**Main GitHub repo:**

* <https://github.com/anirudhchandnani/ISDS577/>

**Helpful GitHub repo:**

* <https://github.com/JifuZhao/Lending-Club-Loan-Analysis>

**Data**

* <https://www.kaggle.com/datasets/wordsforthewise/lending-club/data>

**Glossary**

<https://github.com/anirudhchandnani/ISDS577/blob/main/LCDataGlossary.csv>

**To do tasks:**

* Data Cleaning
* Oversampling/Undersampling
* Research questions (1 to 6)
* Modelling
* Tech to use – Python, Excel, Tableau, Power Bi, R

**Code to refer:**

* For EDA and research questions methodology-<https://www.kaggle.com/code/emmaruyiyang/lending-club-loan-default-prediction-eda>
* For modelling ANN – <https://www.kaggle.com/code/faressayah/lending-club-loan-defaulters-prediction>
* For modelling XGB - <https://www.kaggle.com/code/krishnaraj30/xgboost-loan-defaulters-prediction>
* For oversampling - <https://github.com/JifuZhao/Lending-Club-Loan-Analysis/blob/master/5.%20Machine%20Learning%20with%20Over-sampling.ipynb>
* For undersampling –