LERTKIET LERTCHAYANTEE

DATA ANALYST PORTFOLIO

PROJECTS

01

GameCo

Global video game sales analysis

02

Influenza Season

Providing staff analysis for medical facilities

03

Rockbuster Stealth

Strategy analysis for the new online video service

04

Instacart

Analysis to uncover data on sales patterns

05

Pig E. Bank

Data mining analysis for a global bank

TOOLS USED

- **Microsoft Excel**
- **Pivot Table**
- **Tableau Visualization**
- **PostgreSQL**
- **Python**
- **GitHub**

GAMECO

Global Video Game Sales

Background

GameCo is a new video game company discovering that

“sales for the various geographic regions have stayed the same over time”.

Objectives

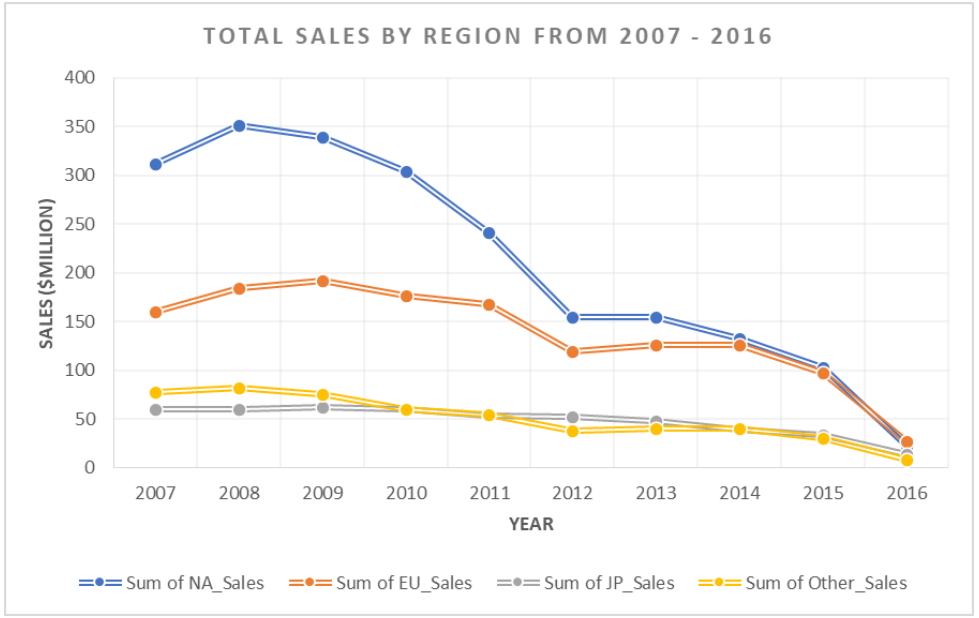
- Analyze global sales trends over time.
- Identify top-selling game genres and the regions with the highest sales.
- Predict the future trends.

Highlighted Skills

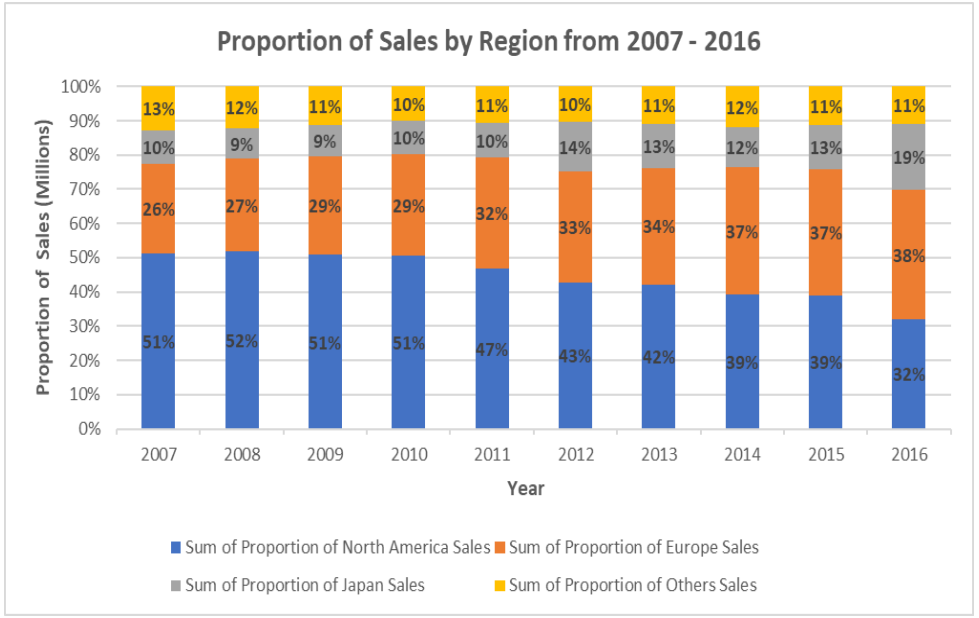
- Grouping/sorting/filtering data
- Descriptive analysis
- Data cleaning
- Data visualization

GameCo

Analysis (Line and Stacked Column Charts)

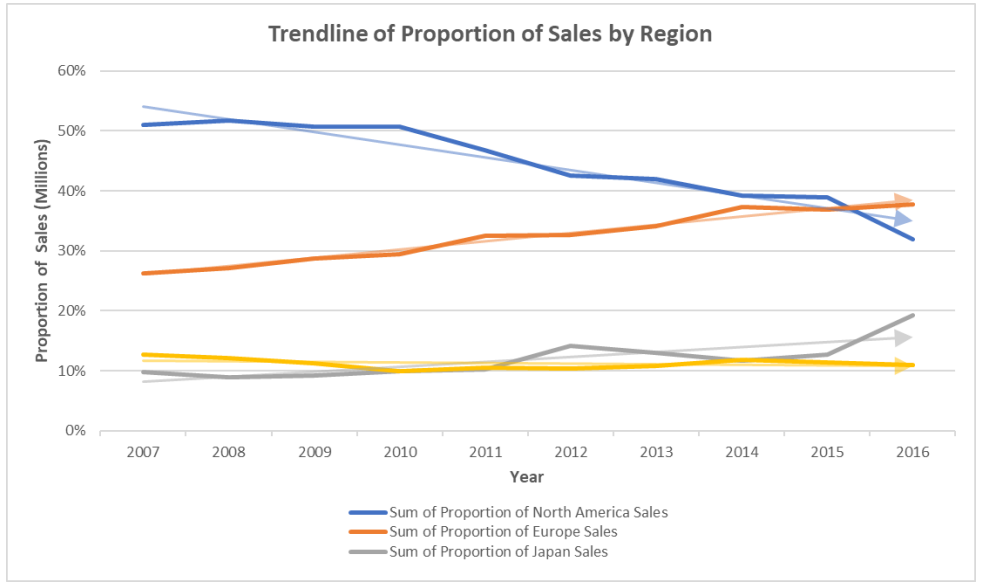


Sales in all regions for the last 10 years (2007-2016) **have decreased over the years.**



The proportion of sales in **North America has taken majority**, but it has **decreased** over the years

Recommendations



Future Trend

1. Focus on distributing marketing budget for **Europe and Japan.**
2. **North America's sales** have decreased and, it continues to **decline in 2017**.
3. **Other regions** is my **last priority** to distribute marketing budget .

Tool Used

Microsoft Excel

INFLUENZA SEASON

CDC Preparation

Influenza Season

Background

The United States has an influenza season where more people than usual suffer from the flu.

“Hospitals and clinics need additional staff to adequately treat these extra patients”.

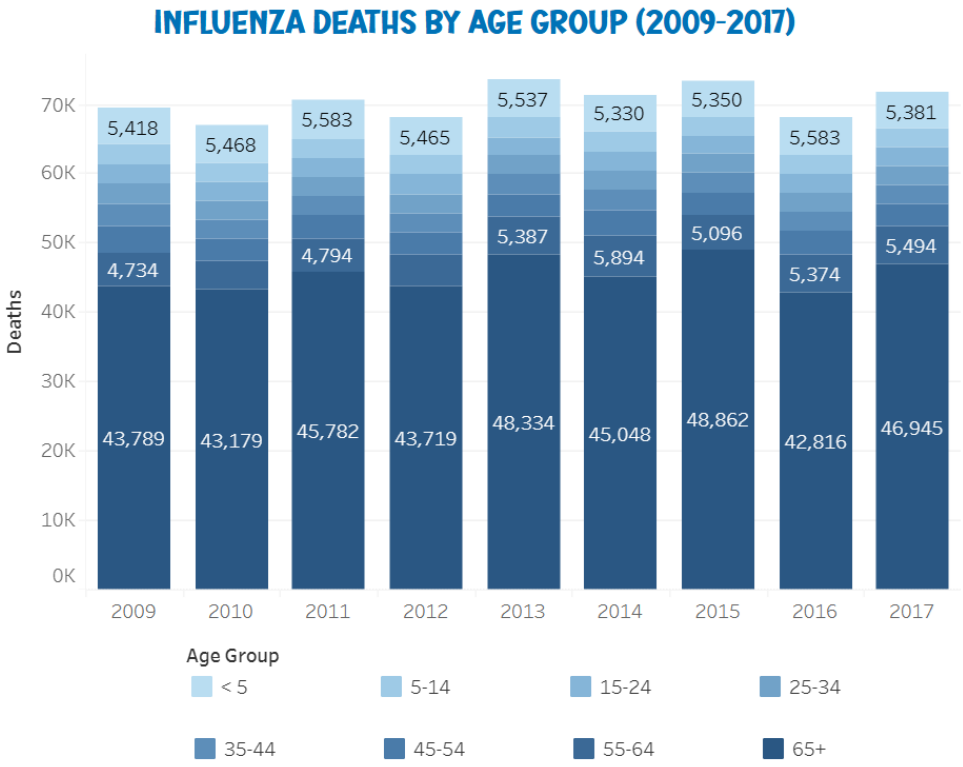
Objectives

- Assist the agency to provides temporary workers to clinics and hospitals.
- Plan for influenza season, when extra staff is in high demand in top sales regions.
- Examine trends in influenza .

Highlighted Skills

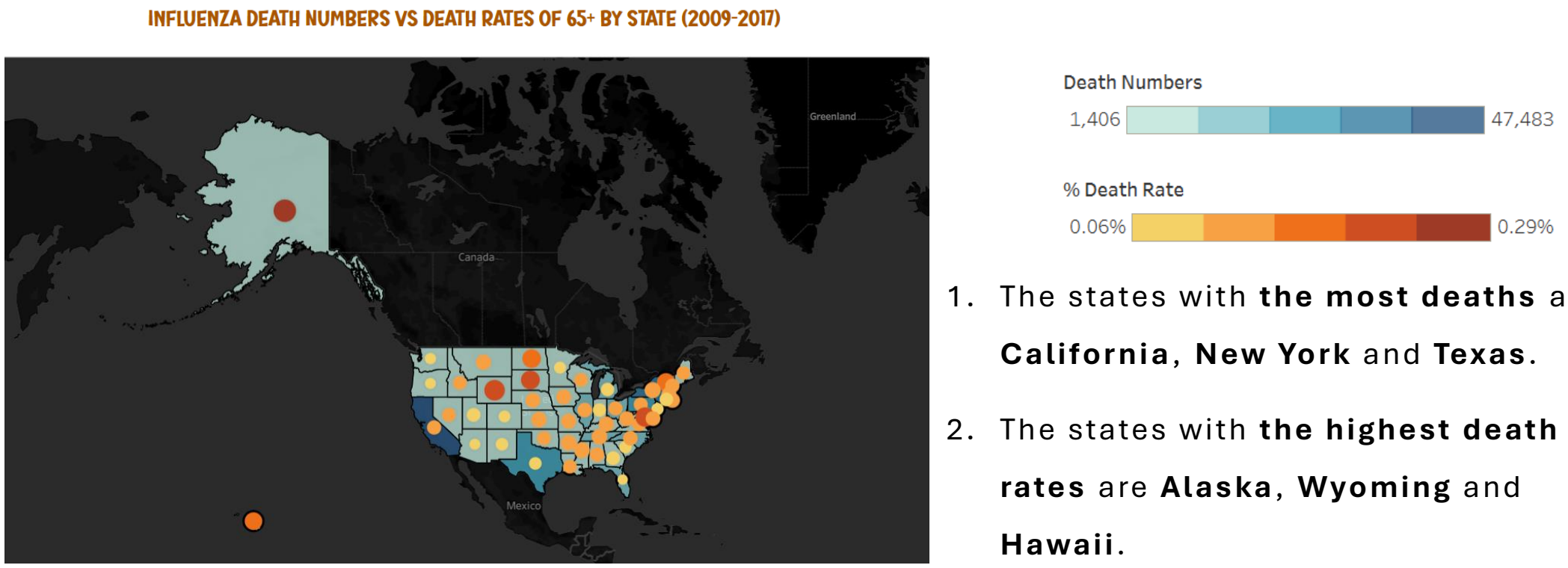
- Data profiling and integrity
- Data Integration
- Statistical hypothesis testing
- Data visualization with Tableau
- Forecasting

Analysis (Stacked Bar Chart)



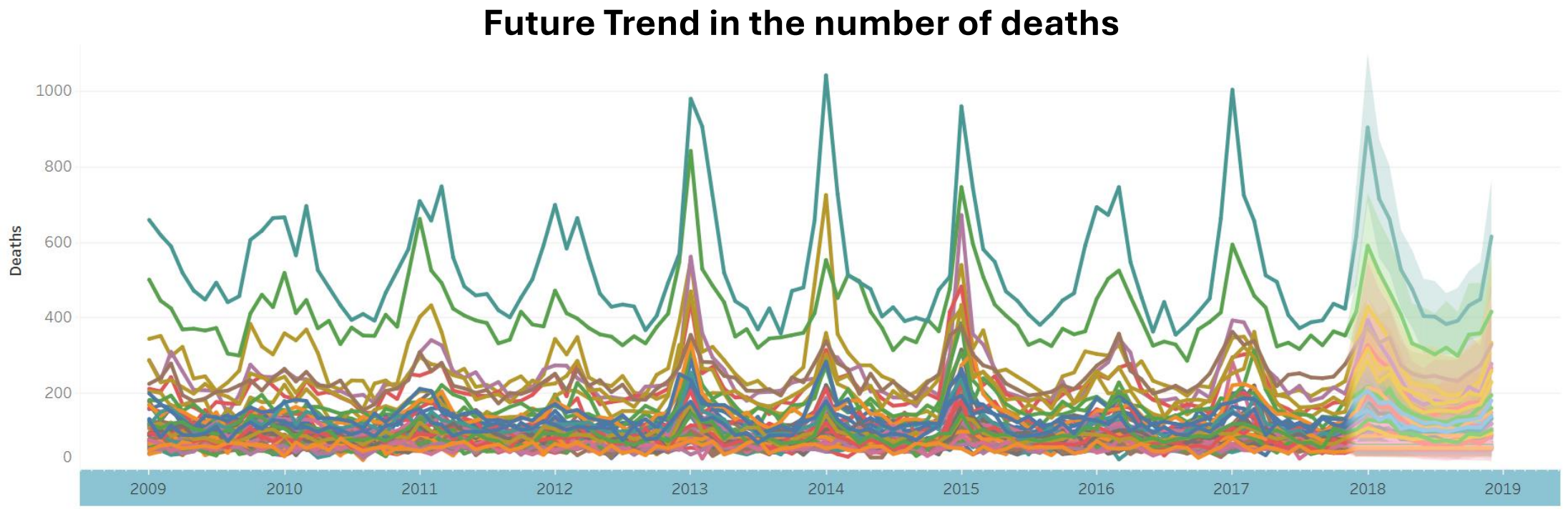
Age groups **under 5** and **over 65** have much **higher total death rates** than other age groups.

Spatial Analysis



1. The states with **the most deaths** are **California**, **New York** and **Texas**.
2. The states with **the highest death rates** are **Alaska**, **Wyoming** and **Hawaii**.

Recommendations



1. Prioritize **vulnerable populations**, especially those **aged 65 and older**.
2. Focus on **states** with both **high death totals** and **high death rates**.
3. Allocate more staff and resources during **peak seasons**.

Click here to view Tableau Storyboard

Tool Used

Microsoft Excel/Tableau

ROCKBUSTER STEALTH

Online Video Streaming Company

Rockbuster Stealth

Background

Rockbuster Stealth is facing stiff competition from streaming services such as Netflix and Amazon Prime.

“The management team is planning to launch an online video rental service in order to stay competitive”.

Objectives

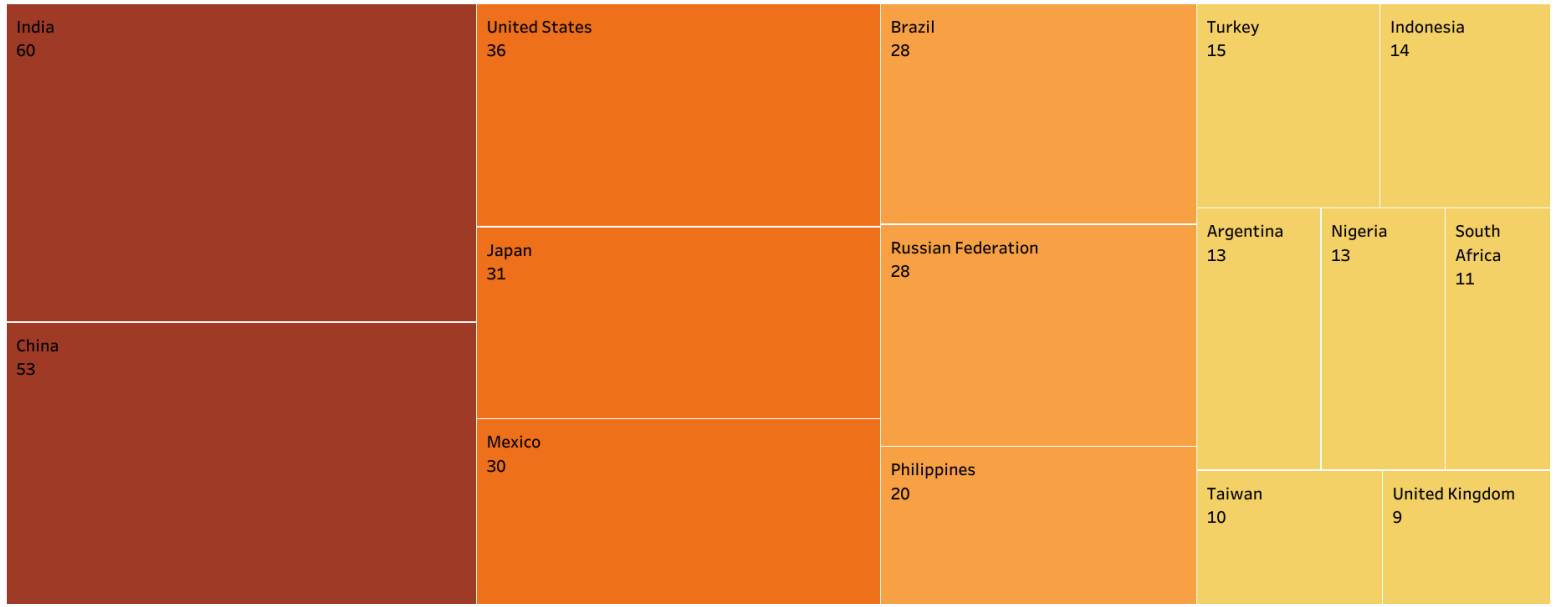
- Answer a series of business questions.
- The answers are expected to be data-driven answers that can be used for the 2020 company strategy.

Highlighted Skills

- PostgreSQL
- Data cleaning in SQL
- Joining Tables
- Subqueries & CTEs

Analysis (Heat Map)

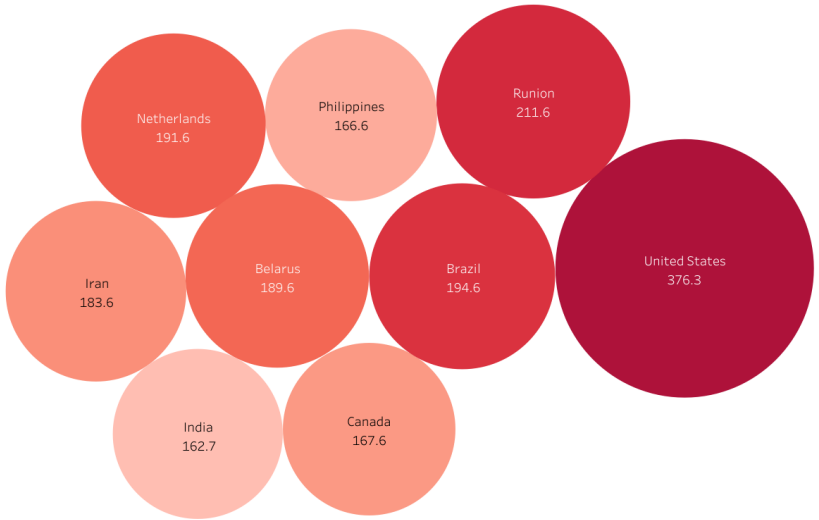
Which countries are Rockbuster customers based in?



India, **China**, **USA**, **Japan** and **Mexico**
with the range of **30 – 60 customers**

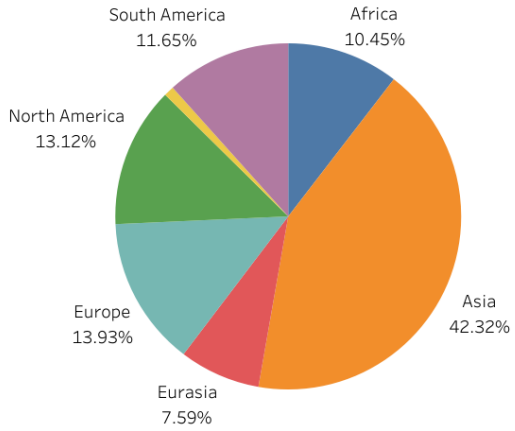
Analysis (Bubble and Pie Charts)

Where are customers with a high lifetime value based?



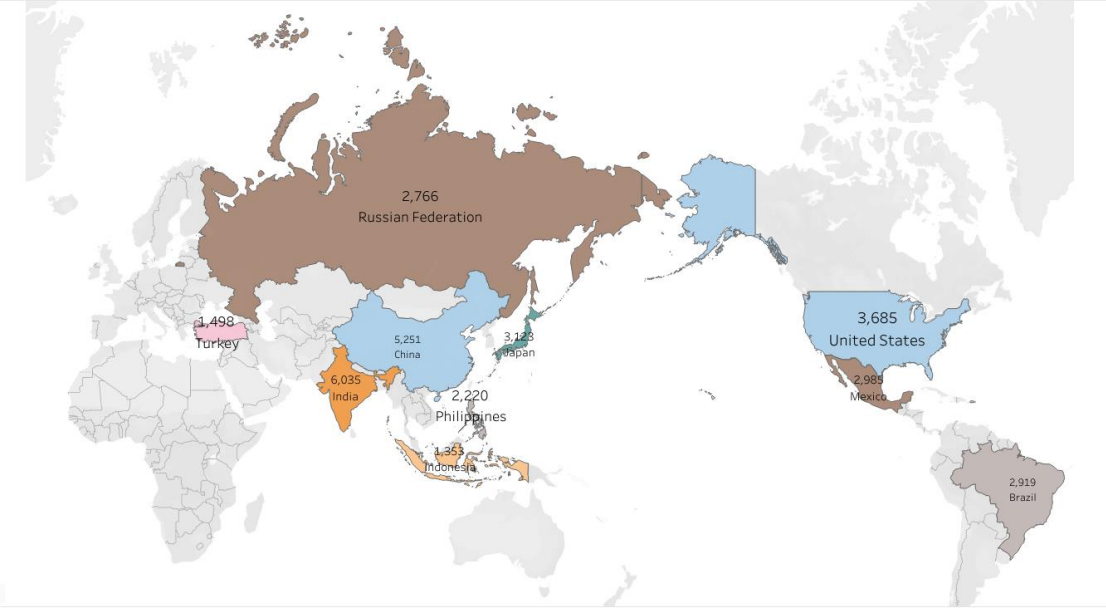
Among top 10 countries,
USA has **the highest total revenue**.

Geographic regions insight



Asia has the **most sales** figures (**42%**).
Oceania has the **least sales** figures (**0.94%**).

Recommendations



Top 10 Countries by Total Sales by Region

1. Consider investing in customer engagement in **Asia.**
2. Boost up revenues in **Oceania** and **South Africa.**
3. Add more inventory in **PG-13** and **NC-17 ratings.**
4. Add more revenues on **long rental duration movies.**

Click here to view the SQL code and additional files

Tool Used

Microsoft Excel/Tableau/SQL/GitHub

INSTACART

Online Video Streaming Company

Instacart

Background

The Instacart stakeholders are interested in the variety of customers and their purchasing behaviors in the database.

“They aim to target different customers with marketing campaigns to assess their impact on product sales”.

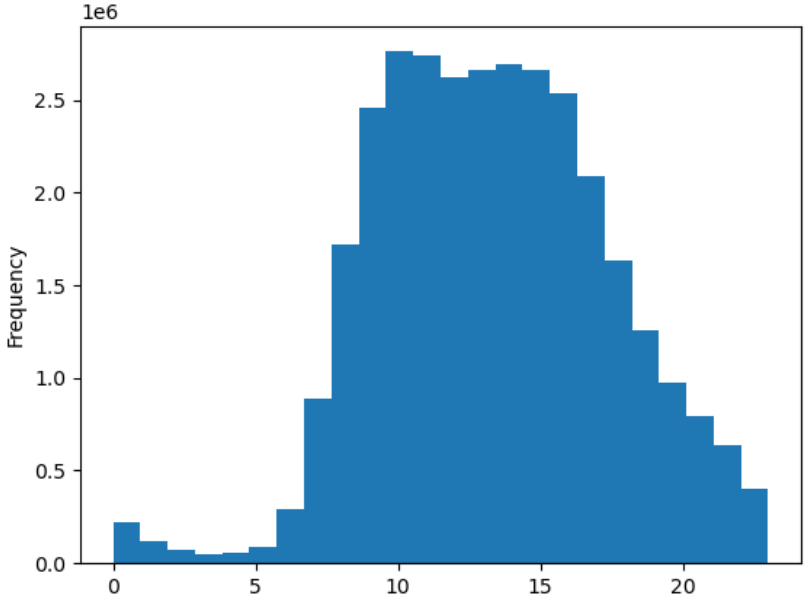
Objectives

- Uncover more information about sales patterns.
- Perform an initial data and exploratory analysis of some of data in order to derive insights and suggest strategies.

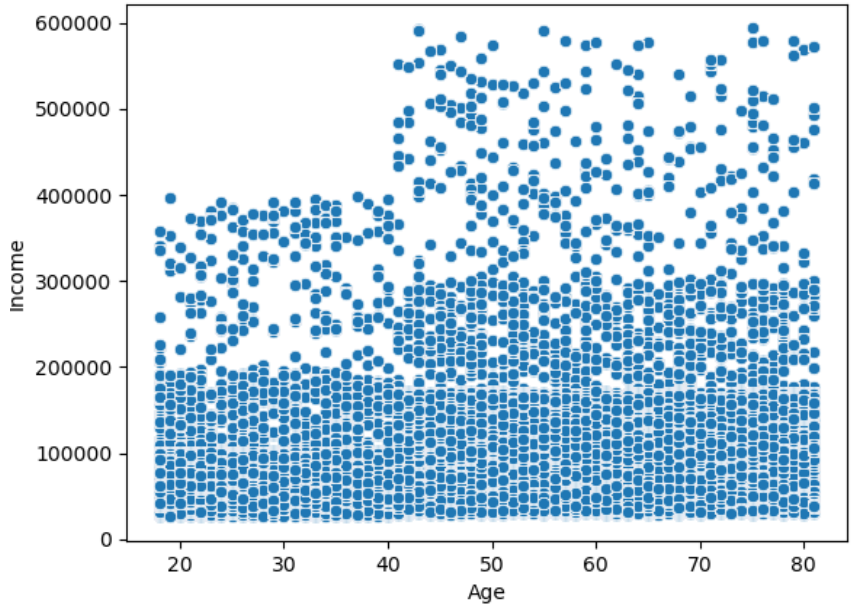
Highlighted Skills

- Data wrangling and subsetting
- Merging dataframes
- Grouping and aggregating data
- Data visualization with Python

Analysis (Histogram and Scatter Charts)

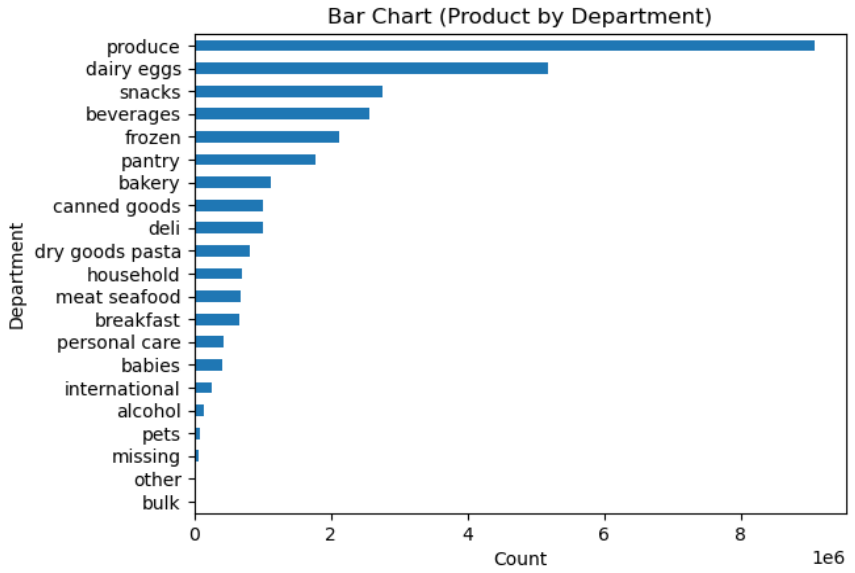


The busiest times of the day are between **9 a.m. and 6 p.m.**

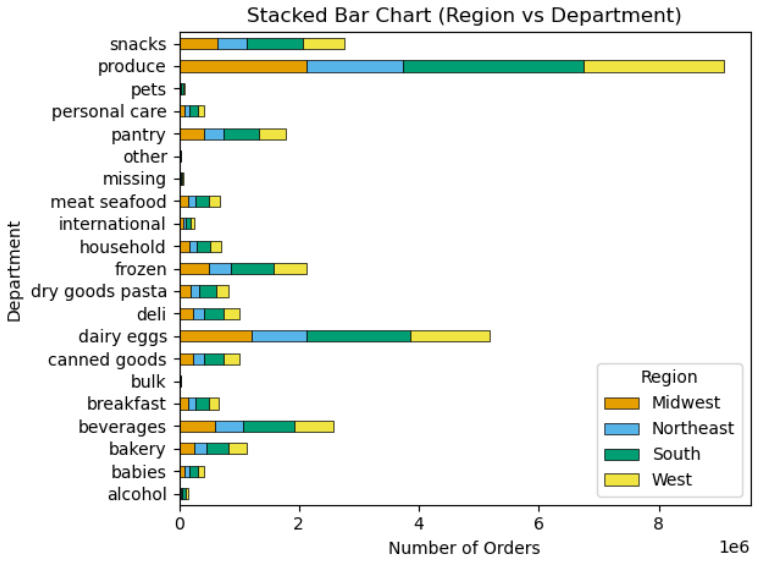


People **aged 40+** are more likely to earn **higher incomes**, especially between 40,000 and 60,000.

Analysis (Column and Stacked Column Charts)



Products from **produce** department leads in frequency, followed by **dairy**, **eggs**, and **snacks**.



Southern region **has the most potential** on purchasing items across all regions.

instacart

Recommendations

1. Launch promotions to upsell on weekdays, targeting **slow hours (7 p.m. to 5 a.m.)**.

2. **Loyal customers** are most likely to return. **A loyalty program** could encourage more shopping.

3. The **new customer group** contributes the least. We should focus on **promoting to them**.

4. **Married people** have greater potential. Instacart should **upsell family-related products**.

Click here to view the Python code and additional files

Tool Used

Microsoft Excel/Python/GitHub

PIG E. BANK

A Global Financial Bank

Pig E. Bank

Background

Pig E. Bank is a well-known global financial institution with an anti-money-laundering compliance department.

“They develop models to flag suspicious transactions indicative of money laundering and fraud”.

Objectives

- Conduct a comprehensive analysis of bank client data to identify the key factors influencing clients' decisions to leave the bank.

Highlighted Skills

- Exploratory data analysis
- Data mining process
- Data modeling
- Time series analysis
- Predictive analysis

Analysis (Descriptive Analysis and Pivot Table)

Loyal clients

	Credit Score	Age	Tenure	Balance	NumOfProducts	Estimated Salary
Min	411	18	0	$ -	1	$ 371.05
Max	850	82	10	$197,041.80	3	$ 199,661.50
Mean	652	38	5	$ 74,830.87	2	$ 98,942.45

Exited clients

	Credit Score	Age	Tenure	Balance	NumOfProducts	Estimated Salary
Min	376	22	0	$ -	1	$ 417.41
Max	850	69	10	$213,146.20	4	$ 199,725.39
Mean	639	45	5	$ 90,101.69	1	$ 96,676.39

Number of Product

Count of NumOfProduc Column Labels		
Row Labels	Loyal Clients	Exited Clients
1	46.76%	69.61%
2	52.60%	15.69%
3	0.64%	13.73%
4	0.00%	0.98%
Grand Total	100.00%	100.00%

Is Active Member?

Count of IsActiveM Column Labels		
Row Labels	Loyal Clients	Exited Clients
0	43.84%	70.10%
1	56.16%	29.90%
Grand Total	100.00%	100.00%

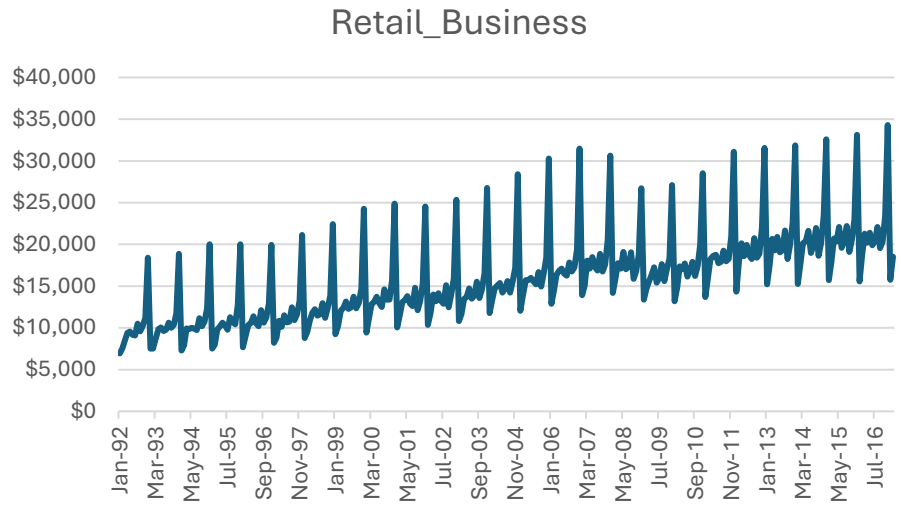
Use **descriptive analysis** (left) and **Pivot Tables** (right) in MS Excel to analyze the data

70.10% of exited clients were **not active members**.

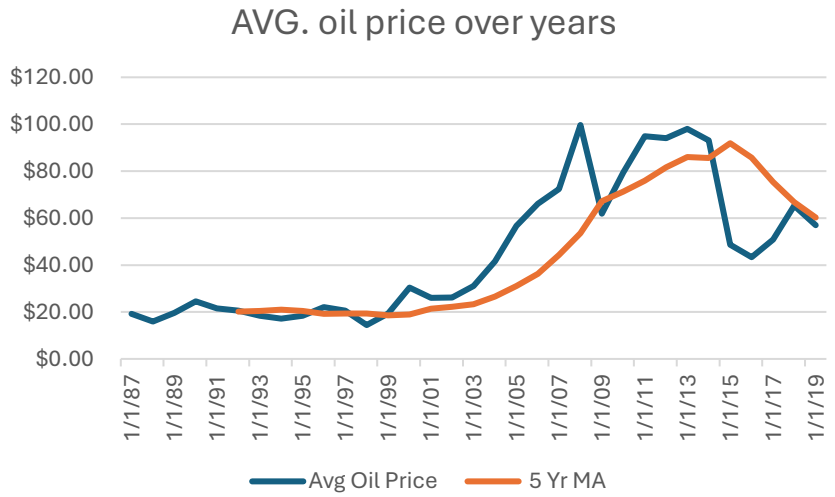
69.61% of exited clients had **only 1 product**.

59.31% of exited clients were **female**.

Analysis (Time-Series with Line Charts)

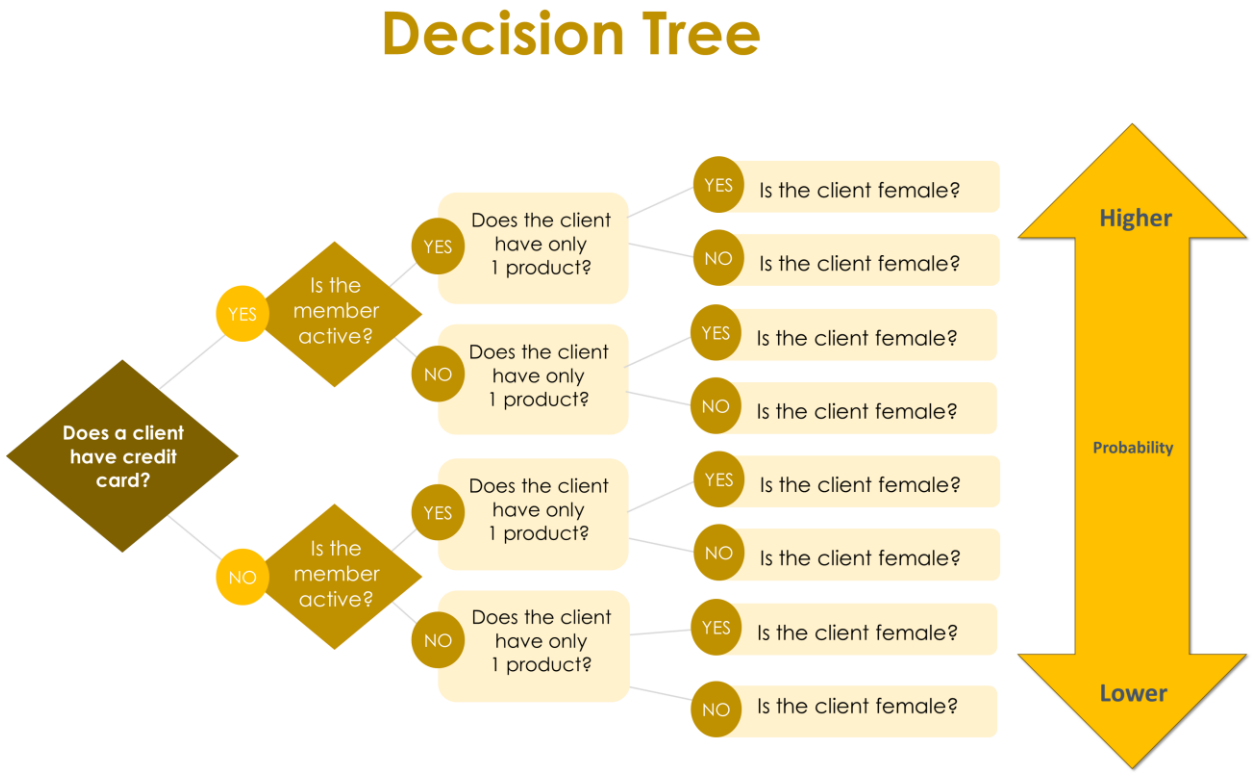


Seasonal pattern with revenue consistently spiking in **November and December.**



Non-stationary with fluctuation. After stability until 1998, it rose, dropped in 2008, and fluctuated until 2019.

Recommendations



The decision tree analyzes data and offers valuable recommendations.

1. Build a loyal base that uses more products and stays engaged.
2. Identify reasons for inactivity and potential product improvements.
3. Examine what competitors offer to understand their advantage over Pig E. Bank.

Tool Used

Microsoft Excel/Pivot Table/GitHub

THANK YOU

Connect with me:

Email

Tableau

LinkedIn

GitHub