

FINAL PROJECT



STATISTICAL ANALYSIS PROVIDER

LOAN APPROVAL PATTERNS THROUGH
DATA ANALYTICS

DECEMBER 20, 2024

OBJECTIVE

- **Overview**

IronSight Analytics is dedicated to transforming the way financial institutions and individuals approach decision-making by leveraging advanced data analytics and predictive modeling. With a commitment to transparency, innovation, and actionable insights, IronSight empowers businesses to optimize their strategies, mitigate risks, and foster lasting relationships with their clients. Our customer-centric approach ensures that every solution is tailored to the unique needs of our partners, driving growth and efficiency in the financial landscape.

- **Main question:**

- Why Analyze Loan Data? Identify key factors that influence loan approval decisions.
- Predict customer likelihood for loan acceptance.
- Improve financial strategies by understanding demographic and behavioral patterns.
- Highlight how data-driven insights can transform decision-making.
- How can loan proposals be made more appealing and readily accepted by current and prospective customers through improved marketing strategies?



DATASET OVERVIEW

Dataset size:

Dataset size: 5000 records and 14 features, including demographics (age, income, family size), financial behavior (credit card usage, online banking), and outcomes (personal loan acceptance).

Data preprocessing steps:

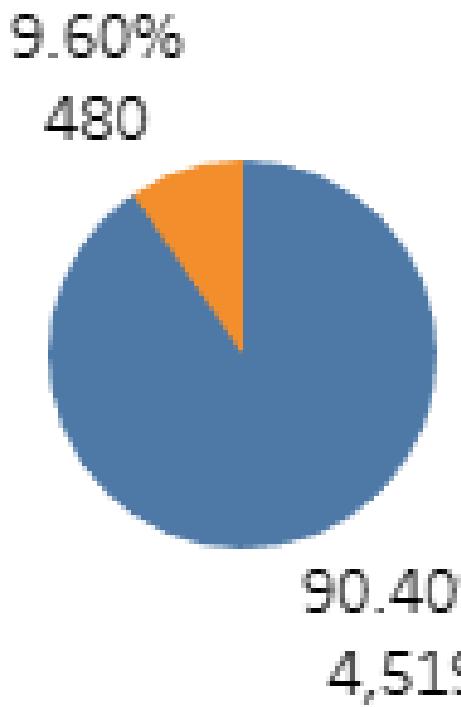
Check shape , data type , frequency , missing values, duplicates, handled categorical data, normalized numerical features and describe the data. Clean the noise , transformed CCAvg from monthly to yearly spending and Converted negative Experience values to positive



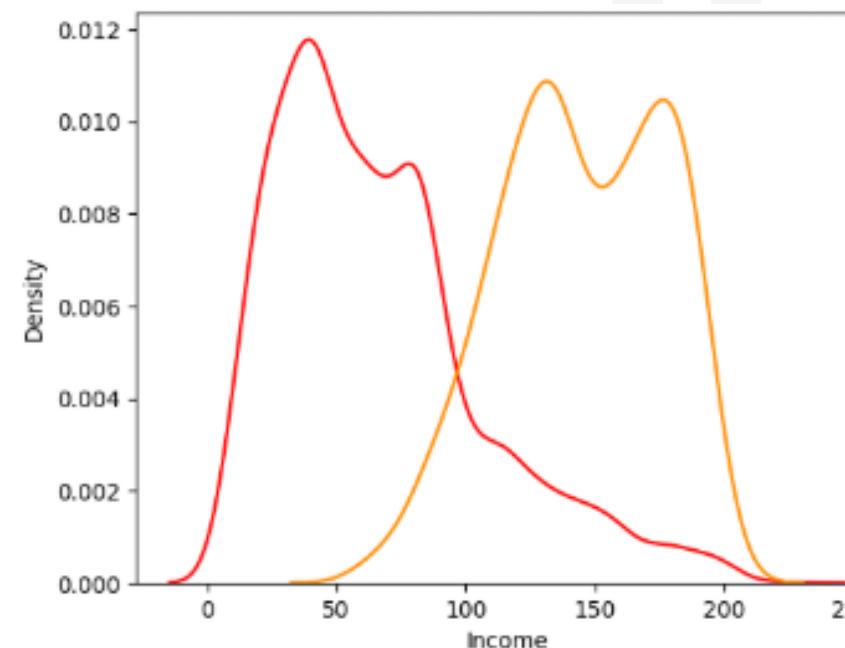
IronSight Analytics

UNDERSTANDING THE DATA

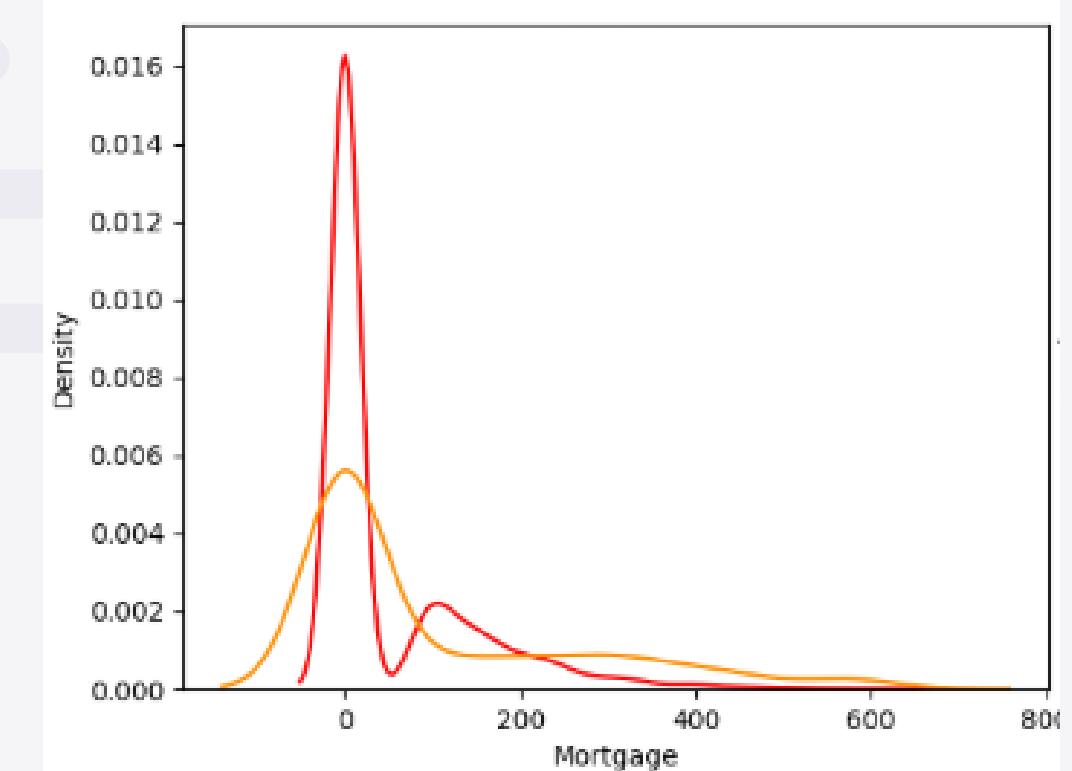
Percentage of customers
that accepted personal
loans in the last campaign



High-income groups tend to accept
loans more frequently



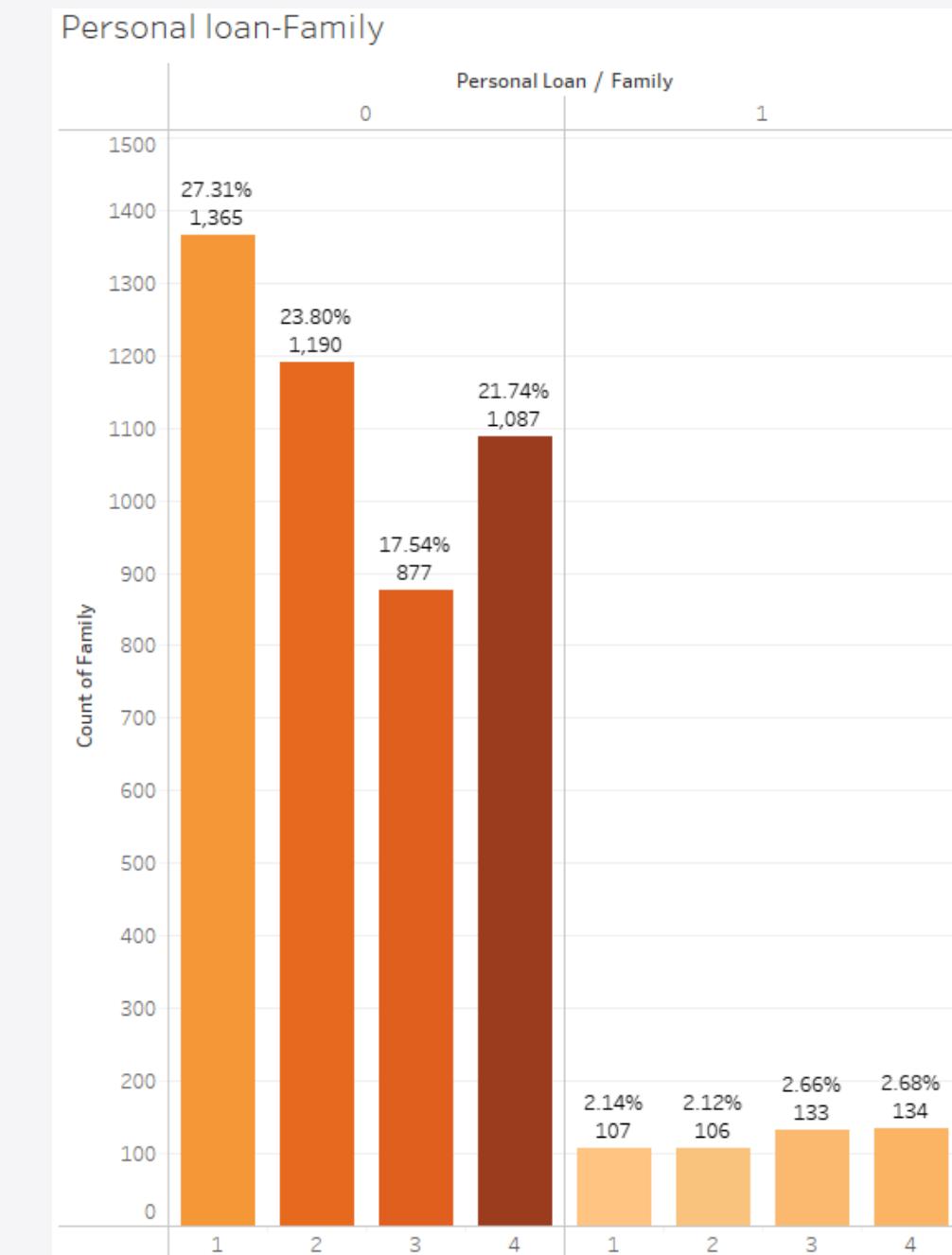
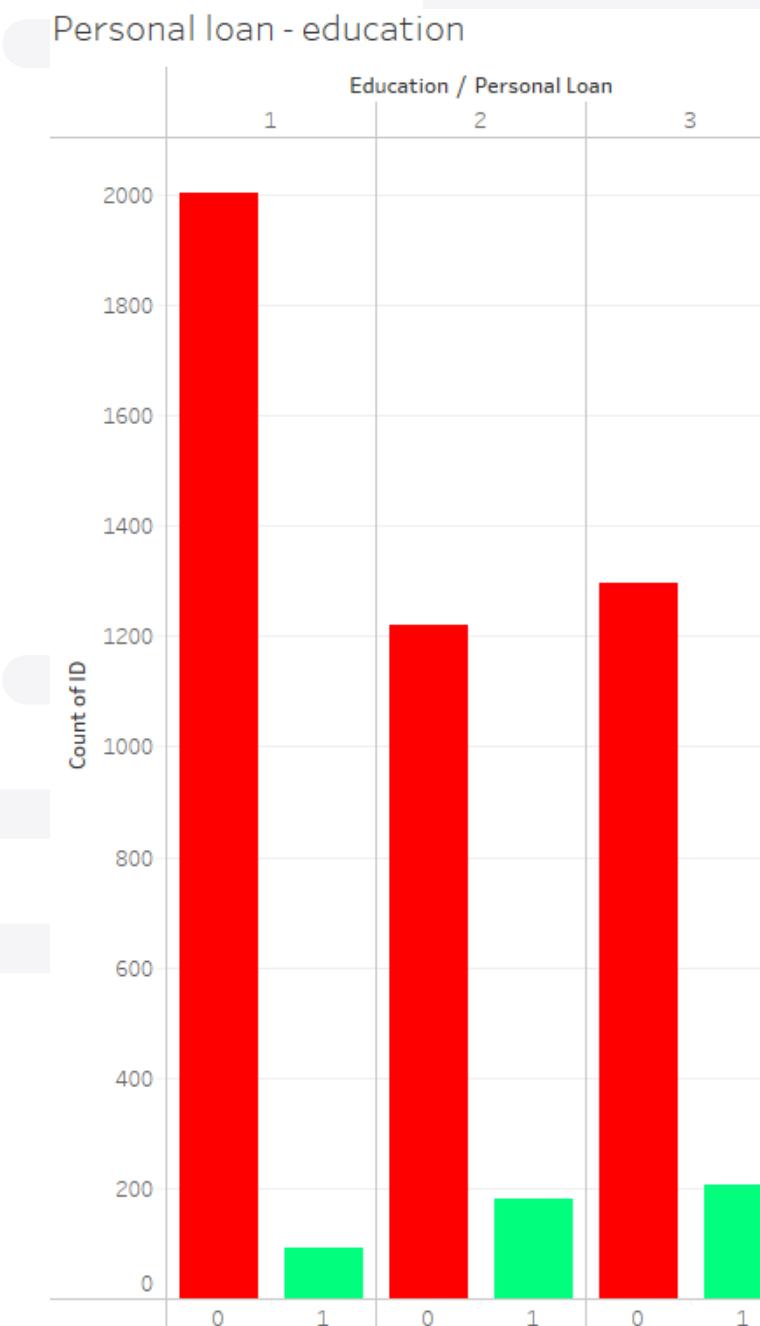
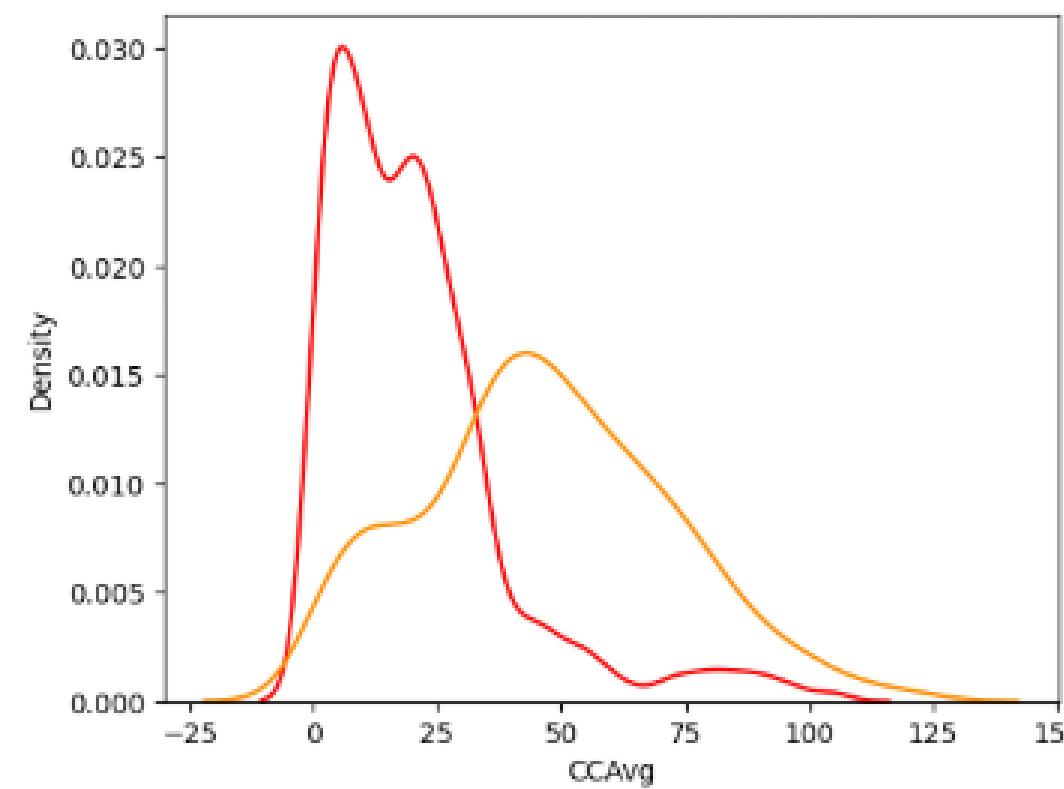
**MOST PEOPLE WHO RECEIVED A LOAN, HAD ZERO
MORTAGAGE**



Correlation Matrix														
	Age	Experience	Income	ZIP Code	Family	CCAvg	Education	Mortgage	Personal Loan	Securities Account	CD Account	Online	CreditCard	Cluster
Age	1.00	0.99	-0.06	-0.03	-0.05	-0.05	0.04	-0.01	-0.01	-0.00	0.01	0.01	0.01	0.02
Experience	0.99	1.00	-0.05	-0.03	-0.05	-0.05	0.01	-0.01	-0.01	-0.00	0.01	0.01	0.01	0.02
Income	-0.06	-0.05	1.00	-0.03	-0.16	0.65	-0.19	0.21	0.50	-0.00	0.17	0.01	-0.00	0.02
ZIP Code	-0.03	-0.03	-0.03	1.00	0.03	-0.01	-0.01	0.00	-0.00	0.00	0.02	0.03	0.02	-0.88
Family	-0.05	-0.05	-0.16	0.03	1.00	-0.11	0.06	-0.02	0.06	0.02	0.01	0.01	0.01	-0.02
CCAvg	-0.05	-0.05	0.65	-0.01	-0.11	1.00	-0.14	0.11	0.37	0.02	0.14	-0.00	-0.01	0.01
Education	0.04	0.01	-0.19	-0.01	0.06	-0.14	1.00	-0.03	0.14	-0.01	0.01	-0.02	-0.01	0.01
Mortgage	-0.01	-0.01	0.21	0.00	-0.02	0.11	-0.03	1.00	0.14	-0.01	0.09	-0.01	-0.01	0.00
Personal Loan	-0.01	-0.01	0.50	-0.00	0.06	0.37	0.14	0.14	1.00	0.02	0.32	0.01	0.00	-0.00
Securities Account	-0.00	-0.00	-0.00	0.00	0.02	0.02	-0.01	-0.01	0.02	1.00	0.32	0.01	-0.01	-0.00
CD Account	0.01	0.01	0.17	0.02	0.01	0.14	0.01	0.09	0.32	0.32	1.00	0.18	0.28	-0.03
Online	0.01	0.01	0.01	0.03	0.01	-0.00	-0.02	-0.01	0.01	0.01	0.18	1.00	0.00	-0.03
CreditCard	0.01	0.01	-0.00	0.02	0.01	-0.01	-0.01	-0.01	0.00	-0.01	0.28	0.00	1.00	-0.02
Cluster	0.02	0.02	0.02	-0.88	-0.02	0.01	0.01	0.00	-0.00	-0.00	-0.03	-0.03	-0.02	1.00



UNDERSTANDING THE DATA



MODELING LOAN Acceptance

Models used

- Tested multiple models:
- Logistic Regression ;
 - Decision Tree ;
 - Random Forest ;
 - K neighbours ;
 - SVC ;
 - K-Means Clustering ;
 - PCA ;
 - Gaussian Naïve Bayes ;
 - AdaBoost.

Method

- Find the best parameters for each model
- Emphasized cross-validation to ensure robust results.

Best Model

Random Forest achieved the highest accuracy of 99,78% . Using the parameters :

- bootstrap=False
- max_depth=10
- random_state=42

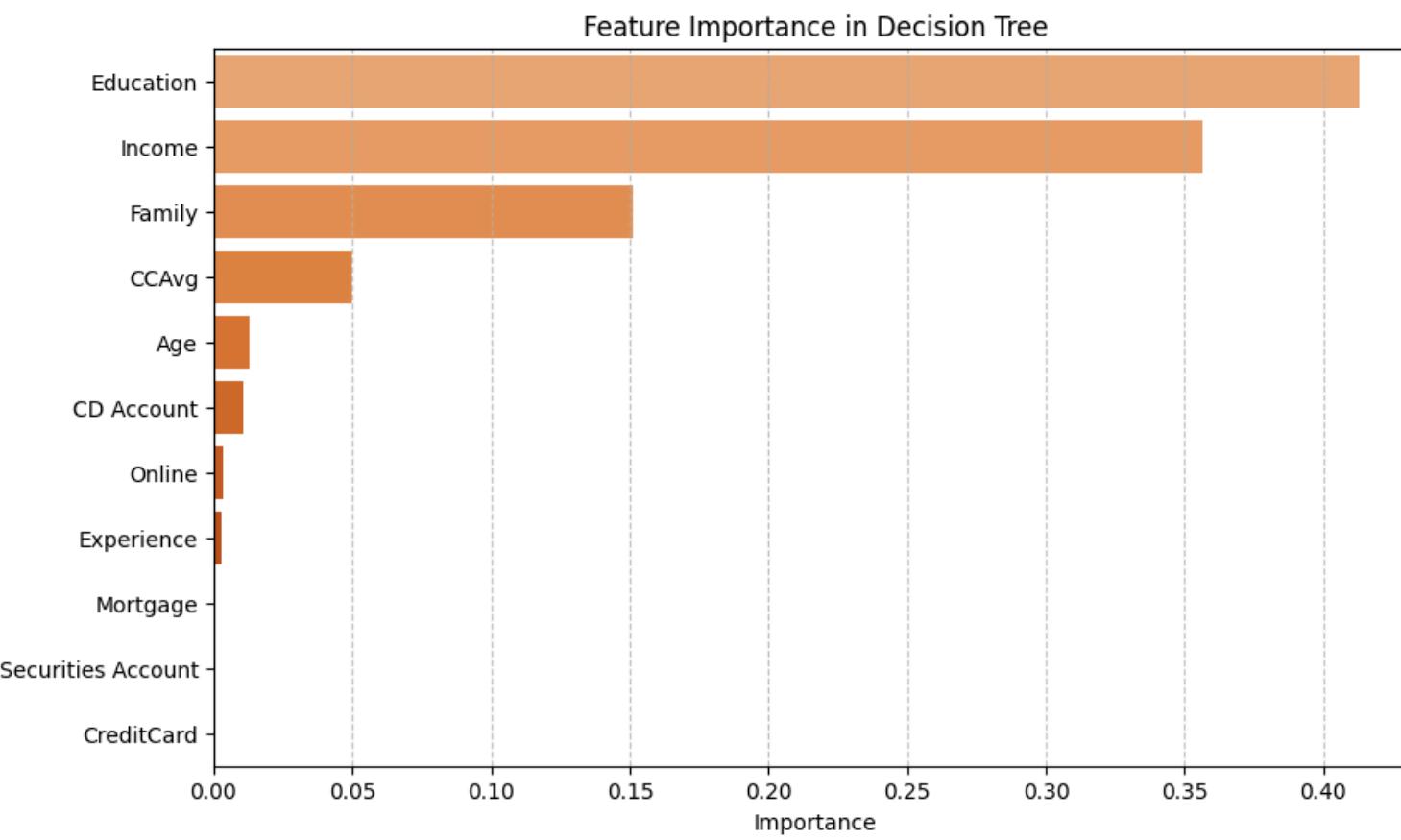
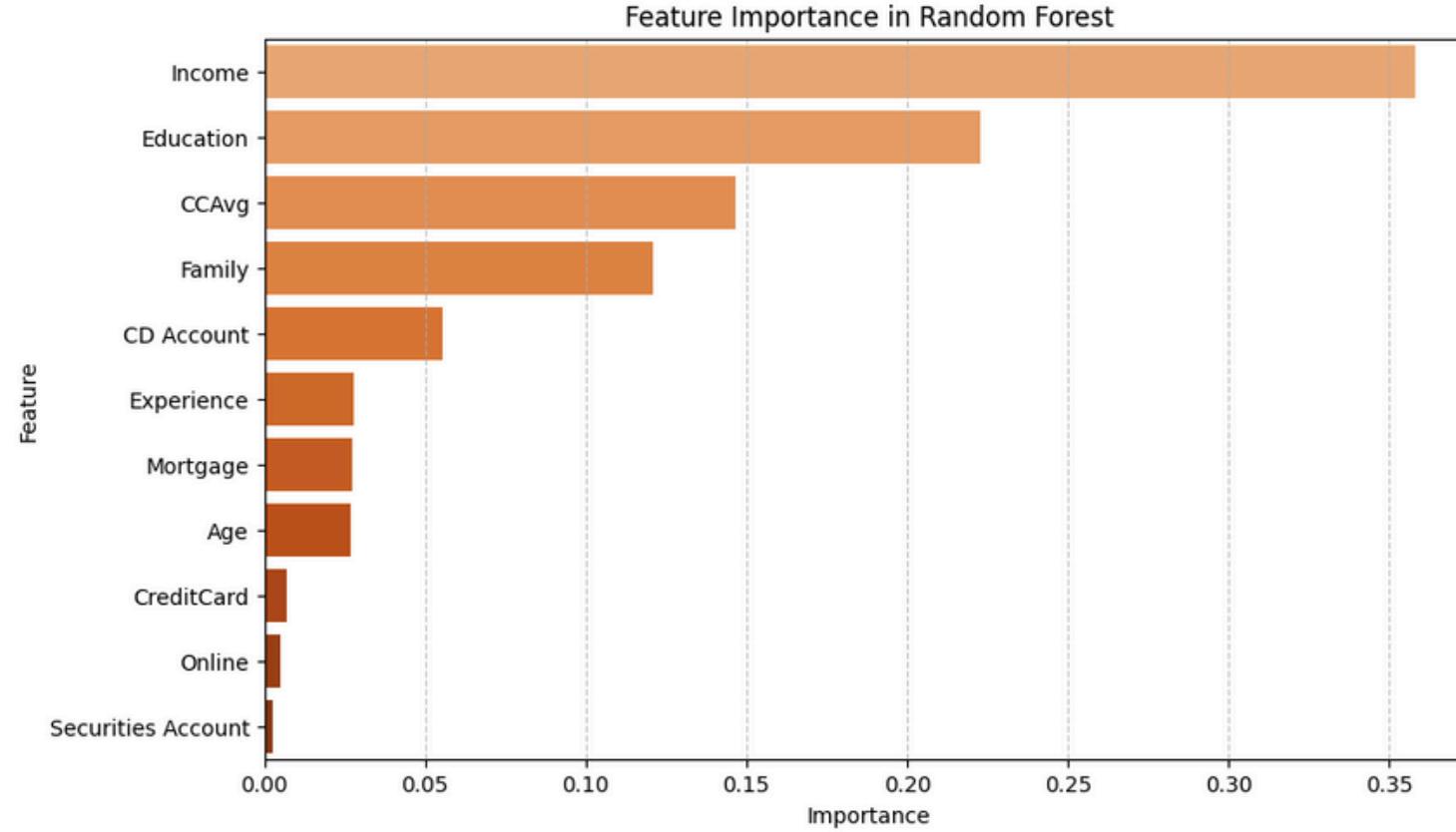


IronSight Analytics

WHAT DRIVES LOAN ACCEPTANCE?

Income, average credit card spending (CCAvg), education and family size emerged as significant predictors.

Experience , CD account ownership, Age and some of others features also influenced decisions even though they are less significants.





- Personal loan acceptance rates by income group, and other major insights

Interactive Insights



From Data to Decisions

- Target high-income individuals with customized loan proposals.
- Emphasize benefits aligned with education level and spending habits.
- Leverage insights on credit card spending (CCAvg) and family size to segment audiences for tailored offers.





From Data to Decisions

Loan markets and customer behaviors are dynamic. Regular monitoring ensures that models stay relevant and effective.

Incorporating fresh data helps:

- Adapt to emerging trends, such as changes in economic conditions or customer preferences.
- Enhance predictive accuracy by reflecting the latest patterns.
- Maintain a competitive edge by evolving alongside market demands.



A photograph of two business professionals in dark suits shaking hands over a wooden desk. On the desk is an open laptop, a pair of glasses, and some papers. The background is slightly blurred.

THANK YOU