# Structure for Final Project Report

The data analysis report is specifically for the client and should address the challenges faced by the client. Write a summary of your overall findings and recommendations to the executives at the bank. Think of this section as your closing remarks of a presentation, where you summarize your key findings, model performance, and make recommendations to improve loan processes at the bank.This needs to be provided in a word document once you knit the document.

Your executive summary must be written in a business tone, with minimal grammatical errors, and should include the following sections:

1. **Introduction**

   An introduction where you explain the business problem and goals of your data analysis

   - What problem(s) is this company trying to solve? Why are they important to their future success?
   - What was the goal of your analysis? What questions were you trying to answer, and why do they matter?

2. **Exploratory Data Analysis**

   Highlights and key findings from your Exploratory Data Analysis section

   - What were the interesting findings from your analysis, and **why are they important for the business**?
   - This section is meant to **establish the need for your recommendations** in the following section

3. **ML Model Building Exercise**

   Your "best" classification model and an analysis of its performance

   - In this section, you should talk about the expected error of your model on future data
     - To estimate future performance, you can use your model performance results on the **test data**
   - Please ensure that you are able to explain the models and list which variables are important in determining the outcome.
   - You should discuss at least one performance metric, such as an F1 or ROC AUC, for your model. However, you must explain the results in an **intuitive, non-technical manner**. Your audience in this case are executives at a bank with limited knowledge of machine learning.

4. **Recommendations**

   Your recommendations to the company on how to reduce loan default rates

- Each recommendation must be supported by your data analysis results
- You must clearly explain why you are making each recommendation and which results from your data analysis support this recommendation
- You must also describe the potential business impact of your recommendation:
    - Why is this a good recommendation?
    - What benefits will the business achieve?

**5. Conclusion**

Wrap up the report with concluding remarks by summarizing the results and your recommendations in two or three paragraphs. You should also comment on whether the initial data analysis based on summary data holds good after ML modeling exercise results. You should also provide detailed recommendations.

**6. Appendix/Appendices**

Include all the code, tables, and plots in this section.

**Timelines**

- **04/18 -** Instructions provided for the final project
- **05/06 -** Students to upload the initial version of the project.
- **05/09 - Writing TA (Andrew White)** and I will provide feedback on the project's writing part. Atthis stage, your R commands or coding standard will not be evaluated.
- **05/13 –** Submit the final version of the project

**Grading Guidelines**

- **<u>Exploratory data analysis – (30 pts)</u>**
    - Summary and Visualization **(15 pts) – R Code including comments**
        - Summary Use group_by and provide summary - 3 summaries. A minimum of 3 summaries needs to be provided.
        - Visualization - Use 3 different visualizations.
    - Written analysis and summarization of data analysis **(15 pts)**
- **<u>ML Model building exercise – (70 pts)</u>**
    - Model 1 – **(8 pts)**
    - Model 2 – **(8 pts)**

- Model 3 with hyperparameter tuning – **(15 pts)**
- Performance metrics and variable importance – **(19 pts)**
- Coding Standards followed, Knit documents **(20 pts)** – Comments need to be provided for each chunk of the code. The code should run properly without any error. The final output should be at least 90% similar to what is knitted from the R Markdown file. You can do some post-hoc formatting of the word document after knitting it from the R Markdown file.

- **<u>Summary document – (50 pts)</u>**
  - Report coverage along with analysis and recommendations – **(20 pts)**
  - Analysis on whether the initial exploratory data analysis holds good - **(10 pts)**
  - Analysis on the best classification model based on the performance – **(10 pts)**
  - Overall writing quality – **(10 pts)**