: Date, Month, Year Granularity: Any chart with user detail metrics should have the ability to click on that metric to view specific information.

Metrics and Charts

Chart 1

Chart Feature	**Your Request**
Chart title	Trip Totals
Chart type *[What type of chart needs to be created? This could include any chart type, including a line chart (timeseries), bar chart, or table.]*	Line
Dimension(s) *[What dimensions does this chart need to include?]*	Date
Metric(s) *[What metrics are relevant to this chart?]*	Trip count

Chart 2

Chart Feature	**Your Request**
Chart title	Trip Counts by Starting Neighborhood
Chart type *[What type of chart needs to be created? This could include any chart type, including a line chart (timeseries), bar chart, or table.]*	Table

Dimension(s) *[What dimensions does this chart need to include?]*	Neighborhood, month
Metric(s) *[What metrics are relevant to this chart?]*	Trip count

Chart 3

Chart Feature	**Your Request**
Chart title	Total Trip Minutes by Destination
Chart type *[What type of chart needs to be created? This could include any chart type, including a line chart (timeseries), bar chart, or table.]*	Bar
Dimension(s) *[What dimensions does this chart need to include?]*	Zip code end, borough end, neighborhood end, user type
Metric(s) *[What metrics are relevant to this chart?]*	Trip minutes

Chart 4

Chart Feature	**Your Request**
Chart title	Average Time to Arrive
Chart type *[What type of chart needs to be created? This could include any chart type, including a line chart (timeseries), bar chart, or table.]*	Table
Dimension(s)	Zip code end, borough end, neighborhood end, start day, grand total

[What dimensions does this chart need to include?]	
Metric(s) *[What metrics are relevant to this chart?]*	Trip minutes

Chart 5

Chart Feature	**Your Request**
Chart title	Seasonal trends
Chart type *[What type of chart needs to be created? This could include any chart type, including a line chart (timeseries), bar chart, or table.]*	Map
Dimension(s) *[What dimensions does this chart need to include?]*	Neighborhood start, neighborhood end, number of rides, average trip duration, weather
Metric(s) *[What metrics are relevant to this chart?]*	Trip minutes, weather, number of rides