

Customer Conversion

Princy Iakov

TABLE OF CONTENTS

01

Data Insights

Understanding the data insights

02

Models Explored

Various models explored along with the metrics

03

Scripts and Web Application

Dive into the python scripts used and web application created

04

Docker

Docker file for deployment

01

Data Insights

KEY NUMBERS

5,000

This is the total
number of
observations

29

Null Values

480

Accepted the
personal loan

PERSONAL LOAN

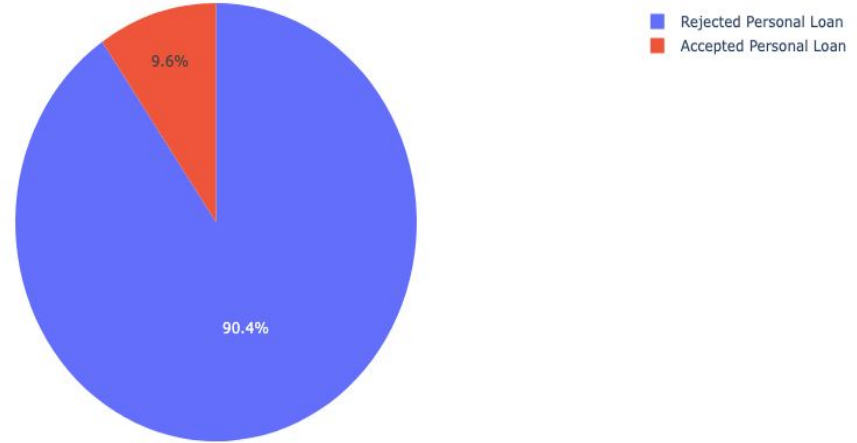
Accepted Personal loan offer

9.6% i.e 480 people
accepted the personal
loan

Rejected Personal loan offer

90.4% i.e 4520 people
rejected the personal
loan

Percentage of Personal Loan

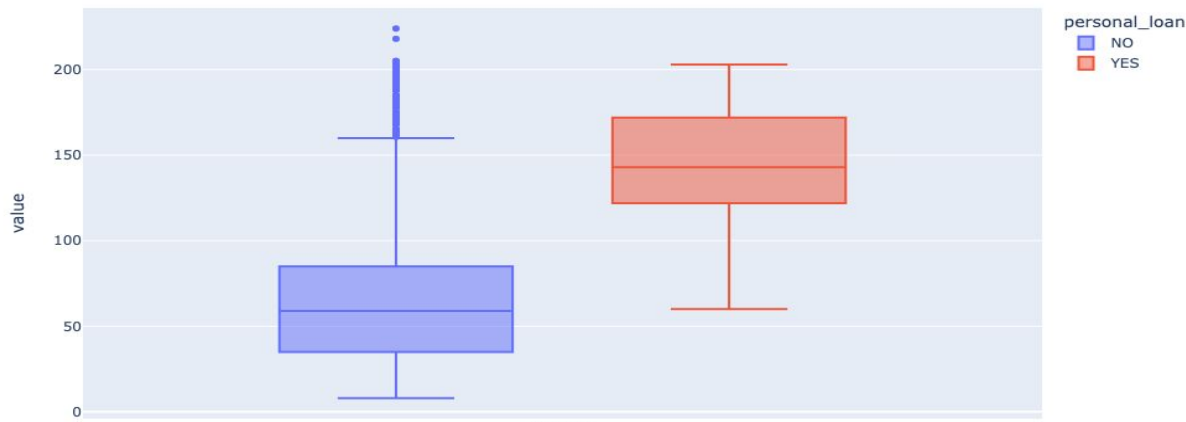
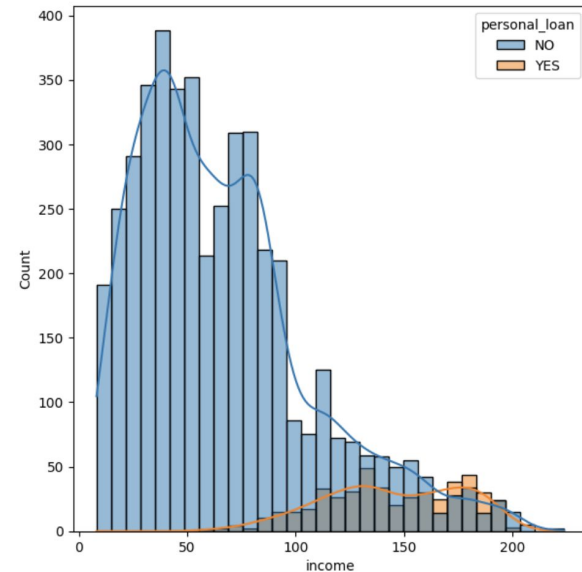
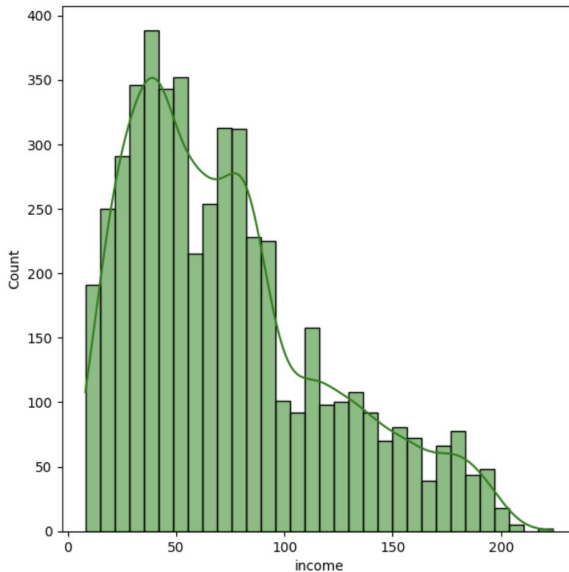


INCOME

Data distribution is right skewed with people with majority earning less than 100k

We see a distinct difference between people accepting and rejecting personal loans in the box plot with median value 143k for personal loan acceptance

People with income ranging between 122k to 172k tend to have higher chance of accepting the loan

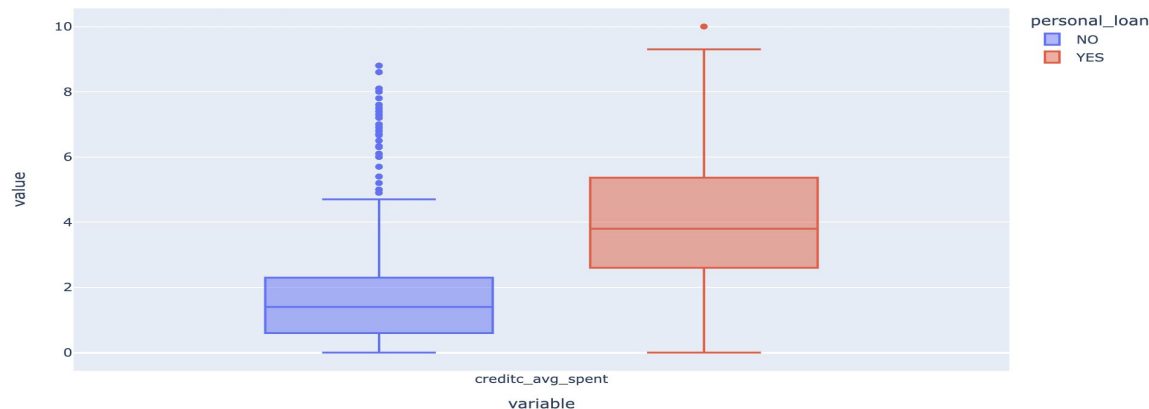
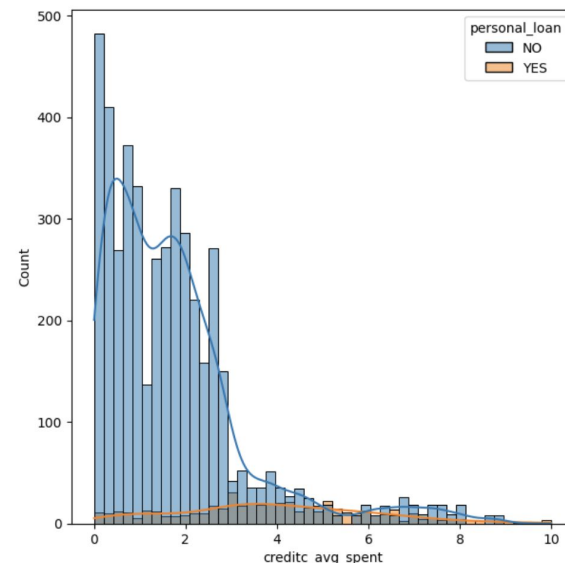
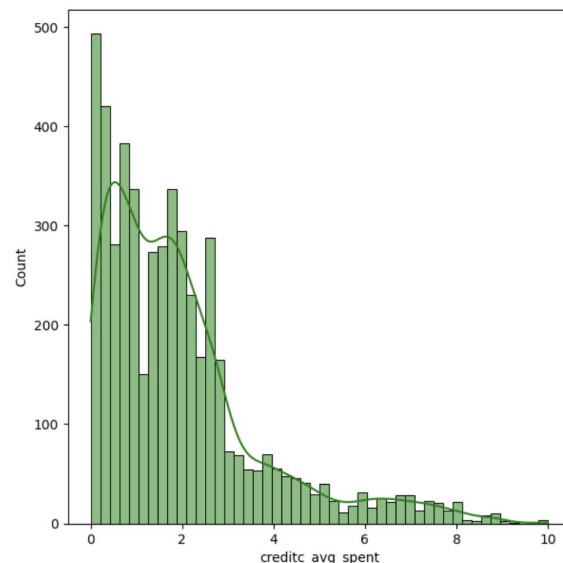


CREDIT CARD AVG SPENDING

Data distribution is right skewed. Majority spending less than 3k per month

We see a difference between people accepting and rejecting personal loans in the box plot with median value 3.8k for personal loan acceptance

People with average spending ranging between 2.6k to 5.4k have higher chance of accepting the loan.



DEPOSIT ACCOUNT

Though the number of people with deposit account is very low, we see a significant difference in loan acceptance rate

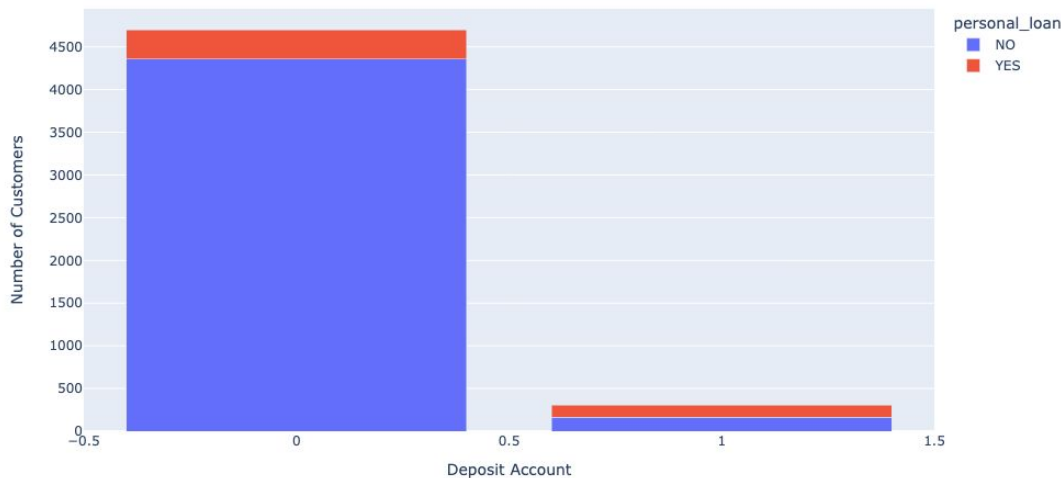
Accepted Personal loan offer

46% of people having deposit account accepted personal loan compared to 7.24 % without an account

Rejected Personal loan offer

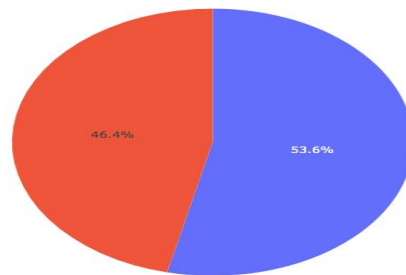
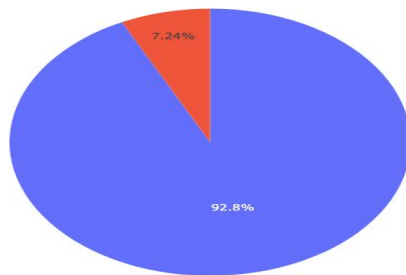
Only 53% of people having deposit account rejected personal loan compared to 92.8% without an account

Personal Loans by Deposit Account



No Deposit Account

Has Deposit Account



EDUCATION

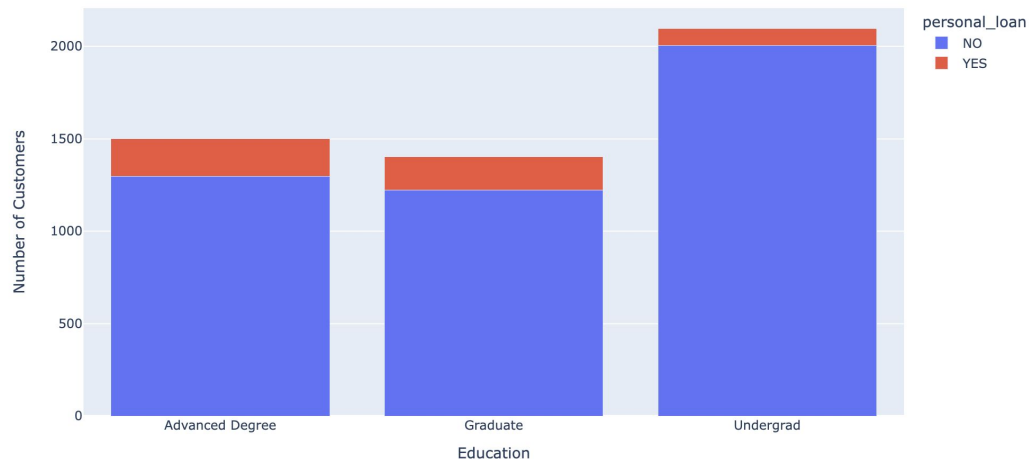
Accepted Personal loan offer

Clients with higher education tend to accept the personal loan more with 13.7% for Advanced degree, 13% Graduate vs 4.44% acceptance for undergraduate

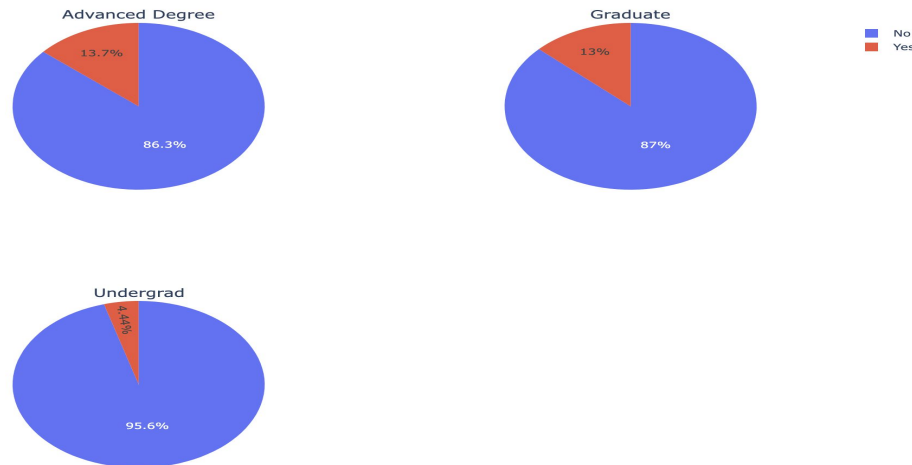
Rejected Personal loan offer

Clients with less education tend to reject the personal loan more with 95.6% rejection for Undergraduate vs 87% for Graduate, 86.3% Advanced degree

Personal Loans by Education Level



Personal Loan Distribution Based On Education

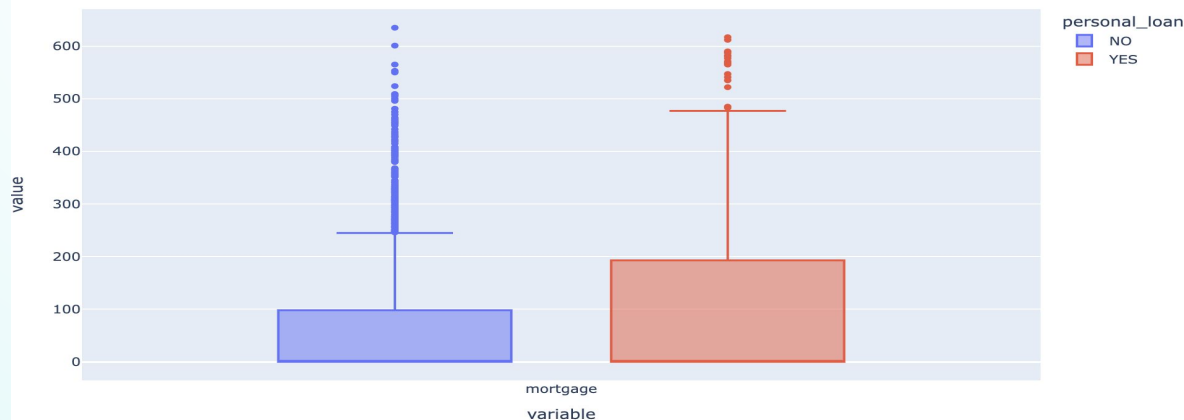
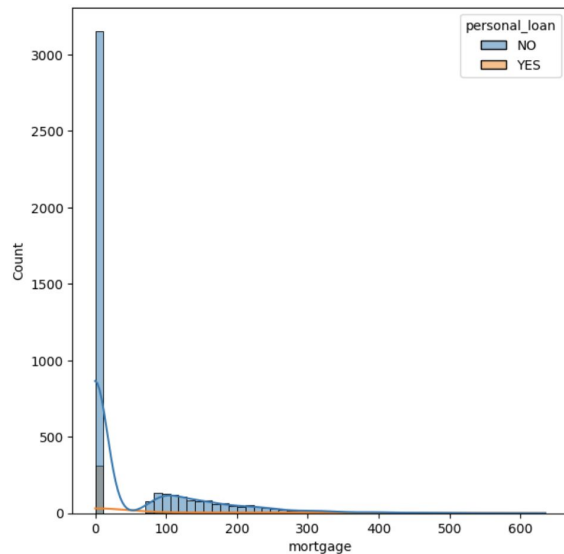
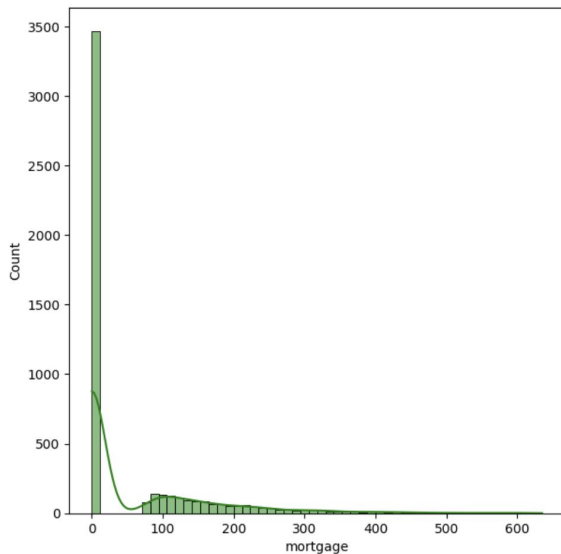


MORTGAGE

Data distribution is right skewed with people with Maximum people having no mortgage

We see a mild difference between people accepting and rejecting personal loans in the box plot with median value 0

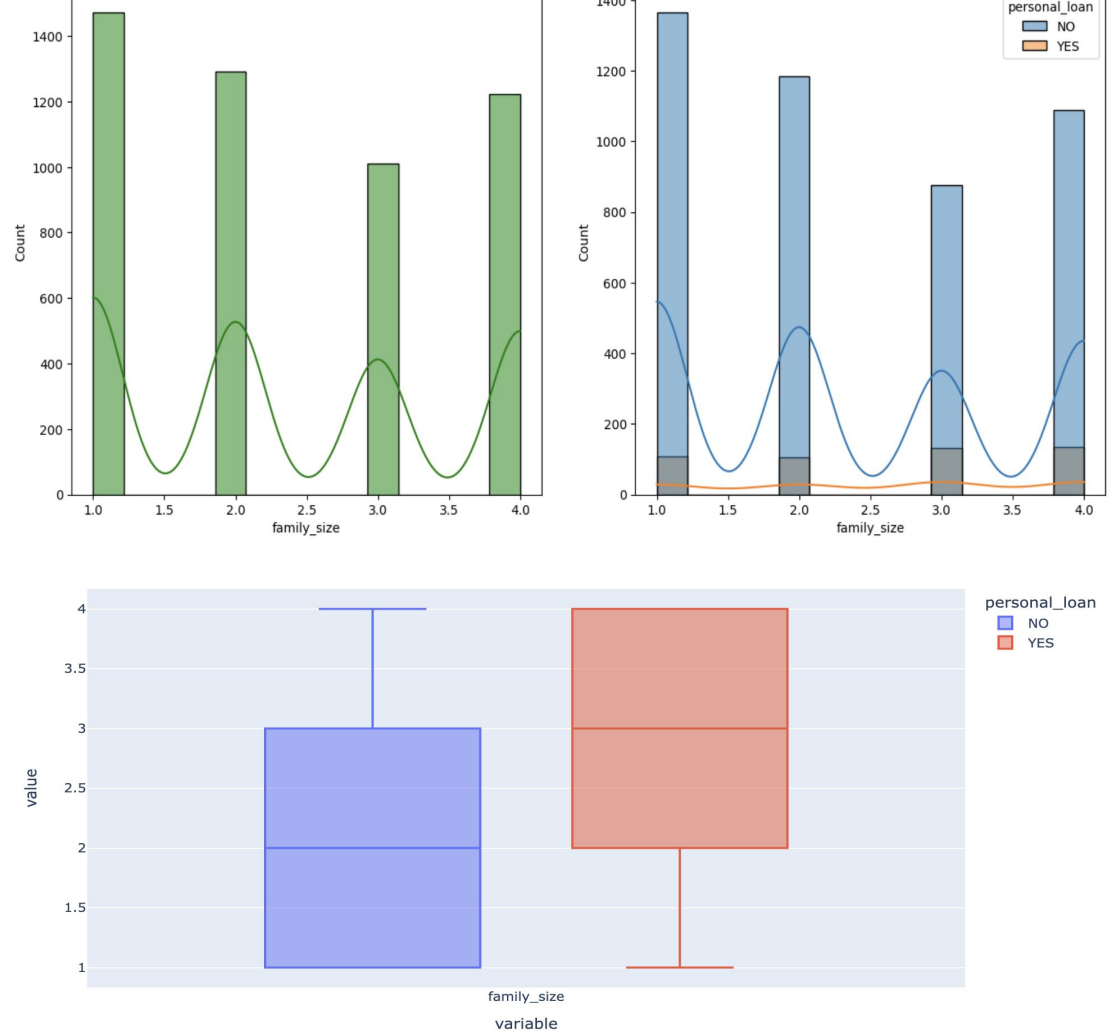
50% of people accepting loan have mortgage ranging from 0 to 193k



FAMILY SIZE

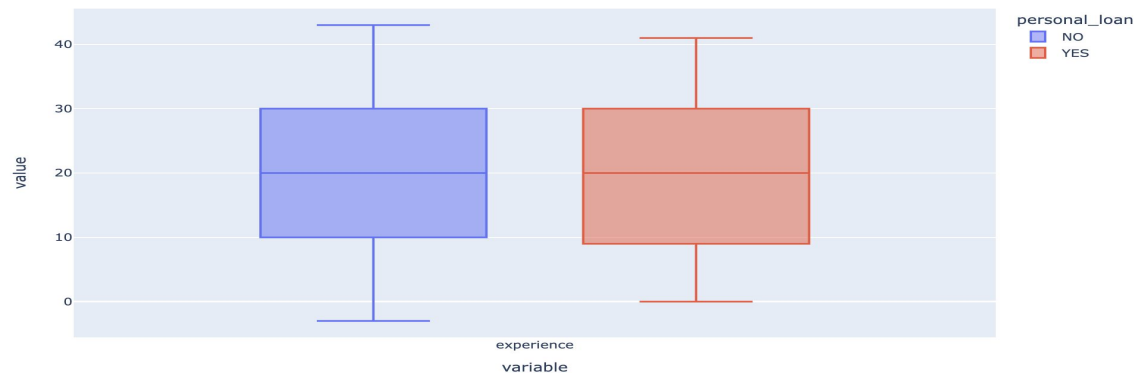
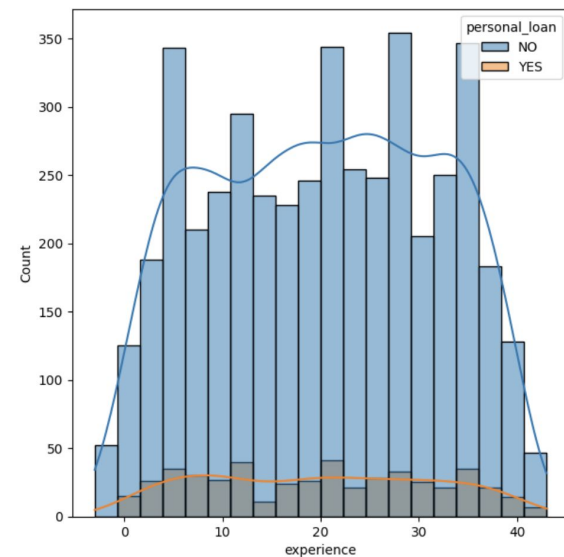
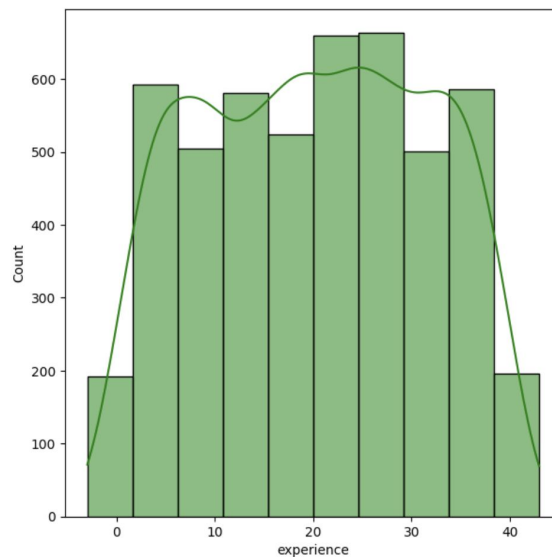
Box plot shows that there is a mild difference in distribution for personal loan acceptance or rejection

Clients with Family size ranging from 2-4 have higher chances of personal loan acceptance with median value 3 for personal loan acceptance



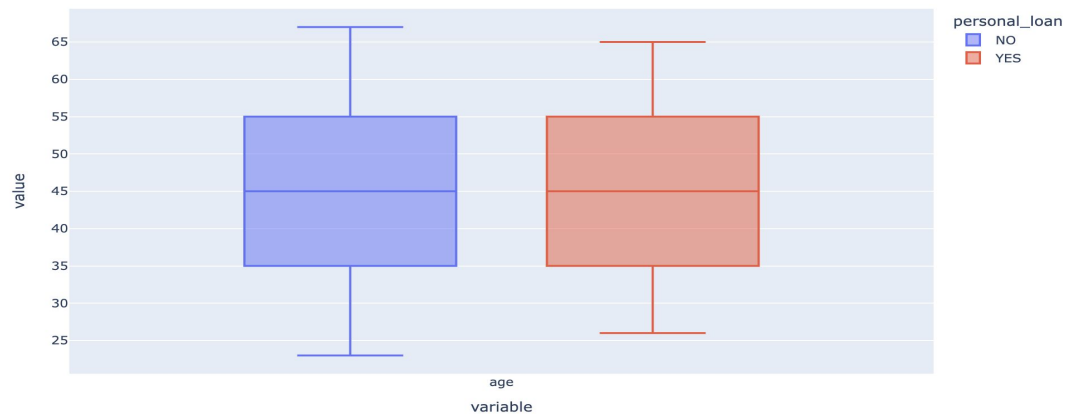
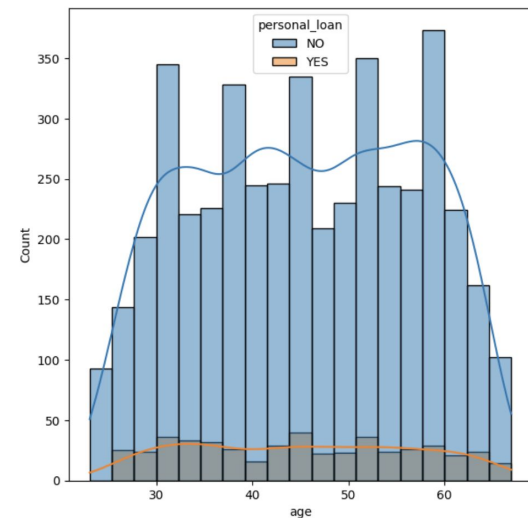
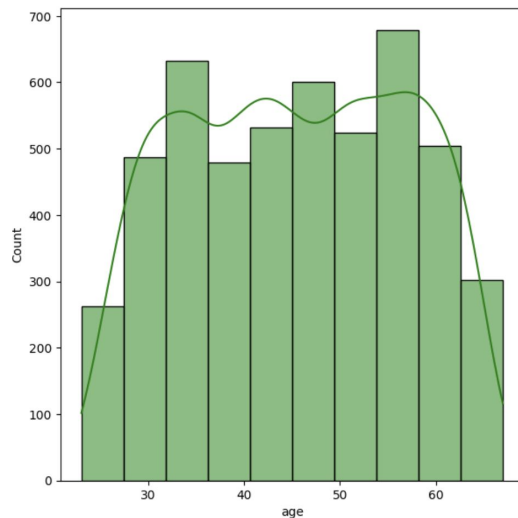
EXPERIENCE

Box plot shows that there is hardly any difference in distribution for personal loan acceptance or rejection



AGE

Box plot shows that there is hardly difference in distribution for personal loan acceptance or rejection





Primary Target Audience

01

Income

Income ranging between 122k to 172k

02

Average Credit Card Spending

Credit card average spending ranging between 2.6k to 5.4k

03

Deposit Account

People with deposit account

PROCESS

EDA

Click [here](#) for Github repository and [here](#) Notebook for EDA

Dockerfile

Dockerfile created for seamless deployment of project



Feature Selection, Modelling

Click [here](#) for Notebook

End to end pipeline Scripts with web application

You can navigate through the scripts by clicking [here](#)

02

Models Explored

MODELS & THEIR METRICS

I explored the following models with and without SMOTE oversampling . Models trained on SMOTE oversampled data performed better.

| MODEL NAME | ROC AUC SCORE | PRECISION | RECALL | F1 SCORE | ACCURACY |
|----------------------|---------------|---|---|---|----------|
| Random Forest | 0.963 | 0 0.99 1 0.90 macro avg 0.95 weighted avg 0.98 | 0 0.99 1 0.94 macro avg 0.96 weighted avg 0.98 | 0 0.99 1 0.92 macro avg 0.95 weighted avg 0.98 | 0.98 |
| Gradient Boosting | 0.969 | 0 0.99 1 0.95 macro avg 0.97 weighted avg 0.99 | 0 0.99 1 0.94 macro avg 0.97 weighted avg 0.99 | 0 0.99 1 0.94 macro avg 0.97 weighted avg 0.99 | 0.99 |
| XGBClassifier | 0.955 | 0 0.99 1 0.95 macro avg 0.97 weighted avg 0.99 | 0 0.99 1 0.92 macro avg 0.96 weighted avg 0.99 | 0 0.99 1 0.93 macro avg 0.96 weighted avg 0.99 | 0.99 |
| KNeighborsClassifier | 0.932 | 0 0.97 1 0.89 macro avg 0.94 weighted avg 0.98 | 0 0.99 1 0.88 macro avg 0.93 weighted avg 0.98 | 0 0.99 1 0.88 macro avg 0.94 weighted avg 0.8 | 0.98 |

03

Scripts & Web Application

SCRIPTS

data_ingestion.py

Script for reading data from source and ingesting data splits and model creation

model_trainer.py

Trains models on different hyperparameters and saves the best model

app.py

Runs web application built on flask for prediction

data_transformation.py

Transform data using sklearn pipeline. Handle missing data, apply Scalar transformation

predict_pipeline.py

Performs prediction using saved model

Web Application

← → ↻ ⓘ localhost:3000/predict

Customer Conversion Prediction

Choose the level of Education:

Has Deposit Account:

How many members are there in the Family:

Has Investment Account:

Annual Income (\$1000):

Credit Card Avg Spending per Month(\$1000):

Value of Home Mortgage if any (\$1000):

FINAL PREDICTION: Will take Personal Loan

04

DockerFile

DOCKERFILE

Created a dockerfile to enable smooth deployment on any platform

Click [here](#) to see the commands to run the docker file

The background of the slide is a light teal color. It features several abstract shapes: a large, faint teal circle on the left, a smaller teal circle on the right, and a teal wavy line at the bottom. Two vertical teal lines are also present, one on the left and one on the right.

THANK YOU!

Do you have any
questions?

princy.iakov@gmail.com

+65 81099121

<https://github.com/principiakov>

CREDITS: This presentation
template was created by Slidesgo,
including icons by Flaticon, and
infographics & images by Freepik.