Cyclistic Bike-Share: A Case Study
-Cedrick C. Sante

Case Study Roadmap - Ask

Guiding questions

● What is the problem you are trying to solve?
-How to convert casual riders to annual subscribers, by analyzing their difference as a user?

● How can your insights drive business decisions?
-By making the team aware of how annual subscribers use our service, which could help in making a marketing strategy to persuade casual riders to upgrade their membership.

Key tasks

1. Identify the business task
-Analyze the difference between annual subscribers and casual riders.

2. Consider key stakeholders
-Marketing team, and executives.

Deliverable
A clear statement of the business task
-Find the difference between annual subscribers and casual riders.

Case Study Roadmap - Prepare

Guiding questions

● Where is your data located?
-It's uploaded on their own website.

● How is the data organized?
-It's saved in CSV file, with columns as ride_id, rideable_type, started_at, ended_at, start_station_name, start_station_id, end_station_name, end_station_id, start_lat, start_lng, end_lat, end_lng, and member_casual.

● Are there issues with bias or credibility in this data? Does your data ROCCC?
-As per initial checking, there are no apparent issues regarding bias, or credibility in this data, as this is shared by the company itself. In a way, this data is not that comprehensive as it has some missing entries in fields regarding the stations, and even geolocation data, but every record has ride_id and membership, as far as initial checking goes.

● How are you addressing licensing, privacy, security, and accessibility?
-It's open-source, so no issues with using it for analysis, and it's stored in a zipped

folder, which makes it a bit secure, at least against changes easily, and it's in a CSV file which makes it interoperable for most data analysis tools.

● How did you verify the data's integrity?
-By making sure that it comes from the given link only, as it's shared by the company directly.

● How does it help you answer your question?
-It can help by analyzing their time of usage, and the concentration of location of the casual users, to focus marketing of annual membership in there.

● Are there any problems with the data?
-There are some missing entries, and it might be too big for spreadsheet, hence using database apps like BigQuery, or programming-based tools like R would be more practical.

Key tasks
1. Download data and store it appropriately.

2. Identify how it's organized.

3. Sort and filter the data.

4. Determine the credibility of the data.

Deliverable
A description of all data sources used
Data is downloaded from the company's website, in CSV file, which is open-source, and contains columns for riding sessions like id, time of usage, stations, geolocation data, and membership.

Case Study Roadmap - Process

Guiding questions
● What tools are you choosing and why?
-I'm planning to use R mainly for analysis, Tableau for some viz, and maybe Bigquery for data manipulation and merging of datasets.

● Have you ensured your data's integrity?
-In a way, yes, by making sure it all comes from just one source.

● What steps have you taken to ensure that your data is clean?
-By checking the columns which can only have a certain set of values like rideable_type, member_casual, and day_of_week.

● How can you verify that your data is clean and ready to analyze?
-By grouping up the data based on the previously mentioned columns, and checking if it follows the desired values.

● Have you documented your cleaning process so you can review and share those results?
-Yes, through screenshots and a notebook in R.

Key tasks
1. Check the data for errors.

2. Choose your tools.

3. Transform the data so you can work with it effectively.

4. Document the cleaning process.

Deliverable
Documentation of any cleaning or manipulation of data

Case Study Roadmap - Analyze

Guiding questions
● How should you organize your data to perform analysis on it?
-By making a pivot table or summary table containing average ride_length based on different members, and day of week, as well as number of trips per day. Additionally, calculating the mean of ride_length, knowing the max ride_length, and mode of day_of_week, of the dataset as a whole also helps.

● Has your data been properly formatted?
-Yes, for R automatically detects the proper data type of a column from imported CSV files.

● What surprises did you discover in the data?
-That there's a customer who's used a bike for 4 days, who probably just wasn't able to return it immediately, as he/she is a PWD, as a docked bike is rented. It's very far from the average usage of only 19 minutes.

● What trends or relationships did you find in the data?
-Bike rents go down as winter/rainy season approaches, and come back up as summer nears. There's also an increase in users during weekends. Bikes are also used longer by casuals as a collective group, which costs us more for the units' maintenance, even though they spend less for our service.

● How will these insights help answer your business questions?
-They would be aware of how we can improve the business's profitability, as the seasonality of the data could help in focusing the times of marketing the service.

● Have you documented your cleaning process so you can review and share those results?
-Yes, through screenshots and a notebook in R.

Key tasks
1. Aggregate your data so it's useful and accessible.

2. Organize and format your data.

3. Perform calculations.

4. Identify trends and relationships.

Deliverable
A summary of your analysis

Case Study Roadmap - Share

Guiding questions
● Were you able to answer the question of how annual members and casual riders use Cyclistic bikes differently?
-Yes, I was.

● What story does your data tell?
-That casual users, as a collective group, have a greater part in our user base, and make use of our bikes in a longer time.

● How do your findings relate to your original question?
-It tells us that casual members use our service differently, specifically longer than annual subscribers.

● Who is your audience? What is the best way to communicate with them?
-Audience are stakeholders such as VP of Sales, Marketing Manager, and some team members of fellow analysts. It's best to communicate with them with less technical things only a programmer or analyst like us would easily understand.

● Can data visualization help you share your findings?
-Yes, as it highlights further the difference and need for change.

● Is your presentation accessible to your audience?

-Yes, through a notebook in R.

Key tasks
1. Determine the best way to share your findings.

2. Create effective data visualizations.

3. Present your findings.

4. Ensure your work is accessible.

Deliverable
Supporting visualizations and key findings

Case Study Roadmap - Act

Guiding questions
● What is your final conclusion based on your analysis?
-It is therefore concluded that casual users, as a collective group, make use of our service longer and are greater in number than the annual subscribers. User base is also greater as summer approaches, during, and a few months after it. Weekends also increase the number of bike renters.

● How could your team and business apply your insights?
-Team could focus marketing around summer times, and make limited time promo offerings during weekends to convert casuals to annual members.

● What next steps would you or your stakeholders take based on your findings?
-We should ramp up marketing around summer times, and offer discounts for annual subs during the weekend.

● Is there additional data you could use to expand on your findings?
-Yes, like the area where to focus advertisements, like those where casual users are concentrated.

Key tasks
1. Create your portfolio.

2. Add your case study.

3. Practice presenting your case study to a friend or family member.

Deliverable

Your top three recommendations based on your analysis

-Marketing should be ramped up as summer approaches, during, and a few months after it.
-Promos like limited time offerings should be given during weekends.
-Though annual subscribers are greater in number, casual riders still needs to be converted, as that group could cost us more in the long run with maintenance of our bikes.