Bank Customer Churn Analysis Portfolio Report

Project Title:  
Predicting and Analyzing Customer Churn for a Bank Using Python, Excel, and Power BI  
  
Objective (Ask)  
Identify factors driving customer churn in a bank, predict churn likelihood using Python, and visualize churn patterns in Power BI to guide strategic business decisions.  
  
Key Business Questions:  
- What is the current churn rate?  
- Which customer segments are churning the most?  
- What demographic, behavioral, or financial factors are linked to churn?  
- Which customer segments should we prioritize for retention efforts, and what targeted strategies should we implement to reduce churn?  
  
Data Overview (Prepare)  
Dataset of 10,000 customers with features including:  
Demographics: Country, Gender, Age  
Financials: Credit Score, Estimated Salary, Balance, Tenure  
Behavior: Number of Products, Has Credit Card, Is Active Member  
Target: Exited (Churn flag)  
  
Derived columns created:  
- Age Group (18-25, 26-35, 36-45, 46-60, 60+)  
- Tenure Group (0-3 Years, 4-6 Years, 7-10 Years)  
- Churn Status (Retained, Churned)  
- Churn Probability Estimate  
- Recommendation  
- Credit Score Range (300-500, 501-600, 601-700, 701-850)  
  
Data Cleaning & Processing (Process)  
Using Excel and Power Query:  
- Cleaned decimals, applied formatting.  
- Created derived columns listed above.  
- Retained flagged outliers for transparency.  
  
Exploratory Data Analysis (Analyze)  
Insights:  
- Current churn rate: 20.37%  
- Highest churn rates observed: Germany, females, 46-60 and 36-45 age groups.  
- Customers with Credit Scores 501-600 and 601-700 exhibited elevated churn counts.  
- High churn numbers in some age brackets driven by customer volume and churn risk combined.  
- Churn by Tenure Group analyzed but omitted due to insignificant variation.  
  
Predictive Modeling (Python)  
- Data prepared: encoded variables, retained outliers, split data.  
- Logistic Regression applied for churn prediction.  
- Performance evaluated with accuracy, precision, recall, F1, ROC-AUC.  
- Churn probabilities generated with predict\_proba(), added as a column.  
  
Dashboard Overview (Share)  
Built in Power BI with:  
- Header Cards: Total Customers, Churned Customers, Overall Churn Rate, Average Credit Score  
- Slicers: Country, Gender, Age Group, Credit Score Range  
- Visuals: Churn Status Distribution, Churn Rate by Country, Churn Distribution by Age Group, Churn Count by Credit Score Range, Customer Churn by Gender, Customer Distribution by Age Group  
  
Color Coding:  
- Churned: Red  
- Retained: Blue  
  
Note: Customer Distribution by Age Group included to contextualize churn counts. Others excluded as slicers and churn charts captured those insights.  
  
Business Recommendations (Act)  
What should we do next?  
- Focus on customers aged 46-60, extend to 36-45.  
- Prioritize German customers.  
- Target 501-600 and 601-700 credit score customers.  
- Design campaigns for females.  
  
Deliverables:  
- Cleaned dataset (.xlsx)  
- Python notebook  
- Power BI dashboard (.pbix)  
- This report  
  
Summary  
End-to-end churn analysis with prediction and visualization using Excel, Power BI, and Python, retaining flagged outliers. Dashboard organized for clarity with dedicated metrics and slicers for churn-focused insights.