

Lab Report 1- CSE 564: Visualization

Credit Card Customers

- Ananna Rahman

Objective: The main purpose of this project is to utilise D3.js to visualise data through various charts.

Dataset: Customer Credit Card dataset is taken from Kaggle. The dataset includes records for 10,000 customers, featuring various details such as age, salary, credit card limit, credit card category, and more. Comprising 23 columns, the dataset contains both numerical and categorical values. Although the dataset originally focuses on attributes that may impact bank churners, for this project, these attributes are not considered, and the dataset is approached as a generalised credit card customer dataset.

Link: <https://www.kaggle.com/datasets/sakshigoyal7/credit-card-customers>

Data Preparation:

Originally Dataset has 23 attributes from which 15 attributes are selected. A random sample of 500 data points from 10,000 has been selected from the dataset for analysis and visualisation.

Python snippet:

```
import pandas as pd

df = pd.read_csv('E:/BankChurners.csv')
random_sample = df.sample(n=500)
random_sample.to_csv('E:/processed.csv', index=False)
```

Categorical Attributes:

Gender: Represents the demographic variable, with 'M' denoting Male and 'F' denoting Female.

Dependents: Indicates the number of dependents each customer has. Although these are numerical values, they are treated as categorical due to their repetitive nature and limited range.

Education: Describes the educational qualification of the account holder (e.g., high school, college graduate, etc.).

Number of Products: Refers to the total number of products held by the customer. Similar to dependents, this attribute is treated as categorical despite being numerical due to its limited and repetitive values.

Income: Represents the annual income category of the account holder, categorised as follows :Less than \$40,000 , \$40,000 - \$60,000, \$60,000 - \$80,000, \$80,000 - \$120,000, Greater than \$120,000,Unknown

Numerical attributes:

Age: Represents the customer's age in years.

Period (Months): Denotes the period of the customer's relationship with the bank.

Credit Limit: Indicates the credit limit on the credit card.

Revolving Balance: Reflects the total revolving balance on the credit card.

Open-to-Buy Credit Line: Represents the open-to-buy credit line for the last 12 months.

Change in Transaction Amount: Represents the total change in transaction amount (Q4 over Q1).

Transaction Amount: Denotes the total transaction amount for the last 12 months.Total

Transaction Count: Indicates the total transaction count for the last 12 months.Change in

Change in Transaction Count: Reflects the change in transaction count (Q4 over Q1).

Average Card Utilisation Ratio: Represents the average card utilisation ratio.

Why this dataset?:

The dataset contains a good mix of both categorical and numerical values, with enough data points for meaningful analysis. I've chosen 10 numerical attributes and 5 categorical attributes to work with. The dataset has clean data, and most categorical variables have at least 6 categories, providing information. The numerical data is distributed well, making it easy to understand and analyse.

Feature:

The main web page features a menu, allowing users to select different options. Upon selecting a menu item, either a bar chart or a histogram will be displayed depending on whether the selected feature is categorical or numerical, respectively. For categorical features, a bar chart will be shown, while for numerical features, a histogram will be displayed. Additionally, for scatter plots, users can select the variables for the x-axis and y-axis using radio buttons. Upon selection, a scatterplot will be generated accordingly. Each chart (bar chart, histogram, scatterplot) will have a toggle button below it, allowing users to switch between horizontal and vertical orientations as per their preference

Running the application:

To run the Angular project with D3.js charts, follow these steps:

1. Ensure you have Node.js version 16.14.2 and Angular version 12.2.17 installed on your system.
2. Navigate to the project directory using the terminal/command prompt. Run the following command to install all dependencies listed in the project's package.json file:

npm install

3. Once all dependencies are installed, start the development server by running the command:

`ng serve`

4. After the server has started successfully, open any web browser and go to the following URL:

<http://localhost:4200/>

This will load the Angular application with the D3.js charts.