

## Rough Breakdown / Timeline - 4 weeks

### Introduction:

- Formulate introduction in Markdown cells **WRITE**
  - provide some relevant background information on the topic so that someone unfamiliar with it will be prepared to understand the rest of your report
  - clearly state the question you tried to answer with your project
  - identify and describe the dataset that was used to answer the question

### Methods and Results:

**\*\* describe in written English the methods you used to perform your analysis from beginning to end that narrates the code**

- Clean data and Wrangle data **CODE**
- Select columns intended for use **CODE**
- Visualize scatter plot with Age vs. Debt on axes and Approval status as colour **CODE**
- Predictive question: which has more effect on approval status
- Propose a new unknown point Age and Debt values **CODE**
- Create a KNN model **CODE**
- Evaluate Model **CODE**
  - Split data into training and testing data subsets
  - Build the KNN model using training data
  - Choose candidate values of K
  - Split training data into two sets - training and validation
  - For each K train the model using training set
  - Evaluate accuracy for each using validation set (cross-validation)
  - Pick the K value that maximizes validation accuracy
- Test random data **CODE**

### Discussion

- Describe findings, how it relates to prediction, what influence does Age vs. Debt ultimately have on approval status based on the model's behaviour **WRITE**
  - summarize what you found **WRITE**
  - discuss whether this is what you expected to find? **WRITE**
  - discuss what impact could such findings have? **WRITE**
    - Considering a given age group, one should aim to have a target level of debt - if any at all, based on the data **WRITE**
  - discuss what future questions could this lead to? **WRITE**
    - Varieties of Debt
    - Parental Financial Support? Wealth Background?

### References

- List References **WRITE**
  - At least 2 citations of literature relevant to the project (format is your choice, just be consistent across the references).
  - Make sure to cite the source of your data as well.