

Data Analytics Portfolio

Bisola Oyeyipo

<https://github.com/Bisola-Oyeyipo>

oyeyipobisola@gmail.com



Professional Background

- Data Entry Officer / Institute of Human Virology, Nigeria.
- IT assistant(Internship) / Computer Professional of Nigeria.
- Bank Cashier / Intercontinental Bank, Nigeria.

Education

- Data Analytics - CareerFoundry(Ongoing)
- Masters in Information Engineering – University of applied sciences, Kiel
- Bachelors in Computer Science – University of Ilorin, Nigeria.

Tools and Techniques

- Python
- Sql
- Tableau
- Microsoft Office (Word, Excel, Powerpoint)
- Numpy

Skills

- Collaboration
- Project Management
- Teamwork
- Human Resource Management
- Effective communication

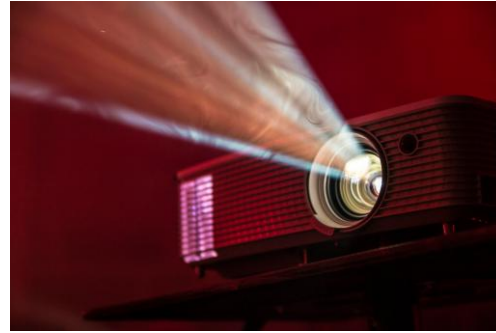
Portfolio Overview



GameCo



US Influenza



Rockbuster
Stealth



Instacart



Pig E. Bank



Chocolate Bar
Rating

Project Details

Project	GameCo	US Influenza	Rockbuster Stealth	Instacart	Pig E.Bank
Sector	Gaming	Public Health	Entertainment	E - Commerce	Banking
Relevance	Sales Planning	Medical Staff planning & allocation	Business Intelligence	Marketing Strategy	Customer Retention & Anti-money Laundering
Skills	Descriptive analysis; data grouping, summarising, visualisation.	Data cleaning, integration, transformation; hypothesis testing; visualisation	Querying, filtering, cleaning, summarising, joining tables, subqeuries, CTE	Data wrangling, merging, grouping, grouping; deriving variables	Predictive analysis, decision tree modelling
Tools	Excel, Powerpoint	Excel, Tableau	SQL, Tableau	Excel, python	Excel, Powerpoint
Links to Related Files	Powerpoint	Youtube	Powerpoint , Tableau , Sql	Github	Excel

Project Details

Project	Chocolate Bar Rating
Sector	Food
Relevance	Rating Analysis
Skills	Exploratory analysis; Linear regression, visualisation.
Tools	Python, Tableau
Links to Related Files	Github , Tableau



GameCo

- **Industry:** Gaming

Objective:

- Gain deep understanding of the current data.
- Use the insight for the next year 2017 marketing plan budget.
- Sales across regions over time.

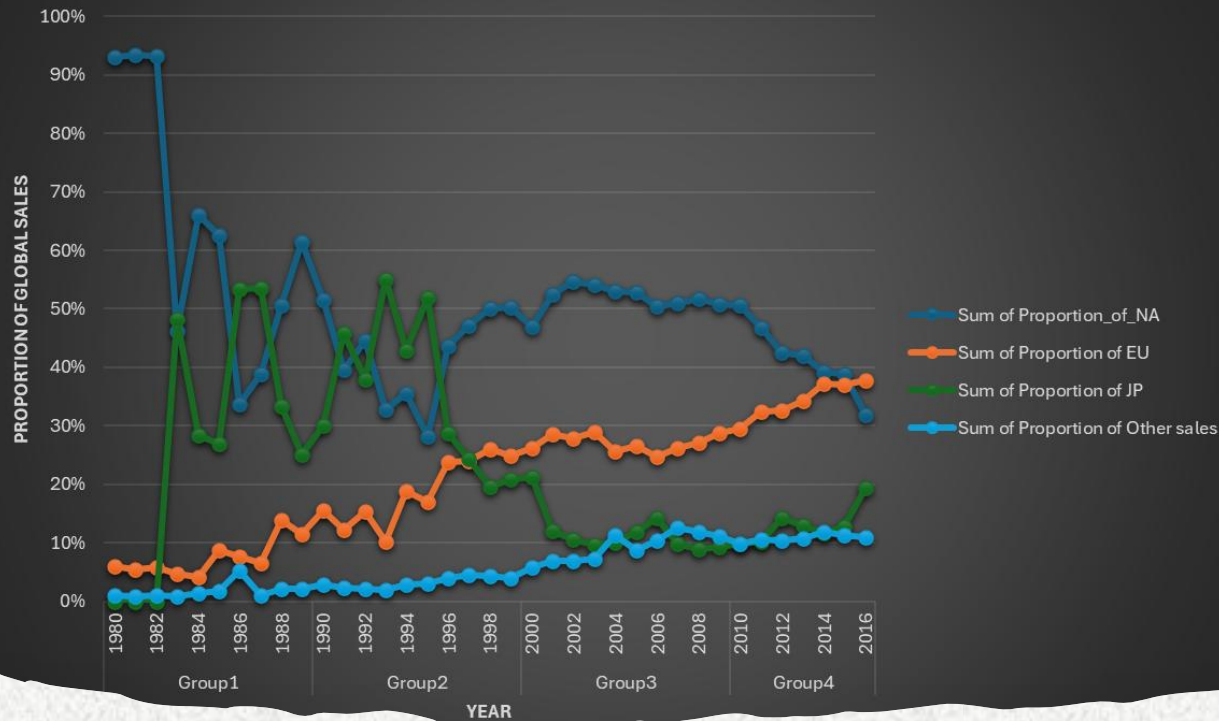
Current Understanding by GameCo: sales for the various geographic regions have stayed the same over time

Method: Descriptive analysis; data grouping, summarising, visualisation.

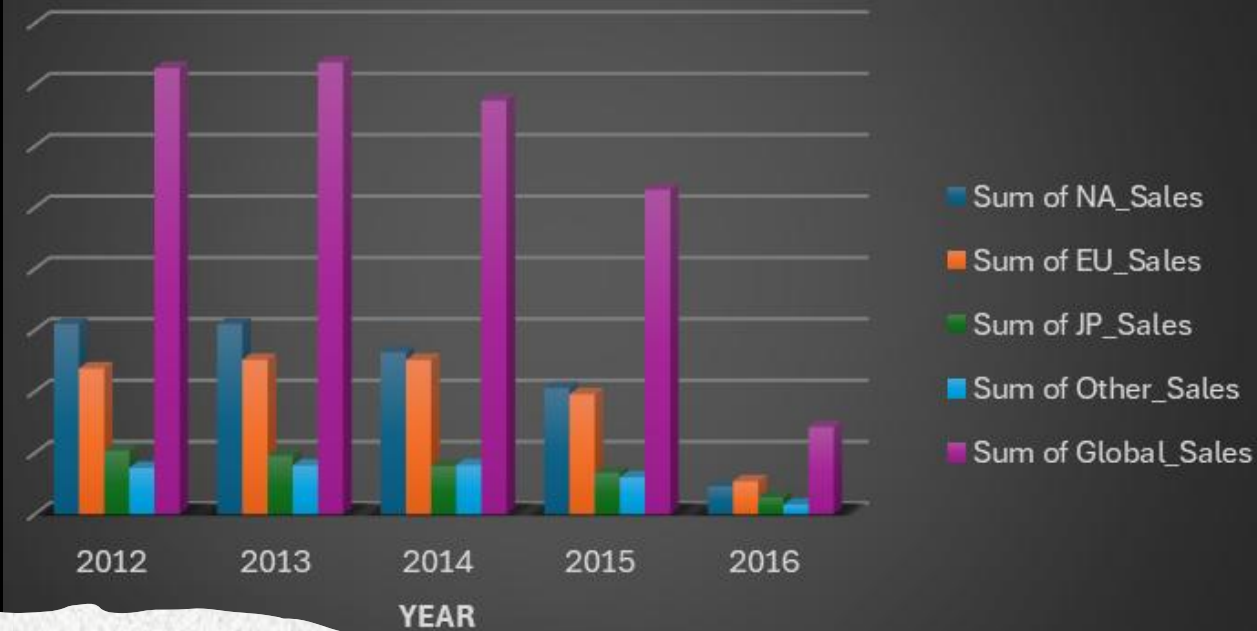
Dataset: GameCo Sale dataset that contains records from 1980 – 2016 and it includes 16,601 rows and 12 columns.



Proportion of Global sales performance from 1980-2016



Global sales in Last 5 Years

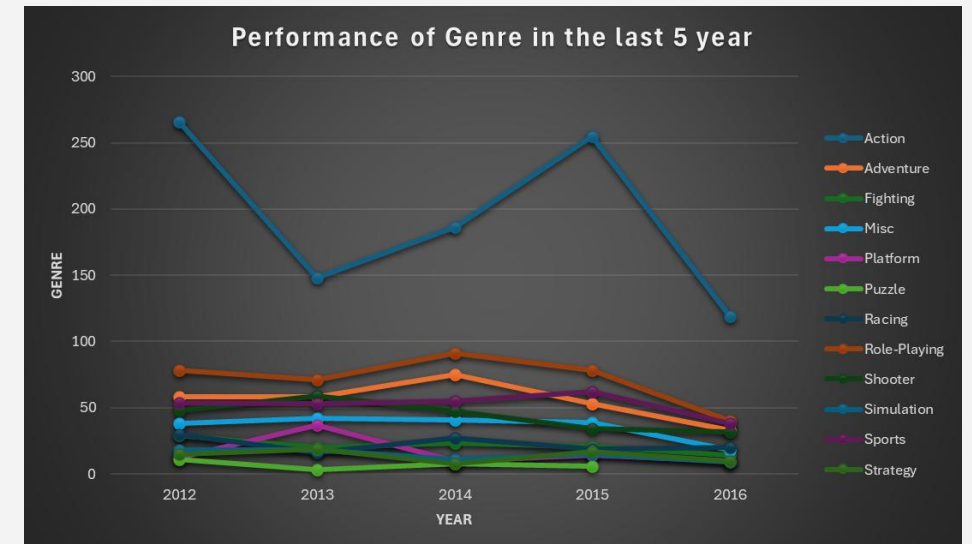
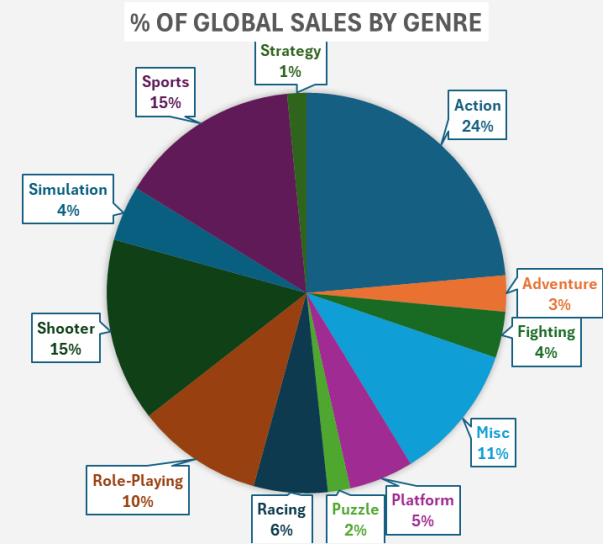


Regional Analysis of GameCo Sales

Global sales in the last 5 years have been decreasing and the sales was at the lowest point in the year 2016 in all the region.

Insight Analysis on Video Games Genre

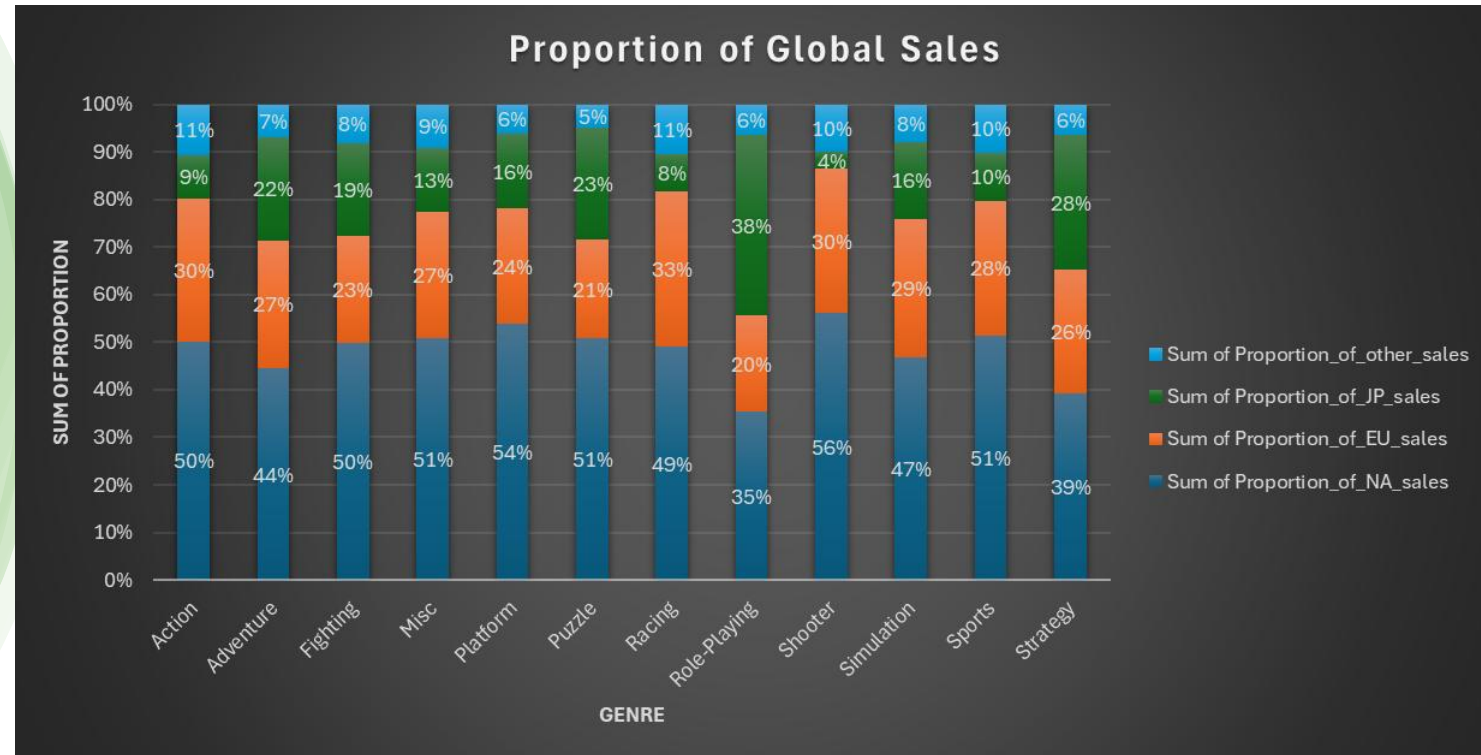
- The highest performing genre globally is the Action(24%), shooter(15%) and sport(15%) game in the last 10 years
- In the last five years, action games, role playing and sports is the top genre in video game performance for GameCo



Top Video Game Genres based on Regions.

The most preferred games in different region.

- North America region: Shooter(56%)
- European Region: Racing(33%)
- Japanese Region: Role Playing(38%)
- Other Region: Action & Racing(11%)



Recommendation

Marketing Actions Plan For GameCo

- Gameco should focus on the major market (NA, EU, JP) for year 2017 budget.
- Most preferred games in each region should be considered when planning the 2017 budget.
- Japan should be given priority in research on how to grow the market so that they do not lose it to other competitors.
- Continuous adjustment of market budget plan based on trends and regions.





US Influenza Season

Industry: Public Health

Objective: Determine when to send staff, and how many, to each state.

Data Source:

- CDC Dataset
- US Census Bureau Dataset
- CDC Flu vaccination Dataset

Tools Used:

Python, Tableau, Excel



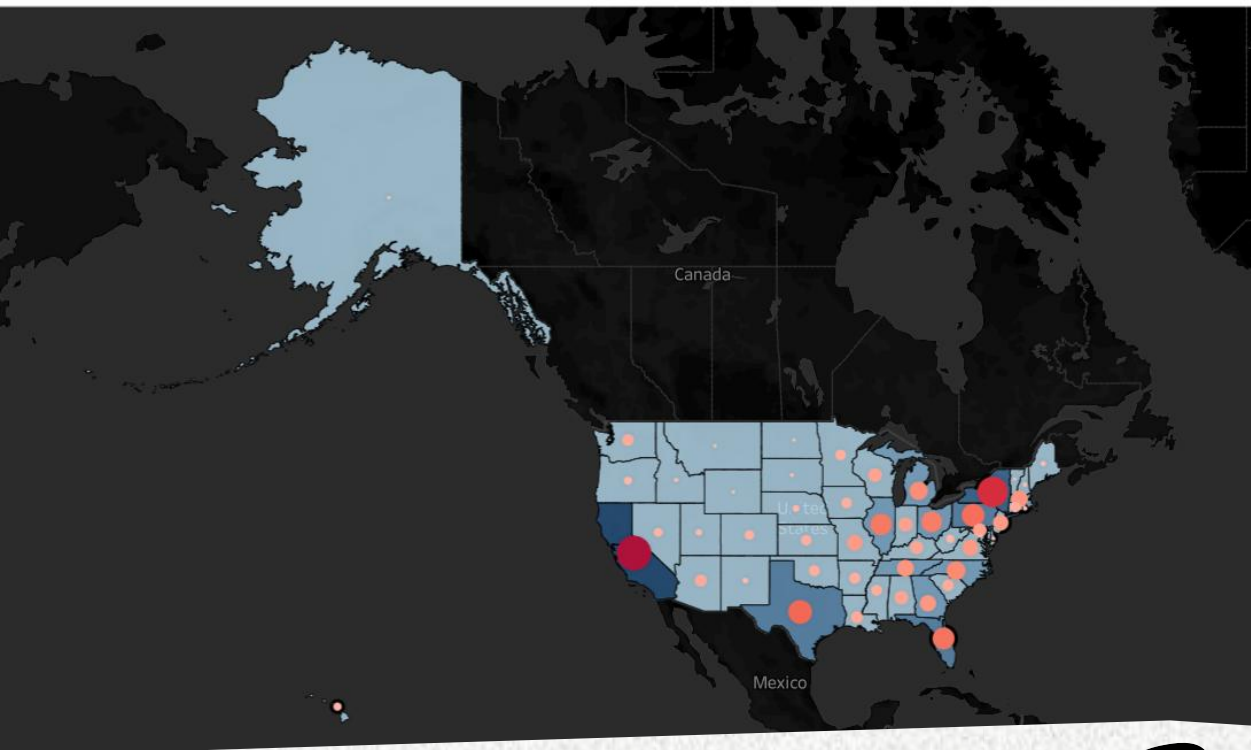
Key Problem

The older population are at higher risk of influenza flu.

Shortage of medical staff during influenza season.

Understaffing and Overstaffing across the states

Influenza deaths of 65+ years & 0-64 years in the US (2009-2017)



Influenza deaths of 65+ years & 0-64 years in the US (2009-2017)

California 47,483 6,519	Pennsylvania 20,596 1,738	North Carolina 13,941 1,656	Michigan 12,947 1,323		Georgia 10,242 1,498
	Florida 19,148 2,981	Virginia 10,004 601	Maryland 7,362 265		Alabama 7,058 687
New York 36,576 4,299	Illinois 18,019 1,985	Missouri 9,536 469	Arizona 5,415		
		New Jersey 9,481 472	Louisiana 5,412		
Texas 22,140 4,619	Ohio 16,866 2,176	Indiana 7,596		Iowa	
				Kansas	

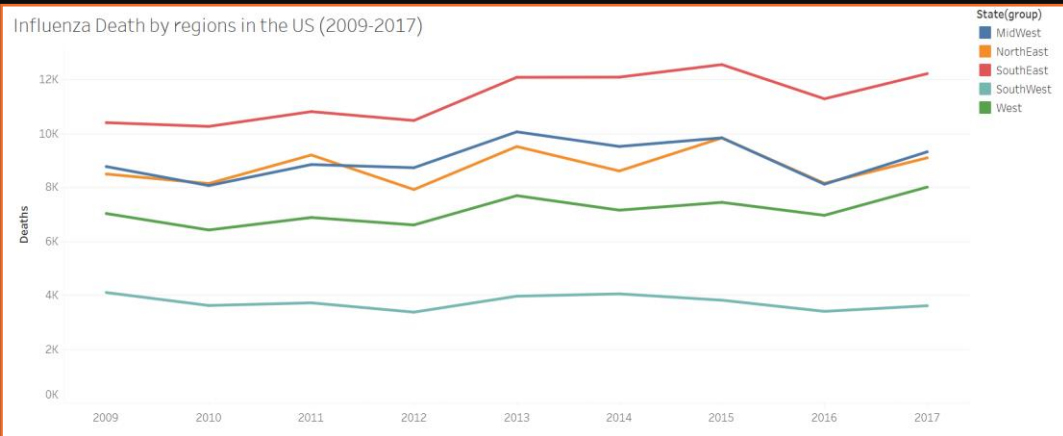
Medical Staffing need by state

The chart display the states with the rate of people affected with influenza flu based on age.

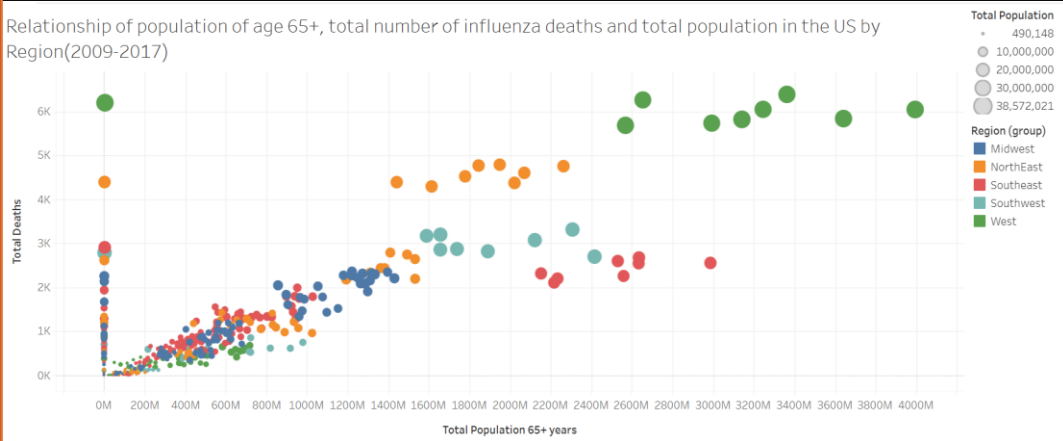
The older age group of 65+ are more affected with the flu and has a higher death rate.

The states with the highest death rate is California, New York and Texas.

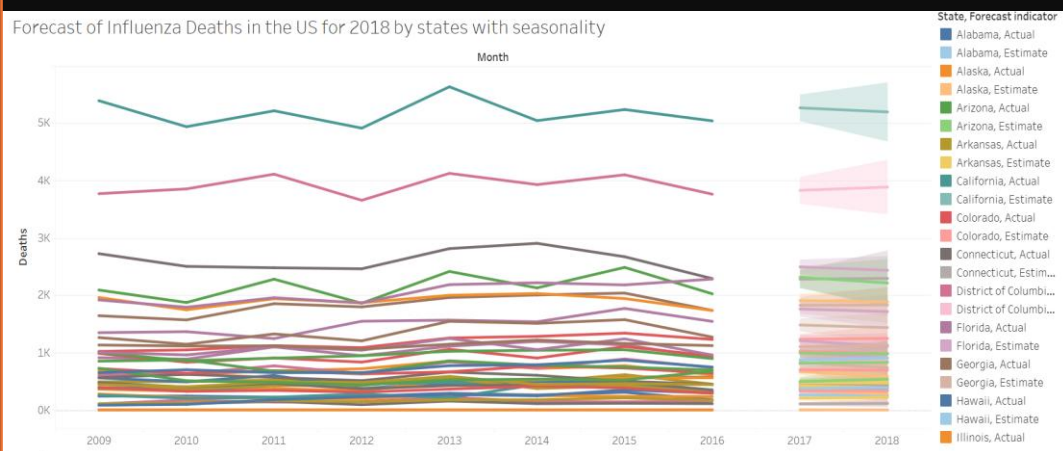
Influenza Death by regions in the US (2009-2017)



Relationship of population of age 65+, total number of influenza deaths and total population in the US by Region(2009-2017)



Forecast of Influenza Deaths in the US for 2018 by states with seasonality



Insight

The Southeast have the overall high rate of Influenza death

The west region of the US has the highest population of old people

The state with the highest demand for medical staff will be California, New York and Texas.

Project Conclusion

- ✓ The older population (age 65+) are more vulnerable to influenza flu.
- ✓ The higher the population of a state, the higher the risk of influenza death in the state. In this project analysis, state like California, New York and Texas are highly populated and they have the highest death rate as a result of influenza.

Recommendation

- ✓ The vulnerable population should be given priority in terms of treatment, quick accessibility to the healthcare centers,
- ✓ Vulnerable population should be encouraged to take influenza vaccine before the start of the season.
- ✓ At the peak of the season, more medical staff should be allocated to highly populated state.
- ✓ Continuous evaluation of medical staff in each state to avoid overstaffing or understaffing in a state.





Rockbuster Stealth Project

Rockbuster Stealth LLC is a movie rental company that used to have stores around the World and looking at the possibilities of operating an online video rental service.

Objectives:

- The most/least movie that contribute to revenue gain and sales figures between region?
- Average rental duration for all videos?
- Rockbuster customer location and customers with high life value.

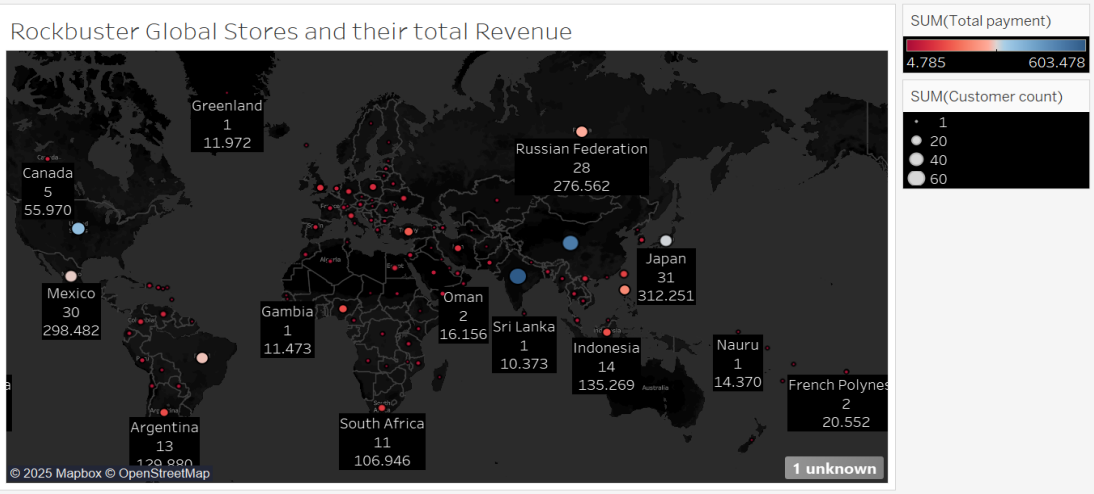
Method & Tools:

SQL, Tableau, Github

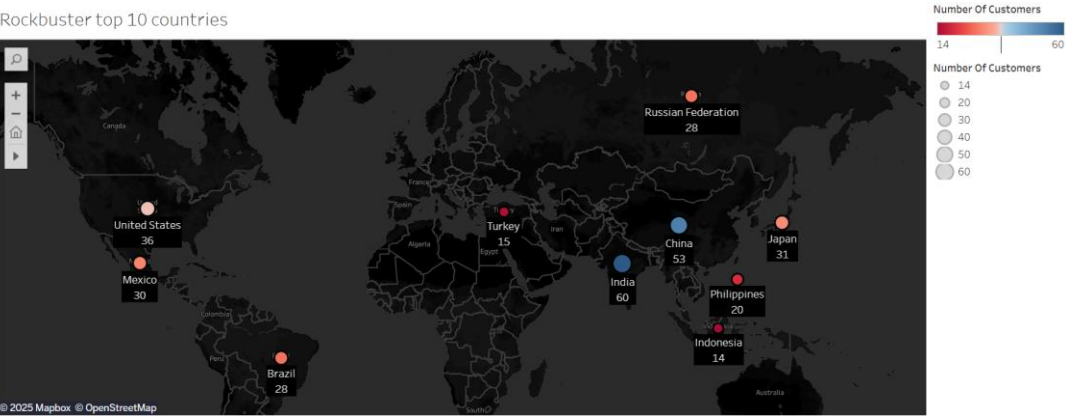


Analysis

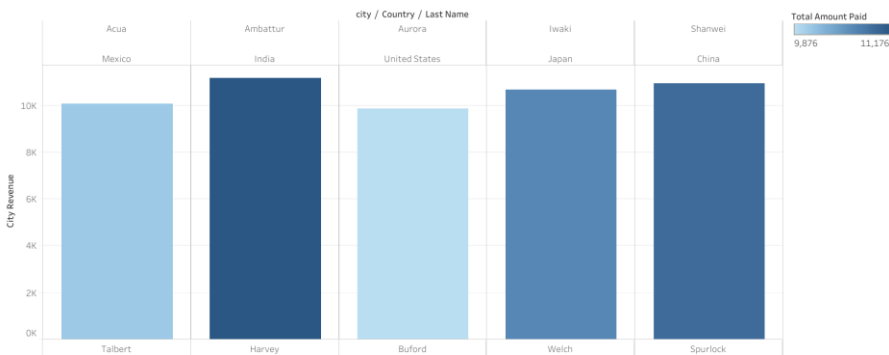
Understanding Rockbuster Global market & top customers



Rockbuster top 10 countries

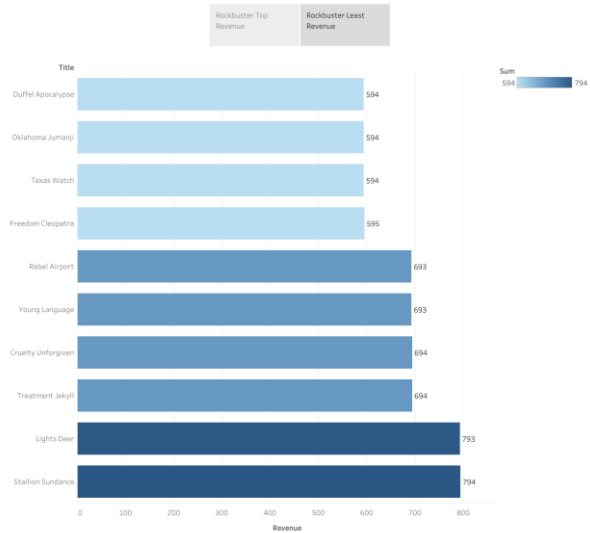


Top 5 customers from the city

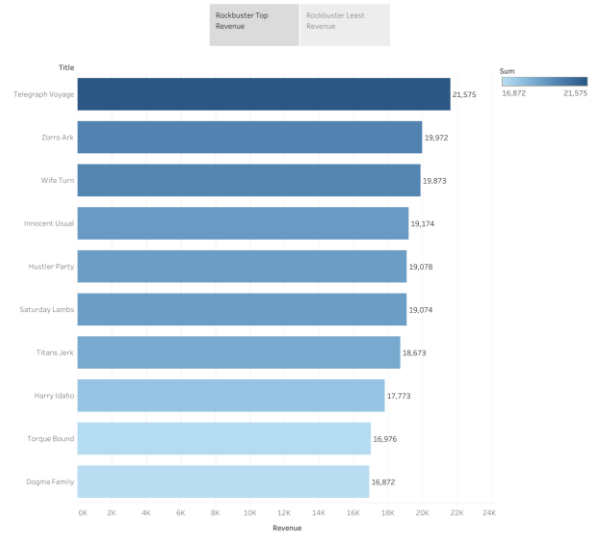


The Least & Top Movie Revenue and Rockbuster leading Market

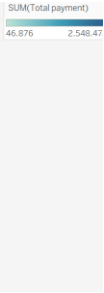
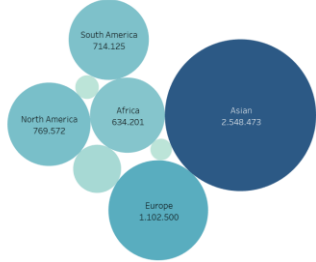
Rockbuster Movie Revenue



Rockbuster Movie Revenue



Rockbuster Sales value across different regions



Conclusion

Key Findings

- The top revenue is generated from multiple movies.
- The top ten countries with the highest customer count are india, china, US, Mexico, Japan, turkey, phillipines, Brazil, Indonesia and Russia.
- The average rental duration for movies is five days.
- The top 5 customers are from the following cities; Aurora, Acua, Ambattur, lwaki, shanwei. And three of the customers are from Asian region.
- The sales value varies across different region and Asian generate the highest sales.

Recommendation

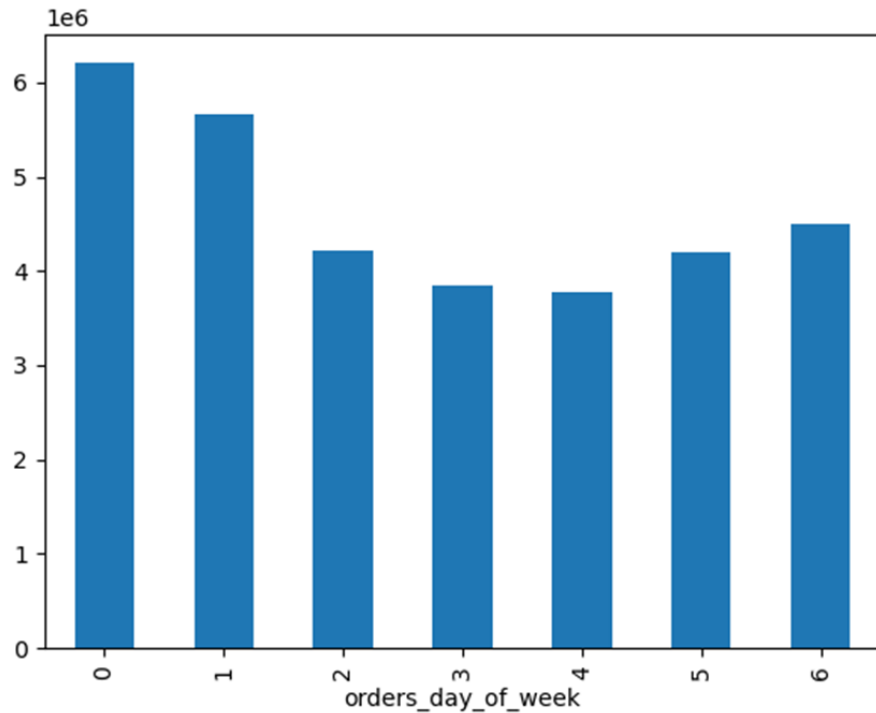
- Give more attention to the high-performing regions like Asian & Europe, to strengthen the current customer base and to gain new customers. These can help in the increase of revenue for Rockbuster.
- Further investigation show be done on movie rental data to understand the type of movies that is peculiar to a region. These will Rockbuster to understand the market more and also know the type of movies to make available.
- Create an environment for customer engagement especially the top cutomers to gain more insights on their preferences.



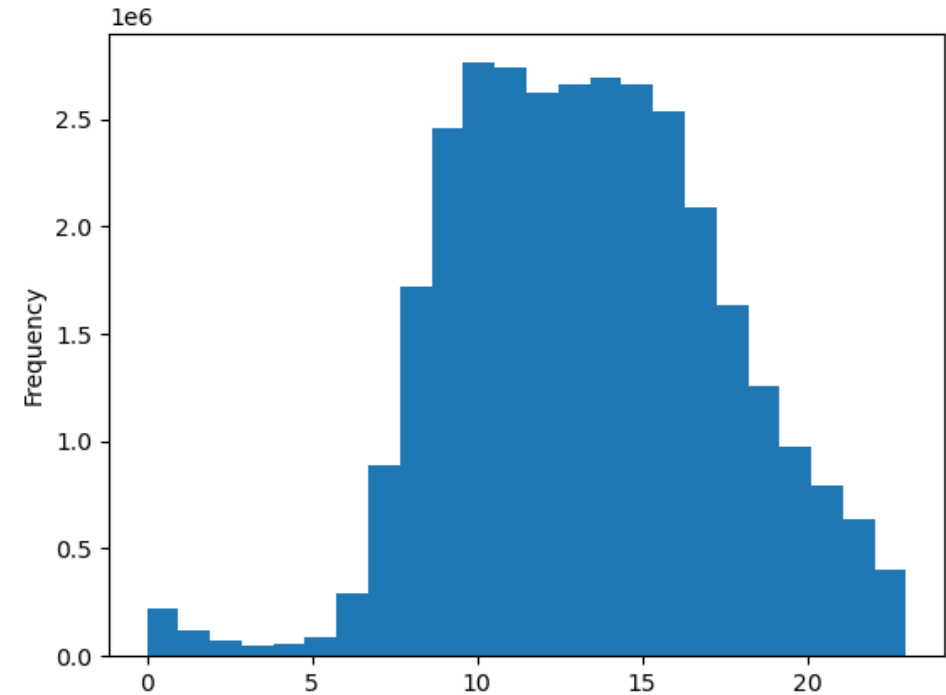
Instacart

- Instacart is an online grocery store that operates through an app. Instacart already has very good sales, but they want to uncover more information about their sales patterns.
- **Objective:** Develop a targeted marketing strategies with the focus on exploratory analysis, customers profiling and Customer segmentation.
- **Methods and Tools Used:**
 - Data wrangling, merging, grouping, grouping; deriving variables
 - Tools used are Python and Excel

Customer Order Time Strategy

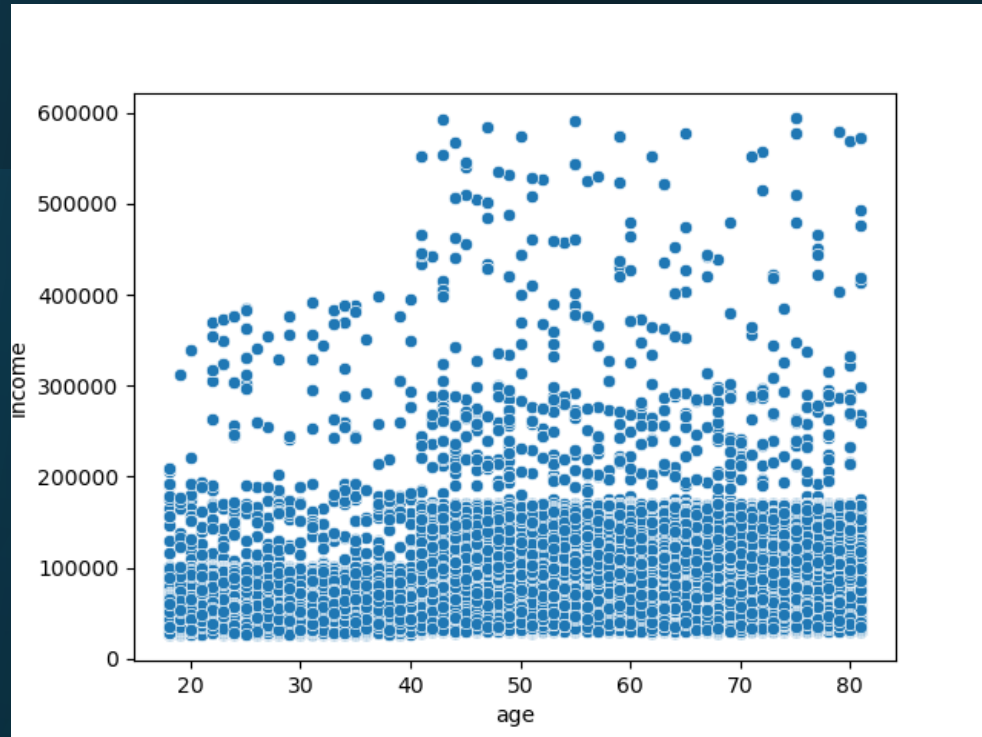


Saturday, Sunday and Friday is the peak shopping days for the week

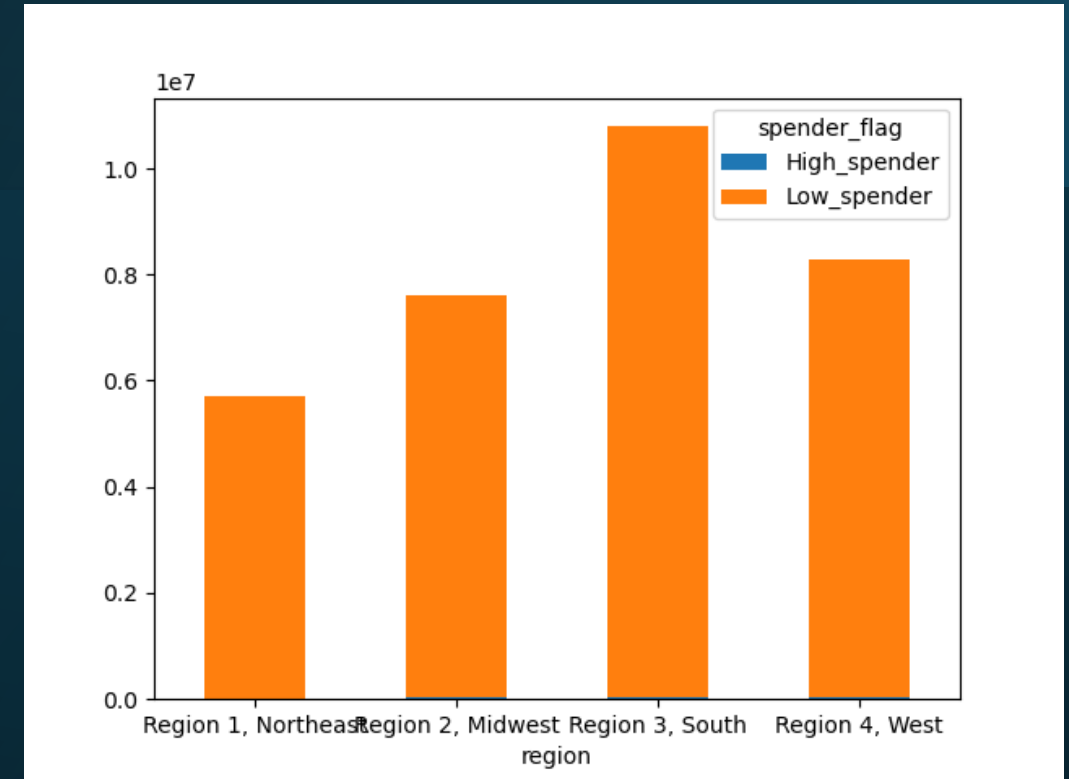


Peak hour for the day is between 10am and 11am.

Customer Profile



Most income earners falls below 200k for all age groups



Most Instacart customers are low spenders across all region and the south have more customers compare to the other regions.

Marketing Strategy

- **Target Audience**

- Focus on mid-income earners because that is Instacart biggest customers.
- Promotion should be aimed on appealing the middle class and low-income earners.

- **Timing Strategy:**

- Since the peak shopping day are between Friday and Sunday, preparations and provision should be made, so as to secure more sales.
- Special offers can be implemented during the week especially Tuesday & Wednesday when sales id the lowest in the week to engage more customers and increase sales for other days of the week.
- The peak shopping hour is between 9am and 11am. With this insight, Instacart can make their working system more efficient which will lead to customer satisfaction.





Pig E. Bank

Pig E. Bank is a fictional global financial institution aiming to strengthen its Anti-Money Laundering risk assessment and fraud detection processes to prevent financial crime.

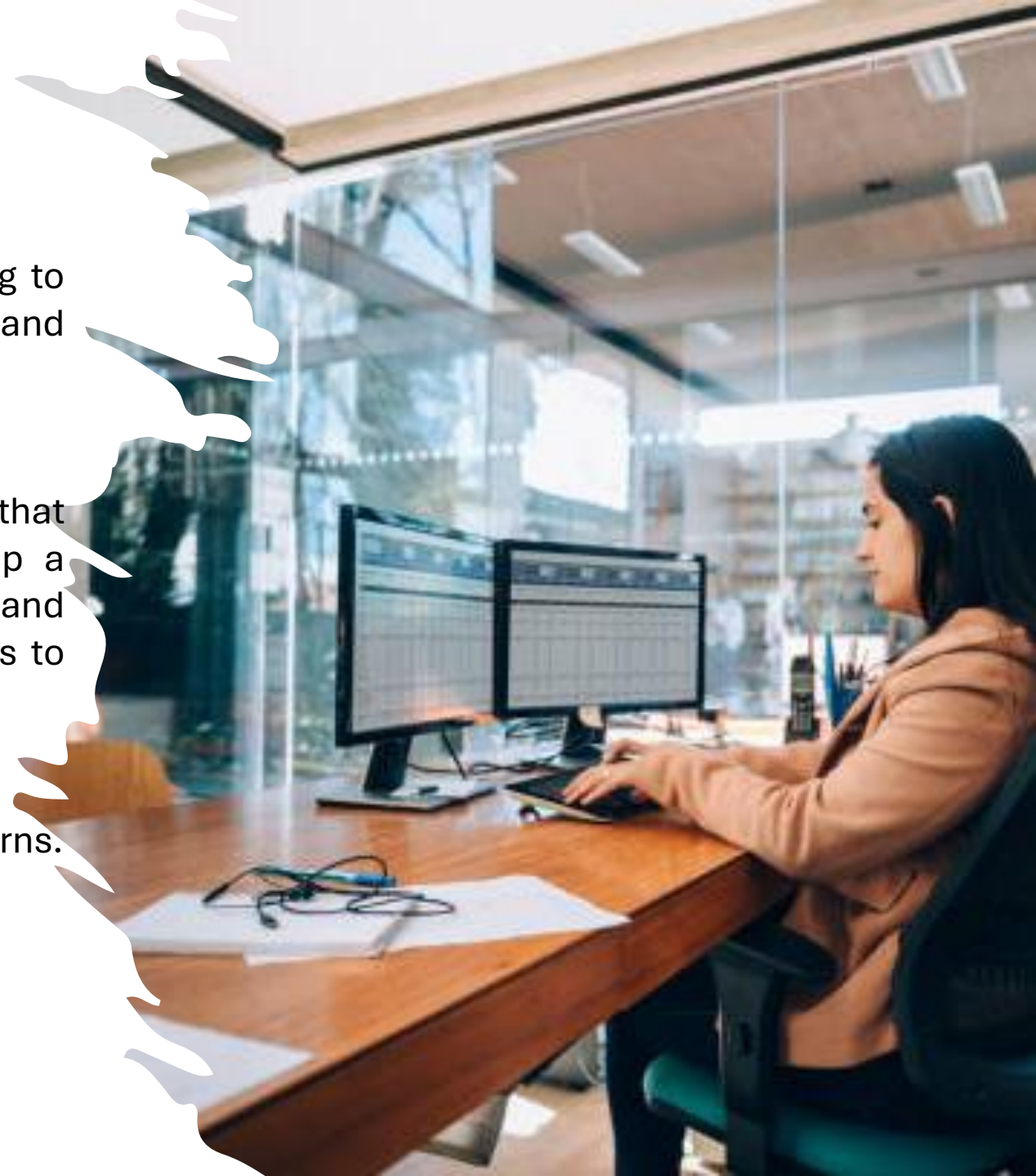
Objective:

Identify high-risk transactions and customer behaviors that indicate potential money laundering activities. • Develop a predictive risk assessment model using regression and classification techniques. • Propose strategic interventions to enhance fraud detection and AML compliance

Methods:

- ✓ Data wrangling & exploratory analysis to detect risk patterns.
- ✓ Regression modelling to quantify fraud likelihood.
- ✓ Classification techniques to segment risky customers.
- ✓ Predictive analytics for real-time fraud prevention

Data: Provided through Pig E. Bank client dataset.



Risk Factor that determined if the Client exit the Bank

- Client activity: Inactive Clients
- Age: Client between the age of 41 – 50.
- Gender: If client gender is female.

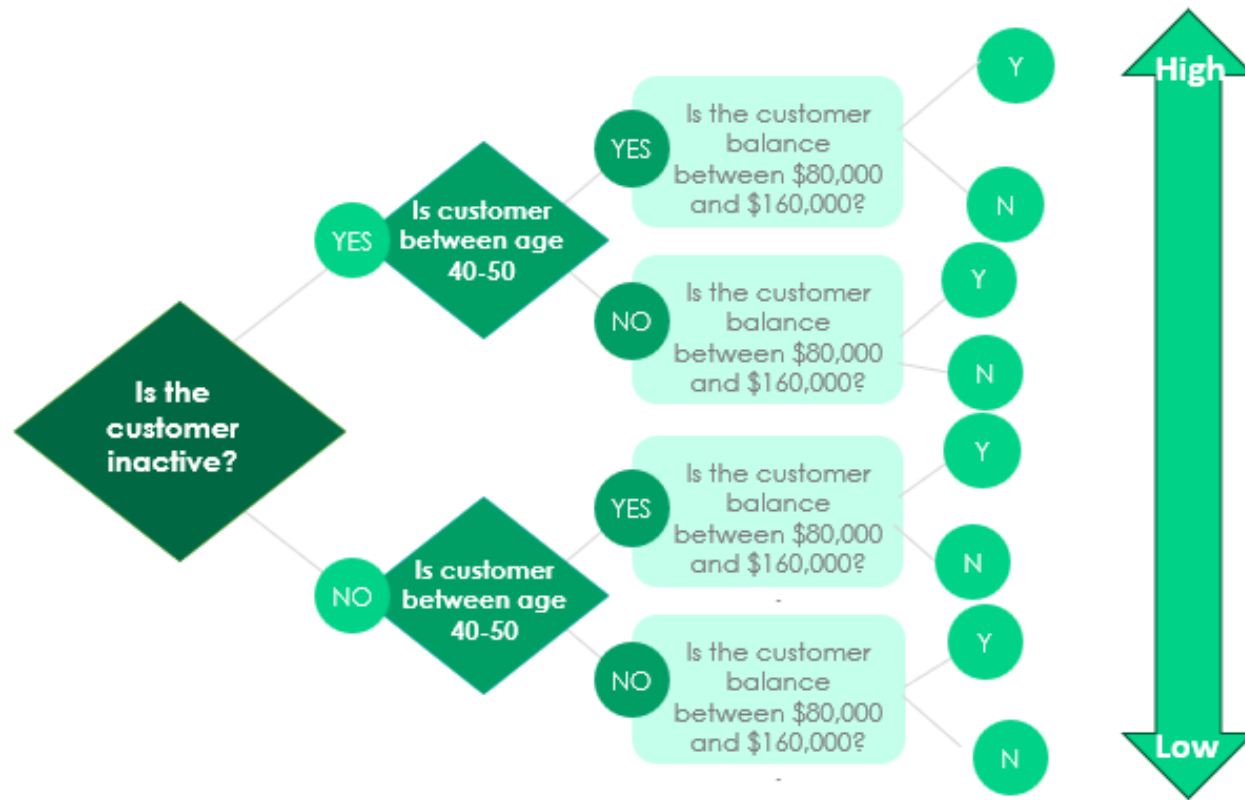
Row Labels	Count of Gender	
Female	121	59,31%
France	51	25,00%
Germany	39	19,12%
Spain	31	15,20%
Male	83	40,69%
France	26	12,75%
Germany	36	17,65%
Spain	21	10,29%
Grand Total	204	100,00%

Row Labels	Count of Age	
21 - 30	12	5,88%
31 - 40	56	27,45%
41 - 50	82	40,20%
51 - 60	38	18,63%
61 - 70	16	7,84%
Grand Total	204	100,00%

Row Labels	Count of IsActiveMember	
0	143	70,10%
1	61	29,90%
Grand Total	204	100,00%

Decision Tree

Factors determining the likelihood of a client leaving Pig E. Bank



The decision tree is an effective tool that can help determine the likely client to leave the bank.

Customer growth & retention Plan

Target mid age client between the age of 41-50 like tailoring different programs to attract this age group to stay like creating a retirement plan, client loyalty incentives etc



Targeted solutions for female client.



Enhancing AML monitoring and controlling.



Chocolate Bar Rating Analysis

Chocolate is one of the most popular candies in the world. The chocolate bar rating was done by experts in the industry and each chocolate is evaluated from a combination of both objective qualities and subjective interpretation.

Objective:

Understand the chocolate bar rating across the world
Gain insight on the factors that give a chocolate bar a high rating .

Methods:

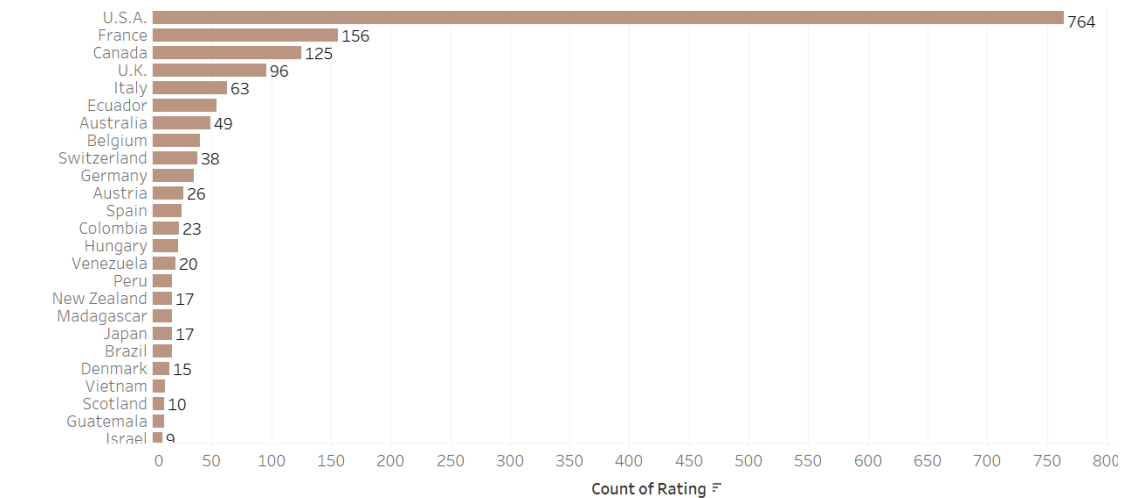
Data Cleaning and transformation
Hypothesis Testing
Exploratory Analysis
Regression Model
Visualization

Dataset & Tools:

Chocolate Bar Ratings from [Kaggle](#)
The tools used are Python and Tableau.

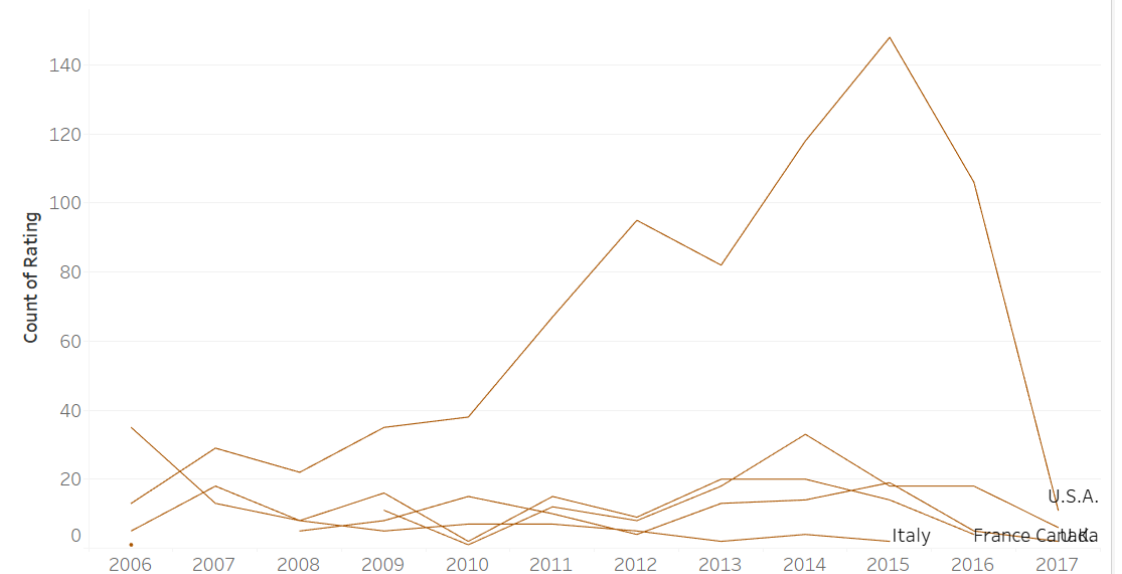
Top Manufacturers of Chocolate Bar

Top Countries with the highest number of Chocolate Bar (2006-2017)



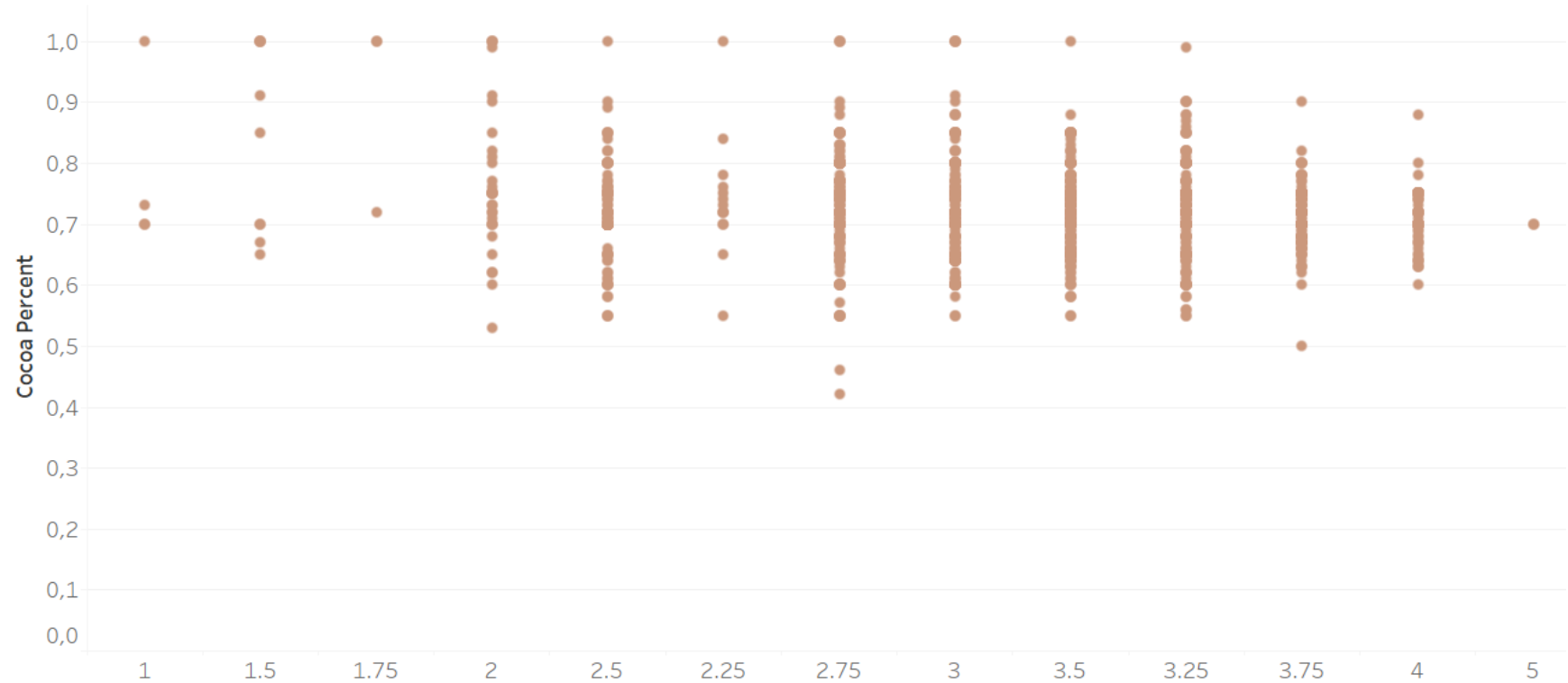
The top manufacturers of the chocolate bar are from USA, France, Canada, UK and Italy.

Rating over the years for the top Manufacturer.



Comparing the top 5 countries that has the highest production of chocolate bar. The USA came top which could mean more chocolate bar where review from the USA because of it popularity compared to other countries.

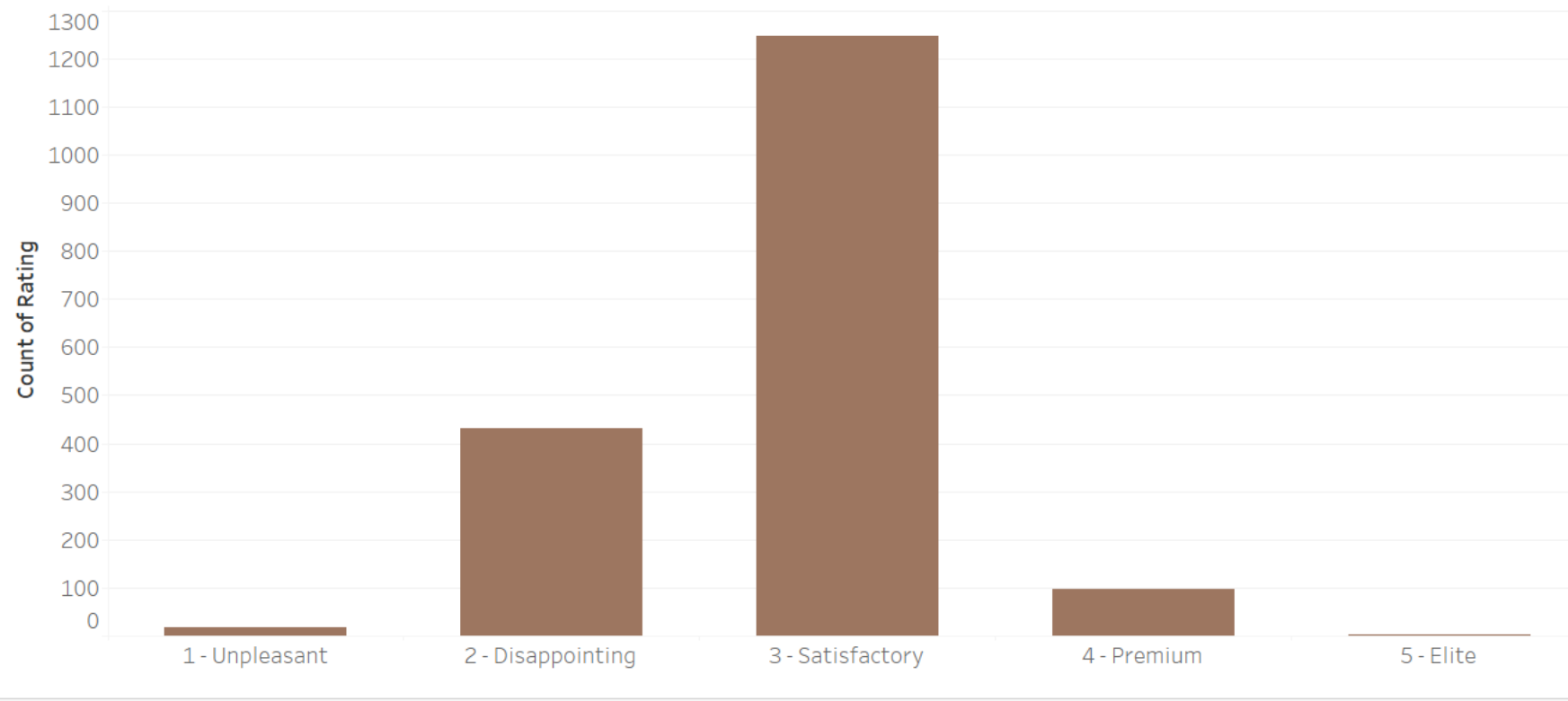
Relationship between the Chocolate bar rating and Cocoa Percent in the Bar



Hypothesis: The higher the chocolate bar rating, the higher the percentage of the cocoa in the chocolate bar.

Observation: The percentage of cocoa in the chocolate bar did affect the rating of the bar. It is observed that most bar have between 60% to 80% percent of cocoa in it and it cuts across all the ratings but more densed from rating of 2 to 4.

Top Rating by Chocolate Experts



Most of the Chocolate bar was rated “3-Satisfactory” by the experts. For the dataset to gain more insight, more observation would be needed like customers review on the chocolate bar to get a broader view of the best chocolate bar.

Conclusion

Key Finding:

1. The rating of the chocolate bar by the expert was done across different countries across the world.
2. The top cocoa bean type used are Trinitaro, Criollo and Forastero. Most chocolate bar contains between 60% to 80% cocoa in it and the percent of cocoa in a chocolate bar is not a strong factor for the bar to get high ratings.
3. Top countries with the highest production of chocolate bar are USA, France, Canada, U.K, Italy. The Top manufacturing company are Soma, Bonnat, Fresco, Pralus and A.Morin.

Recommendation

1. A more robust dataset should be used for analysis to gain more insight for future purpose.
2. The USA market should be studied more to understand the reason why they have the highest manufacturing companies to understand the market.
3. A good strategy should be put in place to get more accurate data from the chocolate industry.
4. Customer ratings can be included in the dataset to also compare the ratings of the experts with the consumers for better insights.





Thank You

