Cyclistic Bike-Share Analysis

1. Ask:

**Scenario:**

Cyclistic is a bike-share program featuring 5824 bicycles and 692 docking stations in Chicago. Cyclistic sets itself apart by also offering reclining bikes, hand tricycles, and cargo bikes, making bike-share more inclusive to people with disabilities and riders who can’t use a standard two-wheeled bike. Most riders opt for traditional bikes; about 8% of riders use the assistive options. Cyclistic users are more likely to ride for leisure, but about 30% use them to commute to work each day. Cyclistic provides 3 types of pricing plans: single-ride passes, full-day passes, and annual memberships. Customers who purchase single-ride or full-day passes are referred to as casual riders. Customers who purchase annual memberships are Cyclistic members. The director of marketing believes the company’s future success depends on maximizing the number of annual memberships. She notes that casual riders are already aware of the Cyclistic program and have chosen Cyclistic for their mobility needs.

**Problem Statement:**

Understand how casual riders and annual members use Cyclistic bikes differently. Design marketing strategies aimed at converting casual riders into annual members. Produce data-driven and actionable results to inform business decision making.

**Key Stakeholders:**

Lily Moreno, Cyclistic marketing analytics team, Cyclistic executive team.

**What is the problem we are trying to solve?**

Cyclistic is faced with an uncertain future and is no longer able to solely rely on their traditional marketing strategies of raising general awareness and appealing to a variety of needs with flexible price points and diverse product choices. In the interest of company growth, the director of marketing believes that Cyclistic should capitalize on the lucrative profit margins of annual subscribers by marketing to existing casual customers and persuading them to become yearly subscribers. If that strategy is plausible, a successfully executed marketing campaign would result in more sustainable long-term revenue. To that end, we need to analyze how and why Cyclistic casual bikers and members differ to weigh any evidence, opportunities, and barriers to any future conversion strategy.

1. Prepare:

**Data Source and organization**

We used Cyclistic’s historical trip data to analyze and identify trends. We downloaded the datasets from [here](https://divvy-tripdata.s3.amazonaws.com/index.html). The data has been made available by Motivate International Inc. under this [license](https://ride.divvybikes.com/data-license-agreement). Note that Cyclistic is a fictional entity and Divvy's open data is used for the purpose of this case study.

This is public data that we can use to explore how different customer types are using Cyclistic bikes. But the data-privacy issues prohibited us from using riders’ personally identifiable information. This means that we didn’t able to connect pass purchases to credit card numbers to determine if casual riders live in the Cyclistic service area or if they have purchased multiple single passes.

We'll be examining a period of 12 months, from the start of December 2020 to the end of November 2021. Each month has a separate comma separated value file with the same headings. Each record consists of a bike trip under the bike-sharing program composed of several features: a unique hash ID serving as the tables primary keys to identify each bike trip, the type of bike used, the type of customer (casual or member), details about the starting and ending docking station (name, ID, latitude, and longitude) and the datetime for when the bike was picked up and dropped off.

## Data Quality Assessment

Reliable: 10% of the records had one or more missing values and there were hundreds of duplicate primary keys, but each record was unique.

Original: The data is sourced by a first-party group (Motivate International Inc.) using their own sources without any third-party organization

Comprehensiveness: This dataset is comprehensive enough for enquiry, and mostly free of human error

Current: The time frame we’re examining gives us an up-to-date view

Cited: Most records are free of errors and represent factual bike trips in Chicago

## Ethical Concerns

* Security: The data is in a cloud repository that is managed by authorized Motivate employees and is secure by a trusted cloud services provider (Amazon Web Services)
* Accessibility: All Cyclistic's bike ride records are open to the public with minimal constraint
* Privacy: Sufficient anonymization has been applied through de-identification, and the data is free of any personally identifiable information
* Licensing: This dataset is constrained by these [license terms](https://www.divvybikes.com/data-license-agreement)
* Credibility: The good faith of the provider of data (Motivate International Inc.) can be relied on to ensure that the data represents what it's supposed to represent, and there is no evidence of intentional misrepresentation

## Data Integrity

The accuracy and consistency of the data over its lifetime has, for the most part, not been compromised. Inconsistencies are present from the time series nature of the data

## Data Purpose

This dataset will assist in finding trends and insights that outlines how Cyclistic bikes are being used by different user types

## Filtering/Sorting

The data is already sorted by the started at and ended at fields. Any records outside of the December 2020 - December 2021 date range have not been included.

## Assumptions:

We assume that the data collection process was accomplished with integrity. The accuracy, bias, and credibility are sufficiently vetted by the first-hand source of the data. The original repository has never been accessed or modified in any unauthorized manner.

1. Process:

**Tools used**

For the pre-processing part we used Microsoft Excel and local storage to store the data sets.

**Transformation**

First, we created a folder to store the .csv files and choose to Save as an Excel Workbook file.

**Pre-processing**

* We opened each file on Excel and removed the duplicates if any and trimmed extra whitespaces.
* Then we added the ride\_length column and the day\_of\_week column.
* We calculated the length of each ride by subtracting the column “started\_at” from the column “ended\_at” (for example, =D2-C2) and format as HH:MM:SS using Format > Cells > Time > 37:30:55.
* We calculated the day of the week that each ride started using the “WEEKDAY” command (for example, =WEEKDAY(C2,1)) in each file. Format as General or as a number with no decimals, noting that 1 = Sunday and 7 = Saturday.
* Then we saved the files as csv formats for further analyzing.

GitHub link: <https://github.com/devoeop/Cyclistic_bike_share_analysis>

1. Share:

**Summary**

1. Casuals primarily use Cyclistic bikes for leisure. We make that assumption since casuals:

* Bike twice as much on Saturdays and Sundays compared to any other day of the week
* Spend significantly longer on average on every bike trip, suggesting that they spend time in-between docking stations doing leisurely activities
* Do not use Cyclistic bikes often enough to warrant paying for an annual membership

1. Members get more out of Cyclistic bikes by using them for leisure and commuting on a consistent basis. We've drawn that conclusion since members:

* Rely on bikes consistently each week and year-round, with no notable preference on a single day of the week.
* Use Cyclistic bikes often during the rush hours on a typical workday
* Are motivated by the economics of an annual membership pass

Tableau Dashboard link: [https://public.tableau.com/views/CyclisticBike-shareAnalysis\_16406314233060/CyclisticBikeshareAnalysis?:language=en-US&:display\_count=n&:origin=viz\_share\_link](https://public.tableau.com/views/CyclisticBike-shareAnalysis_16406314233060/CyclisticBike-shareAnalysis?:language=en-US&:display_count=n&:origin=viz_share_link)

1. Act:

Next Steps: Recommended Paths

A) Reiterating on Data Extraction and Analysis

To confirm that our speculations are true and to uncover any other important behavioural differences, we would need to survey a significant sample of our Cyclistic user population to find out what defines each user group's behaviour. This means relying on qualitative data that would provide insights like opinions and motivations (i.e., what do they use Cyclistic bikes for?) that would provide some much-needed context behind our initial findings. Finding out what can incentivize users will help us determine what to focus on in the marketing campaign. Are they concerned about saving money? Do they care about their carbon footprint impact? Do they want to commute quicker? How attached are they to other more convenient modern forms of transportation? Do they care about exercise? In addition, gathering more data on user demographics would provide more information into your typical casual or member that would give us a better indication of users’ willingness to pay.

The idea here is to get a good sense of the major obstacles and opportunities in front of Cyclistic which could affect any kind of major conversion strategy. For example, if Cyclistic finds out that most Casuals do not have enough disposable income that can be directed to an increase in commitment to the service, then any conceivable conversion strategy is dead in the water.

At the end of the day, a successful conversion relies on information that can be used to provide casual users with enough justification to welcome the change.

B) Forging Ahead with our Initial Findings

Unfortunately, the safest choice often requires the investment of more time and resources. If Cyclistic can tolerate a certain level of risk, it can use the findings in this analysis to kick-start and form the basis of a few potential strategies. To minimize any significant risks, any strategy discussed below requires a lot of due diligence to ensure their viability and test their impact before any full roll outs are approved. This means that the potential impact of any strategy must be scrutinized by considering the benefits, costs, and potential outcomes involved. Enacting techniques like focus groups will be essential if Cyclistic is serious about achieving a high conversion rate.

Top three recommendations moving forward:

1. Explore ways to convey the benefits of biking more frequently

Conveying the most important benefits of bike use by running a campaign on users’ health and on the environment for your typical Casual user and reaching more users as efficiently as possible. If Cyclistic can incentivize Casual users to increase their use of Cyclistic bikes more often, then these users could naturally be persuaded to upgrade their commitment into becoming members. Organizing cycling marathons on special holidays with monetary rewards or limited time ride passes.

1. Explore ways to reward annual passes with additional privileges to make it more lucrative

As it stands, Casuals and Members experience a relatively similar level of service, aside from the difference in bike trip times. They are given the same privileges on a first come first-serve service basis for every available bike type. By minimizing the risk of inconvenience for members, Casuals will bear the brunt of those painful experiences. For example, members could have the privilege of being able to see what bikes are available in real time and can reserve one ahead of time. This does mean that Casuals will lose the first come first-serve experience, and be forced to either deal with it, upgrade their commitment, or use a competing product.

Also giving users discounts on annual passes based on a certain distance covered by them or number of rides taken by them will make it more lucrative.

1. Launching new service and pricing models catering towards Casual users

Designing innovative ways to attract more people to use the bike service can lead to the desired results of executives to drive growth. One model can be like launching monthly passes since casual users use the bike service mainly during autumn and spring months. Also giving them good discount if they opt for the annual pass instead of monthly pass and based on the distance covered by them. Weekend passes is also another good model catering towards Casual users since Casual users use bikes extensively during weekends for long rides all year round.