Greta Lawani

Pig E. Bank Analysis

Customer retention at a fictional global bank

Context

I've been hired as a data analyst by a well-known (and fictitious) global bank. My job includes analytical support to the anti-money-laundering compliance department of data-related projects that help the bank assess client risk and transaction risk, as well as reporting on metrics.

Project Overview

I assess client and transaction risks. I offer guidelines to ensure bias control and provide an overview. Additionally, I focus on building and optimizing models to enhance the efficiency of the compliance program, aiming for a more robust and effective approach.

Data

Fictional client data supplied by CareerFoundry

Skills

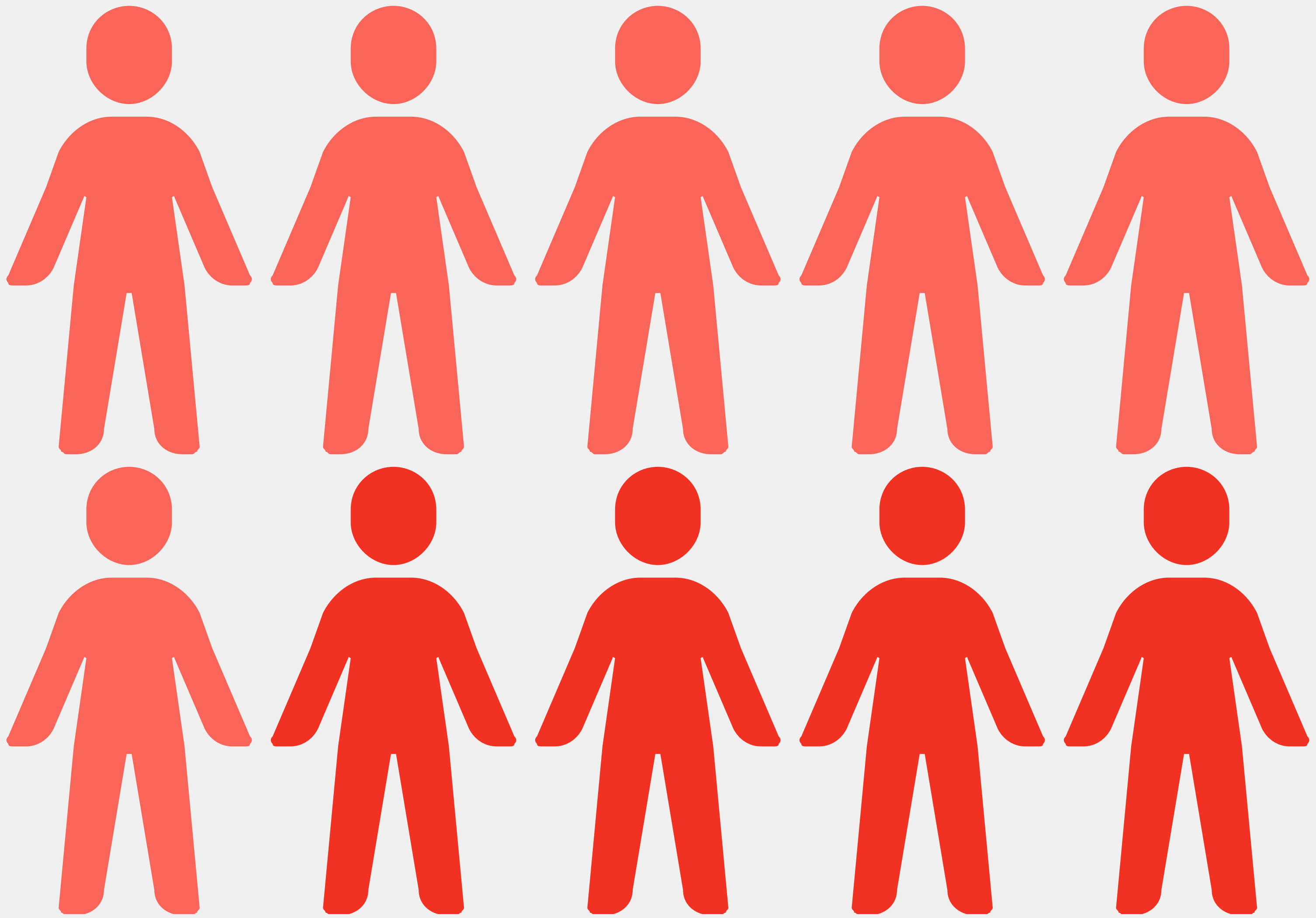
Big data, Data ethics, Data mining, Predictive analysis, Time series analysis and forecasting

Tools

MS Excel, GitHub, Python, Jupyter Notebooks, Canva

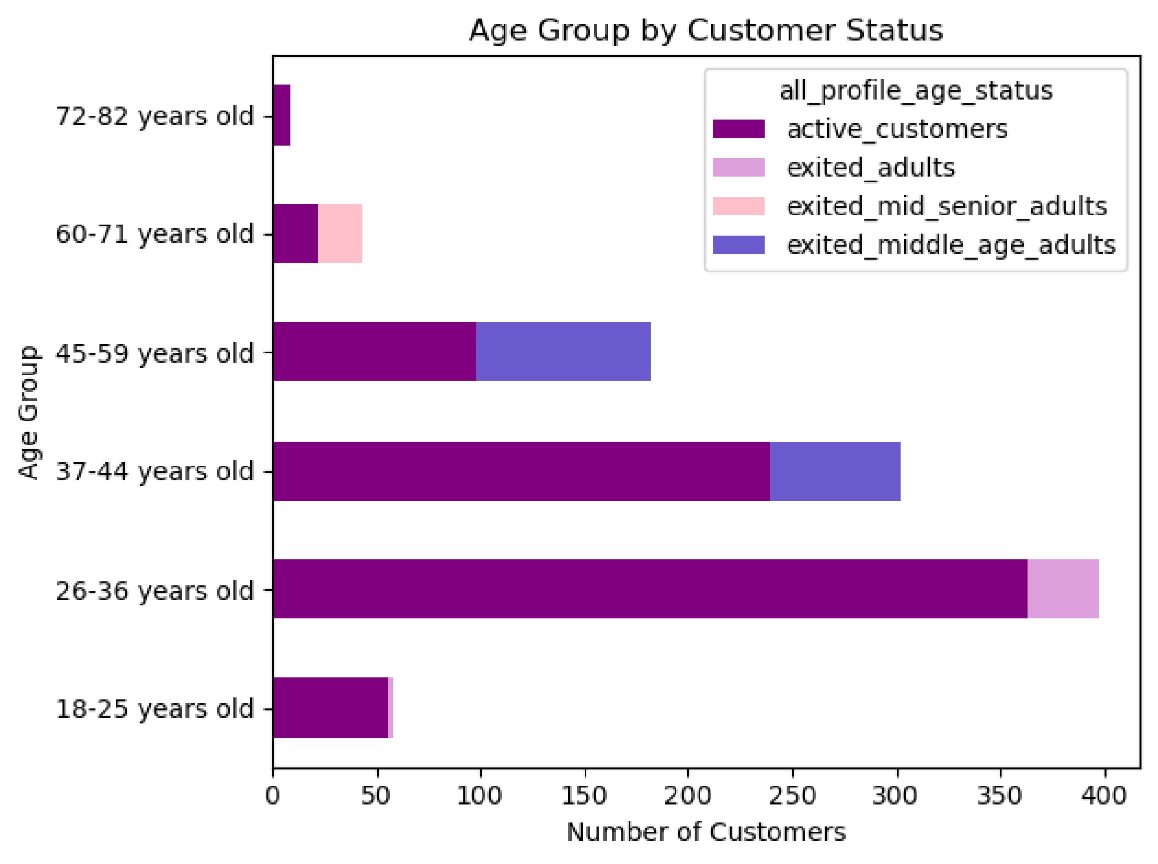
Customers Analysis: Key Insights

- Customers aged 45 and above tend to exit the bank more frequently, with an average departure age of 45
- Former clients demonstrate a notable 70% likelihood of remaining inactive
- Although women comprise less than half of the total client base, they constitute 59% of those who choose to leave the bank



- France represents 48% of the total clients, while 38% of those clients are leaving
    - Clients from France are more likely to stay (52%)
- Germany represents 26% of the total number of customers but has the highest loss percentage with 37%
    - Clients from Germany are more likely to leave
- Spain represents 26% of the total clients, and only 26% of those clients are leaving
    - Clients from Spain are more likely to stay

Customers Status vs. Age Groups



Income Status vs. Customer Groups

Age Distribution

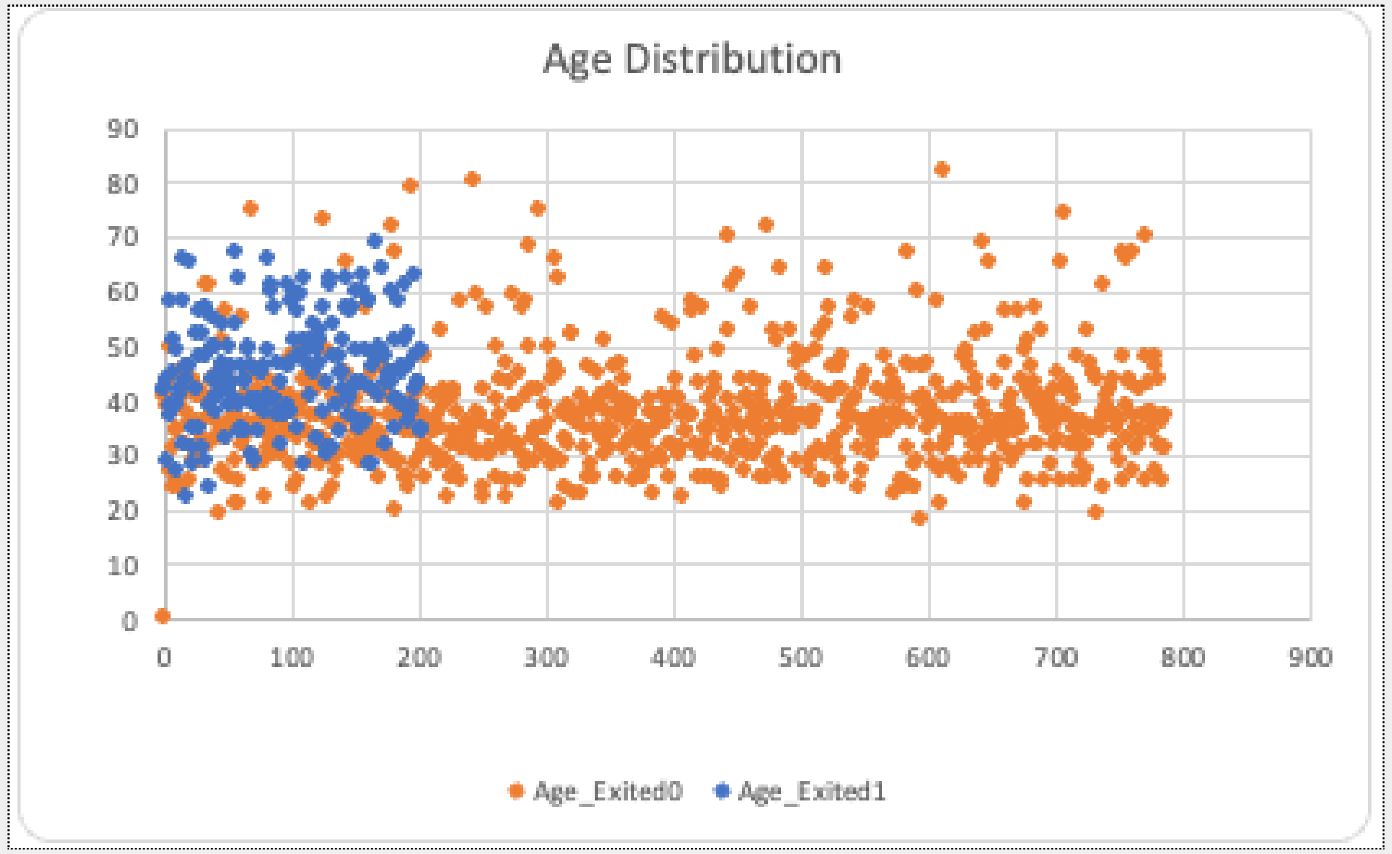
Median age:

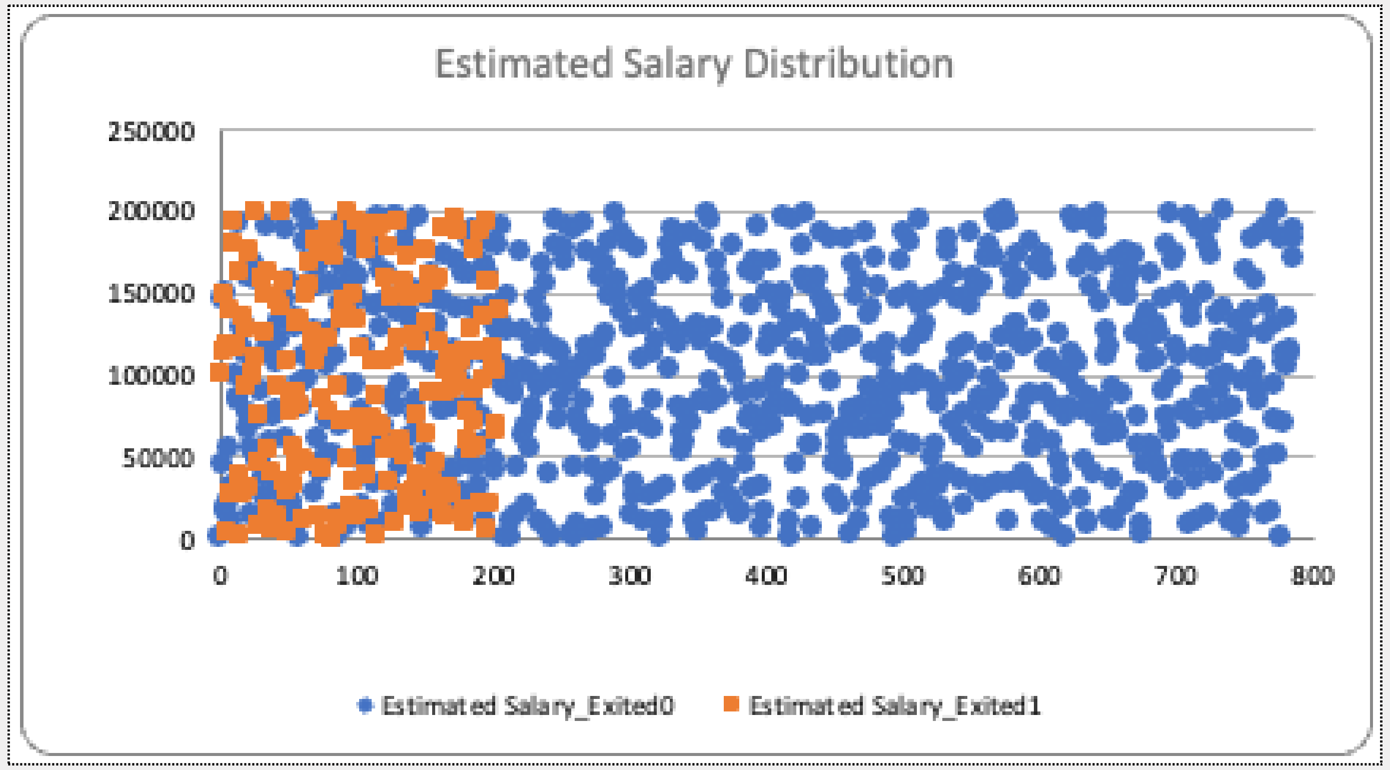
- Active customers > 36 years old
- Exited customers > 45 years old

Salary Distribution

Median salary:

- Active customers > $93.147
- Exited customers > $112.434





Legend:

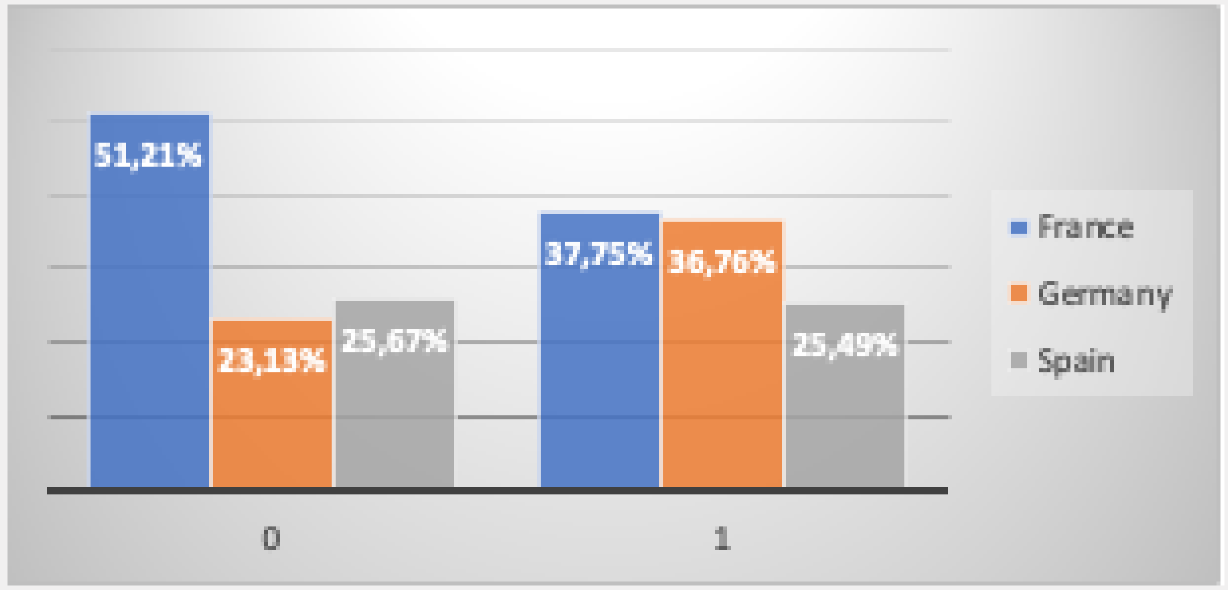
- Value **0** = Active Customers
- Value **1** = Exited Customers

- **Exited0** = Active Customers
- **Exited1** = Exited Customers

Key Insights

- Older clients with higher credit scores
- Clients from France and Spain are more likely to stay
- Clients from Germany are more likely to leave
- No significant difference in the average or distribution of estimated salaries

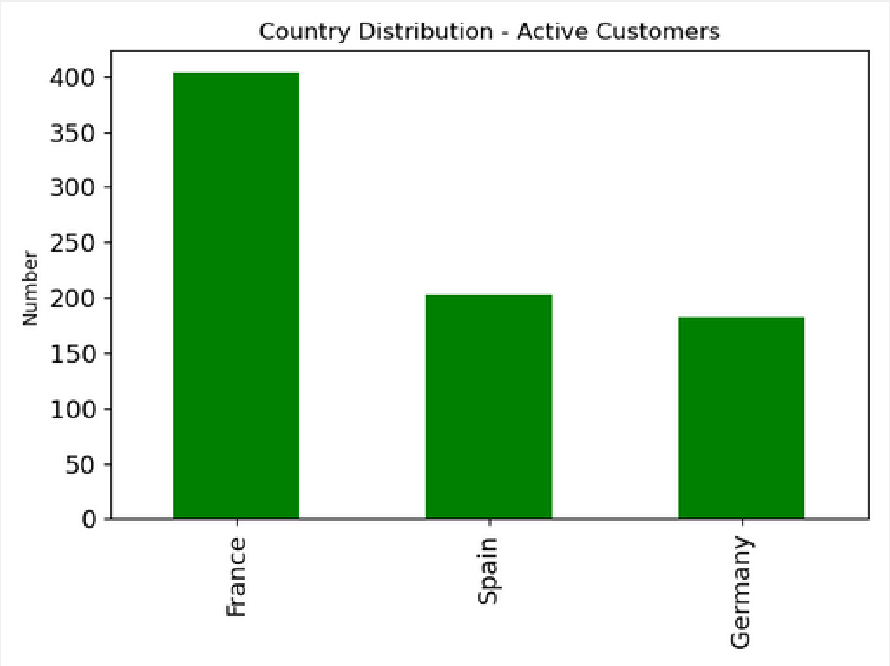
Country Distribution



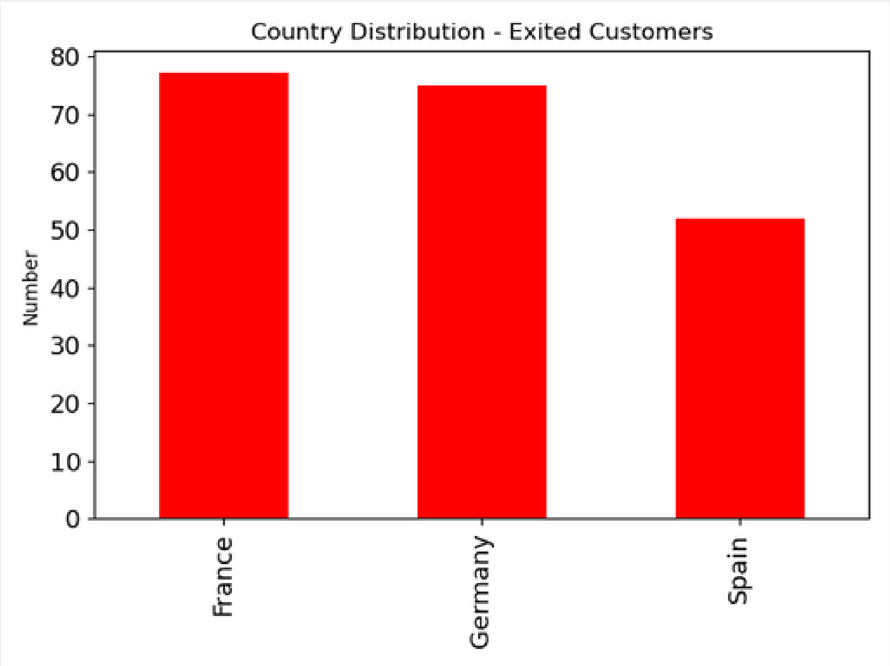
Legend:

- Value **0** = Active Customers
- Value **1** = Exited Customers

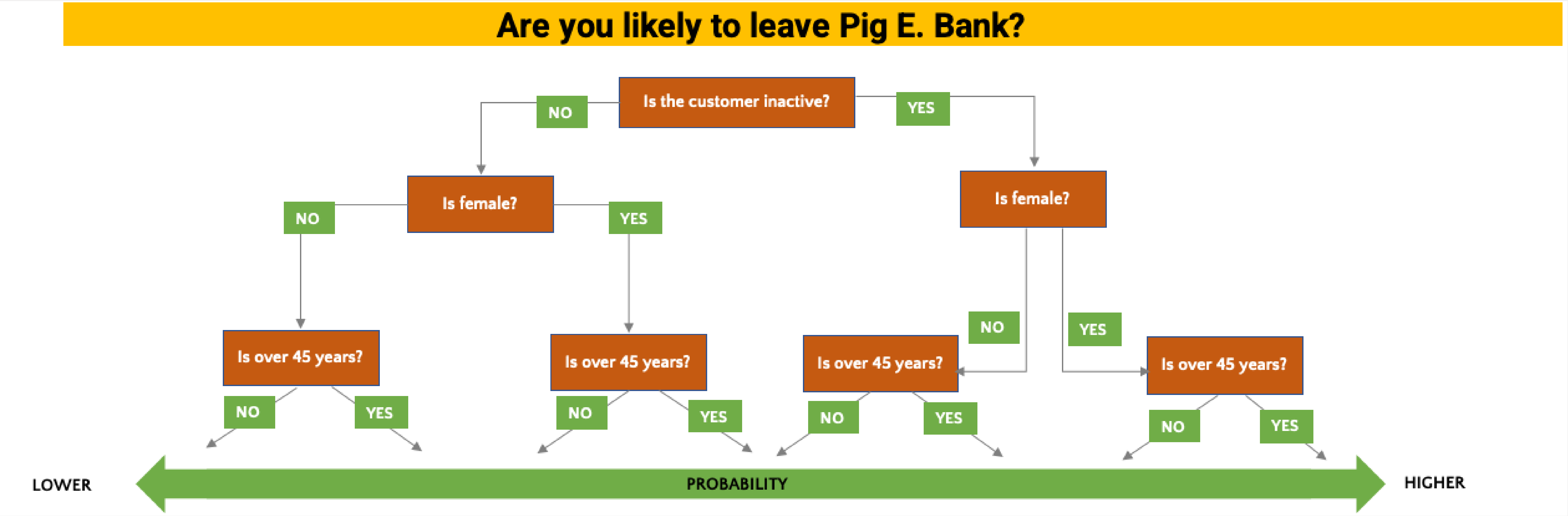
Active Customers



Exited Customers



Decision Tree



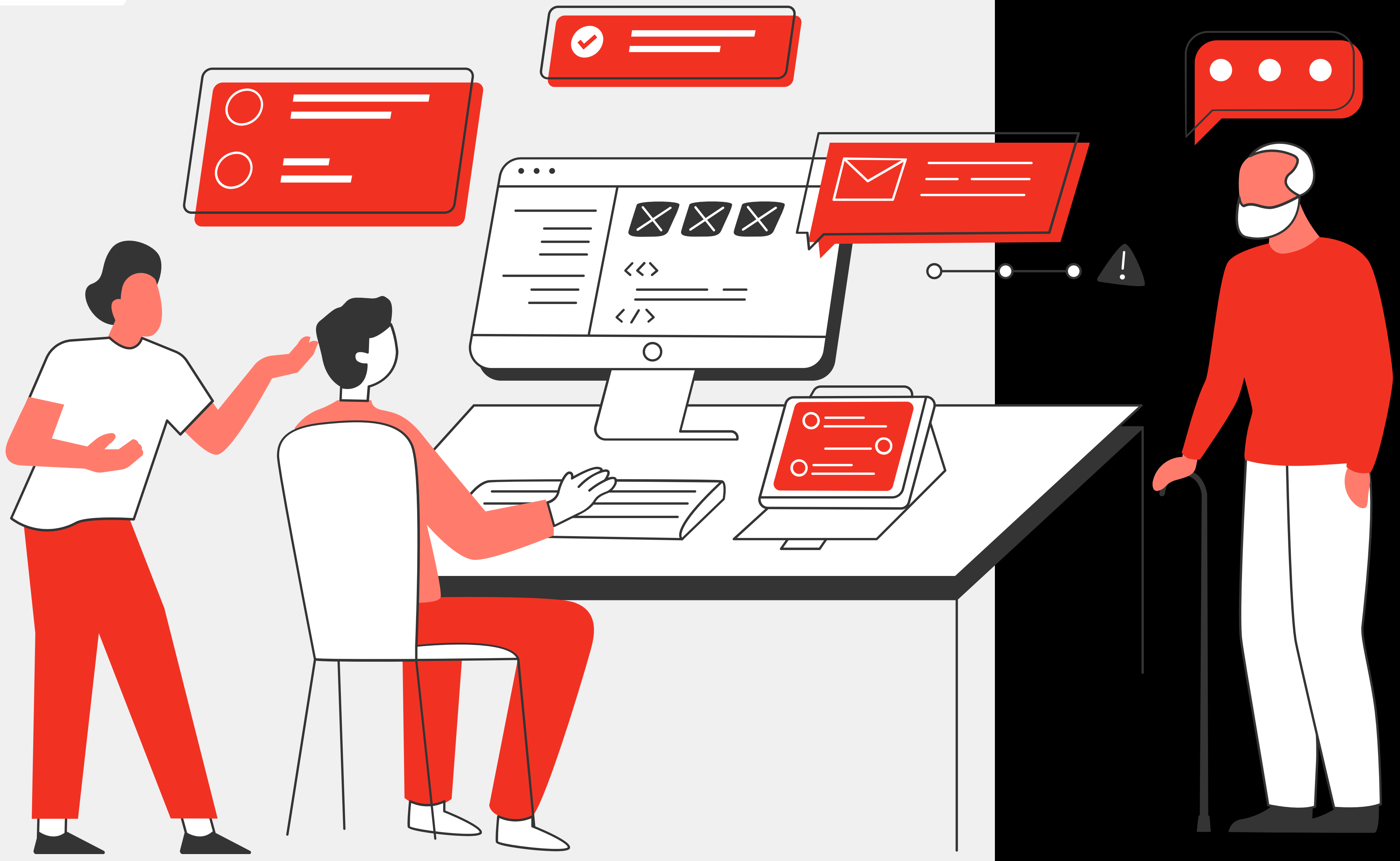
Recommendations

- Tailor retention efforts by developing personalized strategies for customers aged 45 and above, emphasizing incentives and personalized communication. Re-engage inactive clients, particularly in older age groups and among females, through customized promotions or services
- Investigate and address reasons behind higher customer loss in Germany, employing surveys and interviews
- Segment customers based on demographics, crafting targeted campaigns for females and older age groups
- Design country-specific retention strategies, reinforcing positive aspects for higher likelihood-to-stay countries like Spain

GitHub

Next Steps

- Utilize Random Forests for accurate salary prediction by building and merging multiple decision trees, capturing non-linear relationships from historical customer data and financial surveys
- Integrate the model into decision-making processes, ensuring continuous improvement through iterations
- For churn prediction, implement a logistic regression model using longitudinal customer behavior data
- Enhance fraud detection efficiency and strengthen customer engagement with machine learning insights, particularly in personalization like streamlining loan approval decision-making



THANK YOU

Greta Lawani