Reporting: Data Visualization Report on Loan from Prosper Dataset

Summary

First what is Data Visualization?

> Data Visualization is can be defined as representing of information or data in graphic means for better understanding and easy glance through of the data in the dataset.

Two types of Visual Analysis ?

> Exploratory Analysis: This is done to when one is trying to gather insight from a dataset or table and it is usually performing at the start of visualization.

> Explanatory Analysis: This is done at the end of visulization for convey information to readers or stakeholder, it's tells stories about the data and it is usually done at the end of the data visualization process.

Carrying out the Data Visualization Process:

> I chosed the loan prosper dataset to carried out my data visualization project, the dataset is about loan giving to customers of the Prosper organization, the dataset contain 81 distinct variables(columns) and 113937 rows of data or information pertaining to a particular loan giving to a new or old borrower of the orgainzation, the dataset contain details like the total amount borrowed, the length of the loan, the status of the loan like cancelled, charged off, completed, past due etc. I will be exploring the dataset and will be giving answer to some question like the top popular occupation that takes loan the most, the employement status of the borrower and many others that will be listed out while visualizing the dataset.

The exploration visualization are divided into three main section which are:

1. Univariate
2. Bivariate
3. Multivariate

Using the Question-Visualization-Observation framework

> The first visualization for the Univariate is, the distribution of original loan amount,the highest count for loan original amount is bewteen 1000-5000, I checked the term for the loan, the organization have three term which are 36, 60 and 12. The frqeuency of the income range which are arranged in orderly manner,

> For the Bivariate exploration, It can be observed that customers with high income takes the huge amount of loan compare to customers with low income range or not employed customers, Taking huge amount also caused lower interest rate to be charged on the loan, also have better propser score rating compare with other low income customer or not employed customers

For Multivariate exploration, Customers whose amount of loan taking is low tends to have low propser rating compare to customers whose loan amount is high. For customers who have compeleted the payment of their loan, the leander yield have a negative correlation meaning high lender yield is equal to low loan amount and vice versa.

Limitations

1. Coming up with insight to use for the visualization
2. Accessing the dataset variables
3. Making up the observation from the graph