



Zbigniew Wierciński

Data Analyst portfolio



[GitHub_wiercinskizb](https://github.com/wiercinskizb)



[LinkedIn_wiercinskizb](https://www.linkedin.com/in/wiercinskizb)

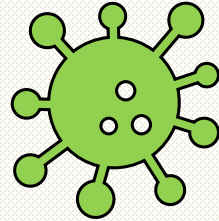
Projects



GAME CO

Analysis of
global video
game sales

Skills



INFLUENZA SEASON

Medical staff
allocation for flu
season preparation

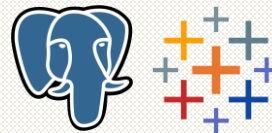
Skills



ROCKBUSTER STEALTH

Rental behavior
analysis for
streaming service

Skills



INSTACART

Customer
demographics and
purchasing
analysis

Skills



PIG E. BANK

Global bank
customer
retention
analysis

Skills





GAME CO

Game Co is a new video game company, which wants to use data analysis to understand the market development to guide strategies and generate growth in the most sustainable markets

Data

The [VG Chart](#) dataset covers global video game sales across platforms, regions, and genres.

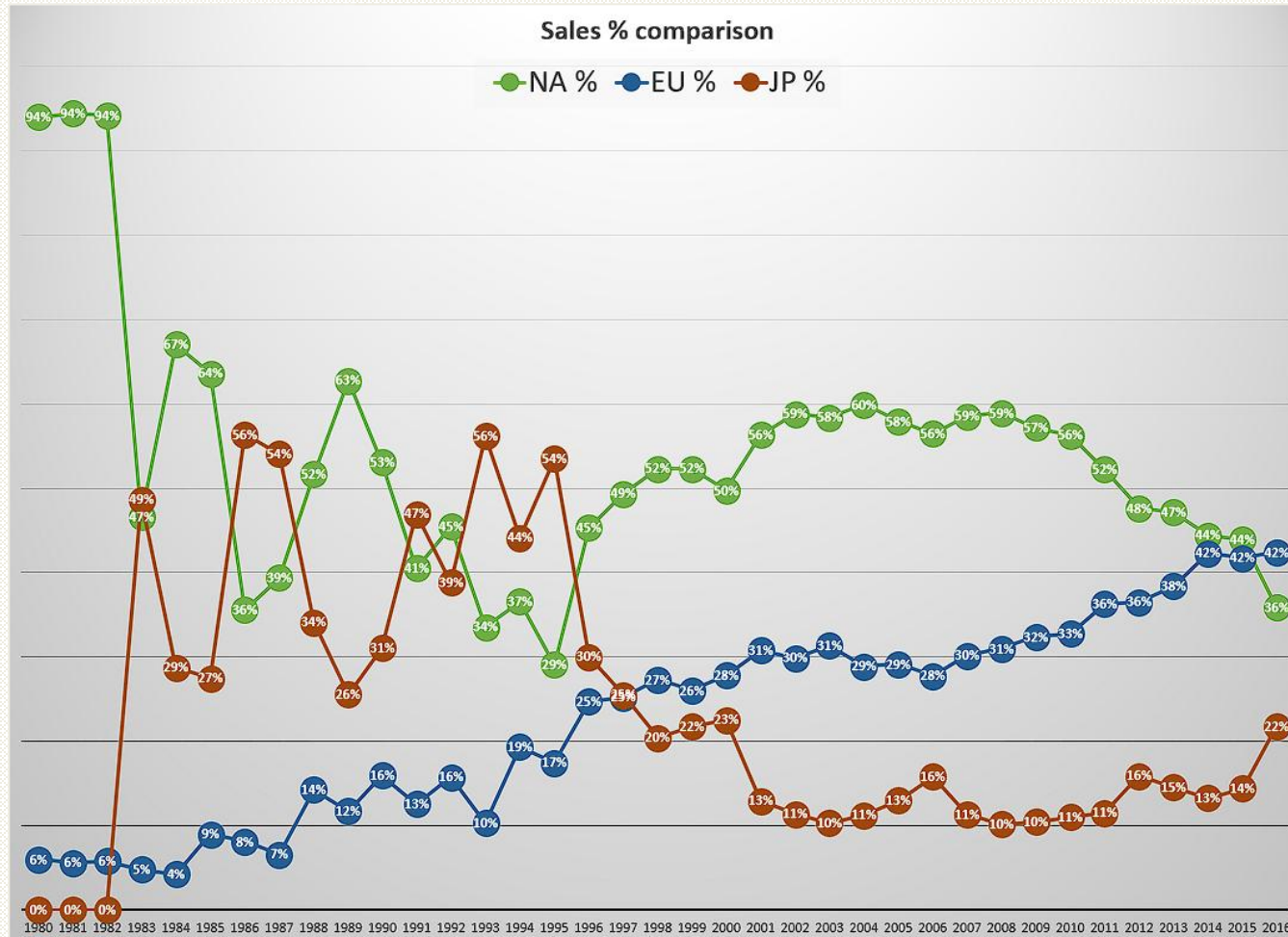
Objectives

- Analyze global sales trends over recent years
- Identify top platforms and publishers to assess market competition
- Deliver insights to shape GameCo's development and strategy

Methods

- Organizing and refining data using Excel (grouping, sorting, filtering)
- Cleaning and preparing data for analysis
- Conducting descriptive analysis to uncover trends
- Visualizing insights through charts and graphs

Regional Sales Trends Over Time: A Comparative Analysis



Description for the Chart: The chart shows the percentage share of video game sales in North America, Europe, and Japan from 1980 to 2016, highlighting key trends in market dynamics.

Key Observations:

North America (Green Line):

Dominates the market, with shares stabilizing at 50%-60% since the late 1990s.

Europe (Blue Line):

Steady growth from 6% in the 1980s to around 30%-36% in the 2000s.

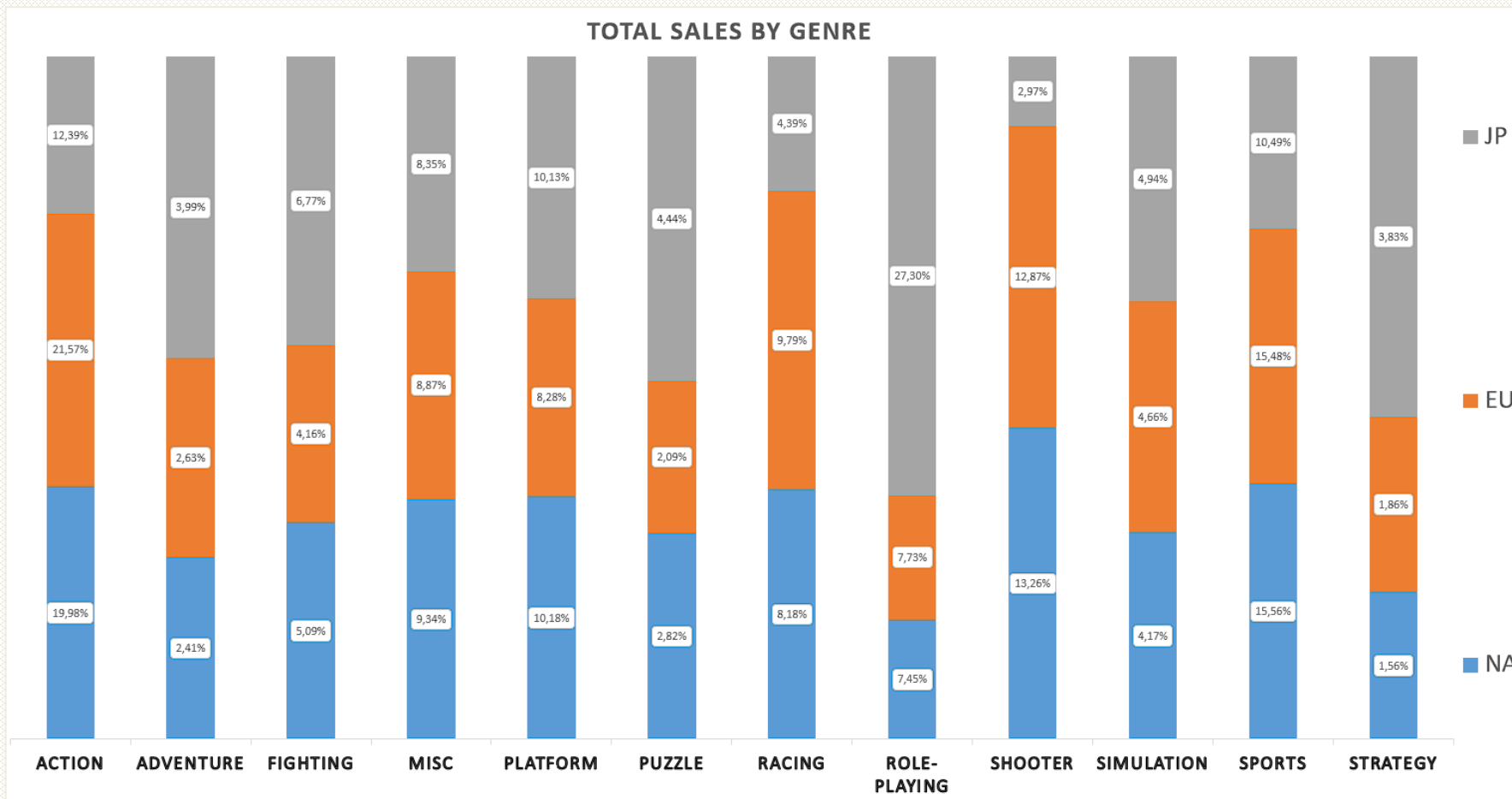
Japan (Orange Line):

Strong in the 1980s (up to 54%) but declines to 10%-16% by the 2000s.

Implications:

Focus on North America and Europe due to sustained growth, while exploring niche opportunities in Japan.

Regional Preferences in Video Game Genres



Key Observations:

North America (Blue Line):

Sports (15.56%) and Shooter (13.26%) lead, with strong interest in Action (19.98%).

Europe (Orange Line):

Similar to NA, Action (21.57%) and Sports (15.48%) dominate.

Japan (Grey Line):

Role-Playing (27.30%) is most popular, followed by Action (12.35%) and Platform (10.13%).

Implications:

Focus on Sports and Shooter games for NA/EU and Role-Playing for Japan.

Recommended Action Plan



1. Actionable Recommendation:

Implement a region-specific marketing and development strategy tailored to the preferences of each major market.

2. Immediate Steps:

- **Pilot Program:** Begin with a focused campaign in North America, emphasizing “Sports” and “Action” games, and evaluate its performance as a model for other regions.
- **Local Partnerships:** Engage local influencers and content creators in each region to strengthen GameCo's appeal through authentic, regionally resonant marketing.
- **Performance Metrics:** Continuously measure sales and engagement data by region to refine and adapt GameCo's strategies for sustained success.

3. Expected Outcomes:

- **Market Growth:** Increased sales by aligning game offerings with regional preferences.
- **Brand Loyalty:** Enhanced customer relationships and loyalty through a locally attuned product experience.



INFLUENZA SEASON

Analyzed trends to optimize staffing for a medical agency, ensuring hospitals receive adequate temporary personnel during peak U.S. flu seasons.

Data

[US Census Data Set](#)

Includes U.S. county and state population (2009–2017).

[CDC Influenza Deaths Data Set](#)

Includes monthly U.S. influenza death counts (2009–2017).

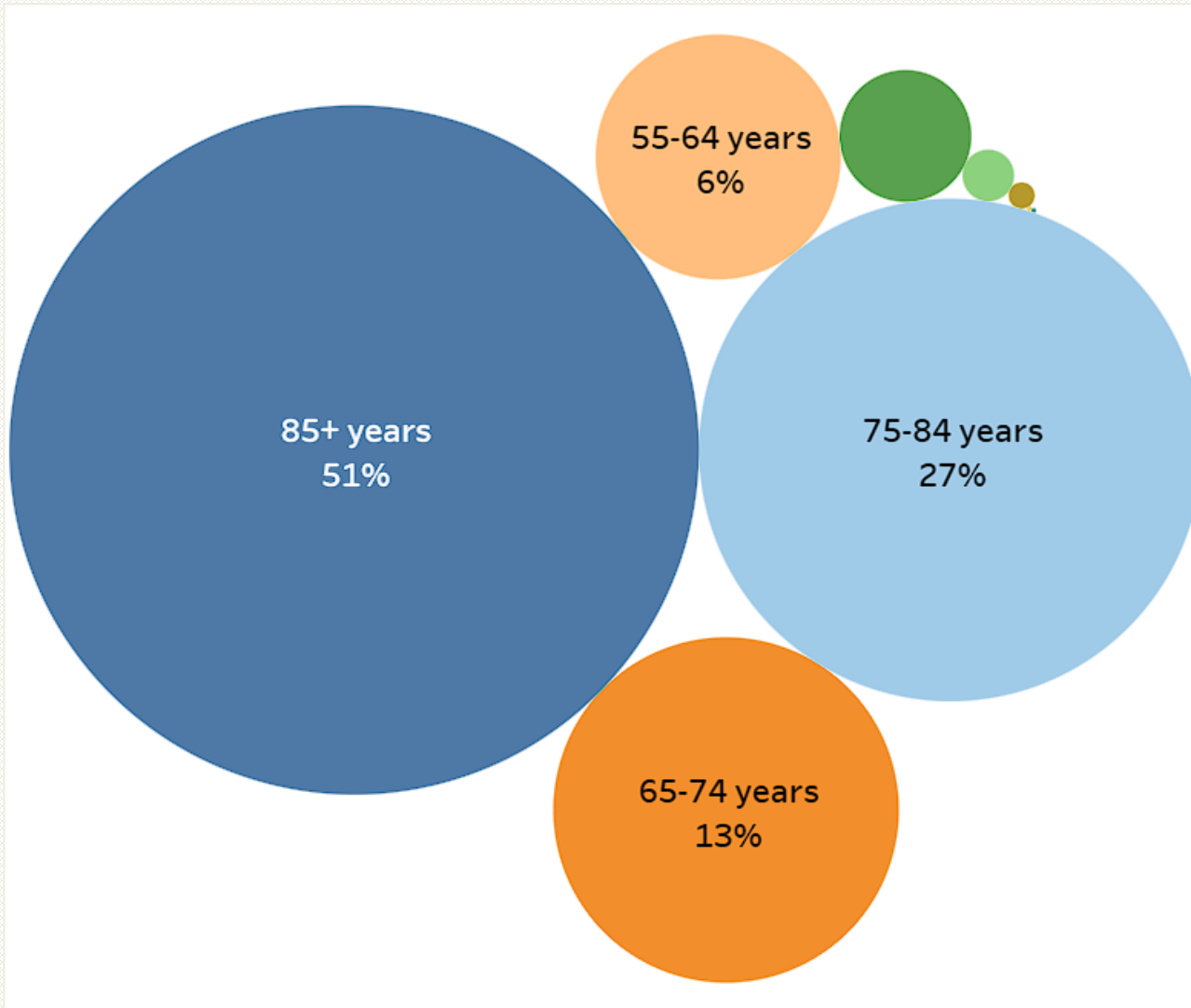
Objectives

- Identify seasonal influenza trends and prioritize high-need states based on vulnerable populations.
- Forecast staffing demand and optimize deployment across U.S. states

Methods

- Data cleaning, grouping, sorting, and filtering
- Integration and transformation of datasets
- Statistical hypothesis testing
- Visual analysis and forecasting
- Storytelling through Tableau

Influenza Death Rate by Age Group (2009-2017)

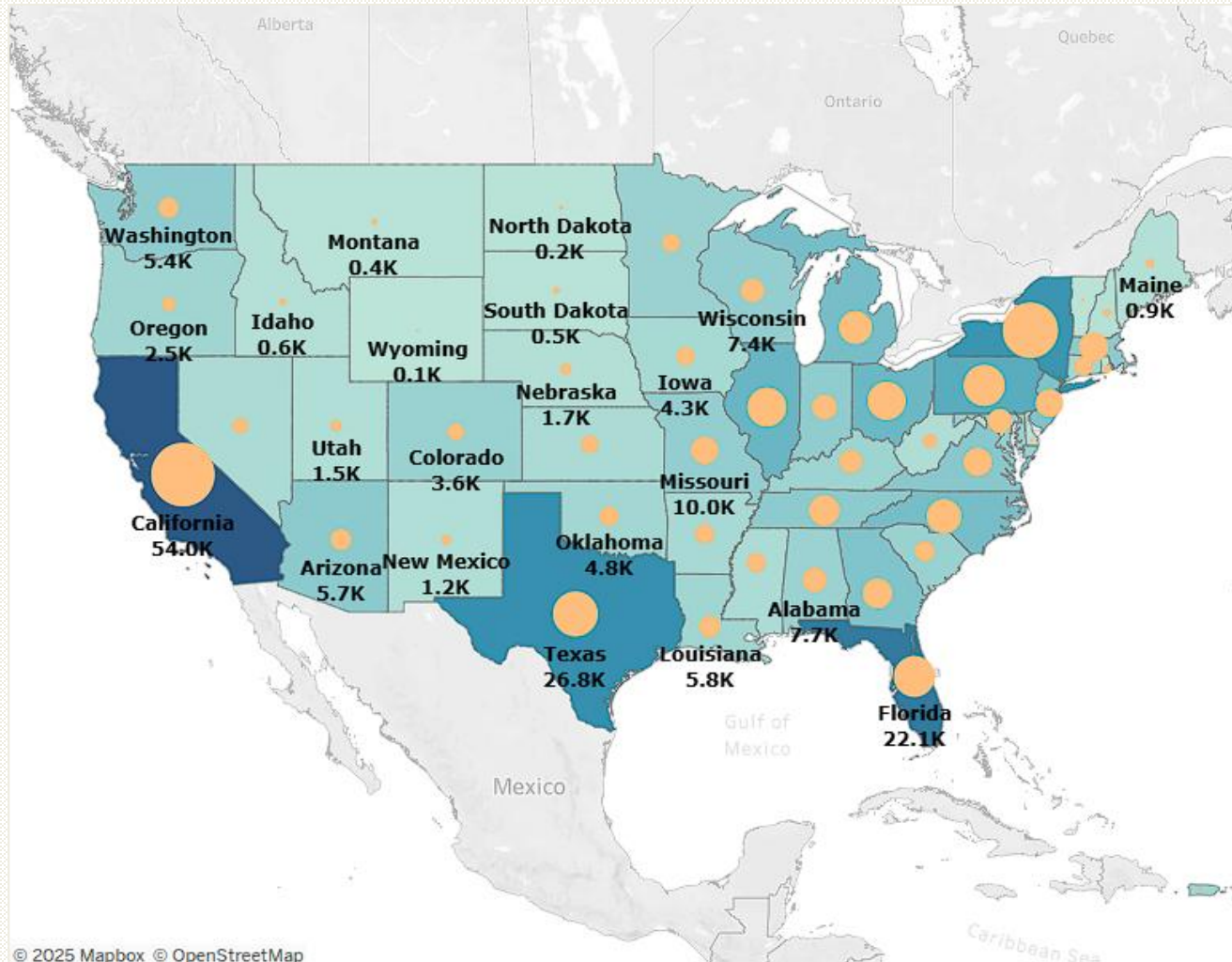


- The chart shows influenza-related deaths by age group (2009–2017).
- Most deaths occurred in the **85+ years** group (**51%**).
- The **75–84 years** group follows with **27%**, and the **65–74 years** group with **13%**.
- Younger age groups, like **55–64 years (6%)**, have significantly lower death rates.
- Highlights the need for targeted healthcare measures for older populations.

Map compares the distribution of the elderly population (65+) and influenza-related deaths across U.S. states over the years 2009–2017.

Source: CDC, Population and Influenza Data (2009–2017)

Larger circles represent higher influenza-related deaths, while darker shades indicate a larger elderly population.



► States with the largest elderly populations, such as **California (5.1M)**, **Florida (3.9M)**, and **Texas (3.2M)**, also show the highest influenza-related deaths, emphasizing a strong correlation between population size and fatalities.

► **California leads with 54K deaths**, followed by **Texas (27K)** and **New York (41K)**, highlighting the need for focused healthcare resources in these states.

► States with smaller elderly populations, such as **Wyoming (0.1M)** and **North Dakota (0.2M)**, report minimal deaths, underlining population density as a key risk factor.



Recommended Action Plan

Prioritize High-Risk States:

- Focus staffing and resources on **California, Florida, Texas, and New York** due to high elderly populations and death rates.

Strengthen Preventive Measures:

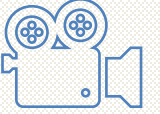
- Expand vaccination campaigns and public awareness programs targeting the elderly.

Enhance Data-Driven Planning:

- Use data to dynamically allocate resources and forecast needs across all states.

Support Smaller States:

- Ensure adequate healthcare resources for smaller but vulnerable states like **Wyoming** and **North Dakota**.



ROCKBUSTER STEALTH

Rockbuster Stealth LLC, once a global leader in movie rentals, is leveraging its existing movie licenses to launch a new online video rental service.

Data

The data includes details on customers, inventory, payments, and other relevant information.

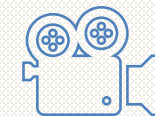
Objectives

- Support the management team with actionable insights to guide the strategic launch of their new online video rental service.
- Analyse critical factors such as revenue, customer distribution, and sales trends to optimize decision-making and market positioning.

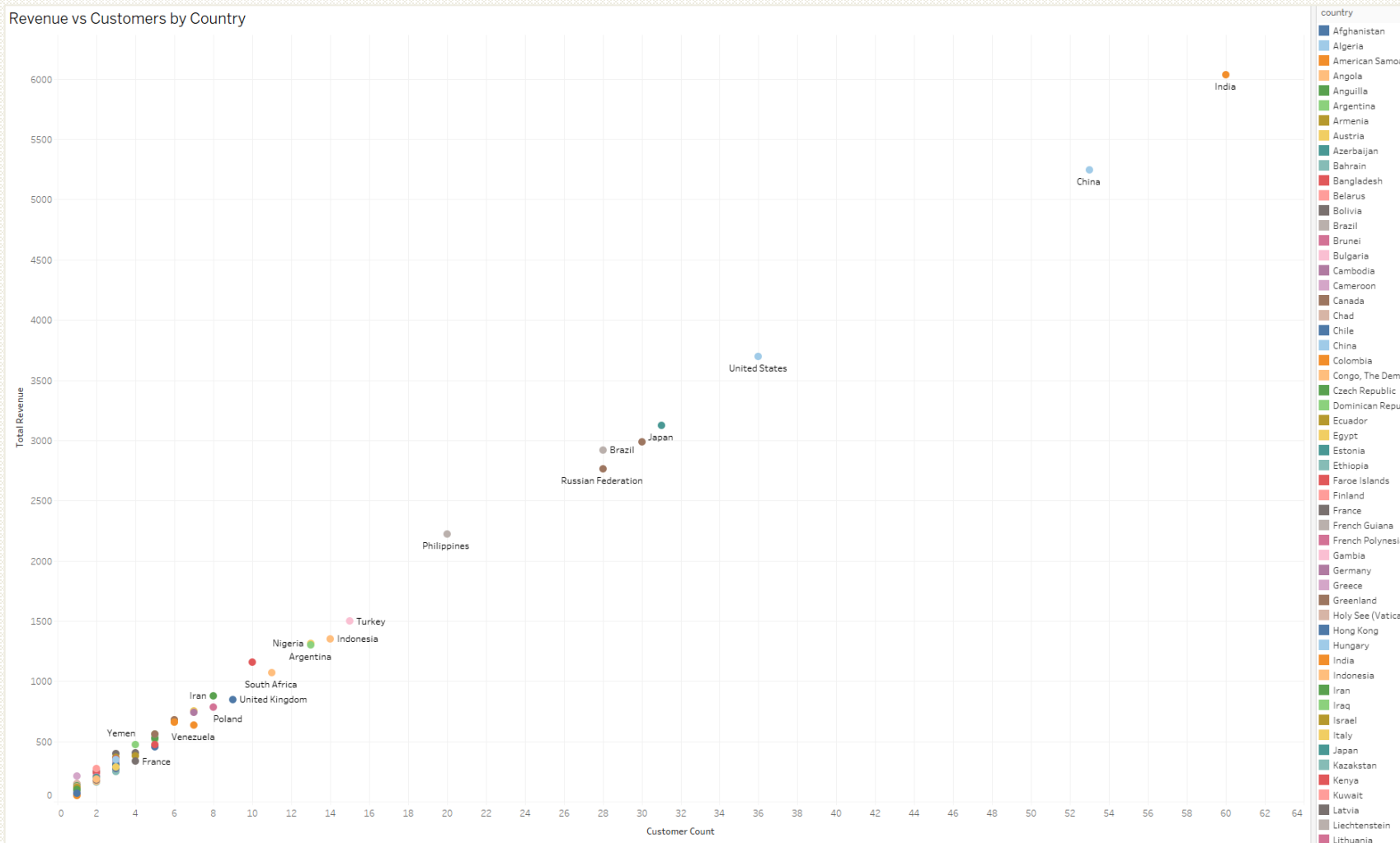
Methods

- Relational databases (SQL) for data management
- Database querying for extracting insights
- Filtering, cleaning, and summarizing data
- Joining tables for comprehensive analysis
- Subqueries and common table expressions (CTEs)

Revenue vs Customers Correlation



Revenue vs Customers by Country



There is a strong correlation between the number of customers and total revenue.

Countries with fewer customers tend to generate proportionally lower revenue.

India and China lead in both customer count and revenue generation.

Top Genre by Country



Top Genre by Country

Visualizing the top-performing genres by country based on total revenue. Bubble sizes represent revenue, providing an intuitive overview of genre popularity.

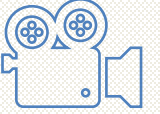


Sports is the most dominant genre globally, especially in countries like India and the United States.

Understanding these patterns can help target marketing strategies more effectively.

Animation and Sci-Fi also perform well in countries like China and the Russian Federation.

Recommended Action Plan



Key Insights:

- **Focus on top-performing regions** (e.g., India, China, USA).
- **Invest in popular genres** (Sports, Animation, Sci-Fi).
- **Target high-value customers** for retention.

Recommendations:

- **Expand Streaming Services:** Prioritize high-revenue regions.
- **Personalized Marketing:** Tailor campaigns to genre preferences.
- **Loyalty Programs:** Retain top customers with exclusive perks.



INSTACART

Instacart is an online grocery store that operates through a mobile app, offering customers the convenience of shopping from home. This project analyzes its sales data to uncover customer purchasing patterns, supporting a targeted marketing strategy.

Data

Customer dataset from CareerFoundry

Instacart online shopping dataset from 2017

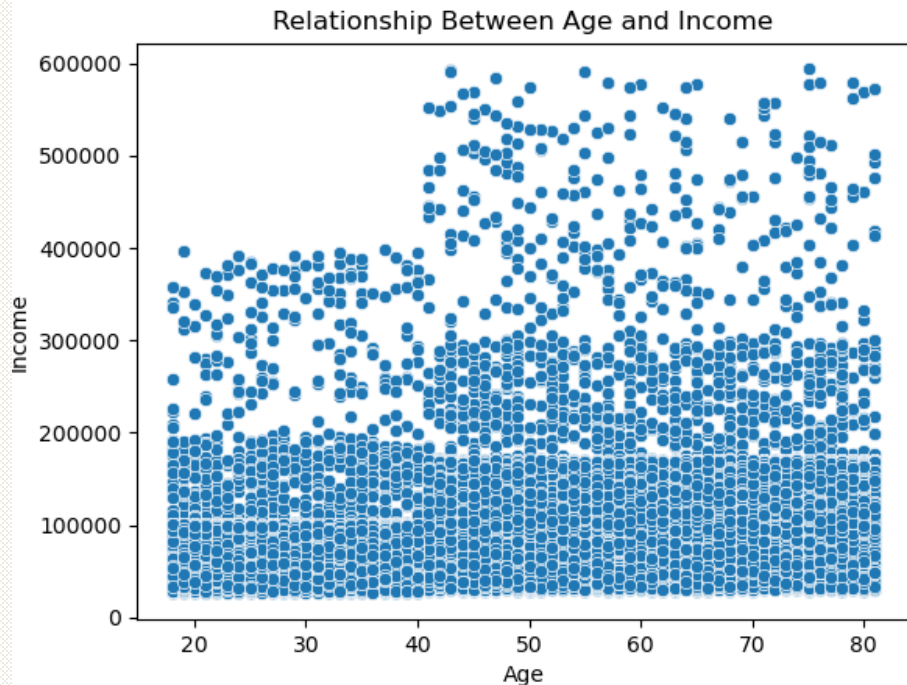
Objectives

- Determine peak and low-order periods to enhance ad scheduling and optimise product marketing strategies.
- Examine customer purchasing behavior, including loyalty, demographics, and regional variations, to refine targeted marketing efforts.

Methods

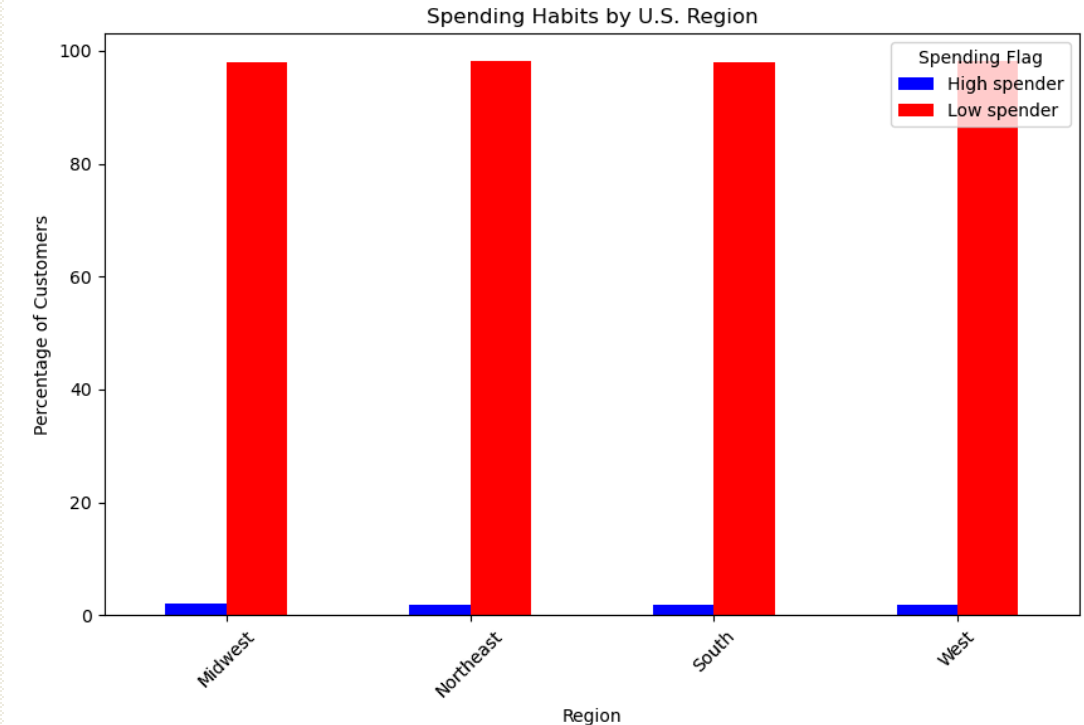
- Python
- Data wrangling & subsetting
- Data merging & variable derivation
- Data grouping & aggregation
- Data visualization using Python

Age and Income & Spending Habits



Key Observations:

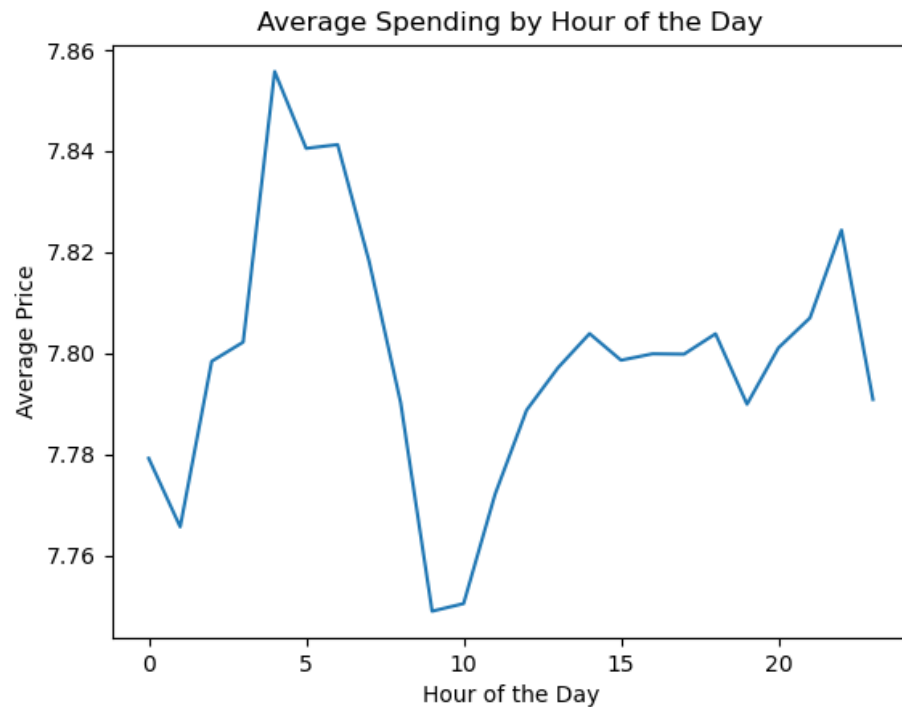
- There is no strong linear correlation between age and income.
- Income levels are highly variable across all ages.
- However, there may be a slight trend where older customers tend to have higher incomes, with a visible clustering around mid-age ranges.
- Distribution suggests that income may be influenced by other factors beyond age alone.



Key Observations:

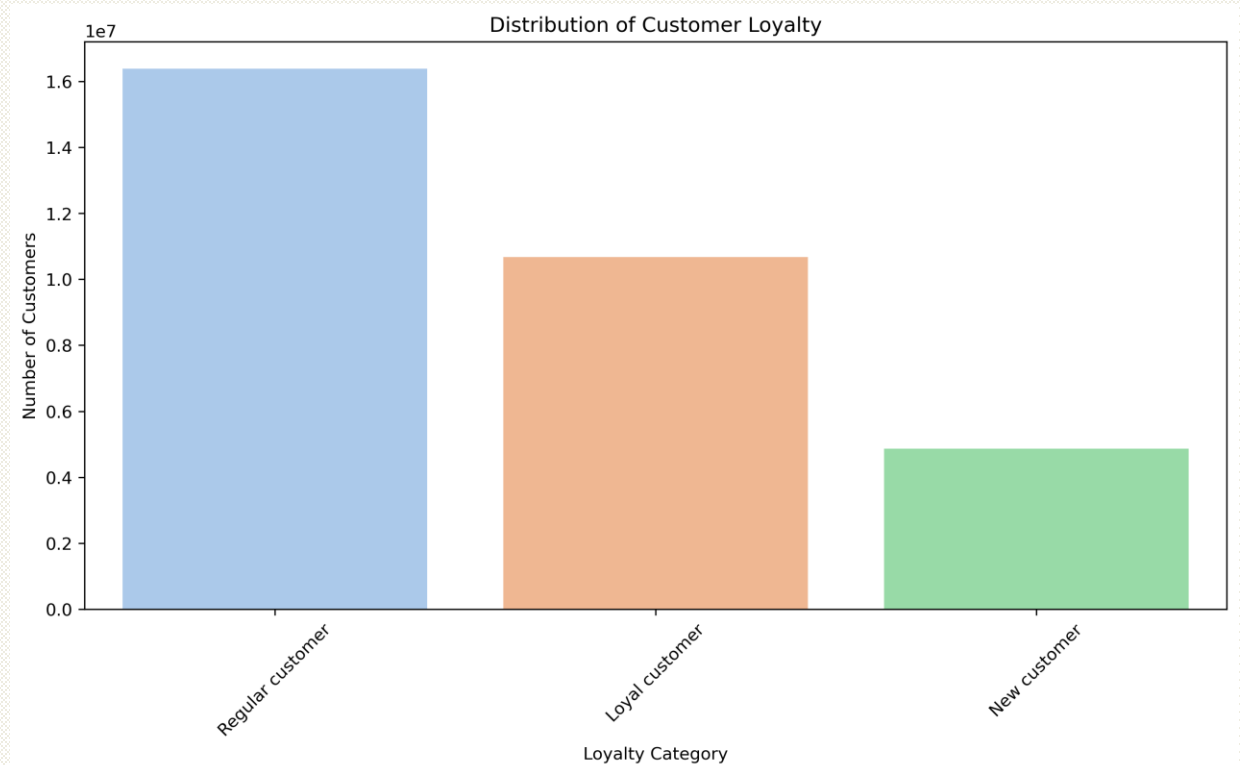
- The majority of customers in all regions are low spenders (~98%), while high spenders make up only a small fraction (~2%).
- There is no significant difference in spending behavior across regions.
- Regional segmentation does not appear to strongly impact spending habits.

Spending by hour & Customer Loyalty



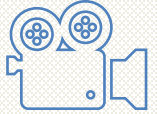
Key Observations:

- There is a noticeable increase in average spending between 2 AM and 6 AM, which might indicate fewer but higher-value purchases during late-night or early-morning hours.
- A dip in spending occurs around 10 AM, potentially suggesting more budget-conscious shopping behavior in the morning.
- From 12 PM to 8 PM, spending stabilizes, with minor fluctuations, before slightly increasing in the late evening.



Key Observations:

1. Regular Customers Dominate → Largest segment, indicating a stable returning base.
2. Loyal Customers (Engaged Shoppers) → Significant proportion, suggesting opportunities for personalized rewards.
3. New Customers → Lowest segment, highlighting the need for onboarding incentives to encourage repeat purchases.



Recommended Action Plan

Optimize Ad Scheduling Based on Peak Shopping Hours

- **Finding:** Most customer activity occurs between **10 AM and 4 PM**, with declining engagement in the evening.
- **Insight:** Digital ads and promotions should be scheduled **before peak hours (8-10 AM)** to capture early customer interest and influence their purchase decisions.
- **Action:** Introduce **morning discounts or flash deals** to boost engagement during off-peak times.

Strengthen Customer Retention & Loyalty Programs

- **Finding:** **Regular customers** form the largest portion of the customer base, followed by loyal customers and a smaller group of new buyers.
- **Insight:** Converting **regular customers into loyal shoppers** can drive long-term revenue growth.
- **Action:** Implement **tiered loyalty programs** (e.g., discounts after multiple purchases) and **exclusive repeat buyer offers** (e.g., "Shop 5 times, get free delivery").

Target High-Demand Product Categories for Promotions

- **Finding:** Specific departments (**4, 16, and 19**) have the highest number of orders.
- **Insight:** Highlighting these **best-selling categories** in marketing efforts can maximize revenue potential.
- **Action:** Feature these high-demand products in **homepage promotions** and bundle them with complementary products to boost sales.



PIG E. BANK

Pig E. Bank, a global financial institution, has a dedicated antimoney laundering compliance department. This project analyzed customer satisfaction data to uncover key drivers of customer loss.

Data

Data provided by CareerFoundry.

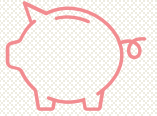
Objectives

- Identify key factors driving client loss using a decision tree model.
- Recommend strategies to improve customer retention.

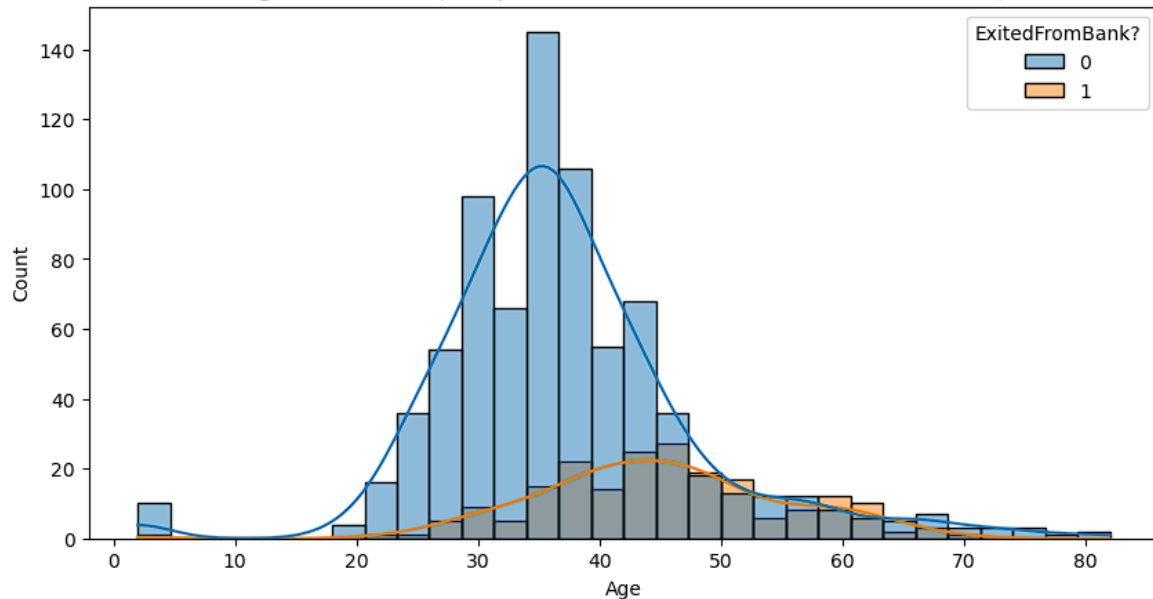
Methods

- Python
- Data wrangling & subsetting & cleaning
- Exploratory Data Analysis
- Data Mining Techniques
- Data Cleaning & Consistency Checks
- Data Modeling Predictive Analysis (e.g., Linear Regression)

Churned vs Retained - Age Distribution & Nr of Products



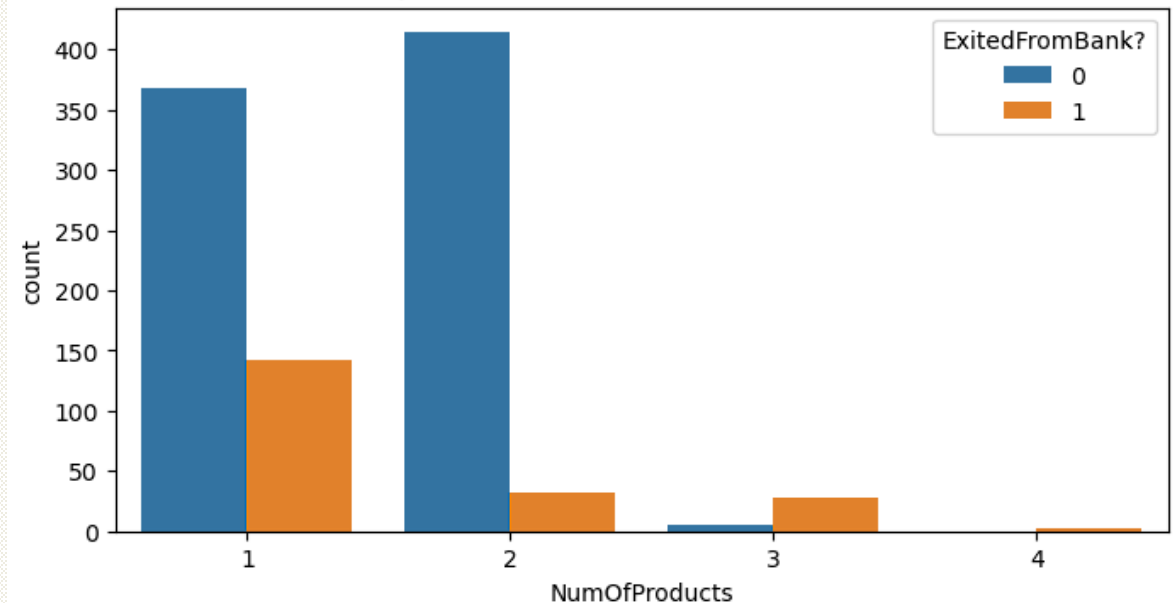
Age Distribution (Comparison of Churned vs. Retained Customers)



Key Observations:

- Majority of Customers are in their 30s
- Churned Customers (Exited) are More Evenly Distributed Across Ages
- Younger Customers (Below 25) and Very Old Customers (Above 60) are Less Represented

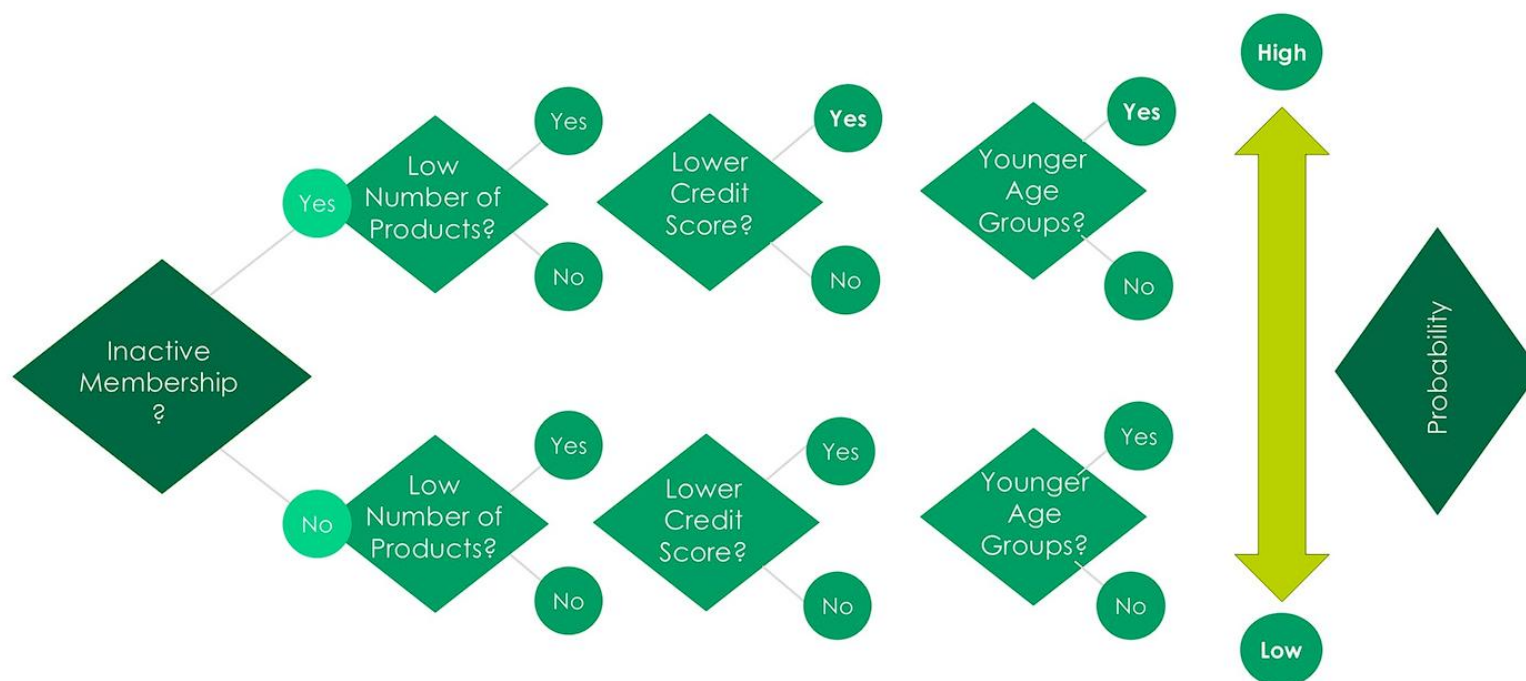
Impact of Number of Products on Churn



Key Observations:

- Customers with 1 Product Have the Highest Churn Rate
- Customers with 2 Products are the Most Loyal
- Very Few Customers Have More Than 2 Products, but Churn is Noticeable

Decision Tree churn leading factors



Key Observations:

- Inactive Membership and Low Number of Products Drive Churn
- Younger Customers are More Likely to Churn
- Having More Products Reduces Churn Risk

Recommended Action Plan



Engage and Retain Inactive Customers

- Implement **automated reminders and personalized offers** to reactivate inactive accounts.
- Offer **loyalty rewards** for returning customers (e.g., discounts on fees, cashback on transactions).

Encourage Multi-Product Adoption

- Target customers with **only one product** by offering **bundled services** or personalized product recommendations.
- Provide **discounts or incentives** for adding a second product (e.g., savings account + credit card).

Focus on Younger Customers with Tailored Offers

- Create **youth-oriented banking solutions** such as **low-fee accounts, financial education programs, or startup-friendly loan options**.
- Develop **gamified loyalty programs** to maintain engagement and long-term retention.

Monitor and Improve Customer Satisfaction for Multi-Product Users

- Conduct **customer feedback surveys** for those with **three or more products** to identify pain points.
- Offer **better user experience and streamlined banking services** to prevent dissatisfaction due to complexity.



THANK YOU



[GitHub_wiercinskizb](https://github.com/wiercinskizb)



[LinkedIn_wiercinskizb](https://www.linkedin.com/company/wiercinskizb)