ANALYZING THE FUTURE IMPACT OF CASHLESS SOCIETY

|  |  |  |
| --- | --- | --- |
| MD.TANJIR HOWLADER  ID:20-43184-1  Dept Name:CSE Institute Name:AIUB City, Country:DHAKA,BANGLADESH email address: tanjirhowlader29@gmail.com  Phone Number: 01795687614 | SHAHRIAR HANNAN  ID: 20-43179-1 Dept Name: CSE Institute Name: AIUB City, Country:DHAKA,BANGLADESH email address: shahriarhfahim@gmail.com  Phone Number: 01533029032 | IFTIKHARUL ISLAM EISHAT  ID: 2043183-1 Dept Name: CSE Institute Name:AIUB City, Country email address  Phone Number |
|  |  |  |

**Abstract**— This research project investigates the transformative impact of digital transactions on the population of Bangladesh, focusing on shopping, bill payments, and rent, while avoiding segmentation by individual age groups. The study aims to provide valuable insights for policymakers and businesses, fostering the development of user-friendly financial services and contributing to a more efficient financial future in Bangladesh.

# **Introduction**

The research project seeks to understand the impact of digital transactions on the population of Bangladesh, with a specific focus on shopping, bill payments, and rent, removing emphasis on individual age groups. Through an exploration of demographics, prediction of user behavior, and an investigation into perceptions of cashless transactions, the study aims to provide valuable insights for policymakers and businesses. These insights will contribute to informed decision-making, fostering the development of user-friendly financial services and shaping a more efficient financial future in Bangladesh.

**Motivation of the Project**: The motivation behind this project arises from a deep interest in the changing dynamics of financial transactions, particularly within the context of Bangladesh's transition toward a cashless society. Several compelling factors drive the exploration of user behavior without singling out specific age groups.

Understanding Societal Change: The swift adoption of cashless transactions reflects a fundamental shift in how individuals manage their finances. Focusing on user behavior, irrespective of age, aims to provide insights into the broader transformation of societal behaviors in response to digital advancements.

Personal and Academic Growth: Engaging with real-world data and machine learning models offers a unique opportunity for personal and academic growth. The project allows for a deeper understanding of theoretical concepts and practical applications in the fields of data science and finance.

Contributions to Society: The commitment to contribute meaningful insights extends beyond individual age groups. The anticipated benefits for society encompass informed policymaking, enhanced business strategies, improved financial inclusion, technological innovation, and overall progress within the unique financial landscape of Bangladesh.

Informed Policymaking: Policymakers can leverage research outcomes to formulate informed policies catering to the diverse needs and concerns of the population, irrespective of age groups.

Business Strategy Enhancement: Businesses, particularly in the financial and retail sectors, can tailor their services to resonate with the preferences and motivations of their target audience without age-specific constraints.

Financial Inclusion: Understanding the impact of a cashless society on individuals may uncover ways to enhance financial inclusion and accessibility, contributing to a more inclusive and equitable financial landscape.

Technological Innovation: Findings from the project may inspire technological innovations and solutions that address the evolving demands of a digitally engaged demographic, fostering a culture of innovation within the financial sector.

Overall Societal Progress: By delving into the intricacies of user behavior without age-specific constraints, the project aspires to be a significant building block in the collective effort toward societal progress, technological adaptation, and financial inclusivity within Bangladesh.

# **Objective of the project**

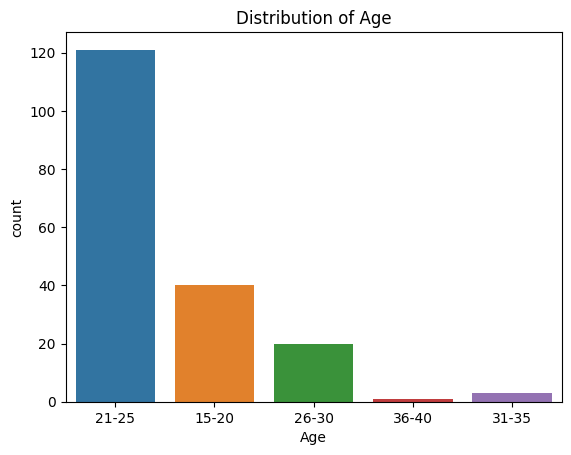
* Demographic Insights:
* Goal: Explore the demographic characteristics of individuals within distinct age groups (e.g., 21-25, 26-30, 31-35) engaging in cashless transactions.
* Outputs: Detailed demographic profiles for each age group, highlighting age, gender, occupation, and income distribution.
* User Behavior Prediction:
* Goal: Develop predictive models for discerning the preferred payment methods for shopping transactions within various age groups.
* Outputs: Machine learning models capable of predicting the payment methods favored by individuals in different age brackets for shopping transactions.
* Perceptions and Motivations Understanding:
* Goal: Explore the perceptions, motivations, and concerns of individuals across different age groups regarding the adoption of cashless payment methods.
* Outputs: Insights into the factors influencing the adoption of cashless transactions within each age group, providing a nuanced understanding of motivations, concerns, and preferences.
* Policy and Strategy Recommendations:
* Goal: Provide tailored recommendations for policymakers, businesses, and financial institutions based on age-specific insights.
* Outputs: Actionable suggestions for developing policies and strategic insights for businesses that cater to the needs of specific age groups.
* Contribution to Academic Knowledge:
* Goal: Contribute to the academic understanding of the impact of a cashless society on user behavior, with a focus on diverse age demographics.
* Outputs: Research papers, articles, or presentations detailing the methodology, findings, and implications for each age group, contributing to academic knowledge.
* Technological Innovation Insights:
* Goal: Identify potential areas for technological innovation and solutions that align with the evolving demands of different age groups.
* Outputs: Insights into features or technologies that could enhance user engagement with digital financial services, considering the preferences of specific age demographics.
* Enhanced Financial Inclusion:
* Goal: Investigate how the transition to a cashless society impacts financial inclusion across various age groups.
* Outputs: Recommendations for enhancing financial inclusion, ensuring that the benefits of cashless transactions are accessible to diverse age segments.
* Understanding Impact on Business Strategy:
* Goal: Provide insights into how businesses can adapt their strategies to align with the preferences and behaviors of specific age groups.
* Outputs: Strategic recommendations for businesses in the financial and retail sectors to optimize services and foster customer engagement tailored to different age demographics.

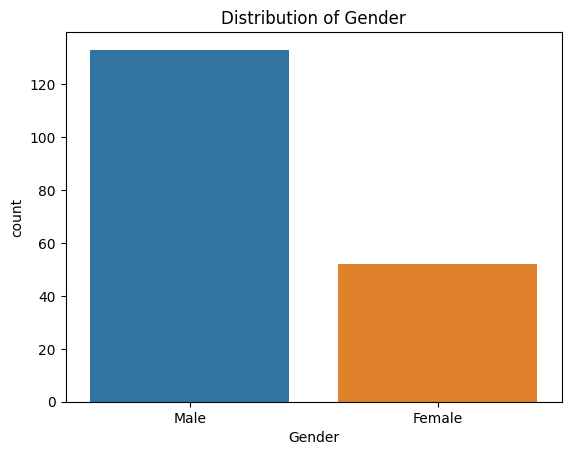
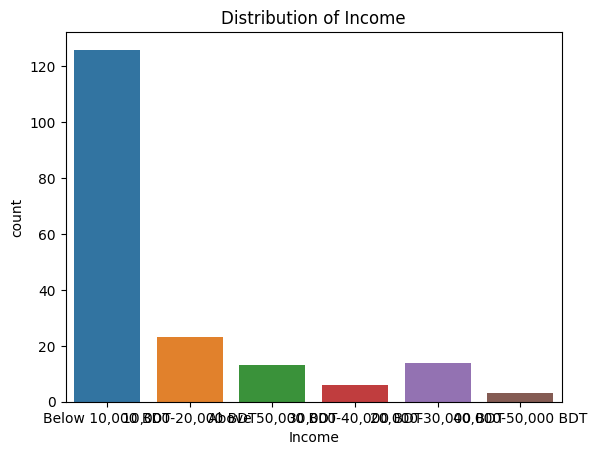
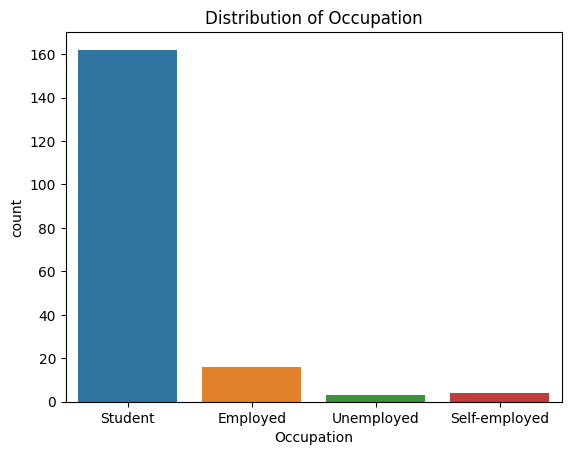
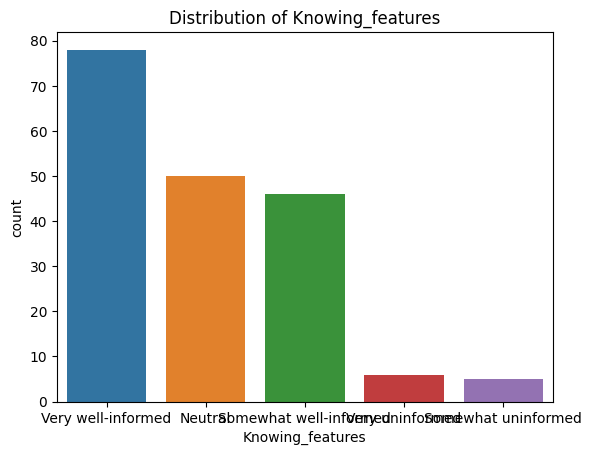
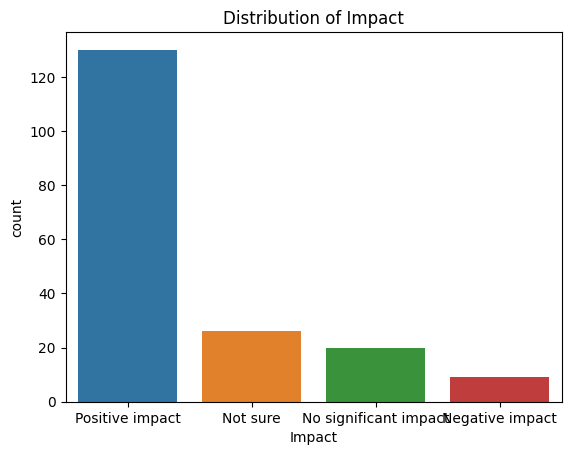
By achieving these goals, the project aims to offer a comprehensive and insightful exploration of the impact of a cashless society on user behavior within the specified demographic, with tangible outputs that contribute to both academic knowledge and practical applications in policymaking and business strategy.

# **Methodology**

Data Loading and Exploration:

The project starts by loading the dataset using Pandas and displaying basic information about the dataset using info(), head(), and describe() functions.





**Data Preprocessing**:

Label Encoding is performed on the target variable ('Shopping\_freq', 'Paybills\_freq', 'Payrent\_freq') using LabelEncoder.

One-hot encoding is applied to categorical features using pd.get\_dummies().

Model Training - Shopping\_freq:

Random Forest, Decision Tree, Logistic Regression, and SVM models are trained to predict 'Shopping\_freq\_encoded'.

Accuracy scores are calculated for each model on the test set.

A new data point is created, encoded, and used for making predictions with each model.

Model Training - Paybills\_freq:

Similar steps are repeated for the 'Paybills\_freq' target variable.

Model Training - Payrent\_freq:

Similar steps are repeated for the 'Payrent\_freq' target variable.

## **Data Collection**

For collect data we have used a goggle form. Collect a dataset named 'cashless\_society.csv' containing 10 columns, including Age, Gender, Occupation, Income, Shopping\_freq, Paybills\_freq, Payrent\_freq, Motivation, Knowing\_features, and Impact.

## **Data processing**

1, First of all we enter necessary Libraries for use the datatset.

2. Then we upload the csv file

3.After that we handle the missing values and ensure the data type was correct.

4.Then we perform one-hot coding for convert the value to numeric.

5.then we check any duplicate row is there or nt if there have then remove it.

6.Finaly after processing the data we saved the clean data.

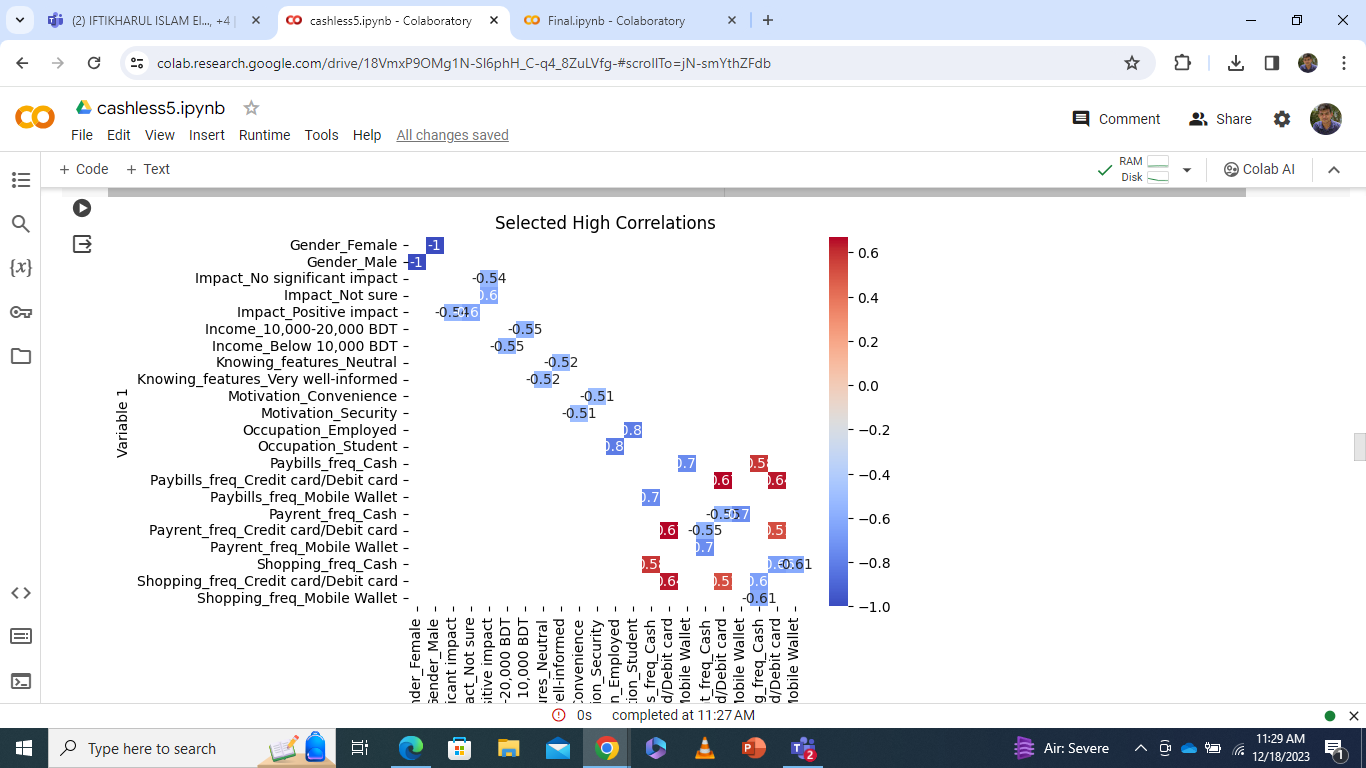
## **Dataset description**

The data set is about cashless society.

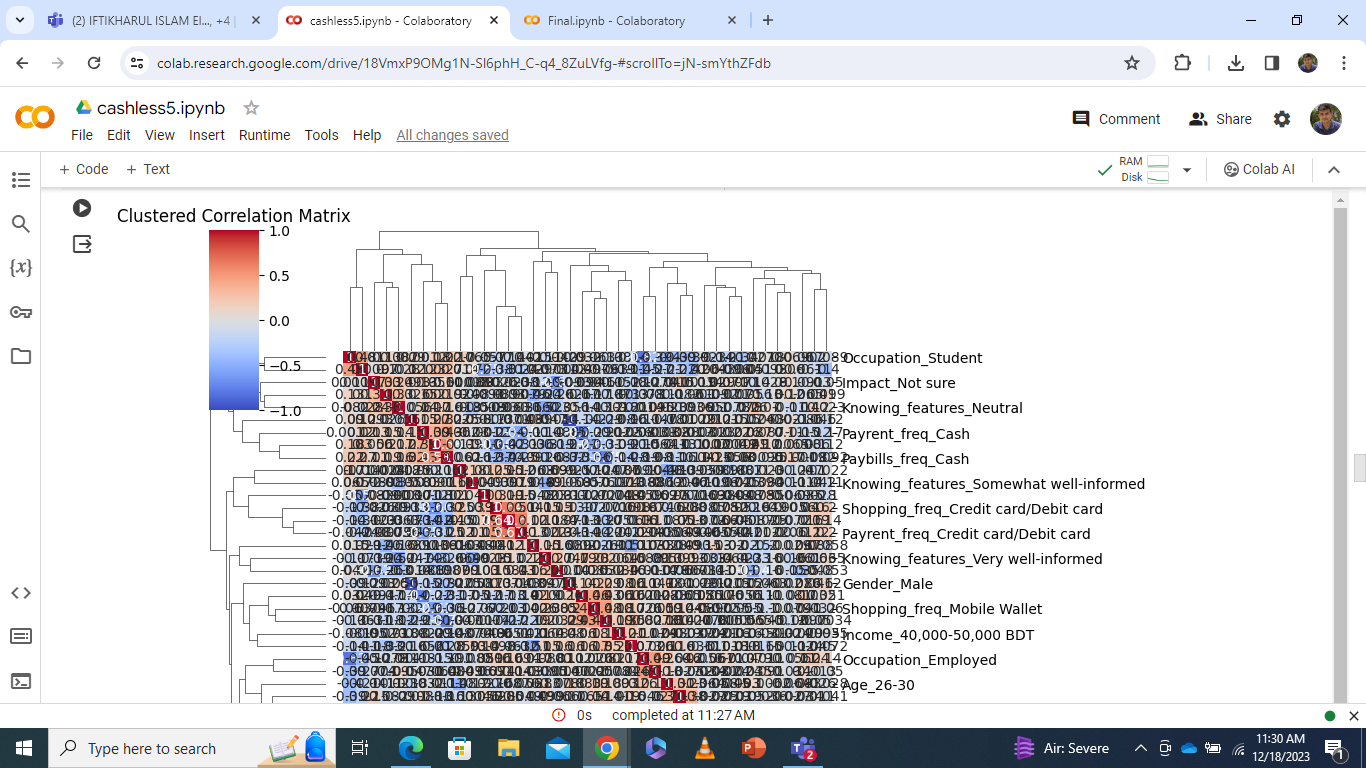
Number of colunm:10

We visualize data in various way such as

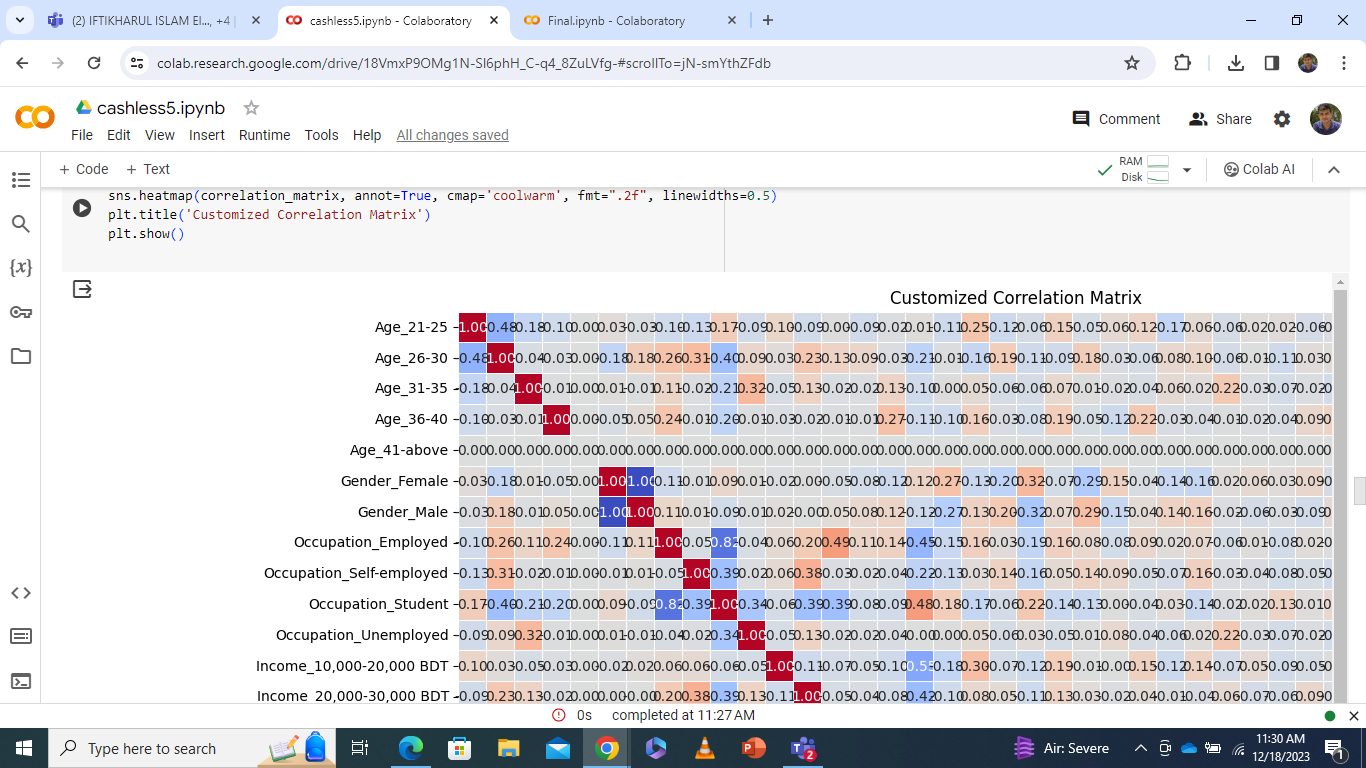
**SELECTED HIGH CORRELATION:**



**CLUSTERD CORELATION MATRIX**

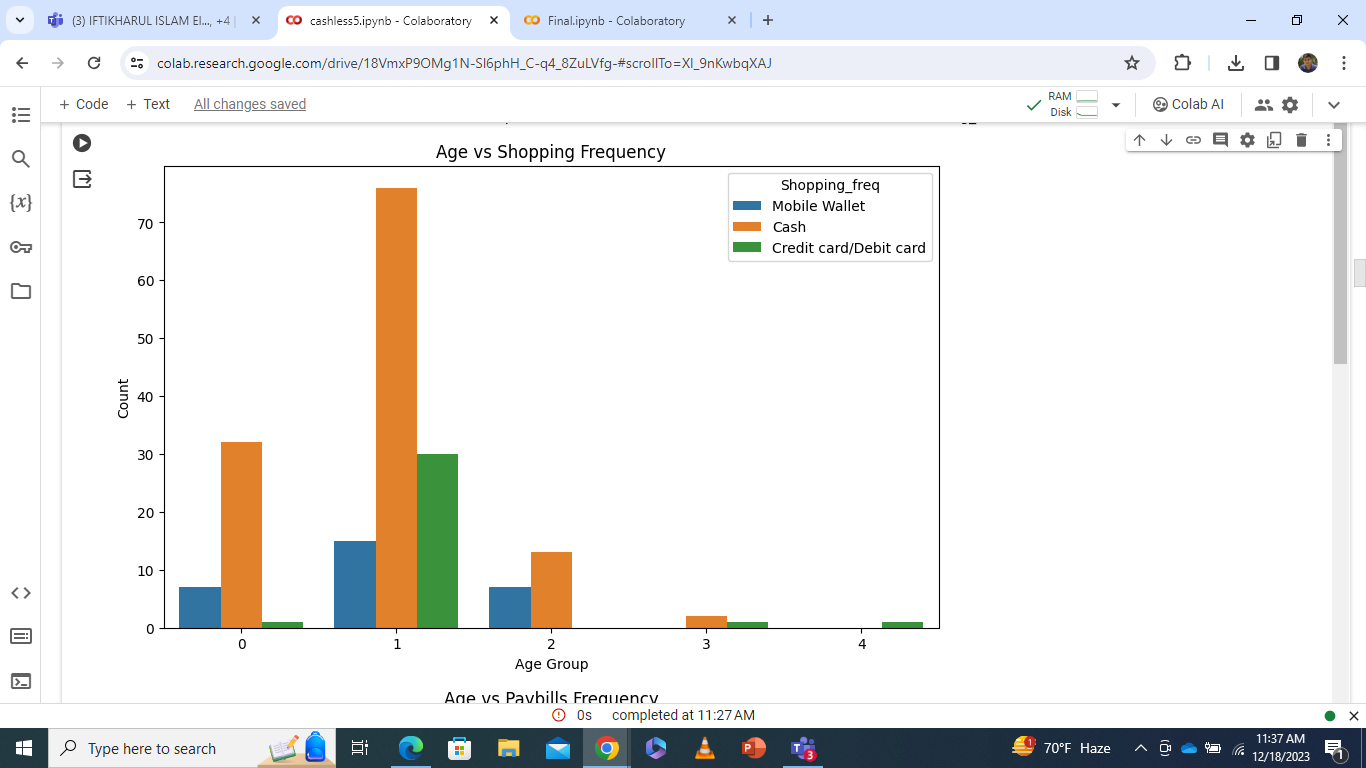


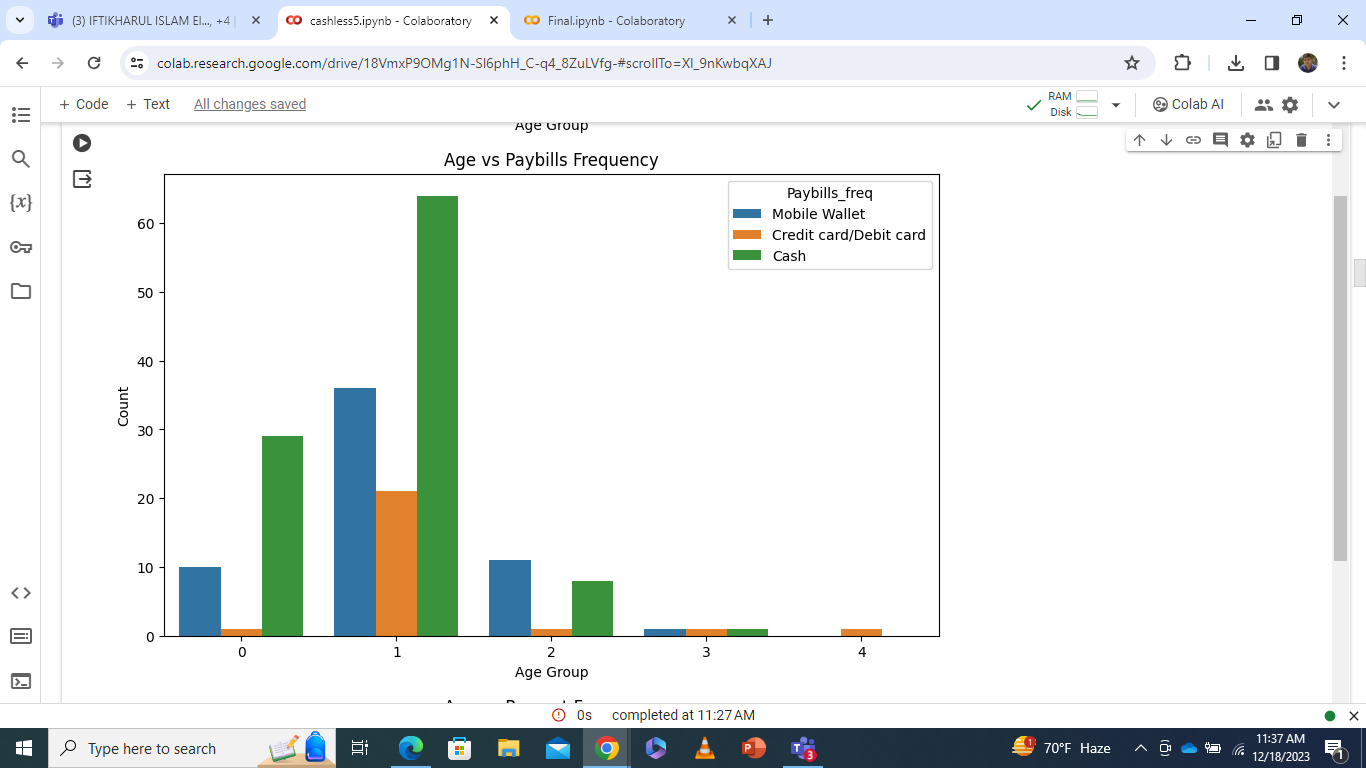
**CUSTOMIZED CO-REALTION MATRIX**

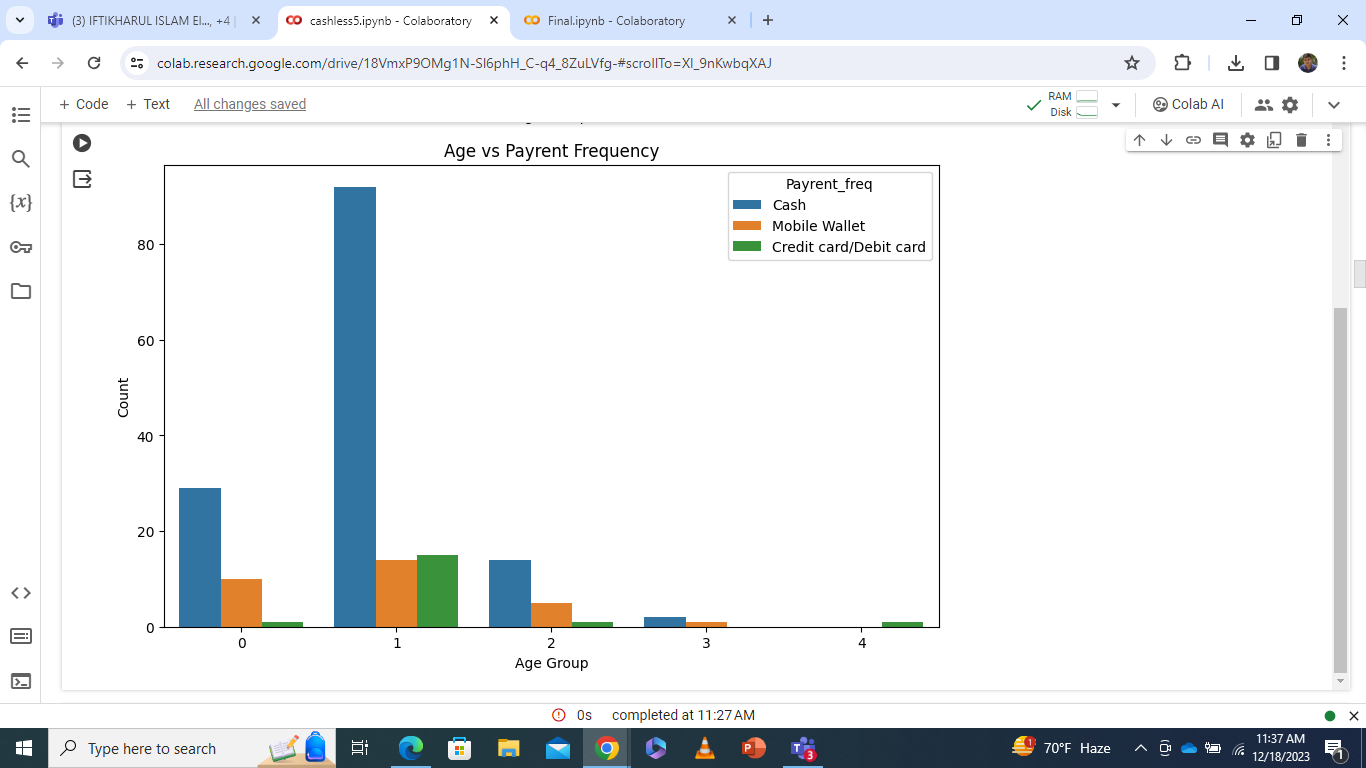


**We also compare the data with one colunm to antother**

**Age vs shopping, paybill, payrent freq:**







**We do the same thing for every columns,**

## **Machine Learning model development and evaluation**

**Machine Learning Model Development:**

Model Selection: You've used Random Forest, Decision Tree, Logistic Regression, and SVM models for your classification tasks.

Data Preprocessing: Features like 'Age,' 'Gender,' 'Occupation,' 'Income,' 'Motivation,' 'Knowing\_features,' and 'Impact' are preprocessed, including label encoding and one-hot encoding for categorical variables.

Training and Evaluation:

For each target variable ('Shopping\_freq', 'Paybills\_freq', 'Payrent\_freq'):

Split the dataset into training and testing sets.

Train the Random Forest, Decision Tree, Logistic Regression, and SVM models.

Evaluate model accuracy using accuracy\_score.

Hyperparameter Tuning:

Random Forest: n\_estimators = 100 (tuned parameter).

SVM: kernel = 'linear' (tuned parameter).

# **Results**

Shopping\_freq

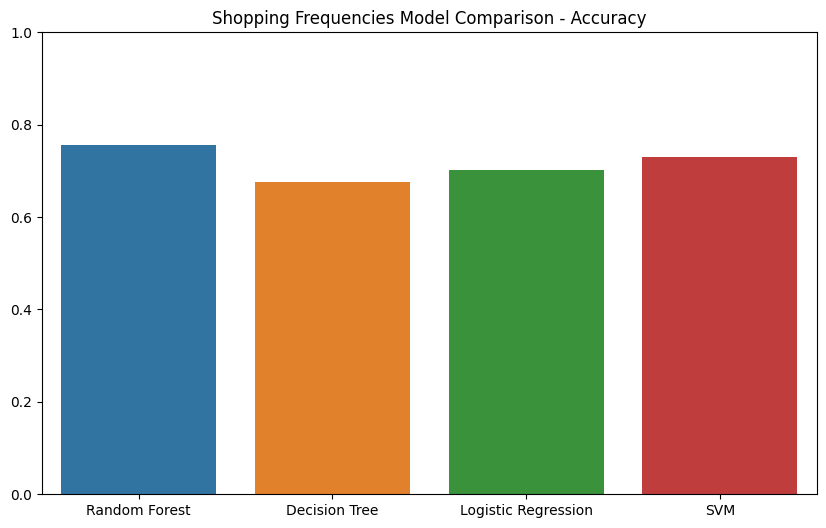
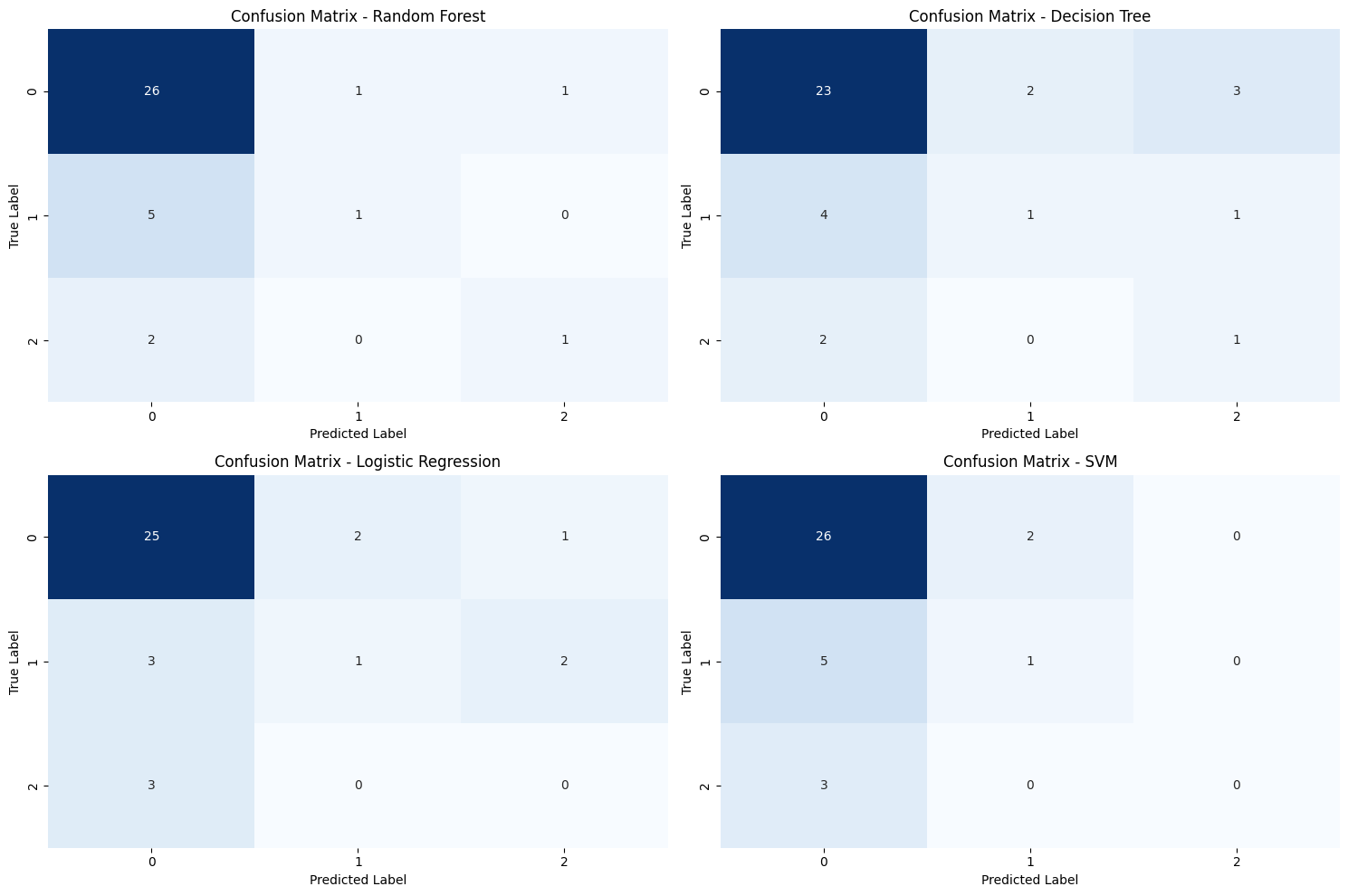
Random Forest Model Accuracy: 0.7568 Predicted Shopping Frequency (Random Forest): ['Cash'] Decision Tree Model Accuracy: 0.6757 Predicted Shopping Frequency (Decision Tree): ['Cash'] Logistic Regression Model Accuracy: 0.7027 Predicted Shopping Frequency (Logistic Regression): ['Cash'] SVM Model Accuracy: 0.7297 Predicted Shopping Frequency (SVM): ['Cash']

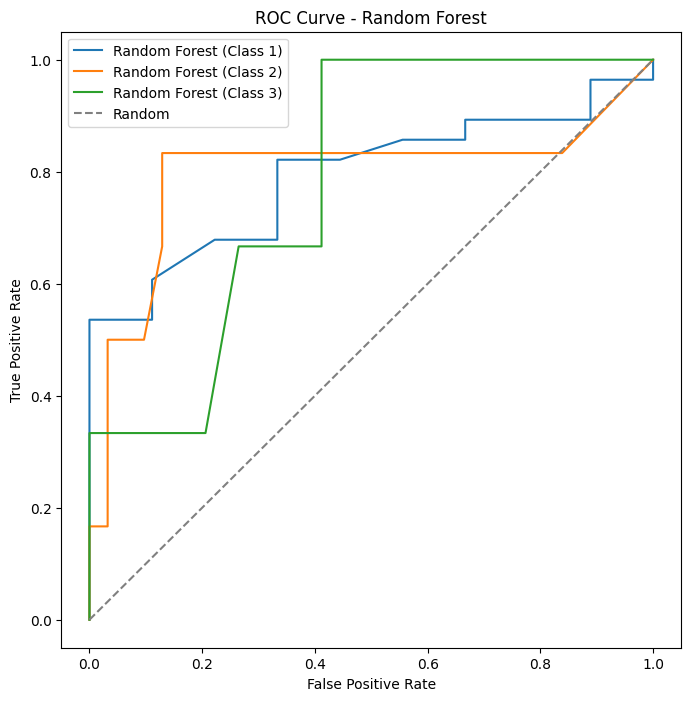
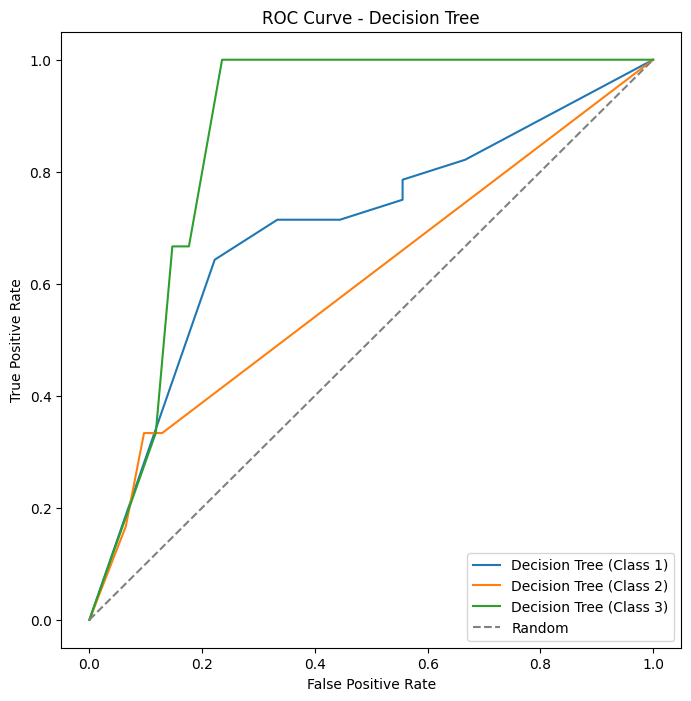
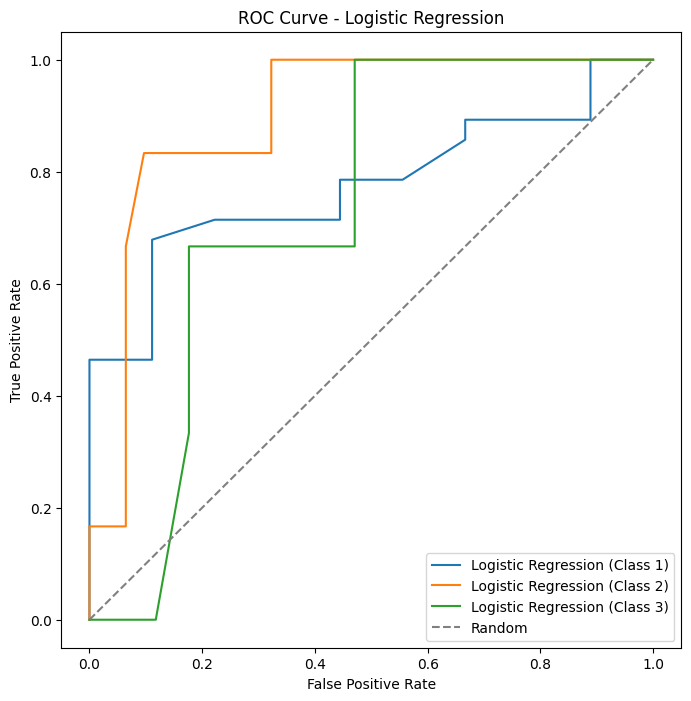
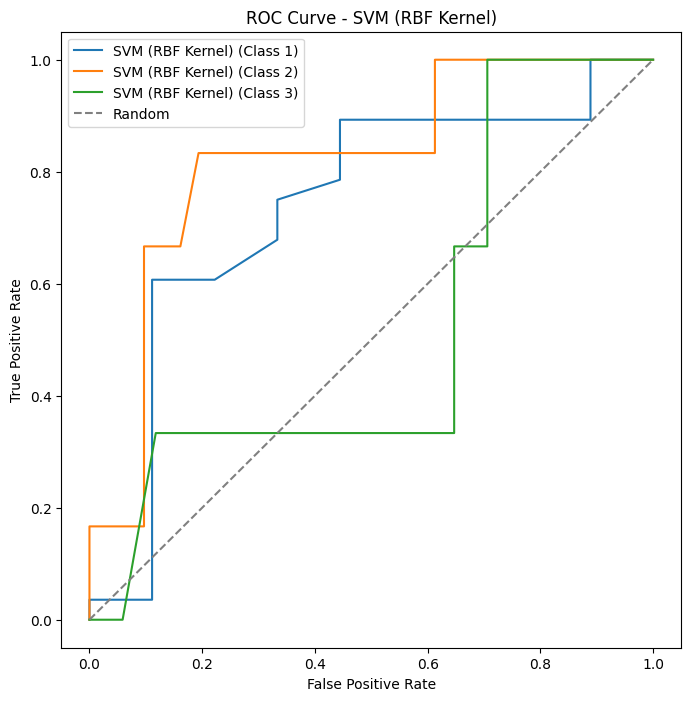
Classification Report for Random Forest: precision recall f1-score support 0 0.79 0.93 0.85 28 1 0.50 0.17 0.25 6 2 0.50 0.33 0.40 3 accuracy 0.76 37 macro avg 0.60 0.48 0.50 37 weighted avg 0.72 0.76 0.72 37

Classification Report for Decision Tree: precision recall f1-score support 0 0.79 0.82 0.81 28 1 0.33 0.17 0.22 6 2 0.20 0.33 0.25 3 accuracy 0.68 37 macro avg 0.44 0.44 0.43 37 weighted avg 0.67 0.68 0.67 37

Classification Report for Logistic Regression: precision recall f1-score support 0 0.81 0.89 0.85 28 1 0.33 0.17 0.22 6 2 0.00 0.00 0.00 3 accuracy 0.70 37 macro avg 0.38 0.35 0.36 37 weighted avg 0.66 0.70 0.68 37

Classification Report for SVM: precision recall f1-score support 0 0.76 0.93 0.84 28 1 0.33 0.17 0.22 6 2 0.00 0.00 0.00 3 accuracy 0.73 37 macro avg 0.37 0.37 0.35 37 weighted avg 0.63 0.73 0.67 37





Paybills\_freq

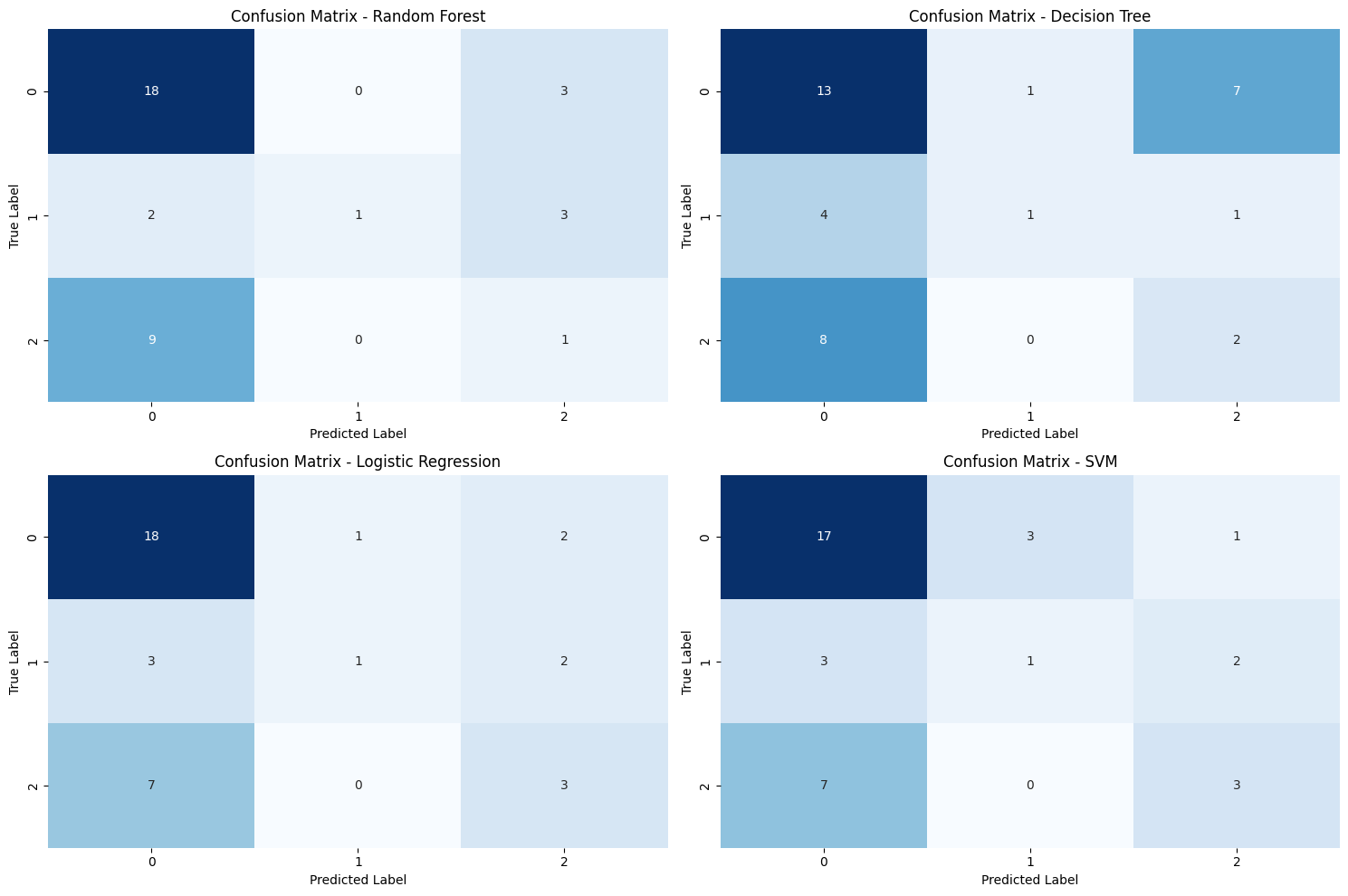
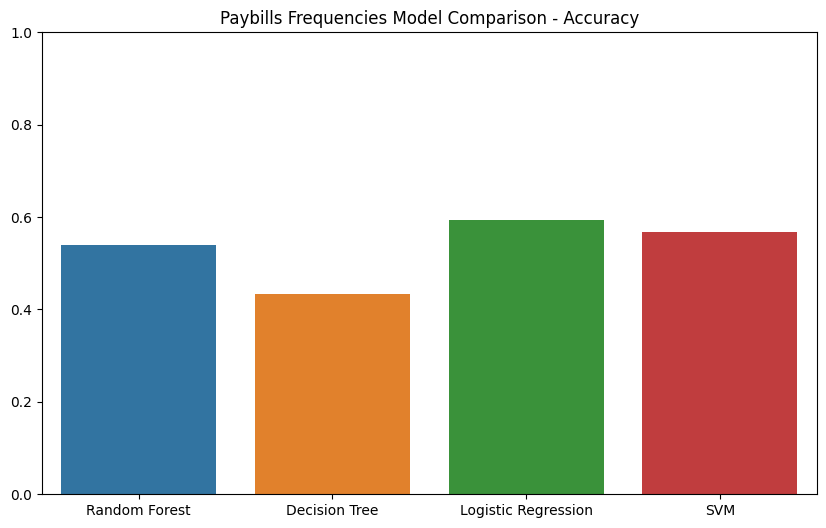
Random Forest Model Accuracy: 0.5405 Predicted Paybills Frequency (Random Forest): ['Mobile Wallet'] Decision Tree Model Accuracy: 0.4324 Predicted Paybills Frequency (Decision Tree): ['Mobile Wallet'] Logistic Regression Model Accuracy: 0.5946 Predicted Paybills Frequency (Logistic Regression): ['Cash'] SVM Model Accuracy: 0.5676 Predicted Paybills Frequency (SVM): ['Cash']

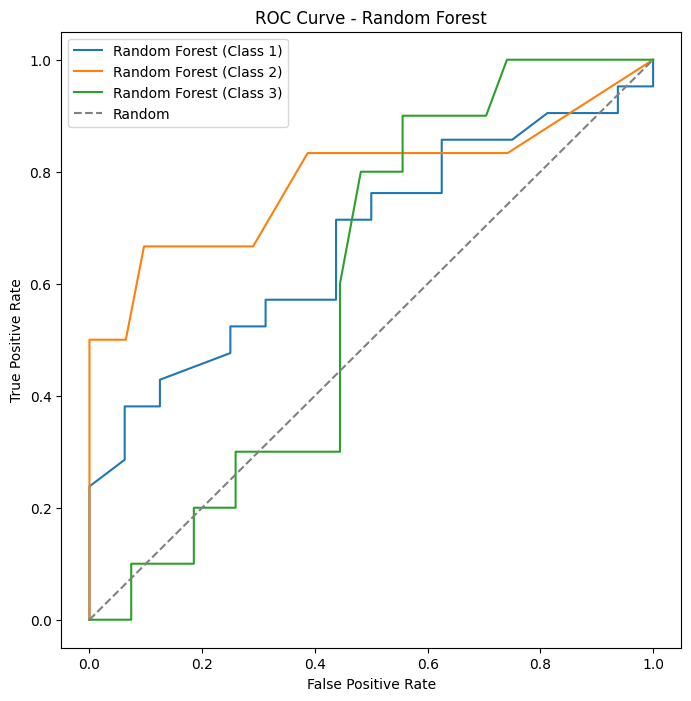
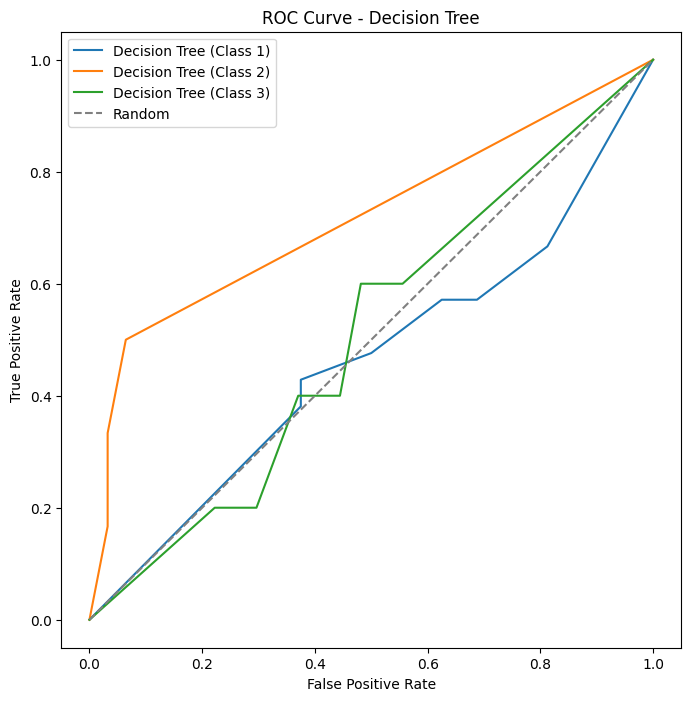
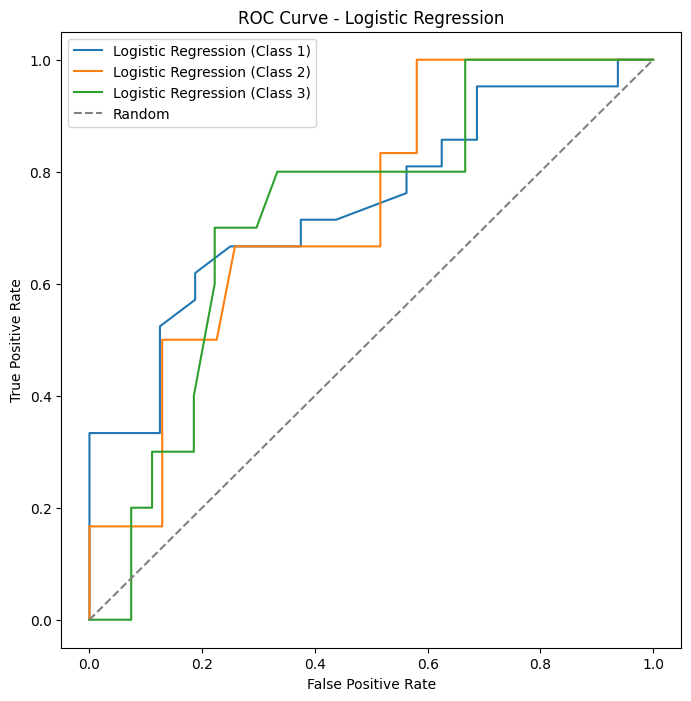
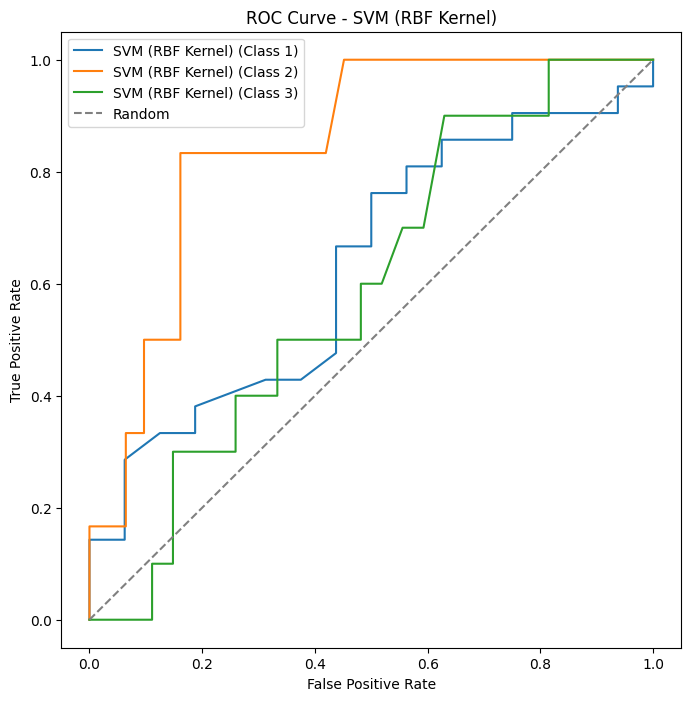
Classification Report for Random Forest: precision recall f1-score support 0 0.62 0.86 0.72 21 1 1.00 0.17 0.29 6 2 0.14 0.10 0.12 10 accuracy 0.54 37 macro avg 0.59 0.37 0.37 37 weighted avg 0.55 0.54 0.49 37

Classification Report for Decision Tree: precision recall f1-score support 0 0.52 0.62 0.57 21 1 0.50 0.17 0.25 6 2 0.20 0.20 0.20 10 accuracy 0.43 37 macro avg 0.41 0.33 0.34 37 weighted avg 0.43 0.43 0.42 37

Classification Report for Logistic Regression: precision recall f1-score support 0 0.64 0.86 0.73 21 1 0.50 0.17 0.25 6 2 0.43 0.30 0.35 10 accuracy 0.59 37 macro avg 0.52 0.44 0.45 37 weighted avg 0.56 0.59 0.55 37

Classification Report for SVM: precision recall f1-score support 0 0.63 0.81 0.71 21 1 0.25 0.17 0.20 6 2 0.50 0.30 0.37 10 accuracy 0.57 37 macro avg 0.46 0.43 0.43 37





Payrent\_freq

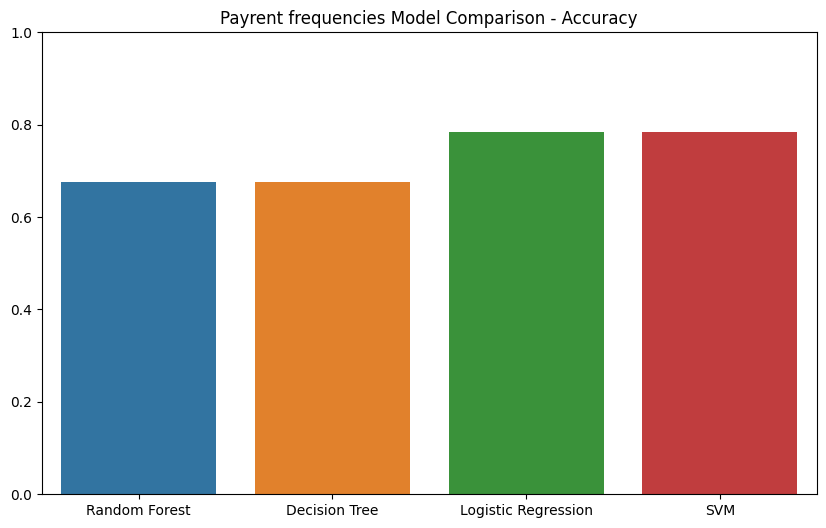
Random Forest Model Accuracy: 0.6757 Predicted Payrent Frequency (Random Forest): ['Cash'] Decision Tree Model Accuracy: 0.6757 Predicted Payrent Frequency (Decision Tree): ['Cash'] Logistic Regression Model Accuracy: 0.7838 Predicted Payrent Frequency (Logistic Regression): ['Cash'] SVM Model Accuracy: 0.7838 Predicted Payrent Frequency (SVM): ['Cash']

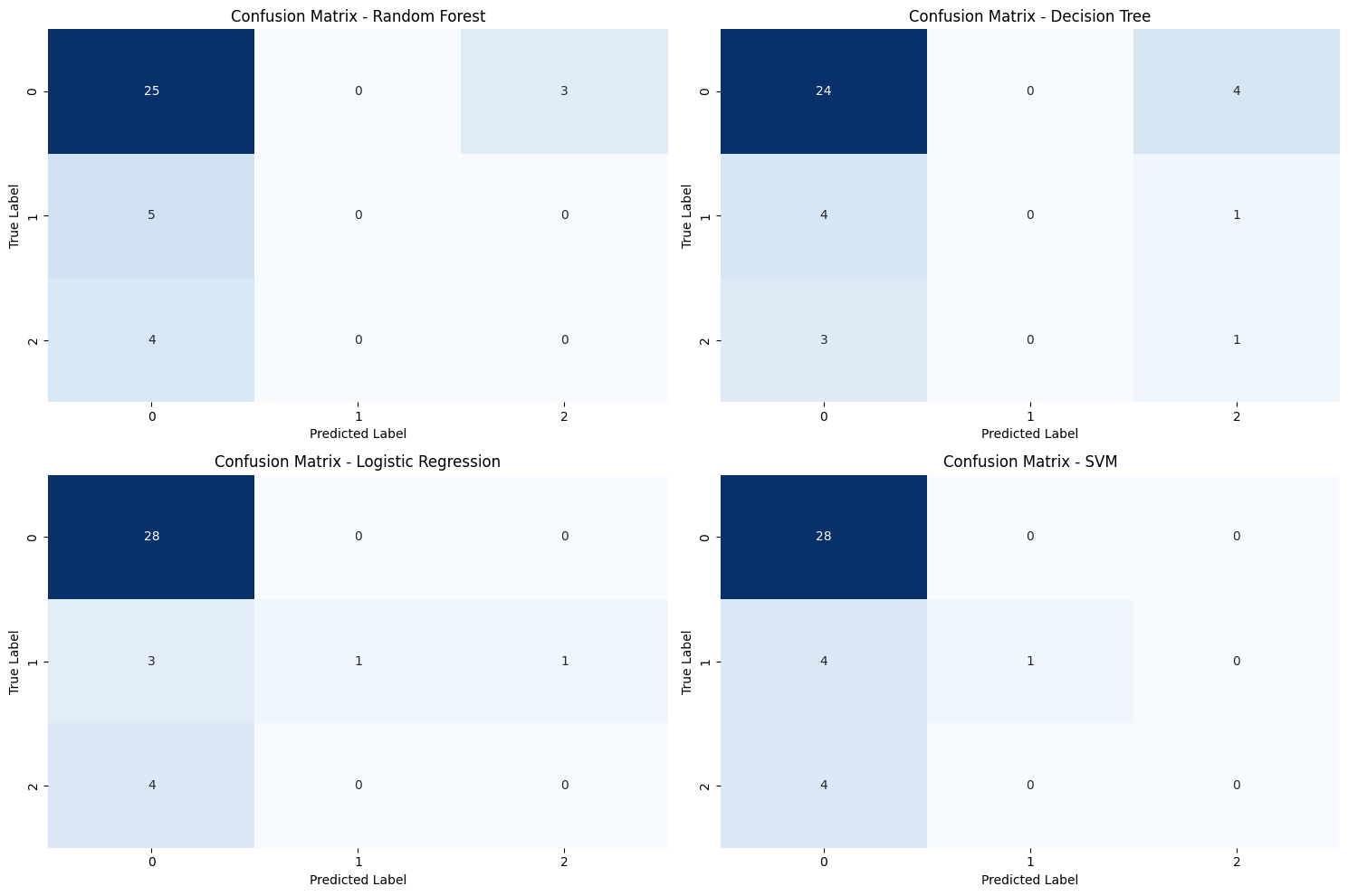
Classification Report for Random Forest: precision recall f1-score support 0 0.74 0.89 0.81 28 1 0.00 0.00 0.00 5 2 0.00 0.00 0.00 4 accuracy 0.68 37 macro avg 0.25 0.30 0.27 37 weighted avg 0.56 0.68 0.61 37

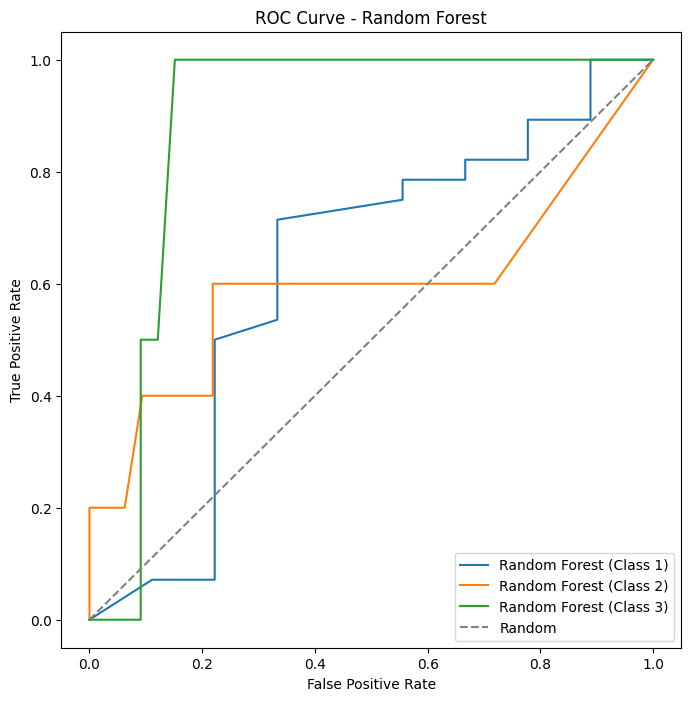
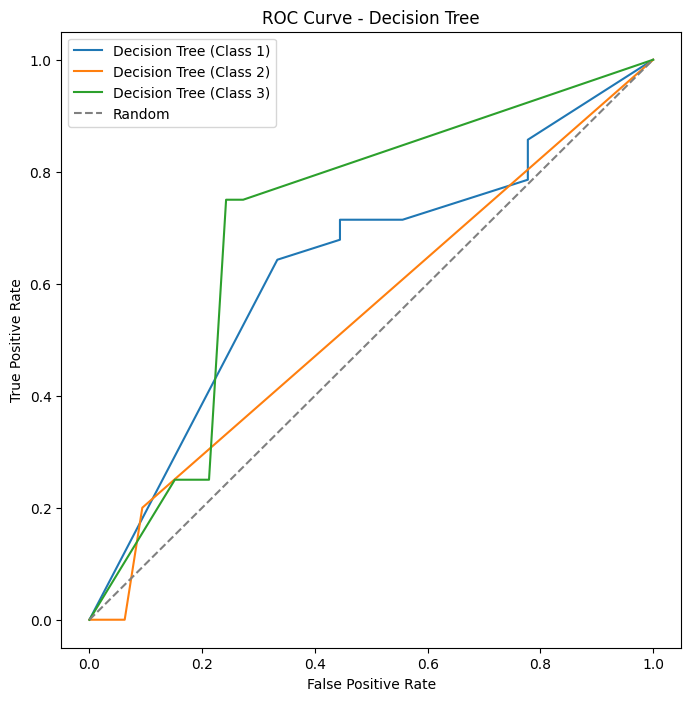
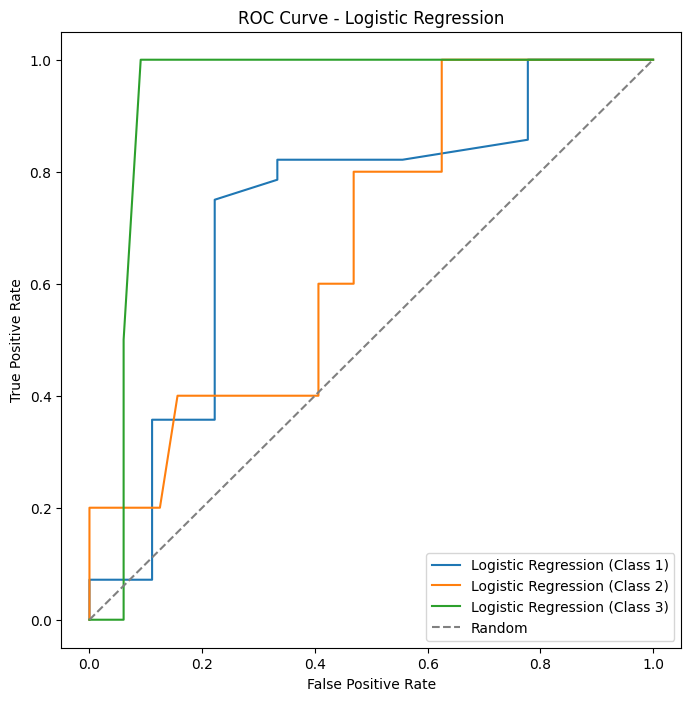
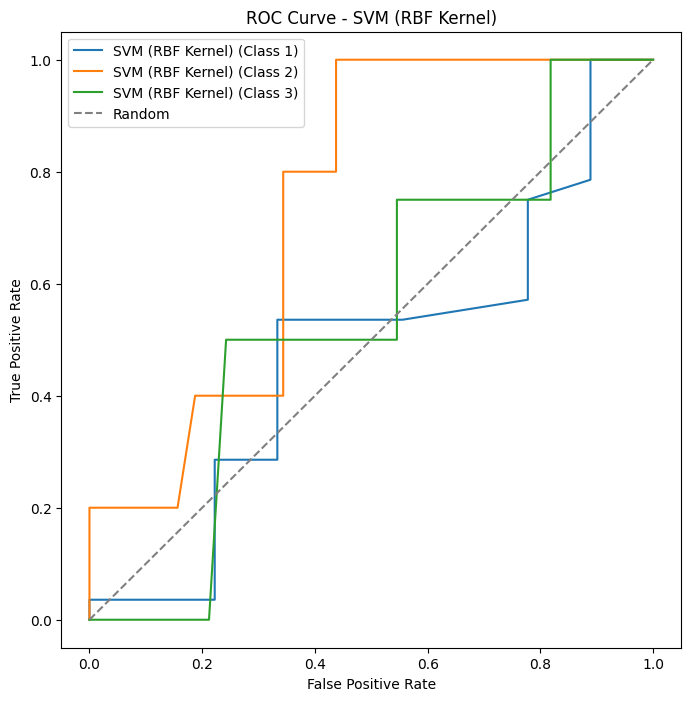
Classification Report for Decision Tree: precision recall f1-score support 0 0.77 0.86 0.81 28 1 0.00 0.00 0.00 5 2 0.17 0.25 0.20 4 accuracy 0.68 37 macro avg 0.31 0.37 0.34 37 weighted avg 0.60 0.68 0.64 37

Classification Report for Logistic Regression: precision recall f1-score support 0 0.80 1.00 0.89 28 1 1.00 0.20 0.33 5 2 0.00 0.00 0.00 4 accuracy 0.78 37 macro avg 0.60 0.40 0.41 37 weighted avg 0.74 0.78 0.72 37

Classification Report for SVM: precision recall f1-score support 0 0.78 1.00 0.88 28 1 1.00 0.20 0.33 5 2 0.00 0.00 0.00 4 accuracy 0.78 37 macro avg 0.59 0.40 0.40 37 weighted avg 0.72 0.78 0.71 37







# **References**

**Notes:**

**Data Sources:**

1. [**https://www.kaggle.com/datasets**](https://www.kaggle.com/datasets)
2. [**https://figshare.com/**](https://figshare.com/)
3. [**https://data.mendeley.com/**](https://data.mendeley.com/)
4. [**https://registry.opendata.aws/**](https://registry.opendata.aws/)
5. [**https://datasetsearch.research.google.com/**](https://datasetsearch.research.google.com/)
6. [**https://www.openml.org/**](https://www.openml.org/)
7. [**https://datahub.io/search**](https://datahub.io/search)
8. [**https://data.gov/**](https://data.gov/)
9. [**https://data.europa.eu/data/datasets?locale=en**](https://data.europa.eu/data/datasets?locale=en)
10. [**https://data.worldbank.org/**](https://data.worldbank.org/)
11. [**https://ukdataservice.ac.uk/**](https://ukdataservice.ac.uk/)
12. [**https://data.nasdaq.com/institutional-investors**](https://data.nasdaq.com/institutional-investors)
13. [**https://www.imf.org/en/Data**](https://www.imf.org/en/Data)
14. [**http://archive.ics.uci.edu/datasets**](http://archive.ics.uci.edu/datasets)