Loan Data Exercise

Exercise (end of Day 1)

1. Based on the *Loan Data* file, determine
   A. the categorical and numerical variables,
   B. the cause variables and effect variables; hence, explanatory analysis.
2. Conduct 'One Categorical Variable' and 'One Numerical Variable' analysis for each variable (data profiling). Explain your insights, if any.
3. Conduct 'Categorical Explanatory and Response Variable' analysis for cause and effect. Explain your insights, if any.

Hint: Focus on analyzing the effect variables, i.e. the 'Loan_Status' and 'Loan_Default' variables to discover which are the major cause variables. Use pie charts, boxplots and scatter plots for quick analyses.

1A)
categorical:
Gender, Married, Dependents, Education, Self_Employed, Loan_Amount_Term, Credit_History, Property_Area, Loan_Status

Numerical:
ApplicantIncome, CoapplicantIncome, CombinedIncome, LoanAmount,

1B)
loan amount term, credit history, loan status,applicant income, coapplicant income, combined income,loan amount.
analysis: correlation
loan status vs __:
i. loan amount term:-0.021928866
ii. credit history:0.595641526
iii. applicant income: -0.004709523
iv. coapplicant income: -0.059187313
v. combined income: -0.031270779
vi. load amount: -0.037317769

2.

Loan Data Exercise

Exercise (end of Day 2)

4. Conduct ‘Categorical Explanatory and Quantitative Response Variable’ analysis for cause and effect. Explain your insights, if any.
5. Conduct ‘Quantitative Explanatory and Quantitative Response Variable’ analysis for cause and effect. Explain your insights, if any.

Hint: Focus on analyzing the effect variables, i.e. the ‘Loan_Status’ and ‘Loan_Default’ variables to discover which are the major cause variables. Use pie charts, boxplots and scatter plots for quick analyses.