**Loan Prediction Problem**

Data set (./Learning/Data/train.csv in BB)

| **VARIABLE DESCRIPTIONS** | |
| --- | --- |
| **Variable** | **Description** |
| Loan\_ID | Unique Loan ID |
| Gender | Male/ Female |
| Married | Applicant married (Y/N) |
| Dependents | Number of dependents |
| Education | Applicant Education  (Graduate/ Under Graduate) |
| Self\_Employed | Self employed (Y/N) |
| ApplicantIncome | Applicant income |
| CoapplicantIncome | Coapplicant income |
| LoanAmount | Loan amount in thousands |
| Loan\_Amount\_Term | Term of loan in months |
| Credit\_History | credit history meets guidelines |
| Property\_Area | Urban/ Semi Urban/ Rural |
| Loan\_Status | Loan approved (Y/N) |

Please prepare a small technical report using the loan train dataset. The number of words for the report is approximate to around 500. You should present the **implementation code, calculation results and visualization artefacts** to support your analysis, which can be made by applying the NumPy, MatPlotLib and Pandas libraries.

The report needs to address the following questions:

1. *How does the applicant’s gender affect the loan application?*
2. *What is the link between the education background and the applicant’s loan?*
3. *Does the applicant’s credit record support the loan application?*
4. *How about the property location?*
5. *The overall analysis and critical thinking on the loan data in terms of your observations.*