**Variable Analysis:**

|  |  |
| --- | --- |
| **Variable** | **Description** |
| ID | The ID variable is a unique identifier for each individual within the dataset. The individual is the customer within this context. |
| Gender | Binary feature that represents the gender of the customer. 0 represents males and 1 represents females. In the US, gender cannot be used when determining whether to extend credit to an applicant (CFPB). This will be an important consideration when creating the machine learning model. |
| Own\_Car | Binary feature that indicates if a customer owns a car. 0 represents no car ownership and 1 represents there is car ownership. In this context, car ownership can show financial stability. |
| Own\_Property | Binary feature that indicates if the customer owns property. 0 represents no property ownership and 1 shows the customer owns property. Property ownership shows financial stability and a solid credit score. |
| Work\_Phone | Binary feature that indicates if a customer owns a work phone. 0 indicates no and 1 indicates yes. Owning a work phone can be a sign of employment stability. |
| Phone | Binary feature that indicates if a customer has a phone. 0 indicates they do not have a phone, while 1 indicates they own a phone. Owning a phone is another indicator of financial stability. |
| Email | Binary feature that indicates whether an applicant provided an email on their application. 0 indicates no email while 1 indicates a provided email. |
| Unemployed | A binary feature that indicates if a customer is unemployed. 0 means the customer is employed and 1 indicates a person is unemployed. Employment is important to prove financial stability and a constant income. |
| Num\_children | A numerical feature that indicates how many children an applicant has. A greater number of children means that a customer will have more expenses. |
| Num\_family | A numerical feature that indicates the total number of family members an applicant has. More family members shows that a customer has more expenses. |
| Account\_length | A numerical feature that indicates the length of a customers account with the bank. Longer account lengths indicates a customer’s stability with the financial institution. |
| Total\_income | A numerical variable showing the total income of a customer. Higher levels of income show more stability and financial security. |
| Age | A numerical variable that shows the age of an individual in years. This variable is correct to multiple decimal points. Older people may represent more financial security. In the US, it is important to understand that age can’t be used to discriminate when making lending decisions but can be considered when the applicant is entering into a binding contract (CFPB). |
| Years\_employed | A numerical variable that indicates the number of years a customer has been employed. Greater years of employment can indicate greater savings and more reliability. |
| Income\_type | A categorical variable that indicates the type of income for a customer (e.g., Working, Commercial Associate, Pensioner, etc.). Different income types indicate different amounts of financial stability. |
| Education\_type | A categorical variable that shows the education level of a customer (e.g., Secondary/secondary special, Higher education, etc.). Higher levels of education can demonstrate greater reliability or potentially student loan debt. |
| Family\_status | A categorical variable that shows the family status of an individual. In the US, it is important to understand that marital status cannot be used to determine whether to extend credit. Marital status can be considered in certain cases such as relying on a spouse for income (CFPB). These are important considerations to avoid discrimination. |
| Housing\_type | A categorical variable that indicates the type of housing that a customer lives in (e.g., House/apartment, With parents, etc.). |
| Occupation\_type | A categorical variable that indicates the type of occupation that an individual is engaged in (e.g., Laborers, Sales Staff, Accountants, etc.). Different occupations indicate different levels of financial security. |
| Target | A binary target variable that indicates if a person is eligible for a credit card. 0 indicates the customer is not eligible, while 1 indicates the person is eligible. |

**Continuous Variables:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **Mean** | **Median** | **Max** | **Min** | **Standard Deviation** |
| Num\_children | 0.423 | 0 | 19 | 0 | 0.767 |
| Num\_family | 2.183 | 2 | 20 | 1 | 0.933 |
| Account\_length | 27.270 | 26 | 60 | 0 | 16.648 |
| Total\_income | 181228.20 | 157500 | 1575000 | 27000 | 99277.31 |
| Age | 43.784 | 42.741 | 68.864 | 20.504 | 11.626 |
| Years\_employed | 5.665 | 3.762 | 43.021 | 0 | 6.342 |

**Discrete Variables:**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Count** | **Proportions** |
| Gender | Males: 6323  Females: 3386 | Males: 0.6513  Females: 0.3487 |
| Own\_car | No Car: 6139  Car: 3570 | No Car: 0.6323  Car: 0.3677 |
| Own\_property | No Property: 3189  Property: 6520 | No Property: 0.3283  Property: 0.6715 |
| Work\_phone | No Work Phone: 7598  Work Phone: 2111 | No Work Phone: 0.7826  Work Phone: 0.2174 |
| Phone | No Phone: 6916  Phone: 2793 | No Phone: 0.7123  Phone: 0.2877 |
| Email | No email: 8859  Email: 850 | No email: 0.9125  Email: 0.0875 |
| Unemployed | Employed: 8013  Unemployed: 1696 | Employed: 0.8253  Unemployed: 0.1747 |
| Income\_type | Commercial Associate: 2312  Pensioner: 1712  State Servant: 722  Student: 3  Working: 4960 | Commercial Associate: 0.2381  Pensioner: 0.1763  State Servant: 0.0744  Student: 0.0003  Working: 0.5109 |
| Education\_type | Academic Degree: 6  Incomplete Higher: 371  Secondary: 6761  Higher Education: 2457  Lower Secondary: 114 | Academic Degree: 0.0006  Incomplete Higher: 0.0382  Secondary: 0.6964  Higher Education: 0.2531  Lower Secondary: 0.0117 |
| Family\_status | Civil Marriage: 836  Married: 6530  Separated: 574  Single: 1359  Widow: 410 | Civil Marriage: 0.0861  Married: 0.6726  Separated: 0.0591  Single: 0.1400  Widow: 0.0422 |
| Housing\_type | Co-op Apartment: 34  House/Apartment: 8684  Municipal Apartment: 323  Office Apartment: 76  Rented Apartment: 144  With Parents: 448 | Co-op Apartment: 0.0035  House/Apartment: 0.8944  Municipal Apartment: 0.0333  Office Apartment: 0.0078  Rented Apartment: 0.0148  With Parents: 0.0461 |
| Occupation\_type | Accountants: 300  Core staff: 877  HR staff: 22  Low-skill Laborers: 53  Other: 2994  Sales staff: 959  Waiters/barmen: 40  Cleaning staff: 146  Drivers: 623  IT staff: 18  Managers: 782  Private service staff: 86  Secretaries: 46  Cooking Staff: 193  High skill tech staff: 357  Laborers: 1724  Medicine Staff: 291  Realty Agents: 16  Security Staff: 182 | Accountants: 0.0309  Core staff: 0.0903  HR staff: 0.0023  Low-skill Laborers: 0.0055  Other: 0.3084  Sales staff: 0.0988  Waiters/barmen: 0.0041  Cleaning staff: 0.0150  Drivers: 0.0642  IT staff: 0.0019  Managers: 0.0805  Private service staff: 0.0089  Secretaries: 0.0047  Cooking Staff: 0.0199  High skill tech staff: 0.0368  Laborers: 0.1776  Medicine Staff: 0.0298  Realty Agents: 0.0016  Security Staff: 0.0187 |
| Target | No Credit Card: 8426  Issue Credit card: 1283 | No Credit Card: 0.8679  Issue Credit card: 0.1321 |