

# DSE5004

Visual Data Exploration-  
Data Preparation and Exploration Project  
Dictionary for *Customer\_Dataset\_Data.csv*

---

## Customer ID

Description: Unique identifier for each customer.

Options:

- Unique alphanumeric string for each customer.
- 

## Region

Description: Region where the customer resides.

Options:

- 1: Northeast
  - 2: Southeast
  - 3: Midwest
  - 4: Southwest
  - 5: West Coast
- 

## Town size

Description: Size of the customer's town.

Options:

- 1: Small
  - 2: Medium
  - 3: Large
  - 4: Extra-Large
  - 5: Mega-City
- 

## Gender

Description: Gender of the customer.

Options:

- 'Male': Customer identifies as male.
  - 'Female': Customer identifies as female.
- 

## Age

Description: Customer's age.

Options:

- Integer value representing years.

---

## Education years

Description: Number of years of formal education completed by the customer.

Options:

- Integer value representing total years of education.
- 

## Job category

Description: Category representing the customer's occupation.

Options:

- 'Professional': Professional-level occupation.
  - 'Sales': Sales-related occupation.
  - 'Labor': Labor-intensive occupation.
  - 'Service': Service-related occupation.
- 

## Union member

Description: Indicates if the customer is a member of a union.

Options:

- 'Yes': Customer is a union member.
  - 'No': Customer is not a union member.
- 

## Employment length

Description: Length of customer's current employment in years.

Options:

- Integer value representing years.
- 

## Retired

Description: Indicates if the customer is retired.

Options:

- 'Yes': Customer is retired.
  - 'No': Customer is not retired.
- 

## Annual Household income

Description: Customer's household income per year.

Options:

- Integer value representing income in dollars.
-

---

**Debt to income ratio**

Description: Ratio of total debt to annual household income, expressed as a percentage.

Options:

- Floating-point value representing percentage.
- 

**Credit debt**

Description: Amount of debt on credit cards.

Options:

- Floating-point value representing debt in dollars.
- 

**Other debt**

Description: Amount of other non-credit debt.

Options:

- Floating-point value representing debt in dollars.
- 

**Loan default**

Description: Indicates if the customer has a history of loan default.

Options:

- 'Yes': Customer has defaulted on a loan.
  - 'No': Customer has not defaulted on a loan.
- 

**Marital status**

Description: Marital status of the customer.

Options:

- 'Married': Customer is married.
  - 'Unmarried': Customer is not married.
- 

**Household size**

Description: Number of people in the customer's household.

Options:

- Integer value representing household members.
- 

**Number of pets**

Description: Total number of pets owned by the customer.

Options:

- Integer value representing the number of pets.

---

**Number of cats**

Description: Number of cats owned by the customer.

Options:

- Integer value representing the number of cats.
- 

**Number of dogs**

Description: Number of dogs owned by the customer.

Options:

- Integer value representing the number of dogs.
- 

**Number of birds**

Description: Number of birds owned by the customer.

Options:

- Integer value representing the number of birds.
- 

**Home ownership**

Description: Indicates whether the customer owns a home.

Options:

- 1: Customer owns a home.
  - 0: Customer does not own a home.
- 

**Vehicles owned**

Description: Number of vehicles owned by the customer.

Options:

- Integer value representing the total number of vehicles.
- 

**Vehicle ownership type**

Description: Type of ownership of the vehicle.

Options:

- 'Own': Customer owns the vehicle(s).
  - 'Lease': Customer leases the vehicle(s).
- 

**Vehicle brand**

Description: Brand type of the customer's vehicle.

Options:

- 'Foreign': Vehicle brand is foreign.
  - 'Domestic': Vehicle brand is domestic.
- 

### Vehicle value

Description: Total value of the customer's vehicles.

Options:

- Integer value representing total vehicle value in dollars.
- 

### Commute time

Description: Average daily commute time in minutes.

Options:

- Integer value representing commute time.
- 

### Political affiliation

Description: Indicates if the customer is politically affiliated.

Options:

- 'Yes': Customer has a political affiliation.
  - 'No': Customer does not have a political affiliation.
- 

### Voting history

Description: Indicates if the customer has a history of voting.

Options:

- 'Yes': Customer has a voting history.
  - 'No': Customer does not have a voting history.
- 

### Credit card type

Description: Type of credit card the customer uses.

Options:

- 'Visa': Customer uses a Visa card.
  - 'MasterCard': Customer uses a MasterCard.
  - 'American Express': Customer uses an American Express card.
  - 'Discover': Customer uses a Discover card.
- 

### Card tenure

Description: Length of credit card ownership in months.

Options:

- Integer value representing months.
- 

### Monthly card items

Description: Number of items purchased per month using the credit card.

Options

40

**You said:**

and in the description, let's add some meat

**ChatGPT said:**

ChatGPT

Here's the data dictionary with expanded descriptions:

---

### Customer ID

Description: A unique identifier assigned to each customer for tracking and data management purposes. This ID helps maintain anonymity while allowing reference to customer-specific data across various analyses.

Options:

- Unique alphanumeric string for each customer.
- 

### Region

Description: Geographic region of residence, useful for analyzing demographic patterns and regional preferences. Regions include major divisions across the United States, each associated with distinct lifestyle and purchasing trends.

Options:

- 1: Northeast
  - 2: Southeast
  - 3: Midwest
  - 4: Southwest
  - 5: West Coast
- 

### Town size

Description: Classification of the town based on

population size, capturing environmental context. Town size can impact access to services, cost of living, and lifestyle, all of which influence spending and service choices.

Options:

- 1: Small
  - 2: Medium
  - 3: Large
  - 4: Extra-Large
  - 5: Mega-City
- 

**Gender**

Description: Gender identity of the customer, an important demographic factor for analyzing trends and preferences across genders in product adoption and service usage.

Options:

- 'Male': Customer identifies as male.
  - 'Female': Customer identifies as female.
- 

**Age**

Description: The customer’s age in years, reflecting generational behavior, lifestyle stages, and typical service preferences, as younger or older customers may have differing needs.

Options:

- Integer value representing years.
- 

**Education years**

Description: Total years of formal education completed by the customer, indicating educational background, often correlated with career type, income, and spending patterns.

Options:

- Integer value representing total years of education.
- 

**Job category**

Description: The type of job held by the customer, providing insights into lifestyle, income levels, and time constraints, each of which can affect product and service use.

Options:

- 'Professional': Occupations in fields requiring specialized knowledge, typically higher income.
  - 'Sales': Jobs focused on selling goods or services, often commission-based.
  - 'Labor': Physically intensive occupations, may vary widely in income.
  - 'Service': Occupations in customer-oriented fields, like hospitality or retail.
- 

**Union member**

Description: Membership in a labor union, indicating potential job security, benefits, and bargaining power, all of which can impact spending behaviors and loyalty.

Options:

- 'Yes': Customer is a union member.
  - 'No': Customer is not a union member.
- 

**Employment length**

Description: Length of employment in the customer’s current position, reflecting job stability and tenure, which often correlate with financial stability and purchasing power.

Options:

- Integer value representing years.
- 

**Retired**

Description: Indicates whether the customer is retired, a key life stage marker that affects income sources, lifestyle, and types of services required.

Options:

- 'Yes': Customer is retired.
  - 'No': Customer is not retired.
- 

**Annual Household income**

Description: Yearly income for the customer’s household, an indicator of economic stability and purchasing power, useful for understanding budget constraints and service affordability.

Options:

- Integer value representing income in dollars.
-

### Debt to income ratio

Description: Percentage ratio of total debt to annual income, providing insight into the customer's financial health, spending flexibility, and debt management practices.

Options:

- Floating-point value representing percentage.
- 

### Credit debt

Description: Total credit card debt amount, an indicator of spending patterns, financial stability, and the potential impact of high-interest debt on disposable income.

Options:

- Floating-point value representing debt in dollars.
- 

### Other debt

Description: Total non-credit card debt amount, covering personal loans, mortgages, or other types of financial liabilities, relevant to overall debt management.

Options:

- Floating-point value representing debt in dollars.
- 

### Loan default

Description: Status of previous loan defaults, if any, which may indicate financial distress, risk behavior, and the likelihood of future default risk.

Options:

- 'Yes': Customer has defaulted on a loan.
  - 'No': Customer has not defaulted on a loan.
- 

### Marital status

Description: The customer's marital status, which can influence household income, lifestyle choices, and spending patterns, especially on family-oriented products and services.

Options:

- 'Married': Customer is married.
  - 'Unmarried': Customer is not married.
- 

### Household size

Description: Number of individuals living in the customer's household, affecting shared income, living expenses, and needs for family-oriented services.

Options:

- Integer value representing household members.
- 

### Number of pets

Description: Total number of pets owned by the customer, providing insights into lifestyle choices and potential spending on pet-related products or services.

Options:

- Integer value representing the number of pets.
- 

### Number of cats

Description: Specific count of cats owned, used in segmentation for targeting pet-related services or products.

Options:

- Integer value representing the number of cats.
- 

### Number of dogs

Description: Specific count of dogs owned, used in segmentation for pet-related product or service targeting.

Options:

- Integer value representing the number of dogs.
- 

### Number of birds

Description: Specific count of birds owned by the customer, often less common but relevant for niche pet-related targeting.

Options:

- Integer value representing the number of birds.
- 

### Home ownership

Description: Homeownership status of the customer, impacting their needs for home-related services or improvements.

Options:

- 1: Customer owns a home.
  - 0: Customer does not own a home.
- 

### **Vehicles owned**

Description: Total number of vehicles owned, indicating the customer's transportation needs and the potential demand for auto-related products.

Options:

- Integer value representing the total number of vehicles.
- 

### **Vehicle ownership type**

Description: Ownership type for the vehicle, reflecting long-term investment versus temporary arrangements.

Options:

- 'Own': Customer owns the vehicle(s).
  - 'Lease': Customer leases the vehicle(s).
- 

### **Vehicle brand**

Description: Type of vehicle brand owned by the customer, capturing domestic vs. foreign preference, often tied to brand loyalty and economic factors.

Options:

- 'Foreign': Vehicle brand is foreign.
  - 'Domestic': Vehicle brand is domestic.
- 

### **Vehicle value**

Description: Total estimated value of all vehicles owned, giving insight into customer investment in transportation.

Options:

- Integer value representing total vehicle value in dollars.
- 

### **Commute time**

Description: Average time spent commuting daily, influencing lifestyle and time allocation for other activities or services.

Options:

- Integer value representing commute time in minutes.

---

### **Political affiliation**

Description: Indicates whether the customer is politically affiliated or engaged, reflecting potential interest in socially responsible services and community-focused products.

Options:

- 'Yes': Customer has a political affiliation.
  - 'No': Customer does not have a political affiliation.
- 

### **Voting history**

Description: Indicates the customer's participation in past voting, providing insights into civic engagement and possibly community-driven spending behaviors.

Options:

- 'Yes': Customer has a voting history.
  - 'No': Customer does not have a voting history.
- 

### **Credit card type**

Description: Type of primary credit card used by the customer, relevant for understanding brand preferences, transaction fees, and rewards programs.

Options:

- 'Visa': Customer primarily uses a Visa card.
  - 'MasterCard': Customer primarily uses a MasterCard.
  - 'American Express': Customer primarily uses an American Express card.
  - 'Discover': Customer primarily uses a Discover card.
- 

### **Card tenure**

Description: Length of time the customer has held their credit card, a possible indicator of brand loyalty and financial stability.

Options:

- Integer value representing months of card ownership.
- 

### **Monthly card items**

Description: Average number of purchases made per month using the primary credit card, providing insights into spending frequency and consumer habits.

Options:

- Integer value representing monthly transactions.
- 

### Monthly card spend

Description: Total monthly expenditure using the credit card, indicating spending level and the customer's purchasing power.

Options:

- Floating-point value representing monthly spend in dollars.
- 

### Active Fitness lifestyle

Description: Indicates if the customer has an active lifestyle, potentially indicating higher interest in health-related products, wearable devices, and fitness services.

Options:

- 'Yes': Customer actively engages in fitness activities.
  - 'No': Customer does not engage in regular fitness activities.
- 

### Telecom provider tenure

Description: Duration the customer has been with their telecom provider, representing customer loyalty and satisfaction with the telecom service.

Options:

- Integer value representing months with the provider.
- 

### Voice service usage

Description: Monthly usage of voice services, measured in hours, indicative of the customer's communication habits and needs.

Options:

- Floating-point value representing hours of voice service use per month.
- 

### Voice service tenure usage

Description: Total lifetime usage of voice services, giving insight into long-term service engagement and communication patterns.

Options:

- Floating-point value representing total lifetime hours of voice service use.
- 

### Equipment subscription

Description: Indicates whether the customer subscribes to additional telecom equipment services, relevant for understanding add-on product adoption.

Options:

- 'Yes': Customer subscribes to equipment services.
  - 'No': Customer does not subscribe to equipment services.
- 

### Equipment monthly usage

Description: Hours per month of telecom equipment usage, useful for tracking product engagement and determining customer needs for upgrades or support.

Options:

- Floating-point value representing monthly usage in hours.
- 

### Equipment lifetime usage

Description: Total hours of equipment usage over the customer's tenure, providing a historical view of usage that may affect equipment life and replacement cycles.

Options:

- Floating-point value representing lifetime usage in hours.
- 

### International plan

Description: Indicates if the customer has an international calling plan, relevant for assessing global communication needs and target marketing.

Options:

- 'Yes': Customer has an international calling plan.
  - 'No': Customer does not have an international calling plan.
- 

### Wireless data plan

Description: Status of wireless data plan, indicating the customer's need for mobile internet connectivity and potential data plan preferences.

Options:

- 'Yes': Customer subscribes to a wireless data plan.
  - 'No': Customer does not subscribe to a wireless data plan.
- 

### Monthly data usage

Description: Average monthly data usage in GB, reflecting the customer's digital activity level and data needs.

Options:

- Floating-point value representing data usage per month in GB.
- 

### Lifetime data usage

Description: Total data usage since plan initiation, useful for understanding long-term usage trends and future data needs.

Options:

- Floating-point value representing lifetime data usage in GB.
- 

### Family plan

Description: Indicates if the customer is part of a family telecom plan, which may affect data sharing, call distribution, and plan customization needs.

Options:

- 'Yes': Customer is on a family plan.
  - 'No': Customer is on an individual plan.
- 

### Cloud storage service

Description: Status of cloud storage service subscription, which reflects the customer's need for digital storage and data security.

Options:

- 'Yes': Customer subscribes to cloud storage services.
  - 'No': Customer does not subscribe to cloud storage services.
- 

### IoT device support

Description: Indicates if the customer has support for Internet of Things (IoT) devices, pointing to

engagement with smart home or connected technology.

Options:

- 'Yes': Customer has IoT device support.
  - 'No': Customer does not have IoT device support.
- 

### Home internet plan

Description: Status and type of internet plan, essential for understanding household connectivity and data demands.

Options:

- 'Yes': Customer subscribes to a home internet plan.
  - 'No': Customer does not subscribe to a home internet plan.
  - Other values (e.g., "4") indicate specific plan types.
- 

### Virtual assistant service

Description: Subscription status for virtual assistant services, showing interest in automation and hands-free technology for convenience.

Options:

- 'Yes': Customer subscribes to virtual assistant services.
  - 'No': Customer does not subscribe to virtual assistant services.
- 

### Mobile security package

Description: Indicates if the customer subscribes to mobile security, highlighting interest in data protection and privacy.

Options:

- 'Yes': Customer has a mobile security package.
  - 'No': Customer does not have a mobile security package.
- 

### Cloud backup service

Description: Indicates if the customer subscribes to cloud backup for data preservation and recovery needs.

Options:

- 'Yes': Customer subscribes to cloud backup services.



- 'No': Customer does not subscribe to cloud backup services.
- 

#### **Video streaming add-on**

Description: Indicates if the customer subscribes to a video streaming add-on, relevant for assessing media consumption habits and potential upsell opportunities.

Options:

- 'Yes': Customer subscribes to video streaming services.
  - 'No': Customer does not subscribe to video streaming services.
- 

#### **Smart home monitoring**

Description: Status of smart home monitoring system subscription, indicating engagement with home automation and security technology.

Options:

- 'Yes': Customer has smart home monitoring.
  - 'No': Customer does not have smart home monitoring.
- 

#### **Streaming hours**

Description: Average weekly hours spent streaming content, useful for understanding media consumption habits.

Options:

- Integer value representing hours per week.
- 

#### **Dedicated home computer**

Description: Indicates if the customer owns a dedicated home computer, reflecting technology adoption and potential software needs.

Options:

- 'Yes': Customer has a dedicated home computer.
  - 'No': Customer does not have a dedicated home computer.
- 

#### **Owns smart device**

Description: Indicates ownership of a smart device, showing the customer's engagement with modern

technology and connectivity needs.

Options:

- 'Yes': Customer owns a smart device.
  - 'No': Customer does not own a smart device.
- 

#### **Owns gaming console**

Description: Indicates whether the customer owns a gaming system, useful for understanding media consumption and entertainment preferences.

Options:

- 'Yes': Customer owns a gaming console.
  - 'No': Customer does not own a gaming console.
- 

#### **Smart home devices**

Description: Status of smart home device ownership, showing interest in home automation and IoT technology.

Options:

- 'Yes': Customer owns smart home devices.
  - 'No': Customer does not own smart home devices.
- 

#### **Financial News streaming subscription**

Description: Status of financial news streaming subscription, indicating interest in financial markets, economic news, and possibly wealth management products.

Options:

- 'Yes': Customer subscribes to financial news streaming.
- 'No': Customer does not subscribe to financial news streaming.