Case study: How does a bike-share navigate speedy success?

Author: Ramit Chutani

Organization: Cyclistic (fictional bike-share company in Chicago)

Last Updated: September 2024

Executive Summary

Introduction

Cyclistic, a bike-share company in Chicago, aims to maximize annual memberships for future success. This report analyzes how casual riders and annual members use Cyclistic bikes differently to inform marketing strategies for converting casual riders into annual members.

Key Findings

1. Usage Patterns: Members (64% of trips) primarily use bikes for short commutes on weekdays, while casual riders (36% of trips) prefer longer, recreational rides on weekends.
2. Trip Duration: Casual riders average 24.58 minutes per trip, compared to 12.8 minutes for members.
3. Popular Locations: Casual riders frequent tourist attractions, while members use bikes in residential and commercial areas.
4. Peak Times: Member usage peaks during commute hours, while casual riders show more distributed usage throughout the day.

Recommendations

1. Targeted Incentives for Recreational Use:
    a. Partner with recreational facilities to offer member discounts
    b. Sponsor local events to engage casual riders
2. Targeted Awareness Campaigns for Commuters:
    a. Collaborate with local businesses to promote membership benefits
    b. Increase advertising in high-traffic areas for casual riders
3. Varied Membership Discounts:
    a. Introduce group and family membership bundles
    b. Implement an individual referral program with membership discounts

Next Steps

1. Review insights and recommendations with the marketing team
2. Develop targeted marketing campaigns based on findings
3. Present to executives for approval and resource allocation
4. Implement pilot marketing initiatives
5. Monitor and evaluate campaign performance

Through implementing these strategies, Cyclistic can effectively target casual riders, highlighting the benefits of membership and potentially increasing conversion rates. This approach aligns with the company's goal of maximizing annual memberships and driving future growth.

ASK Phase

Introduction

This report is an analysis and presentation of a case study as part of the **Google Data Analytics Professional Certificate** on Coursera. This case study **demonstrates the essential skills and knowledge of a data analyst**, from clearly defining a business task through all phases of data analysis, culminating in the effective reporting and publishing of data insights.

Case Study Scenario

*" In this case study, you work for a **fictional company, Cyclistic**, along with some key team members. You are a junior data analyst working on the marketing analyst team at Cyclistic, a bike-share company in Chicago. The director of marketing believes the company's future success depends on **maximizing the number of annual memberships**. Therefore, your team wants to understand how casual riders and annual members use Cyclistic bikes differently. From these insights, your team will design a new marketing strategy to convert casual riders into annual members. But first, Cyclistic executives must approve **your recommendations**, so they must be **backed up with compelling data insights and professional data visualizations**."*

Additional Information

So far, Cylistic's marketing strategy has focused on general awareness and broad consumer segments, offering flexible pricing plans that include single-ride passes, full-day passes, and annual memberships. While casual riders, who use single-ride or full-day passes, are attracted by this flexibility, finance analysts have found that annual members are more profitable. To drive growth, the Director of Marketing aims to convert casual riders into annual members, leveraging their existing familiarity with Cyclistic. To achieve this, the team will analyze historical bike trip data to understand the differences between these groups, explore what motivates casual riders to buy memberships, and determine how digital media can enhance marketing strategies.

Key Stakeholders

The key stakeholders include the **Cyclistic executive team**, who will decide whether to approve the recommended marketing strategies, **Lily Moreno, the Director of Marketing**, who is responsible for developing initiatives to convert casual riders into annual members, and the **Cyclistic marketing analytics team**, who provide the data analysis and insights that will shape and guide the marketing strategy.

These stakeholders are invested in leveraging data-driven insights to guide decision-making and improve profitability, operational efficiency and customer satisfaction.

Scope of Work

There are three guiding questions for the future marketing program:

1. **How do annual members and casual riders differ in their use of Cyclistic bikes?**
2. Why would casual riders buy a membership?
3. How can digital media influence the casual riders to become members based on answers from above?

Finding the answer to the first question is the focus of this case study. The findings of this report will inform the team members responsible for the remaining questions.

Statement of Business Task

"The business task is to analyze how casual riders and annual members use Cyclistic bikes differently to inform marketing strategies for converting casual riders into annual members."

Driving Business Decisions

By understanding the differences in how casual and member riders use Cyclistic bikes, we can tailor marketing campaigns specifically for casual riders, adjust or create pricing plans to better attract them, and optimize services such as bike availability and station placement to enhance overall satisfaction. This improved user experience can naturally lead to an increase in memberships. Additionally, understanding usage patterns allows the company to allocate resources more effectively, focusing on the most profitable and promising customer sub-segments.

PREPARE Phase

Data Source(s)

The data used in this analysis was sourced from the bike-sharing program's ride records. The data is hosted on a public website.This is a **primary internal** dataset containing trips details logged by Cyclistic.

Website link: https://divvy-tripdata.s3.amazonaws.com/index.html

The website provides Cyclistic trip data spanning from 2013 to the present. For this report, we will focus on analyzing the most recent trends by examining the **last 12 months of data, specifically from June 2023 to June 2024**. The data is organized in monthly CSV files, stored in archives.

Archive File Names:
“202306-divvy-tripdata.zip”, “202307-divvy-tripdata.zip”, “202308-divvy-tripdata.zip”, "202309-divvy-tripdata.zip", "202310-divvy-tripdata.zip", "202311-divvy-tripdata.zip", "202312-divvy-tripdata.zip", "202401-divvy-tripdata.zip", "202402-divvy-tripdata.zip", "202403-divvy-tripdata.zip", "202404-divvy-tripdata.zip", "202405-divvy-tripdata.zip", "202406-divvy-tripdata.zip"

The CSV files contain relevant trip details for each unique trip logged by Cyclistic for that time period.

Data Dictionary

The following table outlines the key variables used in this analysis. Each column variable's meaning and format are described below:

Variable Name	Description	Data Type	Data Format	Example
ride_id	Unique identifier for each bike trip	Character		"6F1682AC40EB6F71"
rideable_type	Type of bike used for the trip	Character		"electric bike"
started_at	Date and time when the trip started	Character	YYYY-MM-DD hh:mm:ss	"2023-06-05 13:34:12"
ended_at	Date and time when the trip ended	Character	YYYY-MM-DD hh:mm:ss	"2023-06-05 14:31:56"
start_station_name	Name of starting station for bike trip	Character		"Michigan Ave & 14th St"

start_station_id	ID of starting station for bike trip	Character		TA1307000124
end_station_name	Name of ending station for bike trip	Character		"Woodlawn Ave & 55th St"
end_station_id	ID of ending station for bike trip	Character		TA1307000164
start_lat	Latitude of the trip start location	Numeric		41.9
start_lng	Longitude of the trip start location	Numeric		-87.7
end_lat	Latitude of the trip end location	Numeric		42
end_lng	Longitude of the trip end location	Numeric		-87.6
member_casual	Indicates whether the user is a member or casual rider	Character		"member"

Analysis Objectives

The analysis focuses on understanding the distribution of casual and member users across all other variables, such as bike types. This report examines trip patterns, including start and end times, trip lengths, and round trips, to highlight how members and casual riders use the service differently. The study also assesses the influence of various factors such as seasonality and time of day on bike usage.

The goal is to provide evidence-based insights and recommendations that will inform Cyclistic's marketing strategies for converting casual riders into annual members.

Data Assessment

1. **Data can be considered reliable** because it consists of raw data facts about the trip, the station data is static and the time is logged automatically by the station hardware (assumption). Errors found will be documented in the process phase and corrected or removed.
2. This **data is original** as it is first hand information (internal data) processed by the various bike stations.
3. **Data is not comprehensive** as it is missing key trip data such as trip distance. Knowing the distances traveled by customers who use this service is an important factor when trying to answer the business question on how to differentiate casual users and members. Data does not include previous membership records or past usage history,

which can be important to see trends. It is possible this data is not available due to customer data privacy concerns.

4. **Data is current** and is being updated every month and put into tables. The data has been logged since 2013 up to current year (2024). This allows us to do a historic analysis on the past 12 month trends in bike usage because the data has been kept up to date.
5. **Data is cited** as it is clear where and when it was recorded. What each data point represents is clear using accompanying text files in the archive explaining the columns. The data table structures have changed over time which is reflected in the data dictionary files. The data is in CSV format and only stores data and not the history of changes in each cell.

As this data is sourced internally, and no other sources of data that can answer the business question exist, we proceed with the provided data, while stating that the **data has some validation and comprehensiveness concerns**. It is adequate to start answering the question of how casual and member customers use the service and what are the differences.

Licensing, Privacy, Security, and Accessibility

The data has been made available by Motivate International Inc. under the following license: [Data License Agreement | Divvy Bikes]

This is **public data** and can be used to explore how different customer types are using Cyclistic bikes. Due to data-privacy issues it is prohibited to use riders' personally identifiable information. Therefore it won't be possible to connect past purchases to credit card numbers etc. to determine if casual riders live in the Cyclistic service area or if they have purchased multiple single passes, as an example. No sensitive or personal data is stored, therefore there isn't a concern over customers having access to their data.

PROCESS Phase

Tools Used

The main software tool used for data viewing, cleaning, analysis and visualization is **Rstudio**. It is an IDE for the **R programming language**. The primary library used is the **tidyverse package**.

Other tools used:
1. Google Docs and Obsidian for outlines and note taking
2. ChatGPT and Claude for code debugging and assistance with writing tasks

Data Cleaning

In this step, the quality of the dataset was evaluated to check for errors and evaluate what processing is needed:

1. **Missing values:**
   Missing values were identified, particularly in columns such as start_station_name, end_station_name, and latitude/longitude columns. It was found that a significant portion of the data had empty values for station names and coordinates.
2. **Incorrect data type:**
   Dates and times are stored as character values, and a conversion was necessary to date-time data type to ensure accurate time-based analysis.
3. **Duplicate entries:**
   There are duplicate records found in the ride_id column. To be checked if trip rows were duplicated entirely, and if so these duplicates need to be removed from the dataset.

The following methods were applied to clean the data:

1. **Handling Missing Data:**
   Empty strings in character columns were converted to NA values using the na_if() function. Rows containing null values in critical columns such as start_station_name and end_station_name were removed to improve data quality.
2. **Type Conversion:**
   The columns started_at and ended_at were originally in character format. These were converted to proper date-time objects using the ymd_hms() function from the lubridate package, ensuring correct handling of temporal data.
3. **Removing Duplicates:**
   Duplicate rows were identified based on the ride_id column and removed using the distinct() function, retaining only unique trips in the dataset.

4. **String Trimming:**
   White space was trimmed from all character columns using the str_trim() function from the stringr package to ensure consistency in string-based data.

The following are the data transformation steps taken:

1. **New Data Columns:** The following columns were created to represent relevant information and categories for trips for deeper analysis and categorization:
   a. trip_length (another relevant statistic to analyze across member categories)
   b. day_of_week_started/ended (grouping columns)
   c. month_started/ended (grouping columns)
   d. year, date_started/ended (grouping columns)
   e. trip_duration_category (grouping column)

2. **Filtering Data**:
   To ensure consistency in the analysis, rows with null values in important columns such as latitude and longitude and station names were filtered out, leaving a cleaner dataset for further exploration.

Data Verification and Validation

The following validation steps were performed to ensure the data was accurate and consistent:

1. **Time Data Validation:**
   After converting started_at and ended_at to date-time objects, the absence of null values in these columns confirmed that all entries were valid date-time values.
2. **Location Data Validation:**
   The latitude and longitude columns were checked for feasibility, ensuring all coordinates were valid geographic locations.
3. **Validation of Factor Levels:**
   The rideable_type and member_casual columns were checked for invalid entries, confirming that all values were accurate and free from typos.

Data Processing Documentation

The data cleaning steps are documented in the 'Final R Script.R' file in the following GitHub repository.

Github Link: https://github.com/RamitChutani/BikeShareAnalysisProject/tree/main

ANALYZE Phase

Data Organization for Analysis

First, all the separate monthly data files (last 12 months) were combined into one master table. Before combining, a column consistency check was done to ensure compatibility and correctness. As stated previously in the data transformation section, new columns were created for better analyzing data to answer business questions.

After all data transformation and validation steps, the dataset we are left with has the following structure:

1. **4696424 observations (unique trip IDs/rows)**
2. **19 columns (including newly created columns)**

Data surprises

Here are some surprising data finds in the process of organizing and analyzing the dataset:

There was a lot of null data, mostly for station names and station ID columns:

1. start_station_name (and start_station_id) have 1049262 NA values each
2. end_station_name (and end_station_id) have 1104606 NA values each
3. end_lat/lng have 8808 NA values each
4. start_lat/lng is complete

Out of approximately **original 6.45 million observations,** it is around **16.7% null values** for station_name/id columns.

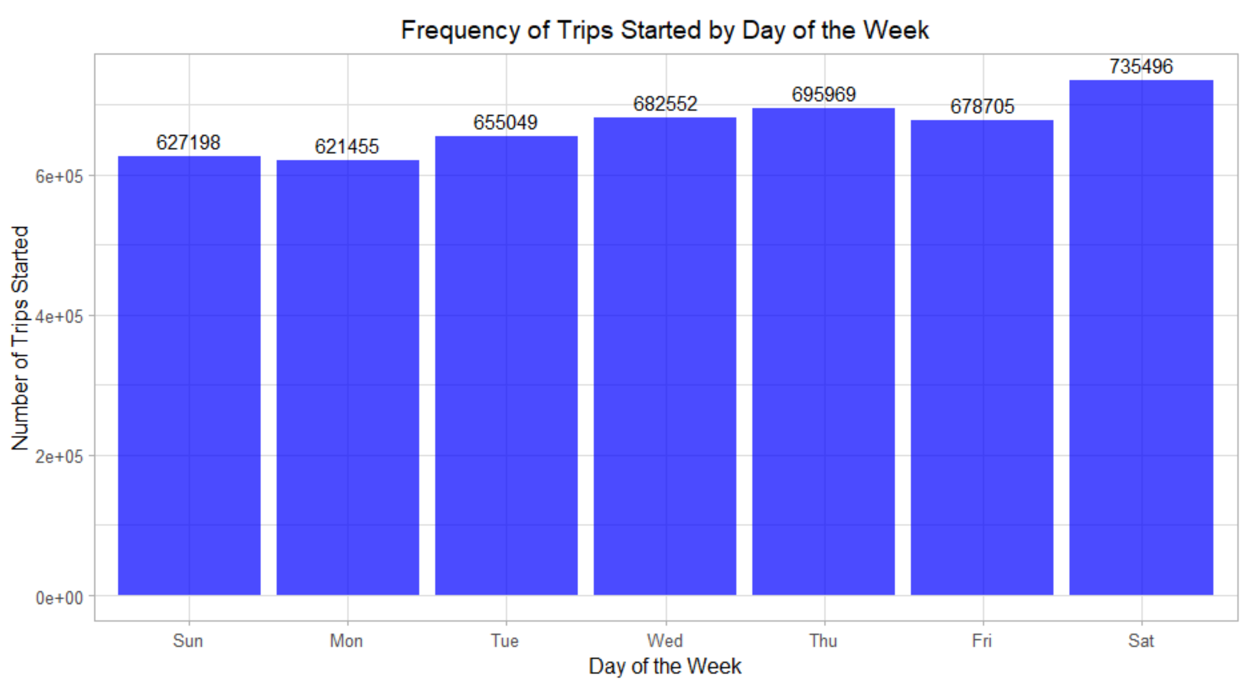
In the previous phase it was considered that the latitude and longitude data columns can be used for identifying the missing data, but surprisingly, the geographical data has varying values for the same station. Analysis found that the **geographical data** and station name data did not produce the same number of unique stations by an order of magnitude. Thus it **could not be used as a proxy for station names**.

Before removing this data, a sample bias check was considered. If the missing data rows are spread unevenly across time, it can lead to sampling error by removing it.

As shown in the following table, NA values are fairly uniformly distributed across each month of analysis, so **sampling bias is unlikely** to be introduced due to removal of null data.

Year Started	Month Started	Total Rows	Start NA Count	End NA Count	Start NA %	End NA %
2023	Jun	719,618	116,259	124,050	16.20%	17.20%
2023	Jul	767,650	122,943	130,304	16.00%	17.00%
2023	Aug	771,693	118,919	125,568	15.40%	16.30%
2023	Sep	666,371	101,312	107,291	15.20%	16.10%
2023	Oct	537,113	84,412	89,253	15.70%	16.60%
2023	Nov	362,518	55,003	57,899	15.20%	16.00%
2023	Dec	224,073	35,710	37,924	15.90%	16.90%
2024	Jan	144,873	19,165	20,749	13.20%	14.30%
2024	Feb	223,164	23,418	25,151	10.50%	11.30%
2024	Mar	301,687	44,255	47,140	14.70%	15.60%
2024	Apr	415,025	74,793	78,519	18.00%	18.90%
2024	May	609,704	109,071	112,811	17.90%	18.50%
2024	Jun	710,510	144,002	147,947	20.30%	20.80%

Another surprise was that the total number of trips did not peak on Sundays but instead peaked on Saturdays then followed by Thursdays. Also, the variance of the number of trips on different days of the week was surprisingly low, with a **peak of 735496 trips on Saturdays** and **minimum of 621455 trips on Mondays**.



Data Insights and Relationships

The dataset was examined using summary statistics to understand the spread and central tendencies of numerical variables such as trip durations and ride counts. Combined with relationship analysis with other column variables the following sections depict the differences in member and casual riders of Cyclistic bikes.

Member Status

As a percentage of all trips, the data shows that **~64% of all trips** recorded in the last 12 months are **member rider trips** and remaining **~36% are casual rider trips**.

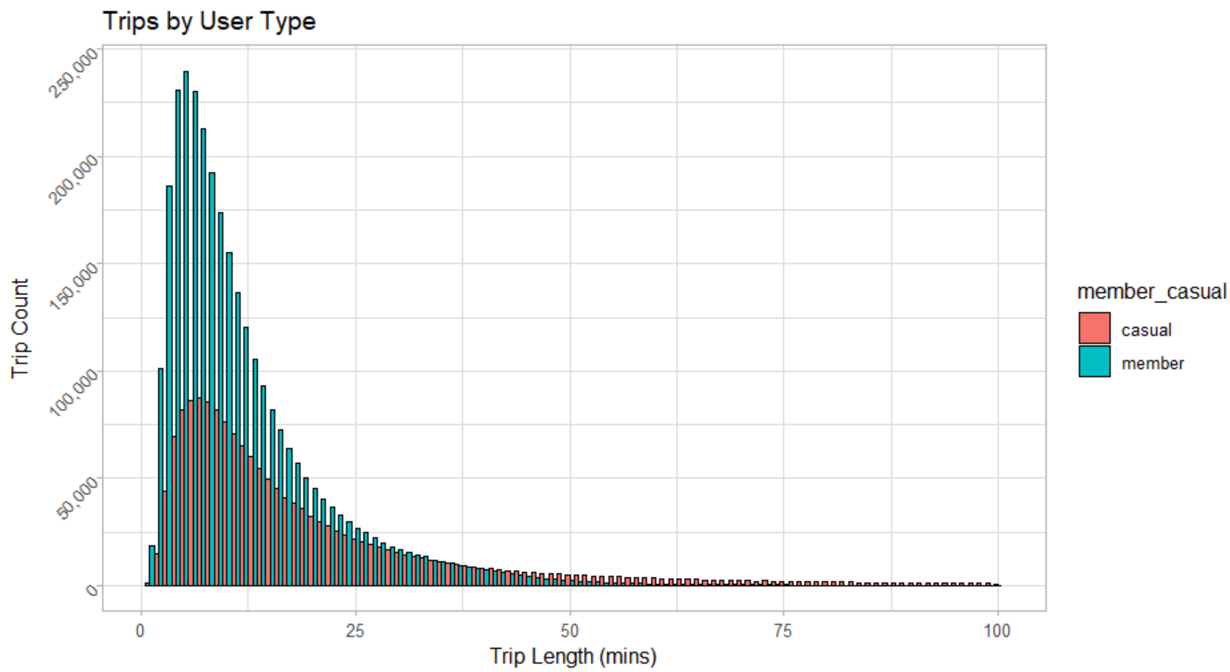
Trip Length

When checking for outliers in trip length, it was found that there were some negative trip lengths, which were removed as **invalid trips** and found to be majority zero min round trips which suggest that false starts were being logged by the system.

On the other hand, the maximum trip length was found to be 7-8 days, which can be explained as people on week-long trips or tourists using the service for their weeklong stay. After removing errors and invalid data, the **average trip length** is shown to be **~17 minutes** for all trips in the last 12 months. The difference in members and casual riders trip lengths is shown in the table below:

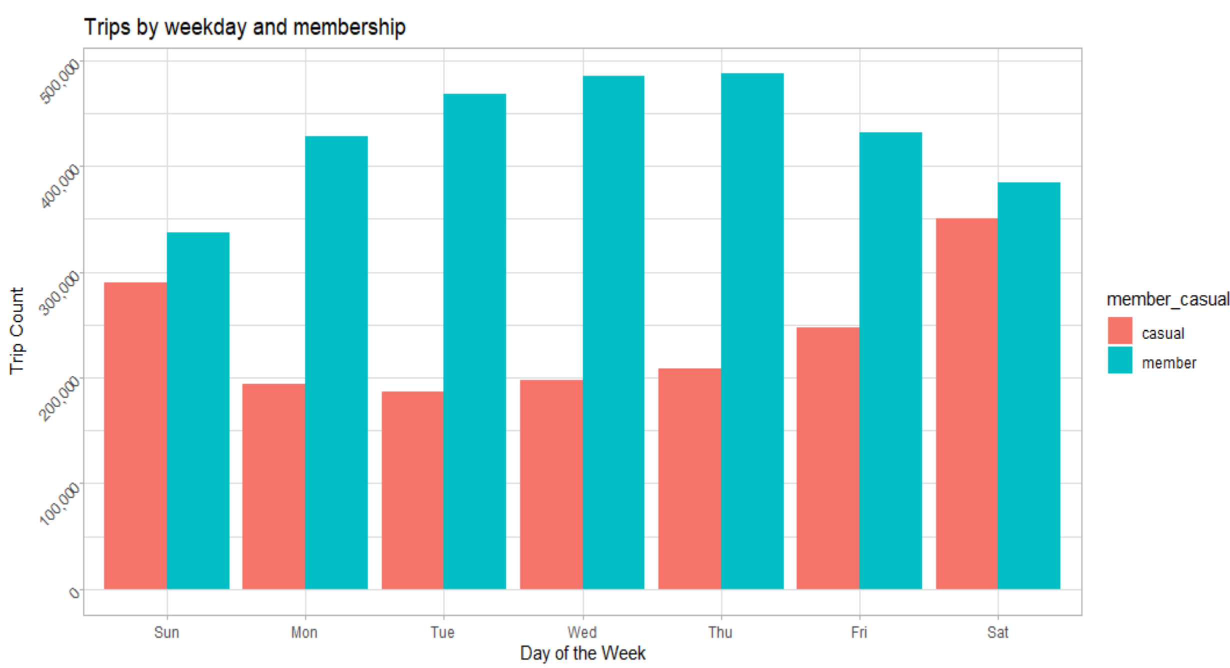
User Type	Trip Count	Mean Trip Length	Median	Min	Max	SD
Casual	1,675,390	**24.58**	13.75	0.52	11,152.27	52.75
Member	3,021,034	**12.8**	9.07	0.5	1,497.65	23.74

Majority of the trips are found to be less than 100 minutes. The following visualization shows these trips split between casual and member riders. It shows the difference in trip counts, and we can see both types of riders **peak trips** are around **8-10 mins**, with member counts peaking at above 250,000 trips and casual rider trips peaking at less than 100,000 trips. Also, **members** seem to use Cyclistic bikes for **mostly shorter trips** (less than 30 min), and the distribution and statistics show that **casual riders** tend to have longer trips on average. **Casual trip counts overtake** the member trip counts for **trip lengths greater than ~40 min**.



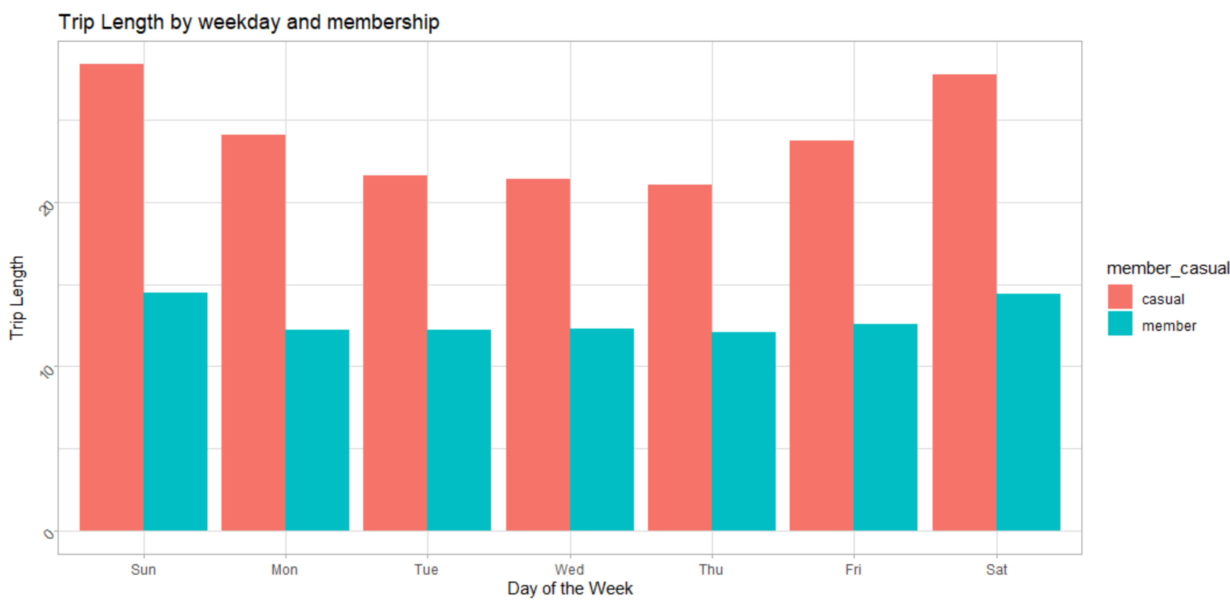
Day of the week

To look for differences in bike usage habits, we can look at trip counts by member status across the days of the week. The following histogram suggests that the trip counts for **members peak on the weekdays** and fall off on the weekends, following an **upside down U-shape distribution** from Sundays to Saturdays. The **casual rider trips show the reverse**, with a **U-shape distribution** and trips peaking in the weekends.



The next histogram shows the relationship of day of the week and average trip lengths, split between member types. This visual suggests that **trip lengths are fairly constant for member trips** at approximately **12 minutes** during weekdays and it bumps up slightly to 14-15 minutes on the weekends.

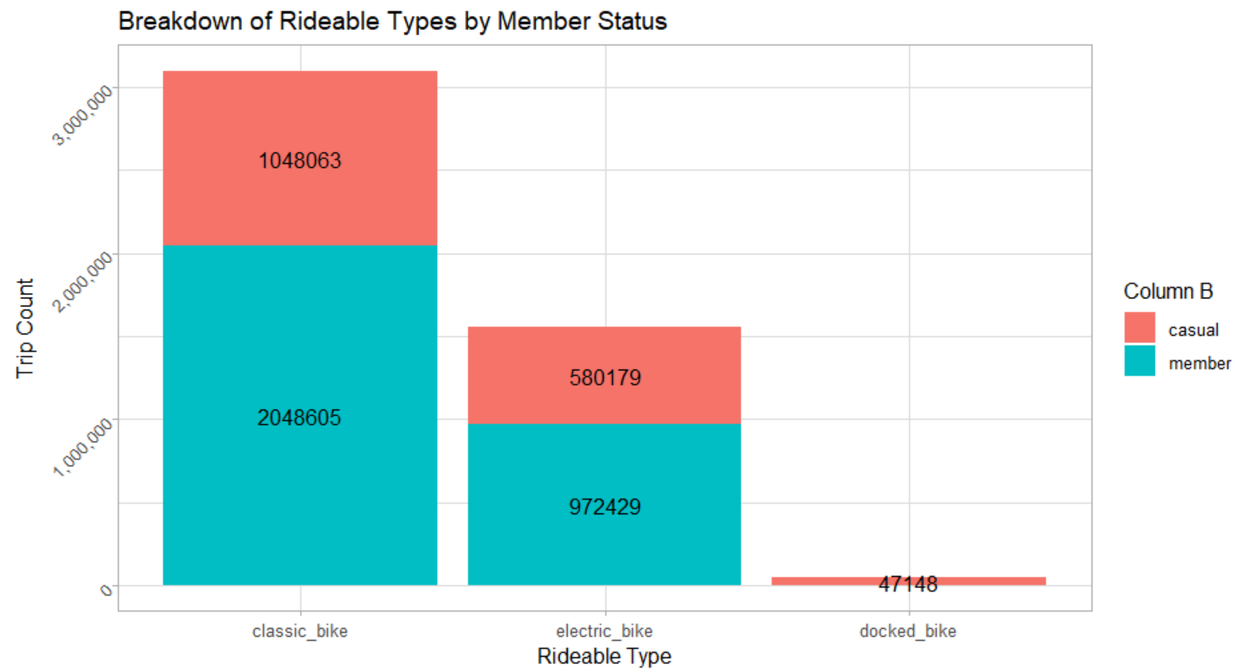
On the other hand, average trip lengths for **casual riders follow a U-shape distribution** from Sunday to Saturday, ranging **from ~21 minutes** in the middle of the week **to just under 30 minutes** on the weekends. This adds to the evidence that casual riders and member riders are using the Cyclistic bike service differently.



Rideable Type

Cyclistic bikes provides three kinds of bikes for users - classic bike, docked bike and electric bike - which is captured by the rideable_type column in the data. The following visual looks at the relationship between rideable type and and the number of trips, separated by user type, to see if any bike type is preferred by either user type.

The stacked bar graph breakdown shows that the **most popular/used bike type is the classic bike** type, with approximately 3 million trips. The second most used bike type is the electric bike, with 1.6 million trips. The last bike type is barely used and only has trips by casual riders, and only counts for ~1% of all trips.

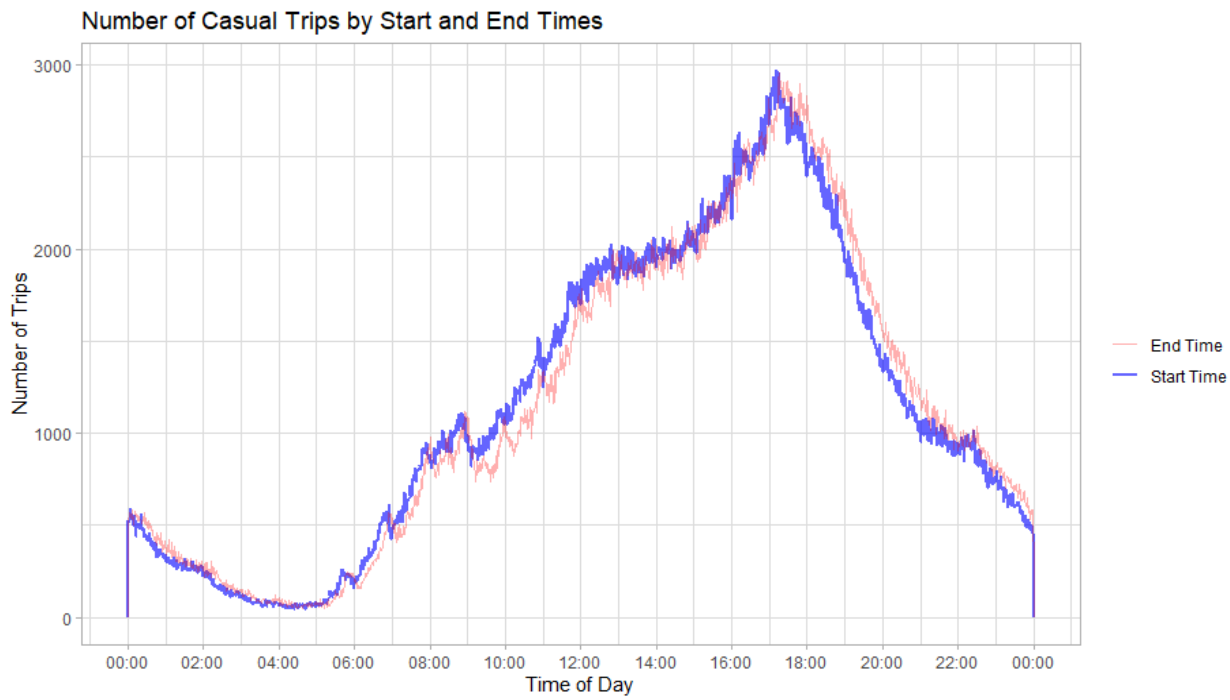


The percentage of casual trips for the two bike types is 33.8% and 37.4% which is consistent with the percentage of casual trips overall (36%). This usage data suggests that there is **no evidence** to say that one bike type is the **preferred bike type** for members or casual riders.

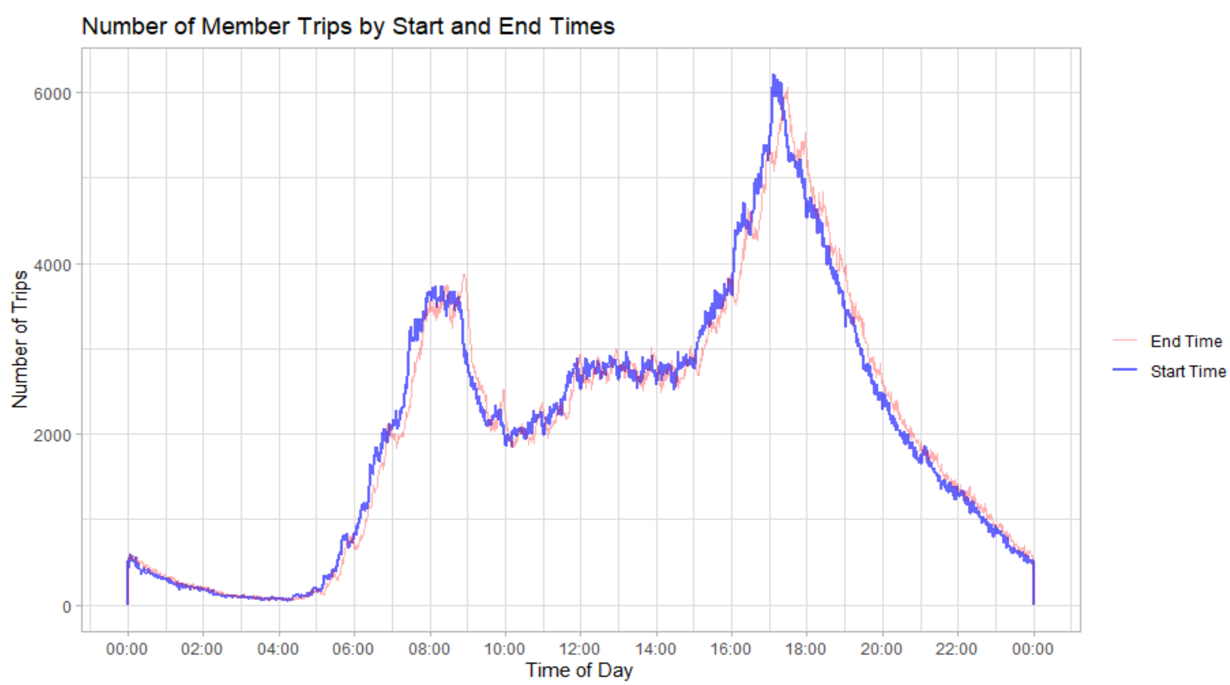
Trip Starting and Ending Times

The data includes time-logs for when each trip started and ended. In the following graphs it can be seen how member trips and casual trips differ in usage times. The following graphs represent the distribution of trip start and end times throughout the day for the two user types, with number of trips on the y-axis and time of day on the x-axis. As shown before, average trip length is around 13 and 25 minutes for member and casual riders respectively, which bears out in the **end-time red frequency line right-shifted from the start-time blue frequency line** by 30 minutes or less for the majority of the day. Also, trip volume for member trips is consistently higher than casual user trips which is likely due to data comprising 65% member trips (casual trips 35%).

Casual trips have a **morning trend upwards** - from 05:00 to 09:00 - and peaks at around 1000 casual trips every minute. There is another **noticeable rise in mid-day activity** which remains constant from 12:00 to 14:00, with around 2000 casual trips being recorded every minute. The **largest peak occurs in the early evening** - around 18:00 - almost reaching 3000 casual trips every minute. Overall, casual trips are spread out throughout the day with a clear preference for afternoon and evening trips. Morning travel is noticeable but less dominant comparatively.



Member trips show a different picture. There is a **sharp increase in morning trips** - around 06:00 - reaching a significant peak of almost 4000 trips every minute by 08:00. However, it is followed by a decrease and then **a plateau** to 3000 trips each minute **during midday**. Another **sharp peak** occurs in the **evening** - around 17:00 - with a volume of 6000 member trips logged each minute. After that the number of trips decreases steadily, similar to casual users. This suggests that member trips **follow normal commuter times**, with distinct peaks in morning and evening rush hours.



Overall, members show a more pronounced morning peak, likely indicating that many are commuting to work during this time, while casual users are likely having more flexible travel schedules. Midday shows a plateau for both user types, a decrease for members from morning peak but an increase for casual users. This possibly indicates recreational, leisure, or lunch-time trips. Both user types show strong evening peaks around 17:00-18:00, which is a common time for workday ending or evening activities.

Station Names, Routes and Round Trips

The following table is generated using data on starting station and ending station names, sorted by the **top ten routes for both casual and member users**, by number of trips. It highlights a significant difference in the average trip length: casual users typically take trips of 30 minutes or longer, while members average between 4 and 8 minutes.

User Type	Route	Trip Count	Avg. trip length (min)
casual	Streeter Dr & Grand Ave to Streeter Dr & Grand Ave	8023	47.25
casual	DuSable Lake Shore Dr & Monroe St to DuSable Lake Shore Dr & Monroe St	6772	41.63
casual	DuSable Lake Shore Dr & Monroe St to Streeter Dr & Grand Ave	5376	27.18
casual	Michigan Ave & Oak St to Michigan Ave & Oak St	4242	57.66
casual	Millennium Park to Millennium Park	2947	52.05
casual	Streeter Dr & Grand Ave to DuSable Lake Shore Dr & Monroe St	2821	28.13
casual	Dusable Harbor to Dusable Harbor	2656	42.12
casual	Montrose Harbor to Montrose Harbor	2564	52.34
casual	Dusable Harbor to Streeter Dr & Grand Ave	2451	26.97
casual	Shedd Aquarium to Streeter Dr & Grand Ave	2416	33.83
member	Calumet Ave & 33rd St to State St & 33rd St	5934	5.36
member	State St & 33rd St to Calumet Ave & 33rd St	5898	4.48
member	Ellis Ave & 60th St to University Ave & 57th St	4150	4.89
member	University Ave & 57th St to Ellis Ave & 60th St	4123	4.92
member	Ellis Ave & 60th St to Ellis Ave & 55th St	3938	5.28
member	Ellis Ave & 55th St to Ellis Ave & 60th St	3673	5.73
member	Loomis St & Lexington St to Morgan St & Polk St	3193	5.48
member	Morgan St & Polk St to Loomis St & Lexington St	2961	5.61
member	MLK Jr Dr & 29th St to State St & 33rd St	2584	8.61
member	State St & 33rd St to MLK Jr Dr & 29th St	2408	8.74

In terms of station usage, **casual** users' most popular routes seem to travel between **tourist or leisure destinations**, with names such as Millennium Park, Streeter Dr & Grand Ave, and DuSable Lake Shore Dr. In contrast, it is likely that **members** frequent stations in more **residential, commercial or academic areas** with station names like State St & 33rd St, Ellis Ave & 60th St, and University Ave & 57th St.

Out of 185,779 **round trips** (~4% of all trips), 123,508 were made by casual users and 62,271 by members, showing a **reversed ratio compared to overall trip patterns**. **Members**, who account for 65% of all trips but only 34% of round trips, likely **use bikes primarily for commuting**, making more one-way trips. Conversely, **casual users**, who make up 35% of total trips but 66% of round trips, could be **using bikes for recreational purposes**, where they begin and end at the same location.

Overall, this data supports the theory that casual users of Cyclistic bikes are more focused on recreation, while members are using the service more for commuting or daily routines.

Answering Business Question

1. **Trip Volume**: Members account for about 64% of all trips, while casual riders make up the remaining 36%.
2. **Trip Length**: Casual riders take longer trips on average (24.58 minutes) compared to members (12.8 minutes). Casual riders are more likely to take trips longer than 40 minutes.
3. **Weekly Usage Patterns**: Member trips peak on weekdays, following an inverted U-shape distribution from Sunday to Saturday. Casual rider trips show a U-shape distribution, peaking on weekends.
4. **Daily Usage Patterns**: Members show distinct peak usage during typical commute hours (around 8:00 and 17:00-18:00). Casual riders have a more spread out usage pattern throughout the day, with a significant peak in the early evening (around 18:00).
5. **Popular Routes and Stations**: Casual riders frequently use stations near tourist attractions and leisure destinations. Members tend to use stations in residential, commercial or academic areas.
6. **Round Trips**: Casual riders make a disproportionately high number of round trips (66% of all round trips, despite making only 35% of total trips). Members make fewer round trips, consistent with commuting behavior.
7. **Bike Type Preference**: There is no significant difference in bike type preference between members and casual riders.
8. **Trip Purpose**: Members' usage patterns suggest they primarily use bikes for commuting and daily routines. Casual riders' patterns indicate more recreational and leisure use.

These differences put together suggest that members primarily use Cyclistic bikes for regular, practical purposes like commuting, while casual riders tend to use them more for leisure and recreational activities.

SHARE Phase

Summary of Findings

Annual members and casual riders of Cyclistic bikes exhibit **distinct usage patterns**. Members account for approximately 64% of trips but tend to take shorter rides, averaging 12.8 minutes, mainly during weekdays and peak commuting hours. In contrast, casual riders, who make up 36% of trips, have longer average trip durations (24.58 minutes) and prefer weekend usage, often for recreational purposes. They are more likely to undertake longer trips and round trips, despite their lower overall trip volume. Members generally utilize bikes for commuting, while casual riders frequent stations near tourist attractions for leisure activities. Overall, the data suggests that **members prioritize practical use**, whereas **casual riders engage in more leisurely biking.**

Data Story

The analysis of Cyclistic bike usage reveals a compelling narrative about how annual members and casual riders differ in their cycling habits. **Members**, constituting the majority of all trips, **primarily use bikes for commuting**, reflected in their short average trip duration, which peaks during weekday rush hours. This pattern illustrates a practical, routine-oriented approach to biking, **as they navigate their daily schedules.**

Casual riders tend to enjoy longer rides, averaging twice as much as average member trips. Their usage peaks on weekends, indicating **a preference for recreational biking**. These writers frequently choose routes that lead to tourist attractions and leisure destinations, **emphasizing a more relaxed and leisurely biking experience**. Moreover, casual riders are notable for taking more round trips, despite representing a third of total trips. This suggests that while they may cycle less frequently, when they do, it's often for longer, more exploratory rides.

Acknowledging Overlaps

It’s essential to recognize that **not all member trips are exclusively for commuting, nor are all casual trips solely for recreation**. Some members may also engage in leisurely rides, and casual riders may use bikes for practical purposes. Understanding this nuance is critical, especially as the analysis aims to find strategies to convert casual riders into members. By identifying the shared motivations and preferences across both groups, Cyclistic can develop targeted marketing strategies and incentives that appeal to casual riders, encouraging them to embrace more frequent and varied bike usage, similar to that of annual members.

ACT Phase

Conclusion

The analysis reveals that annual members and casual riders have distinct usage patterns. A casual rider is more likely using the bike service for longer trips with recreational destinations. A member rider is more likely to be using Cyclistic for short trips concentrated at normal commuting time. This differentiation opens up opportunities to convert casual users to members by addressing what is learned about casual users needs and preferences.

Recommendations

The recommendations to the Cyclistic marketing team on how they can leverage the data insights to convert casual riders into Cyclistic members are as follows:

1. **Targeted Incentives for Recreational Use**: First recommendation is based on incentivizing casual users preference of using Cyclistic for recreation.
    a. **Partnerships with Recreational Facilities -** Collaborate with gyms, parks, and leisure centers to offer discounts or special promotions for Cyclistic members. This could incentivize casual riders who primarily used bikes for leisure to consider membership as a way to save money while enjoying recreational activities.
    b. **Member Events and Sponsorships:** Sponsor local recreational events, such as marathons or community bike rides, to engage casual riders and promote membership benefits, showcasing how Cyclistic can enhance their leisure activities. Introduce member events like guided bike tours on weekends, appealing to casual riders who use Cyclistic for long weekend sightseeing trips.

2. **Targeted Awareness Campaigns for Commuters**: Second method for conversion is to find casual riders who commute but have yet to buy annual membership due to lack of awareness.
    a. **Employer Collaborations:** Partner with local businesses to raise awareness about Cyclistic membership among employees. Offer presentations, brochures, or information sessions about the benefits of cycling for commuting, including potential savings and convenience.
    b. **Targeted Advertising:** Increase advertising in areas with high casual rider traffic, such as commercial and academic stations. Use eye-catching graphics and clear messaging about the advantages of membership, especially for those who may not be aware of Cyclistic's offerings.

3. **Varied Membership Discounts**: Third approach appeals to riders' preference for recreation in groups and financial benefits of discounted membership plans.
    a. **Group and Family Memberships Bundles:** Create grouped membership plans that allow casual riders to sign up with friends or family at a discounted rate. This targets recreational riders who often bike with others, making membership more appealing for group activities. Cyclistic can market this as a "Ride Together, Save Together" campaign, promoting the social and financial benefits of membership. The focus here is on enhancing both cost savings and the group experience.
    b. **Individual Referral Program:** Launch a referral program where any individual rider (whether casual or member) can refer friends or family and earn a discount on their own membership. Possibly, the more people they refer, the greater their discount, making this a personal incentive for casual riders to convert into members and encourage others to do the same.

Additional data to expand findings

To enhance the analysis and provide deeper insights into the behaviors of casual riders and annual members, access to the following additional data would be valuable:

1. User Demographics:
   Knowing the **demographic breakdown of Cyclistic users** could help identify specific segments that are more likely to be casual or member riders. For example, younger users may prioritize recreational trips, while older, working professionals may focus more on commuting. Understanding **where users live** (e.g., proximity to stations or recreational areas) could highlight geographic patterns influencing membership conversion.
2. User Purchase History:
   Data on how often casual riders purchase single-ride or day passes could reveal **spending behaviors**. This would help target users who are already spending enough on passes that converting to membership would be a clear financial benefit. Knowing if casual riders were **previous members who didn’t renew** could provide insights into why they switched back and help refine strategies for retaining members.
3. Weather Data:
   Data on **weather conditions during trips** (e.g., sunny, rainy, temperature) could help understand how external factors influence casual vs. member riders. Casual riders might be more sensitive to weather, whereas members might commute regularly regardless of conditions.

It is essential to address data privacy and storage concerns while collecting such data. Ensuring that users can opt-in and feel secure about how their data is collected and used will be crucial for successful adoption.

Next Steps for Stakeholders

This report presents valuable insights into the differences between casual riders and annual members, along with actionable recommendations to drive membership conversion. These are suggested next steps for stakeholders at Cyclistic:

Review Insights and Recommendations: The marketing team should thoroughly analyze the findings outlined in this report, focusing on the distinct behaviors and preferences of casual riders versus annual members. Understanding these insights will be crucial in shaping targeted marketing strategies.

Develop Targeted Marketing Campaigns: Based on the recommendation for group or family membership options, the marketing team should design a campaign that promotes bundled memberships and highlights exclusive group perks. Implement a referral program that incentivizes individual riders to convert to membership, focusing on financial benefits.

Engage with Executives: Present these insights and recommendations to Cyclistic executives. Highlight the potential for increased profitability by converting casual riders into annual members, and seek their approval to allocate resources for implementing the proposed marketing campaigns.

Pilot Marketing Initiatives: Consider running pilot campaigns for the community membership and referral programs to gauge their effectiveness before a full-scale rollout. Collect feedback and metrics during this phase to refine strategies as needed.

Monitor and Evaluate: Once marketing initiatives are launched, continuously monitor their performance against established metrics. This should include tracking membership conversion rates, user engagement with new offerings, and overall customer satisfaction. Regular evaluations will allow for timely adjustments and enhancements to the marketing strategy.