

1. Do a complete EDA on Loan Defaulter data
State whatever insights you are getting from the graphs you have plotted
And keep that notebook uploaded on kaggle
<https://www.kaggle.com/datasets/gauravduttakiit/loan-defaulter>
2. EDA on Academic Success Dataset
State whatever insights you are getting from the graphs you have plotted
<https://www.kaggle.com/datasets/ravi20076/playgrounds4e06originaldata?select=original.csv>
3. Do an EDA on any dataset
State whatever insights you are getting from the graphs you have plotted

Things u can do in EDA=>

Data shape and size: Get a sense of how big your data is (number of rows and columns) [1].

- **Data types:** Check what data types each variable is stored in (e.g., numbers, text) [2]. This helps ensure data is used correctly in analysis.

Data Cleaning:

- **Missing values:** Identify and address missing data points. You might decide to remove them, fill them in with estimates, or keep them depending on the situation [1].
- **Outliers:** Explore extreme data points that fall way outside the expected range. Decide if they are errors or valid but unusual data points [2].

Understanding the Data:

- **Summary statistics:** Calculate basic statistics like mean, median, standard deviation for numerical variables. You can also use frequency tables for categorical variables [2]. This helps you understand the central tendency and spread of the data.
- **Data visualization:** Create visualizations like histograms, boxplots, and scatter plots to see the distribution of data and relationships between variables [1, 2]. Visualizations can reveal patterns and trends that might be difficult to spot just by looking at numbers.
pen_spark