Case Study 1

**Case Study: How to Achieve Rapid Success of a Bike Share Business?**

# Introduction

Welcome to the Cyclistic bike-sharing analysis case study. In this case study, you'll perform many real-world tasks, typical of a junior data analyst. You'll work for a fictional company called Cyclistic and meet different characters and team members. To answer the company's key questions, you'll follow the steps of the data analysis process: **ask, prepare, process, analyze, share,** and **act**. In this process, **the tables in the case study roadmap**, including guiding questions and key tasks, will help you stay on track.

By the end of this lesson, you'll have a portfolio-ready case study. Download the package and view the details of this case study at any time. So, when you start looking for work, your case study will be a tangible way to demonstrate your knowledge and skills to potential employers.

# Scenario

You're a junior data analyst working on the team of marketing analysts at Cyclistic, a Chicago-based bike-sharing company. The chief marketing officer believes that the future success of the company depends on maximizing the number of annual memberships. Therefore, your team wants to understand what differences exist in the use of Cyclistic bikes between casual riders and annual members. Through this knowledge, your team will design a new marketing strategy to turn casual riders into annual members. Before that, however, Cyclistic executives must approve your recommendations, so you need to back up your proposal with a compelling view of your data and professional visualizations.

# Characters and teams

* **Cyclistic:** A bike-sharing program that includes 5,800 bikes and 600 stations. Cyclistic is notable for also offering recumbent bikes, manual tricycles and cargo bikes that offer a more inclusive use of shared bikes for people with disabilities and cyclists who cannot use a standard two-wheeled bike. Most cyclists choose traditional bicycles, about 8% of cyclists use assisted options. Cyclistic users are more likely to use the bike for recreation, but about 30% use it to get to work each day.
* **Lily Moreno:** The marketing director and your manager. Moreno is responsible for developing campaigns and initiatives to promote the bike-sharing program. Campaigns can include email, social media, and other channels.
* **Cyclistic Marketing Data Computational Analytics Team:** A team of data analysts who are responsible for collecting, analyzing, and reporting data that helps drive Cyclistic's marketing strategy. You joined this team six months ago and have dedicated yourself to not only learning about Cyclistic's mission and business goals, but also seeing how you can help Cyclistic achieve it, from your position as a junior data analyst.
* **Cyclistic** Executive Team**:** The highly detailed executive team will decide whether to approve the recommended marketing program.

# About the company

In 2016, Cyclistic launched a successful bike-sharing offering. Since then, the program has grown to a fleet of 5,824 georeferenced and blocked bikes at a network of 692 stations across Chicago. Bikes can be unlocked from one station and returned to any other station in the system at any time.

Until now, Cyclistic's marketing strategy was based on building overall brand recognition and attracting broad segments of consumers. One of the approaches that helped make this possible was the flexibility of their pricing plans: single-trip passes, full-day passes, and annual memberships. Customers who purchase single-trip passes or full-day passes are called casual cyclists. Customers who purchase annual memberships are called Cyclistic members.

Cyclistic financial analysts concluded that annual members are much more profitable than casual cyclists. Although pricing flexibility helps Cyclistic attract more customers, Moreno believes that maximizing the number of annual members will be key to future growth. Instead of creating a marketing campaign that targets all new customers, Moreno believes there's a good chance of turning casual riders into members. She notes that casual riders are already aware of the Cyclistic program and have chosen Cyclistic for their mobility needs.

Moreno set a clear goal: Design marketing strategies aimed at turning casual cyclists into annual members. However, to do that, the team of marketing analysts needs to better understand how annual members and casual cyclists differ, why casual cyclists would buy a membership, and how digital media might affect their marketing tactics. Moreno and his team are interested in analyzing historical Cyclistic bike travel data to identify trends.

# Ask

Three questions will guide the future marketing program:

1. How are annual members and casual riders different from using Cyclistic bikes?
2. Why would casual cyclists buy annual Cyclistic memberships?
3. How can Cyclistic use digital media to influence casual cyclists to become members?

Moreno assigned you the first question to be answered: How do annual members and casual cyclists differ from using Cyclistic bikes?

You will create a report with the following deliverables:

1. A clear instruction of the business task
2. A description of all data sources used
3. Documentation of all data cleansing and manipulation
4. A summary of your analysis
5. Supporting visualizations and key findings
6. The three most important recommendations based on your analysis

Use the following case study roadmap as a guide. Note: Completing this case study in a week is a good goal.

|  |
| --- |
| **Case Study Roadmap - Ask** |
| **Guiding questions**   * What is the problem you are trying to solve? * How can your insights drive business decisions? |
| **Key tasks**   1. Identify the business task 2. Consider key stakeholders |
| **Deliverable**  A clear instruction of the business task |

# Prepare

You'll use Cyclistic's historical travel data to analyze and identify trends. [Download the last 12 months of Cyclistic](https://divvy-tripdata.s3.amazonaws.com/index.html) travel data  [here.](https://divvy-tripdata.s3.amazonaws.com/index.html) (Note: Datasets have a different name because Cyclistic is a fictitious company. For the purposes of this case study, the datasets are appropriate and will allow you to answer your business questions. The data has been provided by Motivate International Inc. under this [license](https://www.divvybikes.com/data-license-agreement).) This is public data you can use to explore how the types of customers who use Cyclistic bikes differ. Please note, however, that due to data privacy concerns, you are prohibited from using riders' personally identifiable information. This means you won't be able to connect pass purchases with credit card numbers to determine if casual cyclists live in the Cyclistic service area or if they purchased multiple single-ride passes.

Now, prepare your data for analysis using the following case study roadmap as a guide:

|  |
| --- |
| **Case Study Roadmap - Prepare** |
| **Guiding questions**   * Where is your data located? * How is the data organized? * Are there problems with the bias or credibility of this data? [Is your data reliable, original, comprehensive, current and cited (ROCCC)?](https://www.coursera.org/learn/data-preparation/lecture/lHirM/what-is-bad-data) * How are you addressing authorization, privacy, security, and accessibility? * How did you verify data integrity? * How does it help you answer your question? * Is there a problem with the data? |
| **Key tasks**   1. Download the data and store it properly. 2. Identify how they are organized. 3. Sort and filter the data. |

|  |
| --- |
| 4. Determine the credibility of the data. |
| **Deliverable**  A description of all data sources used |

# Process

Now, process your data for analysis using the following case study roadmap as a guide:

|  |
| --- |
| **Case Study Roadmap - Process** |
| **Guiding questions**   * What tools do you choose and why? * Have you ensured data integrity? * What steps did you take to ensure your data is clean? * How can you verify that your data is clean and ready to analyze? * Did you document your cleaning process so you could review and share these results? |
| **Key tasks**   1. Check for errors in the data. 2. Choose your tools. 3. Transform data so you can work with it effectively. 4. Document the cleaning process. |
| **Deliverable**  Documentation of all data cleansing and manipulation |

## Follow these steps:

1. [Download the last 12 months of Cyclistic](https://divvy-tripdata.s3.amazonaws.com/index.html) travel data.
2. Unzip the files.
3. Create a folder on your desktop or in Drive to host your files. Use appropriate file naming conventions.
4. Creates subfolders for . CSV and .XLS or Sheets to have a copy of the original data. Moves the downloaded files to the appropriate subfolder.
5. Follow these instructions for Excel (a) or Google Sheets (b):
   1. Start Excel, open each file, and choose Save as an Excel Workbook file. Place it in the subfolder you created for .XLS files.
   2. Open each . CSV in Google Sheets and save it in the appropriate subfolder.
6. Open your spreadsheet and create a column called "ride\_length." Calculate the extent of each trip by subtracting the "started\_at" column from the "ended\_at" column (for example, =D2-C2) and use the HH:MM:SS format using Format > Cells > Time > 37:30:55.
7. Create a column named "day\_of\_week" and calculate the day of the week on which each trip started using the "WEEKDAY" command (for example, =WEEKDAY(C2,1)) in each file. Use the General format or number without decimals, notice that 1 = Sunday and 7 = Saturday.
8. Move to the analysis step.

If you want, keep working with the data to become better acquainted and maybe even identify new approaches to answering business questions.

# Analyze

Now that your data is properly stored and ready for analysis, start getting it up and running. Use the following case study roadmap as a guide:

## Case Study Roadmap - Analyze

|  |
| --- |
| **Guiding questions**   * How should you organize your data for analysis? * Is your data formatted correctly? * What surprises did you discover in the data? * What trends or relationships did you find in the data? * How will this knowledge help you answer your business questions? |
| **Key tasks**   1. Consolidate your data to make it useful and accessible. 2. Organize and format your data. 3. Perform calculations. 4. Identify trends and relationships. |
| **Deliverable**  A summary of your analysis |

**Follow these steps to use spreadsheets**

Open your spreadsheet app, then complete the following steps:

1. Where appropriate, make columns consistent and combine them into a single worksheet.
2. Clean and transform your data to prepare it for analysis.
3. Perform a descriptive analysis.
4. Run some calculations on a file to get a better idea of the layout of the data. Options:
   * Calculate the average ride\_length
   * Calculate the maximum ride\_length
   * [Calculate the day\_of\_week](https://support.google.com/docs/answer/3094029?hl=en) mode
5. Create a pivot table to quickly calculate and visualize data. Options:
   * Calculate the average ride\_length for members and occasional cyclists. Try rows = member\_casual; Values = Average

of ride\_length.

* + Calculate the average ride\_length for users per day\_of\_week. Test columns = day\_of\_week; rows = member\_casual; Values = Average of ride\_length.
  + Calculate the number of trips for users per day\_of\_week by adding the trip\_id count to Values.

1. Open another file and perform the same descriptive analysis steps. Explore different seasons to make some initial observations.
2. Once you've worked for some time with individual spreadsheets, merge them into a year-round view. Do this with the tool you've chosen to perform your final analysis, whether it's a spreadsheet, database, and SQL or R Studio.
3. Exports a summary file for further analysis.

## Follow these steps to use SQL

Open the SQL tool of your choice, then complete the following steps:

1. Import your data.
2. Explore your data, perhaps by looking at the total number of rows, different values, maximum, minimum, or average values.
3. Where appropriate, use the JOIN statements to combine your respective data into a table.
4. Create summary statistics.
5. Research interesting trends and save that information to a table.

## Follow these steps to use R

Open R Studio and [use this script](https://docs.google.com/document/d/1TTj5KNKf4BWvEORGm10oNbpwTRk1hamsWJGj6qRWpuI/edit) to complete the following steps:

1. Import your data.
2. Organize columns and merge them into a single data frame.
3. Clean and add data to prepare it for analysis.
4. Perform a descriptive analysis.
5. Exports a summary file for further analysis.

# Share

Since you've already done your analysis and gained some insights into your data, now create the visualizations to share your findings. Moreno reminded you that data must be sophisticated and well-polished in order to effectively communicate to the executive team. Use the following case study roadmap as a guide:

|  |
| --- |
| **Case Study Roadmap - Share** |
| **Guiding questions**   * Were you able to answer the question how annual members and occasional cyclists differ from Cyclistic? * What story does your data tell? * How do your findings relate to your original question? * What is your audience? What's the best way to communicate with her? * Can data visualization help you share your findings? * Is your performance accessible to your audience? |
| **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 |

## Follow these steps:

1. Take a piece of paper and pen and sketch out some ideas about how you'll visualize the data.
2. When you've chosen a visual shape, open the tool of your choice and create your visualization. Use presentation software, for example,

PowerPoint or Google Slides; your spreadsheet program; Tableau or R.

1. Create your data visualization, remember to use contrast to capture the audience's attention and direct them towards the most important concepts. Use artistic principles that include size, color, and shape.
2. Ensure clear meaning through the proper use of common elements, such as titles, subheadings, and tags.
3. Perfect the visualization of your data through a special attention to detail.

# Act

Now that you've finished creating your visualizations, act on your findings. Prepare the deliverables Moreno asked you to create, including the top three recommendations that are based on your analysis. Use the following case study roadmap as a guide:

|  |
| --- |
| **Case Study Roadmap - Take Action** |
| **Guiding questions**   * What is your conclusion based on your analysis? * How could your team and your company apply your findings? * What next steps might you or stakeholders take based on your findings? * Is there additional data you could use to expand on your findings? |
| **Key tasks**   1. Create your portfolio. 2. Add your case study. 3. Practice presenting your case study to a friend or family member. |
| **Deliverable**  The three most important recommendations based on your analysis |

## Follow these steps:

1. If you don't already have a portfolio, create one online. (Use Create an interactive portfolio with Google Sites or [Build a portfolio with Google Sites](https://applieddigitalskills.withgoogle.com/c/middle-and-high-school/en/build-a-portfolio-with-google-sites/build-a-portfolio-with-google-sites/introduction-to-build-a-portfolio-with-google-sites.html).)
2. Consider how you want to present your case study in your portfolio.
3. Upload or link your case study findings to your portfolio.
4. Write a short paragraph describing the case study, its process, and your findings.
5. Add the paragraph to present your case study in your portfolio.

# Closure activity

Congratulations on finishing the Cyclistic bike share case study! If you want, complete one of the other case studies to further develop your portfolio. You can also use the **ask**, **prepare**, **process**, **analyze**, **share** , and **act** steps in the case study roadmap to create a new project of your own. Good luck with your job search!