Credit Card Eligibility Insights: Understanding Approval Factors

Summary

In this data analysis project, I explored the factors that determine credit card eligibility using a comprehensive dataset covering demographic, financial, and personal attributes. The analysis was performed to uncover patterns and relationships that financial institutions might consider when approving or rejecting credit card applications.

The primary goals of the analysis were to:

- Identify key trends in credit card approval based on housing, income, employment, and digital accessibility.
- Analyze correlations between various attributes (e.g., employment stability, housing type, and communication tools).
- Provide actionable insights to enhance understanding of credit card approval criteria and improve decision-making processes.

This analysis provides insights into the impact of housing, income levels, employment duration, and communication access on credit card approval rates.

Dataset

The dataset used in this analysis is the Credit Card Eligibility Dataset: Determining Factors, which is available on Kaggle, a popular platform for data science and machine learning projects (dataset link).

This dataset provides a comprehensive collection of demographic, financial, and personal attributes for individuals applying for credit cards. It contains multiple variables that financial institutions commonly consider when evaluating a person's creditworthiness.

Each row represents a unique individual identified by a distinct ID, with a combination of attributes that offer insights into their background, financial situation, and personal characteristics. These attributes help paint a holistic picture of the individual, enabling analysis to determine the key factors influencing credit card approval.

Key Variables in the Dataset:

ID:

- ID: A unique identifier for each individual (customer) in the dataset. This variable ensures that each record is distinct and allows for the proper referencing of individuals across the dataset.

Demographic Information:

- Gender: The individual's gender.
- Age: The age of the individual.
- Family Status: Marital or familial status, indicating whether the person is single, married, divorced, etc.
- Number of Children: How many children the individual has.
- Number of Family Members: Total number of individuals in the family unit, which includes children and dependents.

Financial Data:

- Total Income: The individual's total income, providing an indication of their financial standing.
- Account Length: How long the individual has held an account with a financial institution.
- Years Employed: The number of years the person has been employed, offering insights into employment stability.
- Income Type: Type of income (e.g., employed, self-employed, pensioner, etc.), giving a snapshot of the individual's income source.

Housing Information:

- Own Car: A binary feature indicating whether the individual owns a car.
- Own Property: A binary feature showing whether the person owns a property, which is often considered an indicator of financial stability.
- Housing Type: The type of housing the individual resides in (e.g., house, apartment, living with parents), which can be used to assess the person's living situation and associated financial responsibilities.

Occupation and Education:

- Occupation Type: The field or industry in which the individual is employed (e.g., high-skill tech staff, laborers, managers), providing insights into income stability and potential career growth.
- Education Level: The highest level of education completed by the individual (e.g., secondary, higher, incomplete higher education), which could be correlated with earning potential.

Communication Tools:

- Phone: A binary feature indicating whether the individual has a personal phone.
- Work Phone: A binary feature that denotes whether the individual has a work phone, which may indicate employment stability.
- Email: A binary feature showing whether the person has an email address, which could be used for communication with financial institutions.

Employment Status:

- Unemployed: A binary feature that indicates whether the individual is unemployed, which directly influences financial stability and creditworthiness.

Target Variable (Outcome):

- Target: The key binary feature that indicates whether the individual was approved for a credit card (1 = approved, 0 = not approved). This is the dependent variable used for classification in predictive models.

Data Dictionary

Column Name	Description	Type	Indicator	Notes
ID	An identifier for each individual (customer)	Number	#	No Null
Gender	The gender of the individual	Number	Binary (0, 1) 0 = male 1 = female	No Null
Own_car	A binary feature indicating whether the individual owns a car.	Number	Binary (0, 1) 0 = no 1 = yes	No Null
Own_property	A binary feature indicating whether the individual owns a property.	Number	Binary (0, 1) 0 = no 1 = yes	No Null
Work_phone	A binary feature indicating whether the individual has a work phone.	Number	Binary (0, 1) 0 = no 1 = yes	No Null
Phone	A binary feature indicating whether the individual has a phone.	Number	Binary (0, 1) 0 = no 1 = yes	No Null
Email	A binary feature indicating whether the individual has provided an email address.	Number	Binary (0, 1) 0 = no 1 = yes	No Null
Unemployed	A binary feature indicating whether the individual is unemployed.	Number	Binary (0, 1) 0 = no 1 = yes	No Null
Num_children	The number of children the individual has.	Number	#	No Null
Num_family	The total number of family members.	Number	#	No Null
Account_length	The length of the individual's account with a bank or financial institution.	Number	#	No Null

Total_income	The total income of the individual.	Number	#	No Null
Age	The age of the individual.	Number	#	No Null
Years_employed	The number of years the individual has been employed.	Number	#	No Null
Income_type	The type of income (e.g., employed, self-employed, etc.).	Text	T	No Null
Education_type	The education level of the individual.	Number	T	No Null
Family_status	The family status of the individual.	Number	T	No Null
Housing_type	The type of housing the individual lives in.	Number	T	No Null
Occupation_type	The type of occupation the individual is engaged in.	Number	T	No Null
Target	The target variable for the classification task, indicating whether the individual is eligible for a credit card or not (e.g., Yes/No, 1/0).	Number	Binary (0, 1) 0 = no 1 = yes	No Null

Data Processing and Preprocessing

The dataset was generally very clean, and only minimal preprocessing was required to ensure consistency and accuracy. The following adjustments were made to facilitate analysis:

- Formatting: Minor adjustments were made to standardize certain columns. For instance:
- The Total Income column was formatted as currency for better readability.
- Columns such as Age and Years Employed were rounded up to the nearest whole number to simplify analysis and visualization.

Insights

Employment Stability and Occupation Type

Hypothesis: Occupations with longer employment durations may indicate higher creditworthiness and thus a greater likelihood of credit card approval.

Findings:

- Employment duration does not strongly correlate with credit card approval. In fact, 15 out of 19 occupations where credit card approval was granted had applicants with below-average employment durations.
- Only 4 occupations, such as Realty agents and High-skill tech staff, showed above-average employment durations for approved applicants.
- Longer employment durations do not guarantee approval, as occupations with shorter employment durations (e.g., IT staff, waiters) also saw approvals.

Conclusion: Employment duration alone is not a reliable predictor of credit card approval. Other factors such as income, credit history, and occupation type may play more critical roles in the approval process.

Housing Status and Financial Stability

Hypothesis: Owning a property indicates financial stability and may increase credit card eligibility.

Findings:

- The vast majority of approved credit card holders (1,131 individuals) live in houses or apartments, making it the dominant housing type among those approved.
- Surprisingly, 64 individuals living with their parents were also approved, indicating that financial independence is not strictly required for approval.
- Other housing types, such as municipal apartments and rented apartments, made up a much smaller portion of approved applicants.

Conclusion: Credit card approvals are heavily skewed toward those living in houses or apartments, but individuals living with parents also have notable approval rates. This suggests that homeownership or independent living is not the sole determinant of creditworthiness.

Income Distribution by Housing Type

Key Observations:

- Homeowners tend to have higher incomes, particularly those earning above $300,000, with several individuals exceeding $500,000.
- Non-homeowners typically earn less, with income levels rarely exceeding $1,000,000, suggesting a potential link between higher incomes and property ownership.

Conclusion: Homeownership is more common among higher-income individuals, but a large proportion of both homeowners and non-homeowners have incomes clustered in the $100,000–$300,000 range. The presence of high-income outliers among homeowners suggests that higher income may increase the likelihood of owning property.

Digital Accessibility and Financial Behavior

Hypothesis: Individuals with access to modern communication tools (phone, work phone, and email) might be better candidates for credit due to improved financial management and accessibility.

Findings:

- No Access:
    - A significant portion of both approved (56%) and denied (55%) applicants had no access to any communication tools such as work phone, phone, or email. This shows that lack of access to communication tools is not a decisive factor in credit card approval or denial.
- Phone Access:
    - 14% of approved applicants and 16% of denied applicants had phone-only access. The higher percentage of denied applicants in this category suggests that having phone access alone may not significantly improve the chances of approval.
    - Additionally, work phone + phone access is slightly higher for denied applicants (11%) than approved applicants (10%).
- Work Phone Only: 1
    - 0% of approved applicants had work phone only access, while 9% of denied applicants had the same level of access. This indicates that access to a work phone alone does not strongly correlate with a higher likelihood of approval.

- Multiple Communication Tools:
  - A small percentage of both approved (2%) and denied (2%) applicants had access to both phone and email, suggesting that access to multiple tools is relatively rare and does not heavily influence the outcome.
  - All Methods Access: This is also relatively rare, with 1% in both groups, indicating that comprehensive access to communication tools does not strongly impact credit card approval.

Conclusion: The data suggests that access to communication tools, whether it's phone, work phone, email, or a combination, does not play a significant role in determining credit card approval. The high percentage of applicants with no access or phone-only access in both the approved and denied groups indicates that other factors, such as income or credit history, likely carry more weight in the decision-making process.

Overall Conclusion

The analysis of credit card eligibility factors has revealed that traditional assumptions about financial stability and communication access may not be reliable predictors of approval. While factors such as employment stability and housing status play a role in determining eligibility, they are not the sole determinants of approval, as other key elements like income levels and credit history seem to carry greater weight.

Employment Stability: Long-term employment does not guarantee credit card approval. The data shows that individuals from various occupations, including those with shorter employment durations, are still approved for credit cards, challenging the notion that longevity in a job is essential for creditworthiness.

Housing Status: Owning a home is more common among higher-income individuals, but surprisingly, living with parents does not preclude credit card approval. This suggests that homeownership, while a signal of financial stability, is not a decisive factor.

Income Distribution: Although homeowners tend to have higher incomes, the majority of both homeowners and non-homeowners have income levels clustered between $100,000 and $300,000. Therefore, while higher income may influence the likelihood of homeownership, it does not guarantee credit approval.

Digital Accessibility: Access to communication tools (work phone, phone, email) does not play a significant role in credit card approval decisions. The large percentage of approved and denied applicants with no communication access highlights that factors such as income and credit history are likely to be more crucial.

Ultimately, this analysis underscores the importance of multi-dimensional factors in determining creditworthiness. Employment duration, housing status, and digital accessibility alone are insufficient to predict credit card approval, suggesting that credit history and financial health metrics, such as debt-to-income ratios or existing credit obligations, may have a larger influence on approval outcomes. Financial institutions would benefit from incorporating a more comprehensive view of applicants' profiles when assessing creditworthiness.