

# Reflective Log

## Overview

Working on this data analysis project provided me with a valuable opportunity to apply the tools and techniques gleaned throughout the module. Utilizing **R programming** and delving into libraries such as **readr**, **dplyr**, and **ggplot2** became an immersive experience. This reflective log aims to encapsulate my personal journey, delving into the challenges faced and my thoughts on the approach taken during the module and the completion of the project.

## Module Experience

The module equipped me with a solid foundation in data analysis techniques, emphasizing the significance of a structured approach. Notably, learning about the CRISP-DM methodology proved to be a valuable asset. Its systematic stages, from comprehending the business problem to deploying insights, offered a clear roadmap for navigating data-driven projects. This approach significantly influenced my personal strategy, ensuring a more organized workflow.

## Personal Experience with Tools and Techniques

1. **Library Familiarity:** The hands-on application of R libraries, including **readr** for data import, **dplyr** for manipulation, and **ggplot2** for visualization, stood out as a module highlight. This personal experience strengthened my understanding of these tools, rendering them more accessible for future projects.
2. **Data Pre-processing Challenges:** Navigating through a substantial dataset of 52 CSV files posed challenges in data cleaning and preparation. My personal journey through these processes was transformative, highlighting the importance of careful data handling. Addressing missing values emerged as a significant challenge during the data pre-processing phase. Missing values can occur for various reasons, such as non-response in surveys, data entry errors, or intentional omissions. My approach to handling missing values involved a careful examination of the nature of the missingness, understanding whether it was completely random or correlated with other variables. After investigation of the data, I decided to remove null value based on focus of each report, opting for a strategy that balanced data completeness with the goal of ensuring data quality and reliability in subsequent analyses.
3. **Effective Data Visualization:** The utilization of **ggplot2** for data visualization became a pivotal aspect of my personal experience in the project. Crafting meaningful and visually engaging plots evolved into a skill honed during the project. The emphasis on effective data communication through visualization profoundly influenced my approach to presenting insights.

# Approach to the Project

1. **CRISP-DM Methodology Implementation:** My personal journey involved the implementation of the CRISP-DM methodology, ensuring a structured and thorough analysis. The clear delineation of stages facilitated a systematic progression, contributing to the overall success of the project. This personal approach is poised to shape my strategy in future data projects.
2. **Challenges Faced and Lessons Learned:** Despite facing difficulties in tackling the size of the dataset, these challenges evolved into transformative learning experiences for me. They underscored the importance of adaptability, attention to detail, and problem-solving skills in real-world data projects. My task included investigating the data and deciding which datasets and fields are valuable and aligned with the business questions I wanted to answer. From this experience, I learned essential lessons. Adaptability proved paramount in the ever-changing landscape of real-world data, demanding flexibility and strategic adjustments. Attention to detail emerged as a critical aspect, emphasizing the impact of small details on the overall analysis. Effective problem-solving skills were reinforced as integral to overcoming challenges in data projects. Furthermore, mastering data investigation techniques, specifically identifying valuable datasets and fields aligned with business objectives, became a key skill. In summary, these challenges not only provided me with valuable experiences but also imparted crucial lessons that will guide my approach to future data projects.

## Conclusion

In conclusion, this reflective log encapsulates my personal journey through the module and the completion of the data analysis project. The practical application of tools and techniques, coupled with the structured CRISP-DM methodology, has significantly shaped my approach to data-driven tasks. As I anticipate similar endeavors in the future, I carry with me the lessons learned and the confidence gained from this enriching personal experience.