

CSE 564 Visualization Mini project 1 Report

Dataset chosen :

Credit Card customer details dataset - this dataset was used to analyze customer attrition/customer churn which means loss of clients or customers.

Dataset URL : <https://www.kaggle.com/sakshigoyal7/credit-card-customers>.

This dataset contains details of every credit card holder. I have chosen 16 attributes for this assignment. Among these 6 are categorical and 10 are numerical.

CATEGORICAL VARIABLES

	VARIABLE NAME	Description
1	Attrition Flag	If the customer's account is closed then its value is 1 else 0.
2	Gender	Gender of the account holder. M = Male, F = Female
3	Education Level	Educational Qualification of the account Holder. (Ex: High School, college graduate etc.)
4	Marital Status	Marital Status of the account holder. Married, Single, Divorced, Unknown
5	Income Category	Annual Income Category of the account holder.(<40K, 40K-80K, 80K-120K, >120K)
6	Card Category	Type of Card.(Blue, Silver, Gold, Platinum)

NUMERICAL VARIABLES

	VARIABLE NAME	Description
1	Customer Age	Age of the account holder.
2	Months on Book	Period of relationship with the bank.
3	Credit Limit	Credit Limit on the Credit card.
4	Total Revolving Bal	Total Revolving Balance on the credit card.
5	Avg Open To Buy	Open to buy credit line.(Average of last 12 months)
6	Total Amt Chng Q4 Q1	Change in transaction amount.(Q4 over Q1)

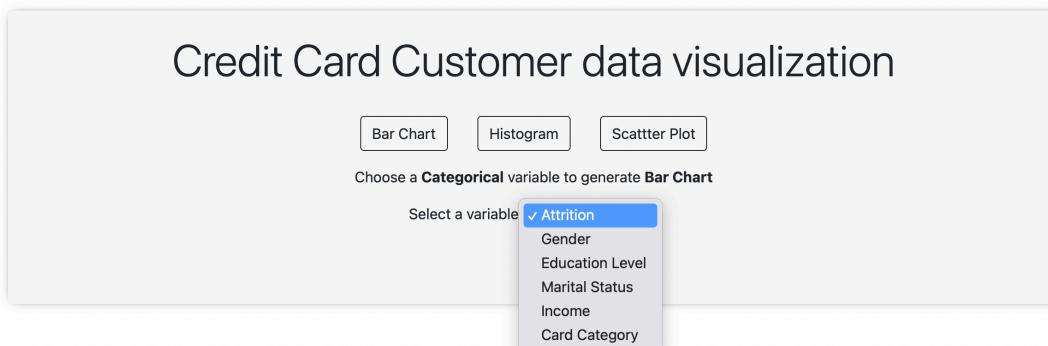
	VARIABLE NAME	Description
7	Total Trans Amt	Total transaction amount.(For last 12 months)
8	Total Trans Ct	Total transaction count.(For last 12 months)
9	Total Ct Chng Q4 Q1	Change in transaction count.(Q4 over Q1)
10	Avg Utilization Ratio	Average card utilization ratio.

Why this dataset?

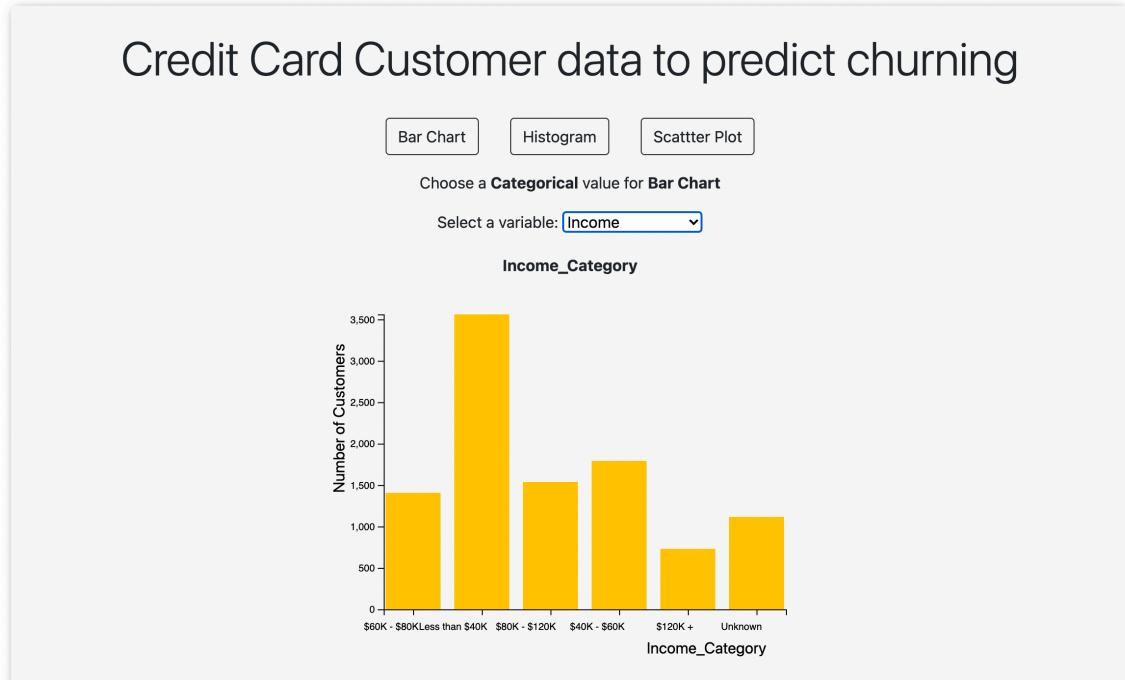
- Good mix of categorical and numerical variables.
- Helps in addressing a real-word business problem
i.e., customer attrition by analyzing the data of the customers. With these statistics, business decisions can be made.

The D3-based visual interface has the following :

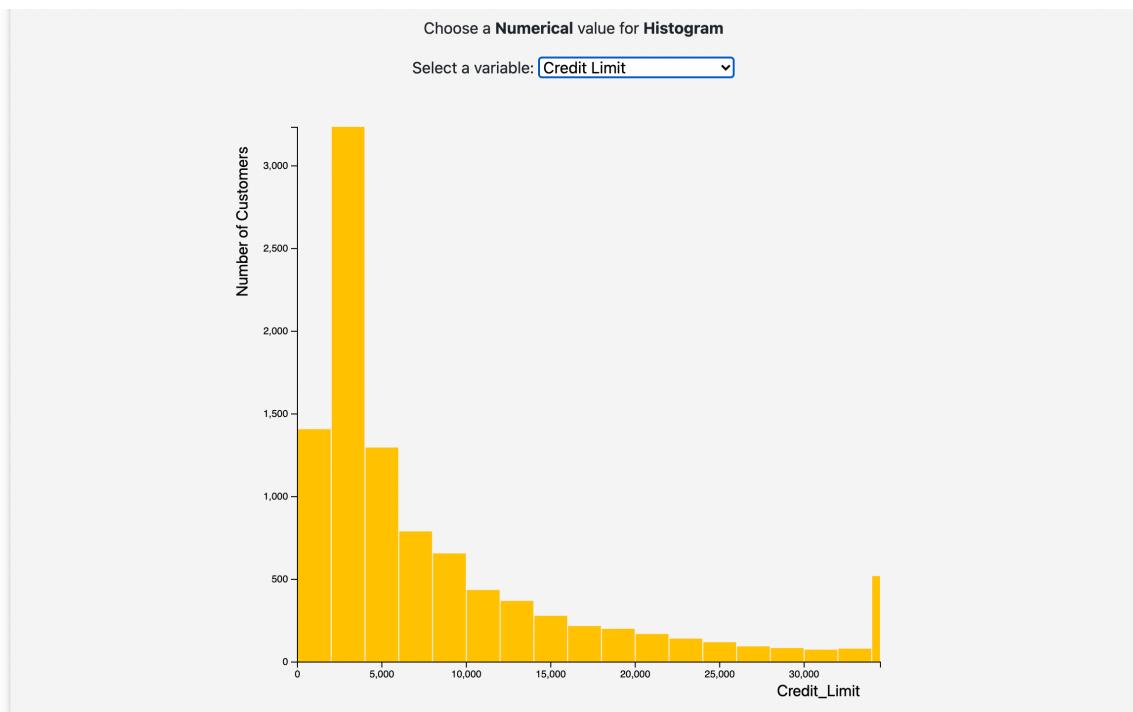
1. a menu to allow users to select a variable and update chart.



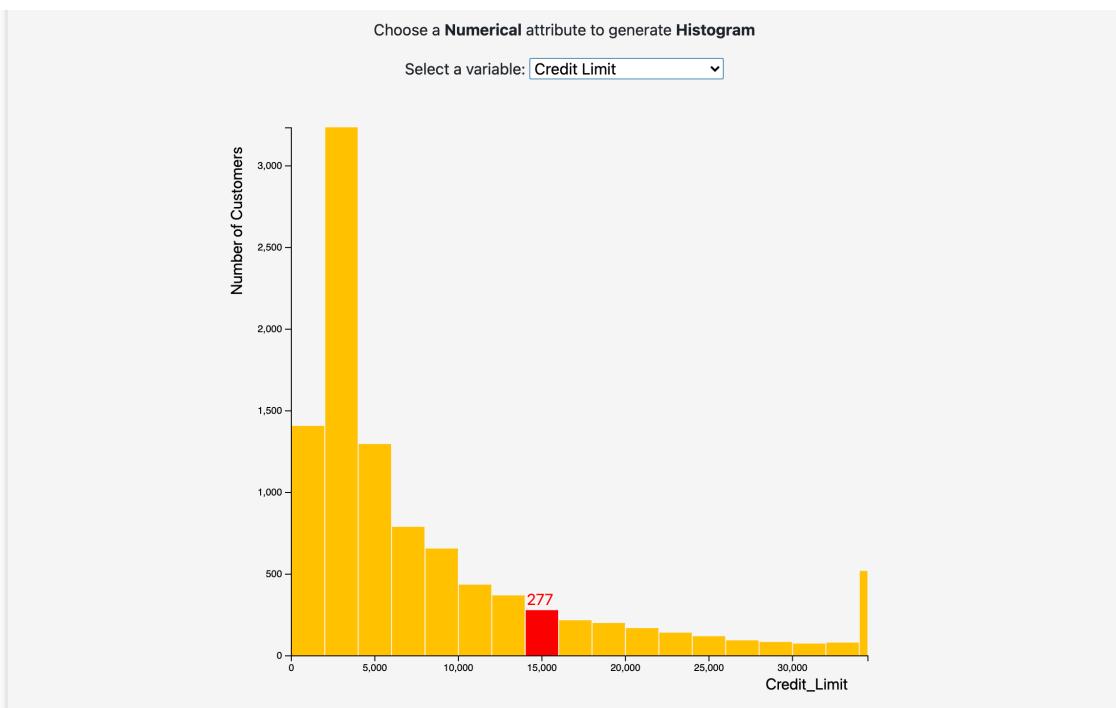
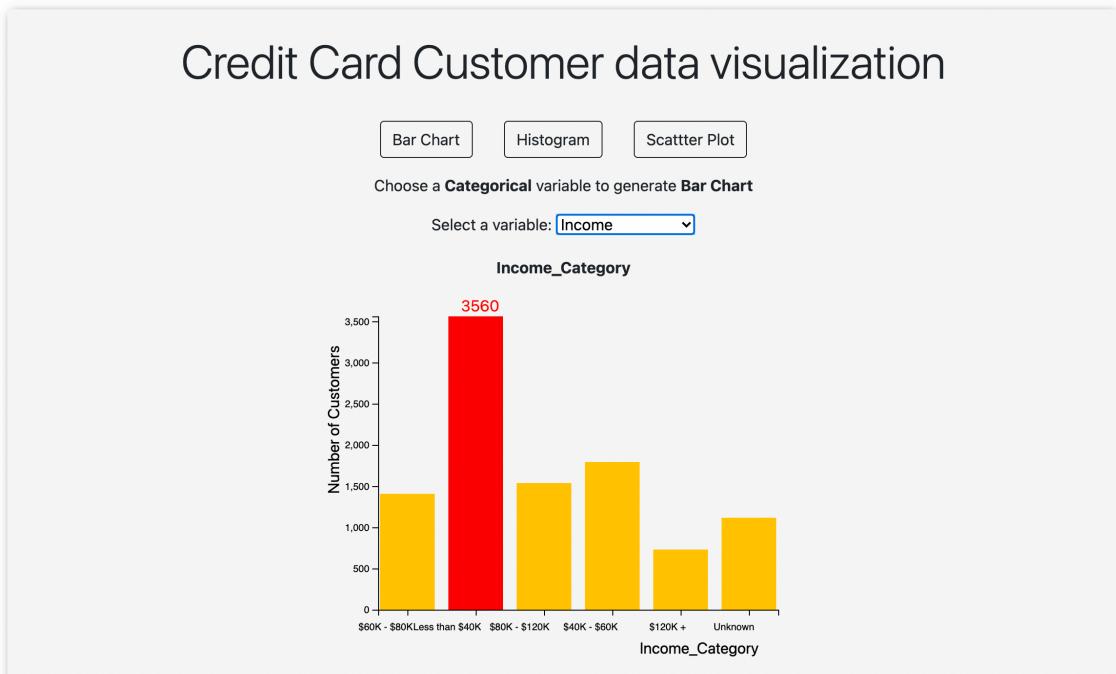
2.A bar chart if a categorical variable is selected



3. a histogram if a numerical variable is selected (bin it into a fixed range (equi-width) of your choice)

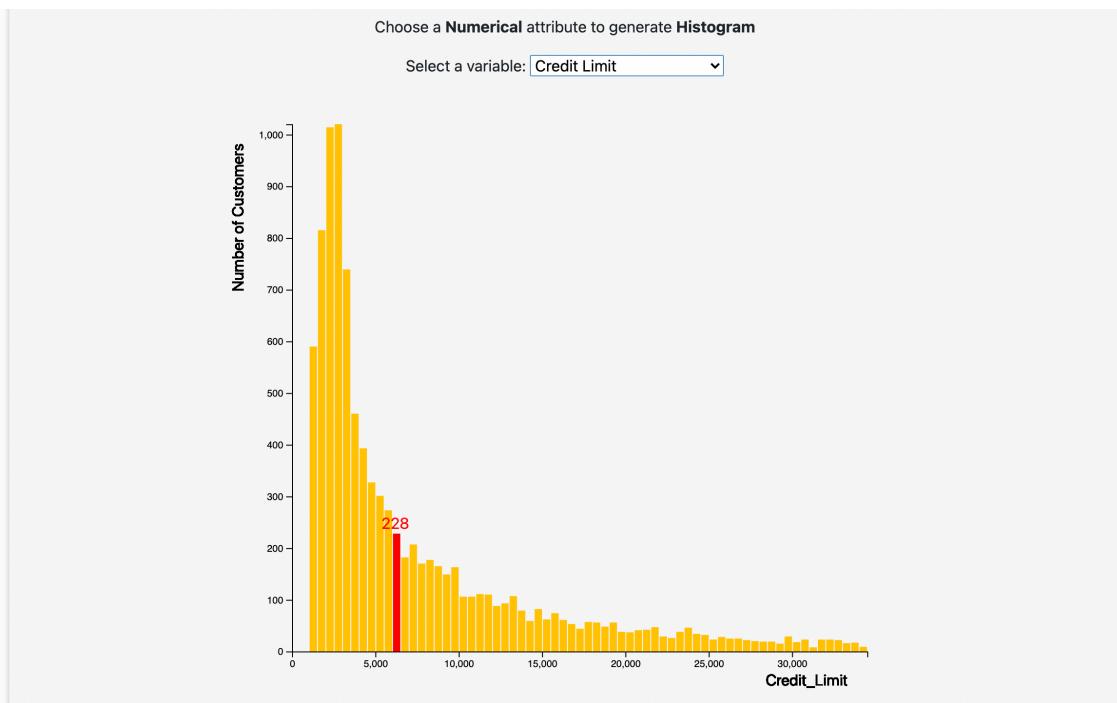


4. on mouse-over displays the bar's value above the bar in red color and 5. on mouse-over also colors the bar in red to highlight it.

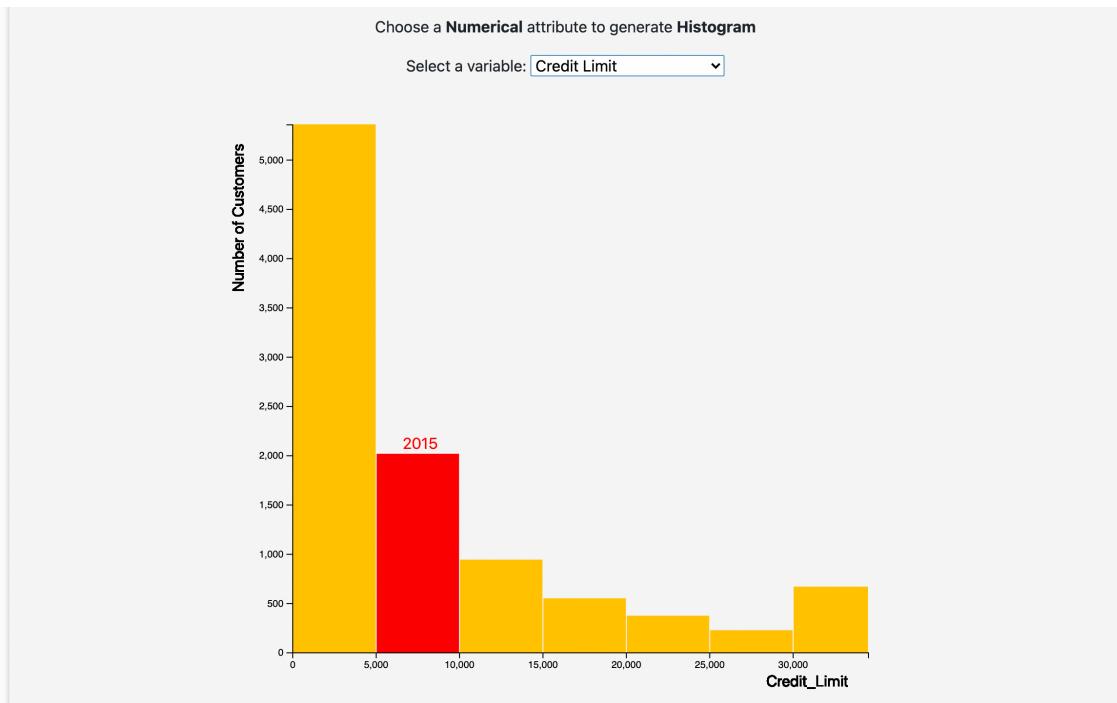


6. mouse (with left mouse button down) move left (right) decreases (increases) bin width/size (for numerical variables only)

When mouse moved to the right - increasing bin size.



When mouse moved to the left - decreasing bin size.



7. Produces a scatterplot of two selected variables (using a radio button to determine which of the two variable axes is to be loaded).

