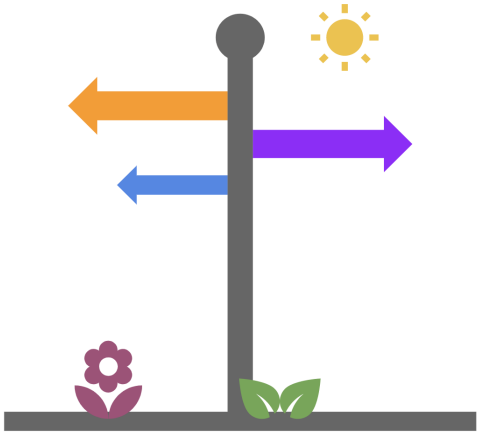
Case Study 3: Follow Your Own Case Study Path



Introduction

Welcome to your very own case study! This is different from all the other cases in this course, as it gives you the opportunity to choose a business question that is of particular interest to you. As a result, finding datasets and conducting analysis will be more complex and will require you to make many decisions on your own.

By the end of this lesson, your goal is to have a portfolio-ready case study. To make sure you are set up for success, it is recommended that you complete one of the more structured case studies first. In addition, be sure to read the included scenario and follow the Case Study Roadmap, which details the steps of the data analysis process: **ask**, **prepare**, **process**, **analyze**, **share**, and **act.**

Once your case study is complete, download the packet and reference the details of your work anytime. Then, when you begin your job hunt, your case study will be a tangible way to demonstrate your knowledge and skills to potential employers.

Scenario

You are a junior data analyst working for a business intelligence consultant. You have been at your job for six months, and your boss feels you are ready for more responsibility. He has asked you to lead a project for a brand new client — this will involve everything from defining the business task all the way through presenting your data-driven recommendations. You will choose the topic, ask the right questions, identify a fresh dataset and ensure its integrity, conduct analysis, create compelling data visualizations, and prepare a presentation.

Ask

Five questions will guide your case study:

1. What type of company does your client represent, and what are they asking you to accomplish?
2. What are the key factors involved in the business task you are investigating?
3. What type of data will be appropriate for your analysis?
4. Where will you obtain that data?
5. Who is your audience, and what materials will help you present to them effectively?

You will produce a report with the following deliverables:

1. A clear statement of the business task you have selected to investigate
2. A description of all data sources used
3. Documentation of any cleaning or manipulation of data
4. A summary of your analysis
5. Supporting visualizations and key findings
6. Based on what you discover, a list of additional deliverables you think would be helpful to include for further exploration
7. Your top high-level insights based on your analysis

Note: Completing this case study within a week is a good goal.

To begin, spend some time brainstorming ideas. An effective presentation should be about 30 minutes long, so make sure your business task isn't too simple or too complex. Below are three examples of business tasks that you can use for inspiration.

- **Rank university sports teams**: Imagine a business wants to make endorsements for players for various sports teams. Explore data over the past few years, consider which teams are ranked in the top five at the beginning of their season and how rankings shift over time. Which teams show potential for endorsements?
- **Expand the vacation rental market**: Imagine a real estate management company wants to determine if they should expand into a new geographic area. Use public data from Airbnb to answer questions about the vacation rental market in a specific area. Investigate how neighborhoods or amenities influence Airbnb prices, what listings are rented most frequently, and where super hosts are located.
- **Consider consumer purchase patterns and weather patterns:** Weather can have an impact on consumer interest in a variety of products. Does colder weather grow demand for scarves and soup? Does warmer weather drive sales of fans and flip flops? And does severe weather increase consumption of bottled water and batteries? Use public data from the National Oceanic and Atmospheric Administration to discover insightful weather-related purchasing patterns.

Note: These examples don't have an immediate connection to the datasets so as to not influence your choice of business tasks based on data.

Use the following Case Study Roadmap as a guide:

Case Study Roadmap - Ask
Guiding questions ● What topic are you exploring? ● What is the problem you are trying to solve? ● What metrics will you use to measure your data to achieve your objective? ● Who are the stakeholders? ● Who is your audience? ● How can your insights help your client make decisions?
Key tasks 1. Identify the business task 2. Determine key stakeholders 3. Choose a dataset

4. Establish metrics
Deliverable ☐ A clear statement of the business task you have selected to investigate

Prepare

Find data that is appropriate for your analysis from any credible dataset. There are numerous public datasets that you can use. You have already encountered some of them throughout this program:

- World Happiness Report (made available by Sustainable Development Solutions Network under a CC0 license): Regional data about overall happiness that contains interesting insights into the relationship between happiness, money, health, and many other metrics.
- Avocado Prices (made available by Justin Kiggins under a CC0 license): Historical data about avocado prices and sales in grocery stores throughout the United States.
- Movies Dataset (made available by Rounak Banik under a CC0 license): Metadata about 45,000 movies, with data points including cast, crew, plot keywords, budget, revenue, ratings, release dates, languages, production companies, and more.
- Amazon Top 50 Best Selling Books (made available by Souter Saalu under a CC0 license): Data about Amazon's bestselling books from 2009 to 2019, categorized into fiction and nonfiction.

Once you have decided on a dataset, gather and store the data appropriately, clean the data, and make sure it is reliable and error-free. Document your process, as clients often ask to see both raw and cleaned data. 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 issues with bias or credibility in this data? Does your data ROCCC?
- How are you addressing licensing, privacy, security, and accessibility?
- How did you verify the data's integrity?
- How does it help you answer your question?
- Are there any problems with the data?

Key

- Download data and store it appropriately.
- Identify how it's organized.
- Sort and filter the data.
- Determine the credibility of the data.

Deliverable

- A description of all data sources used

Process

Then, process your data for analysis using the following Case Study Roadmap as a guide:

Case Study Roadmap - Process
Guiding questions • What tools are you choosing and why? • Have you ensured your data's integrity? • What steps have you taken to ensure that your data is clean? • How can you verify that your data is clean and ready to analyze? • Have you documented your cleaning process so you can review and share those results?
Key tasks • Check the data for errors.

• Choose your tools. • Transform the data so you can work with it effectively. • Document the cleaning process.
Deliverable Documentation of any cleaning or manipulation of data

Follow these steps:

1. Download the dataset.
2. Create a folder on your desktop or Drive to house the files. Use appropriate file-naming
3. conventions. Create subfolders for .CSV or .XLS files.
4. Follow these instructions for either Excel (a) or Google Sheets (b):
    a. Launch Excel, open your file, and choose to Save As an Excel Workbook file. Put it in the subfolder you created for .XLS files.
    b. Open each .CSV file in Google Sheets and save it to the appropriate
5. subfolder. Open your spreadsheet and appropriately label the columns.
6. Proceed to the analyze step.

If you like, continue working with the data to better familiarize yourself and perhaps even identify new approaches to answering the business questions.

Analyze

Now that your data is stored appropriately and has been prepared for analysis, start putting it to work. Use the following Case Study Roadmap as a guide:

Case Study Roadmap - Analyze
Guiding questions

- How should you organize your data to perform analysis on it?
- Has your data been properly formatted?
- What surprises did you discover in the data?
- What trends or relationships did you find in the data?
- How will these insights help answer your business questions?

Key tasks

- Aggregate your data so it's useful and accessible.
- Organize and format your data.
- Perform calculations.
- Document your calculations to keep track of your analysis steps.
- Identify trends and relationships.

Deliverable

- ☐ A summary of your analysis

Follow these steps for using spreadsheets

Open your spreadsheet application, then complete the following steps:

1. Where relevant, make columns consistent and combine them into a single worksheet.
2. Clean and transform your data to prepare for analysis.
3. Conduct descriptive analysis.
4. Run a few calculations to get a better sense of the data layout.
5. Create a pivot table to quickly calculate and visualize the data.
6. Once you are working with several individual spreadsheets, merge them using the tool you have chosen to use to perform your final analysis, either a spreadsheet, a database and SQL, or R Studio.
7. Export a summary file for further analysis.

Follow these steps for using SQL

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

1. Import your data.
2. Explore your data, perhaps looking at the total number of rows, distinct values, maximum, minimum, or mean values.
3. Where relevant, use JOIN statements to combine your relevant data into one table.
4. Create some summary statistics.
5. Investigate interesting trends and save that information to a table.

Follow these steps for using R

Open R Studio and follow this script for inspiration. Note: The script is from another case study, but provides clear instructions that can guide your own analysis. Then complete the following steps:

1. Import your data.
2. Make columns consistent and merge them into a single dataframe.
3. Clean up and add data to prepare for analysis.
4. Conduct descriptive analysis.
5. Export a summary file for further analysis.

Share

Now that you have performed your analysis and gained some insights into your data, create visualizations to share your findings. You will be presenting to your client and other stakeholders, so visuals should be sophisticated and polished in order to effectively communicate your insights. Use the following Case Study Roadmap as a guide:

Case Study Roadmap - Share
Guiding questions • Were you able to answer the business question? • What story does your data tell? • How do your findings relate to your original question? • Who is your audience? What is the best way to communicate with them? • Can data visualization help you share your findings?

• Is your presentation accessible to your audience?
Key tasks • Determine the best way to share your findings. • Create effective data visualizations. • Present your findings. • Ensure your work is accessible.
Deliverable ☐ Supporting visualizations and key findings

Follow these steps:

1. Take out a piece of paper and a pen and sketch some ideas for how you will visualize the data.
2. Once you choose a visual form, open your tool of choice to create your visualization. Use a presentation software, such as PowerPoint or Google Slides; your spreadsheet program; Tableau; or R.
3. Create your data visualization, remembering that contrast should be used to draw your audience's attention to the most important insights. Use artistic principles including size, color, and shape.
4. Ensure clear meaning through the proper use of common elements, such as headlines, subtitles, and labels.
5. Refine your data visualization by applying deep attention to detail.

Act

Now that you have finished creating your visualizations, act on your findings. Organize the deliverables you created, including your top high-level insights based on your analysis. Use the following Case Study Roadmap as a guide:

Case Study Roadmap - Act
Guiding questions • What is your final conclusion based on your analysis?

- How could your team and business apply your insights?
- What next steps would you or your stakeholders take based on your findings?
- Is there additional data you could use to expand on your findings?

Key tasks

- Create your portfolio.
- Add your case study.
- Practice presenting your case study to a friend or family member.

Deliverable

- ☐ Your top high-level insights based on your analysis
- ☐ Based on what you discover, a list of additional deliverables you think would be helpful to include for further exploration

Wrap-up

Congratulations on finishing your very own case study! If you like, complete one of the other case studies to continue growing your portfolio. Best of luck on your job search!