

2. Data Quality and Cleaning Process

The inconsistencies in the dataset, along with data quality and cleaning processes, were analyzed in the following Excel sheets:

- Data Profile of Raw Data
- Data Integrity of Raw Data
- Cleaned Data
- Data Profile of Cleaned Data
- Data Integrity of Cleaned Data

3. Data Organization

The cleaned and processed data is available in the following Excel sheets:

- Cleaned Data Pivot Tables
- Remaining Clients
- Remaining Clients Pivot Table
- Exiting Clients
- Exiting Clients Pivot Table
- Statistical Analysis

Key Observations from Statistical Analysis of Remaining and Exiting Clients

- Customer Inactivity: 70% of customers who exited the bank were listed as inactive.
- Credit Score: Exiting customers had poor credit scores.
- Geographic Influence: Germany had a disproportionately high exit rate (36.76%) compared to total customers (25.93%) and remaining customers (23.13%).
- Age Factor: Exiting customers had an overall higher average age (45 years).
- Gender Disparity: More women (59.31%) than men exited the bank.
- Income Influence: A higher percentage of customers earning less than \$10,000 left the bank (7.35%) compared to those staying (5.21%).
- Balance Trends:
 - Customers exiting the bank had a higher average balance (\$90,239.22) compared to those remaining (\$74,830.87) and the total customer base (\$78,002.72).

- Most exiting customers had balances are between \$90,000 to \$130,000.

Factors Influencing Customer Exit from Pig E. Bank

The following characteristics increase a customer's likelihood of leaving the bank:

- Listed as inactive
- Age between 38 and 47
- Resident of Germany
- Balance between \$90,000 and \$130,000
- Female

4. A detailed decision tree analysis is shown below:

The biggest discrepancy is in Germany, where a higher proportion of exited customers (36.76%) compared to remaining ones (23.13%) indicates that customers from Germany are leaving at a disproportionate rate. The discrepancy is quite significant for gender, with female customers exiting at a much higher rate (59.31%) compared to their presence in the remaining customer base (43.33%). While both factors show significant discrepancies, Gender appears to be a stronger predictor of customer exit than Country, as the difference in retention vs. exit rates is slightly larger for females than for German customers.

