

# Reflective Summary



Submittive Assignment

*Jake Bonnici*

November 18, 2022

---

## 1 Reflection on my work

In this assignment I have learned to apply the process of CRISP-DM and use project template in order to have reproducible work for myself or others working on it in the future. I have used Rstudio in a previous assignment, therefore that was relatively familiar, especially the ggplot library. The first thing i setup was the project template which took me some time to understand which environments I should be using and for what. After some more time i got used to that and started to shift my focus more closely on the CRISP-DM procedure. Seeing the CRISP-DM cycle it seemed really intuitive to look at and all the steps made sense to take sequentially. Applying the CRISP-DM method proved to be slightly more challenging. The part which had me most confused was when applying it towards multiple cycles and creating a good flow between each cycle. I then understood that this cycle is not a strict procedure that has to be followed from start to finish at all times, but rather more of a guide to use for parts which are necessary. For example, when starting to work with the second cycle i've had to understand which question would makes sense to ask next for the stakeholder and then proceeded further to prepare the data, etc. I realised that the first question to ask might have been incorrect in some cases, and therefore the focus shift from modeling the data, all the way to the start, business understanding. However, that first cycle of realising that the question is not correct is as important to the cycle as the one that is. Additionally, some phases of the CRISP-DM were a bit shorter with regards to explanation in my Rmarkdown report. I would think that some of the processes, for example; deployment was a less relevant step to us, since we're simply submitting our assignment as our deployment. My focus was more posed towards

the first phase of the CRISP-DM cycle. I found that to be the hardest point of our assignment, even though we had the freedom to choose any question. The data wrangling part was simpler in a sense that I knew what the problem within the data was and I fixed what was needed. When it came to writing the R Markdown report, the latex commands were familiar to me, however to begin with I was confused by the different environments that were being used at the same time, for example; The code chunk, my `*Analysis.R*` code and how to connect them all together. Additionally, the knit button was giving me countless errors, which were eventually googled out. Now, I feel more confident and am able to see the benefits of working with Project Template and R Markdown, together. If I had to restart the assignment today, I'd be able to elaborate my analysis further since it consisted mainly of bar charts and possibly choose a more realistic business scenario.