Problem Statement –

Develop a basic understanding of risk analytics in banking and financial services and understand how data is used to minimise the risk of losing money while lending to customers.

Solution –

With our dashboards which are created using Power BI latest tools helps the company to make a decision based on the applicant's profile like if the applicant is likely to repay the loan then approving the loan otherwise not.

About The Data set:-

This dataset contains detailed financial and demographic information for 3,000 banking clients. Each record includes attributes such as age, income, occupation, nationality, loyalty classification, number of credit cards, savings, loans, and account balances. The dataset also captures customer behaviors across various banking services like superannuation savings, business lending, foreign currency accounts, and risk weighting. It can be used for customer segmentation, risk profiling, and financial product analysis.

Data Cleaning :-

1. Check for duplicate rows using the "Client ID" column as a unique identifier.
2. Check for nulls in key columns like Age, Occupation, Nationality, or financial figures.
3. Fill missing values appropriately (e.g., use average/median for numeric fields, "Unknown" for text).
4. Convert "Joined Bank" column to proper date format.
5. Ensure numeric columns (e.g., Deposits, Lending, Risk Weighting) are properly typed as numbers.
6. Trim and clean names and occupations (remove leading/trailing spaces, fix inconsistent casing).
7. Normalize values in "Loyalty Classification" or "Fee Structure" (e.g., unify spelling).
8. Replace IDs like GenderId, BRId, or IAId with meaningful labels if mappings are known or provided.

EDA Using Python:-

Exploratory Data Analysis (EDA) summary based on Banking dataset using pandas, matplotlib, and seaborn

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sbs
import numpy as np
```

```python
df=pd.read_csv('/content/Banking.csv')
df.head()
```

	Client ID	Name	Age	Location ID	Joined Bank	Banking Contact	Nationality	Occupation	Fee Structure	Loyalty Classification	...	Bank Deposits	Checking Accounts	Saving Accounts	Foreign Currency Account	Business Lending	Properties Owned	Risk Weighting	BRId	GenderId	IAId
0	IND81288	Raymond Mills	24	34324	06-05-2019	Anthony Torres	American	Safety Technician IV	High	Jade	...	1485828.64	603617.88	607332.46	12249.96	1134475.30	1	2	1	1	1
1	IND65833	Julia Spencer	23	42205	10-12-2001	Jonathan Hawkins	African	Software Consultant	High	Jade	...	641482.79	229521.37	344635.16	61162.31	2000526.10	1	3	2	1	2
2	IND47499	Stephen Murray	27	7314	25-01-2010	Anthony Berry	European	Help Desk Operator	High	Gold	...	1033401.59	652674.69	203054.35	79071.78	548137.58	1	3	3	2	3
3	IND72498	Virginia Garza	40	34594	28-03-2019	Steve Diaz	American	Geologist II	Mid	Silver	...	1048157.49	1048157.49	234685.02	57513.65	1148402.29	0	4	4	1	4
4	IND60181	Melissa Sanders	46	41269	20-07-2012	Shawn Long	American	Assistant Professor	Mid	Platinum	...	487782.53	446644.25	128351.45	30012.14	1674412.12	0	3	1	2	5

5 rows × 25 columns

```python
df.describe()
```

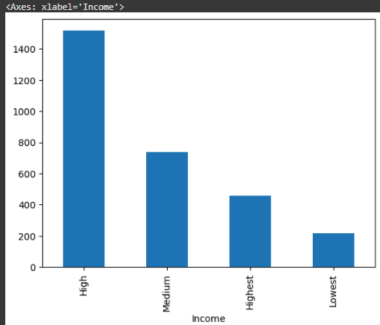
	Age	Location ID	Estimated Income	Superannuation Savings	Amount of Credit Cards	Credit Card Balance	Bank Loans	Bank Deposits
count	3000.000000	3000.000000	3000.000000	3000.000000	3000.000000	3000.000000	3.000000e+03	3.000000e+03
mean	51.039667	21563.323000	171305.034263	25531.599673	1.463667	3176.206943	5.913862e+05	6.715602e+05
std	19.854760	12462.273017	111935.808209	16259.950770	0.676387	2497.094709	4.575570e+05	6.457169e+05
min	17.000000	12.000000	15919.480000	1482.030000	1.000000	1.170000	0.000000e+00	0.000000e+00
25%	34.000000	10803.500000	82906.595000	12513.775000	1.000000	1236.630000	2.396281e+05	2.044004e+05
50%	51.000000	21129.500000	142313.480000	22357.355000	1.000000	2560.805000	4.797934e+05	4.633165e+05
75%	69.000000	32054.500000	242290.305000	35464.740000	2.000000	4522.632500	8.258130e+05	9.427546e+05
max	85.000000	43369.000000	522330.260000	75963.900000	3.000000	13991.990000	2.667557e+06	3.890598e+06

	Checking Accounts	Saving Accounts	Foreign Currency Account	Business Lending	Properties Owned	Risk Weighting	BRId	GenderId	IAId
count	3.000000e+03	3.000000e+03	3000.000000	3.000000e+03	3000.000000	3000.000000	3000.000000	3000.000000	3000.000000
mean	3.210929e+05	2.329084e+05	29883.529993	8.667598e+05	1.518667	2.249333	2.559333	1.504000	10.425333
std	2.820796e+05	2.300078e+05	23109.924010	6.412303e+05	1.102145	1.131191	1.007713	0.500067	5.988242
min	0.000000e+00	0.000000e+00	45.000000	0.000000e+00	0.000000	1.000000	1.000000	1.000000	1.000000
25%	1.199475e+05	7.479440e+04	11916.542500	3.748251e+05	1.000000	1.000000	2.000000	1.000000	5.000000
50%	2.428157e+05	1.640866e+05	24341.190000	7.113147e+05	2.000000	2.000000	3.000000	2.000000	10.000000
75%	4.348749e+05	3.155750e+05	41966.392500	1.185110e+06	2.000000	3.000000	3.000000	2.000000	15.000000
max	1.969923e+06	1.724118e+06	124704.870000	3.825962e+06	3.000000	5.000000	4.000000	2.000000	22.000000

```python
df['Estimated Income'].min()
```

```python
bins=[30000,50000,100000,300000,float('inf')]
labels=['Lowest','Medium','High','Highest']
df['Income']=pd.cut(df['Estimated Income'],bins=bins,labels=labels)
df.head()
df['Income'].value_counts().plot(kind='bar')
```

<Axes: xlabel='Income'>



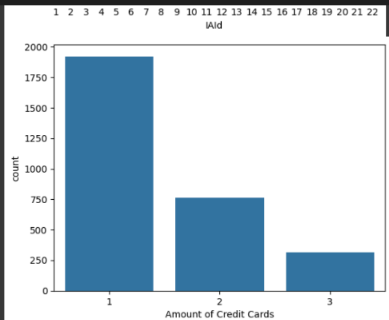

```python
categories=df[["BRId","GenderId","IAId","Amount of Credit Cards","Loyalty Classification","Occupation","Nationality","Properties Owned","Income"]]
for col in categories:
  print(f"value counts for'(col)':")
  display(df[col].value_counts())
```

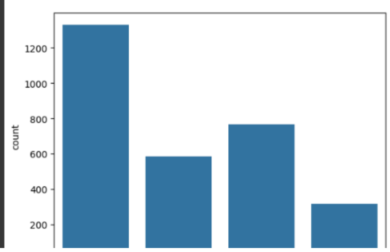
value counts for'(col)':

BRId	count
3	1352
1	660
2	495

Univariete Analysis: Analysis on All Columns

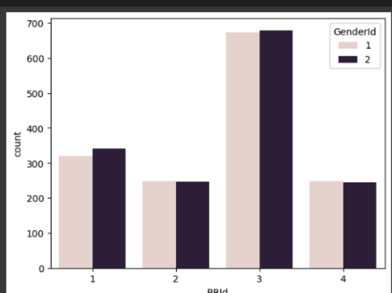
```
[ ] for i, predictor in enumerate(df[["BRId","GenderId","IAId","Amount of Credit Cards","Loyalty Classification","Occupation","Nationality","Properties Owned","Income"]]):
        plt.figure(i)
        sbs.countplot(data=df, x=predictor)
```



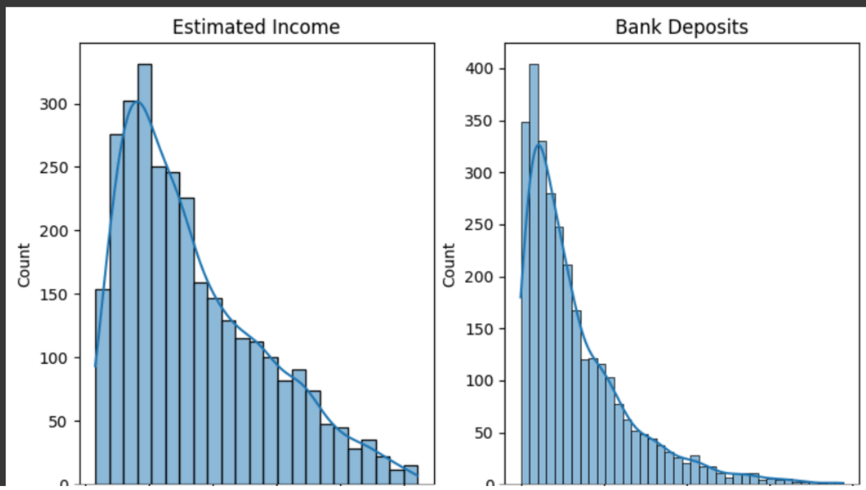


Bivariate Analysis: Analysis on Specific Columns to get data more in detail in visual charts

```
[ ] for i, predictor in enumerate(df[["BRId","GenderId","IAId","Amount of Credit Cards","Loyalty Classification","Occupation","Nationality","Properties Owned","Income"]]):
        plt.figure(i)
        sbs.countplot(data=df, x=predictor, hue='GenderId')
```



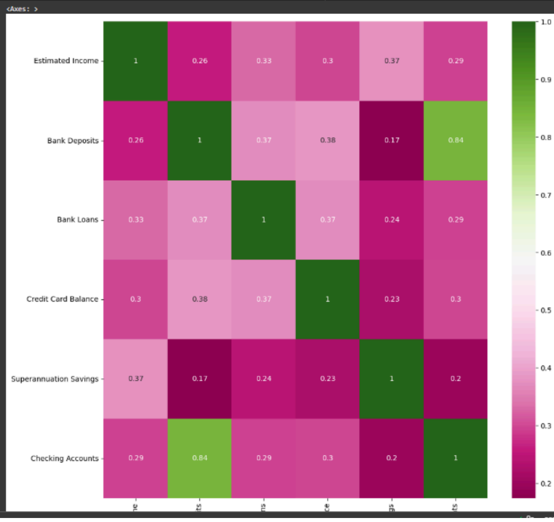
```
numerical=['Estimated Income','Bank Deposits','Bank Loans','Credit Card Balance','Superannuation Savings','Checking Accounts']
plt.figure(figsize=(9,36))
for i, col in enumerate(numerical):
  plt.subplot(6,2,i+1)
  sbs.histplot(df[col],kde=True)
  plt.title(col)
plt.show()
```



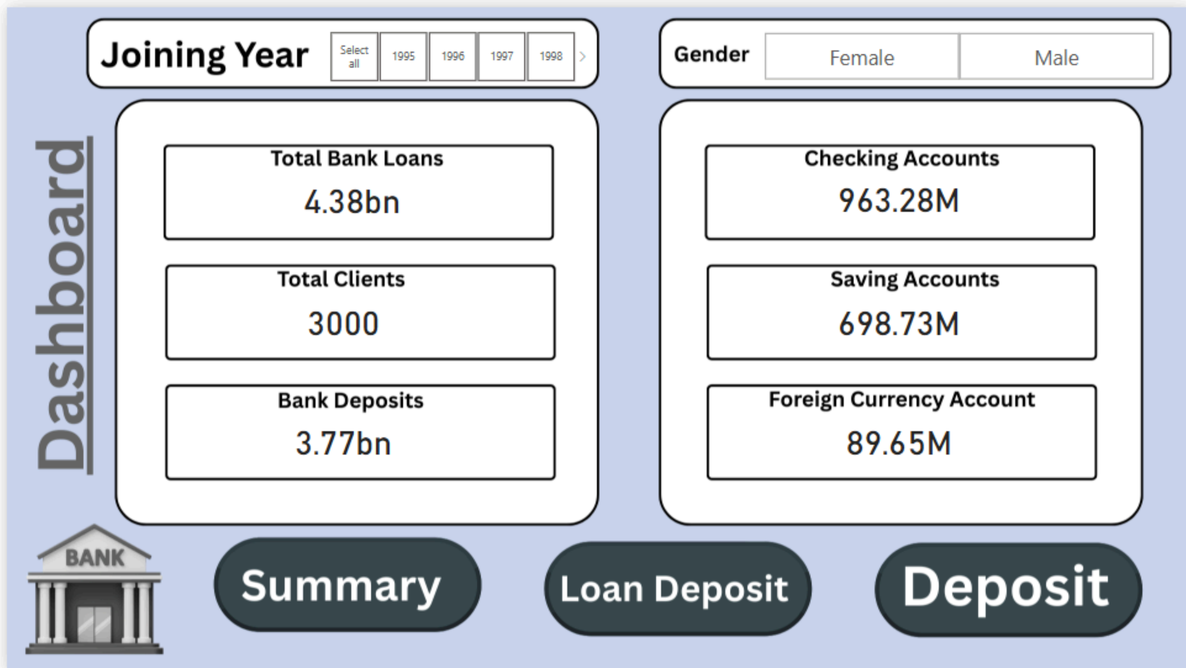
Heatmaps

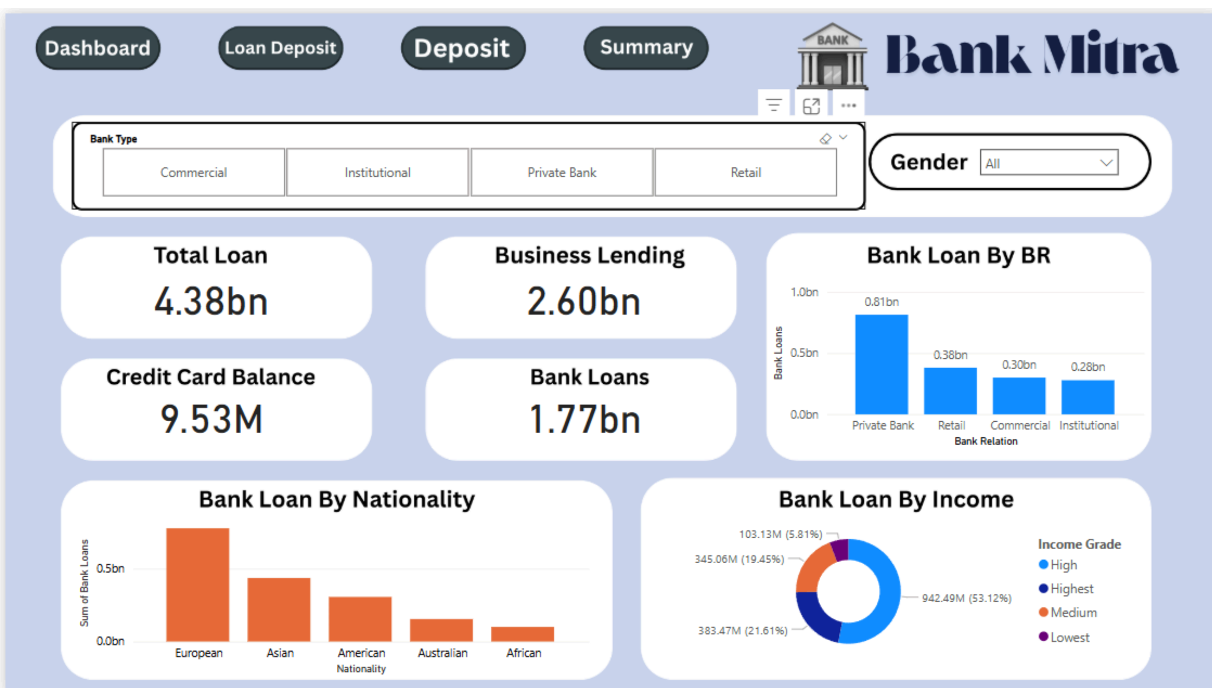
```
numerical=['Estimated Income','Bank Deposits','Bank Loans','Credit Card Balance','Superannuation Savings','Checking Accounts']
heatmap=df[numerical].corr()
plt.figure(figsize=(12,12))
sbs.heatmap(heatmap,annot=True,cmap='PiYG')
```

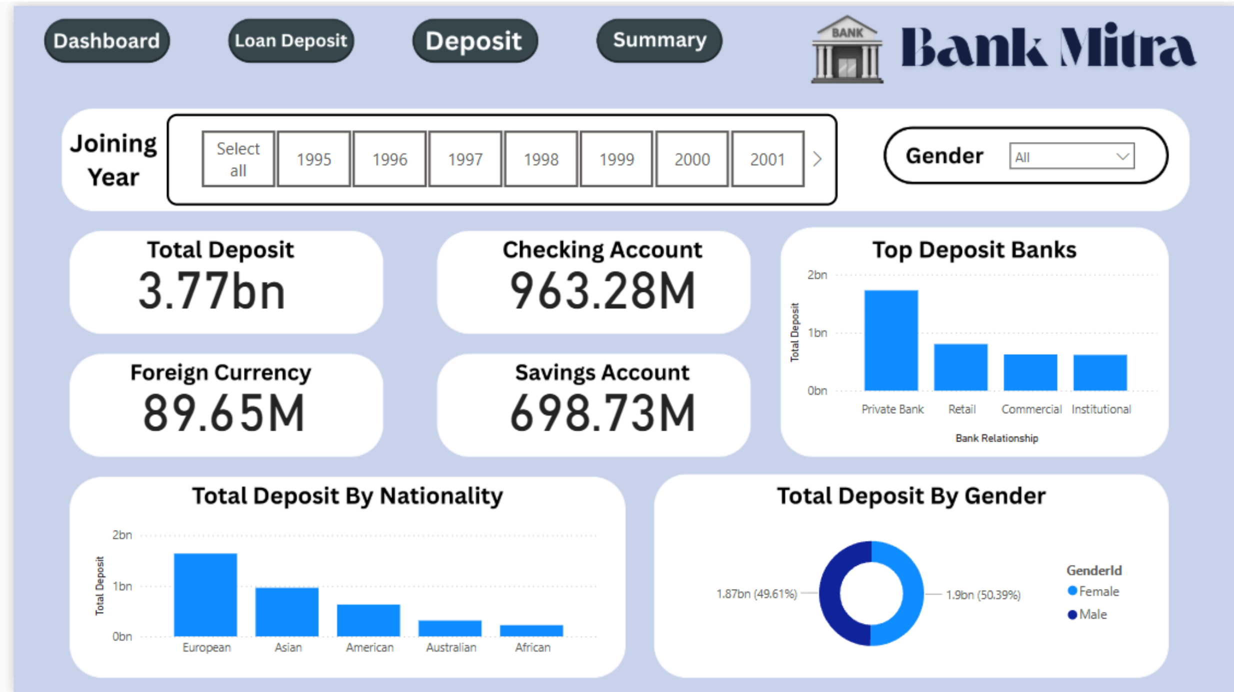
<Axes: >

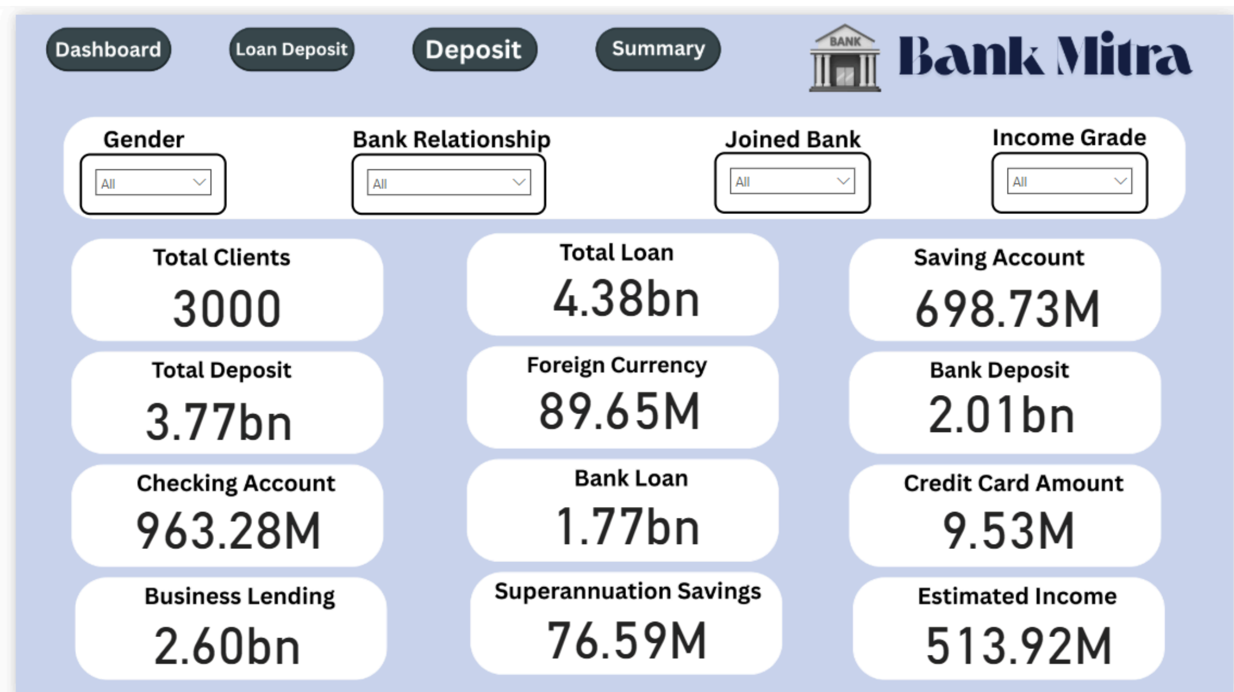


Visualization Using Power Bi :


Joining Year
Select all
1995
1996
1997
1998
Gender
Female
Male
Dashboard
Total Bank Loans
4.38bn
Total Clients
3000
Bank Deposits
3.77bn
Checking Accounts
963.28M
Saving Accounts
698.73M
Foreign Currency Account
89.65M
BANK
Summary
Loan Deposit
Deposit


Dashboard
Loan Deposit
Deposit
Summary
BANK
Bank Mitra
Bank Type
Commercial
Institutional
Private Bank
Retail
Gender
All
Total Loan
4.38bn
Business Lending
2.60bn
Credit Card Balance
9.53M
Bank Loans
1.77bn
Bank Loan By BR
1.0bn
0.5bn
0.0bn
Bank Loans
0.81bn
0.38bn
0.30bn
0.28bn
Private Bank
Retail
Commercial
Institutional
Bank Relation
Bank Loan By Nationality
Sum of Bank Loans
0.5bn
0.0bn
European
Asian
American
Australian
African
Nationality
Bank Loan By Income
103.13M (5.81%)
345.06M (19.45%)
942.49M (53.12%)
383.47M (21.61%)
Income Grade
High
Highest
Medium
Lowest





Conclusion :-

Empowered by the latest data visualization techniques, Power BI dashboards are among the most effective resources for using in banking sector. As outlined in this write-up, a banking operations dashboard in Power BI can be developed with key banking related metrics and KPIs.