



---

# ANDRES GONZALEZ

---

Data analyst | Portfolio

andd011@yahoo.com





# TABLE OF CONTENTS



1

## Video game popularity

Data analysis for **GameCo**

2

## Preparing for influenza season

Data analysis for a **medical staf agency** in the US

3

## Online video rental service

Data analysis for **Rockbuster**

4

## Grocery basket app

Data analysis for **Instacart**

5

## Risk factors that have contributed to churn

Data analysis for **PigE Bank**

6

## Credits

Thanks to SlideGo

# About me

Hi, I'm Andrés — a data analyst with a deep curiosity about how the world works. I love uncovering the “why” behind trends, behaviors, and decisions. My passion for logic, problem-solving, and storytelling naturally led me to data analysis, where I get to explore meaningful questions using tools like SQL, Python, and data visualization.

Before pivoting to data, I worked in translation, customer support, and teaching — always focusing on clear communication and understanding people's needs. I've been working remotely since 2014, and I'm now excited to bring my analytical mindset and cross-functional experience into new data-driven roles.



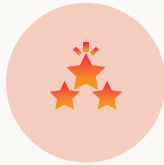
---

# VIDEO GAME POPULARITY

---

**Company: GameCo**

# SUMMARY



## Goal

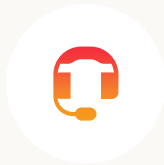
Perform a descriptive analysis to foster a better understanding of how GameCo's new games might fare in the market



## Datasets

Historical sales of video games (1980 - 2016).

**Some limitations:** No data from 2017 onwards / No info about video games downloads via mobile devices and apps



## Techniques

Descriptive analysis:

- Exploratory data analysis
- Measures of central tendency
- Shape of the data
- Spread of the data



## Tools



# THOUGHT PROCESS

## Recommendations

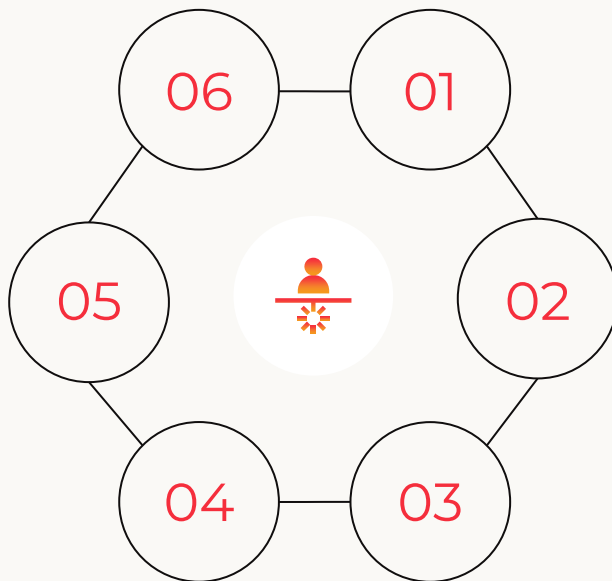
Suggestions about strategy, budget, and investment and/or further analysis

## Customers' preferences

This one is linked to sales – if sales have significantly changed over time, what customers' preferences might be related?

## Sales

Have sales significantly changed over time? If so, what factors can be related to such changes?



## Identify relevant elements

Competitors, most popular video games, sales status, customers' preferences and trends.

## Competitors

What companies have the biggest share? How big is it? Is the same company dominating in two or more regions?

## Games popularity

What video games do users prefer in each region?

# GAMES POPULARITY & COMPETITORS (2016)



## North America

Main competitor: EA | Market share: 16%

Most popular genre: Shooter | 33%



## Europe

Main competitor: EA | Market share: 26%

Most popular genre: Shooter | 29%



## Japan

Main competitor: Namco | Market share: 17%

Most popular genre: Action | 42%



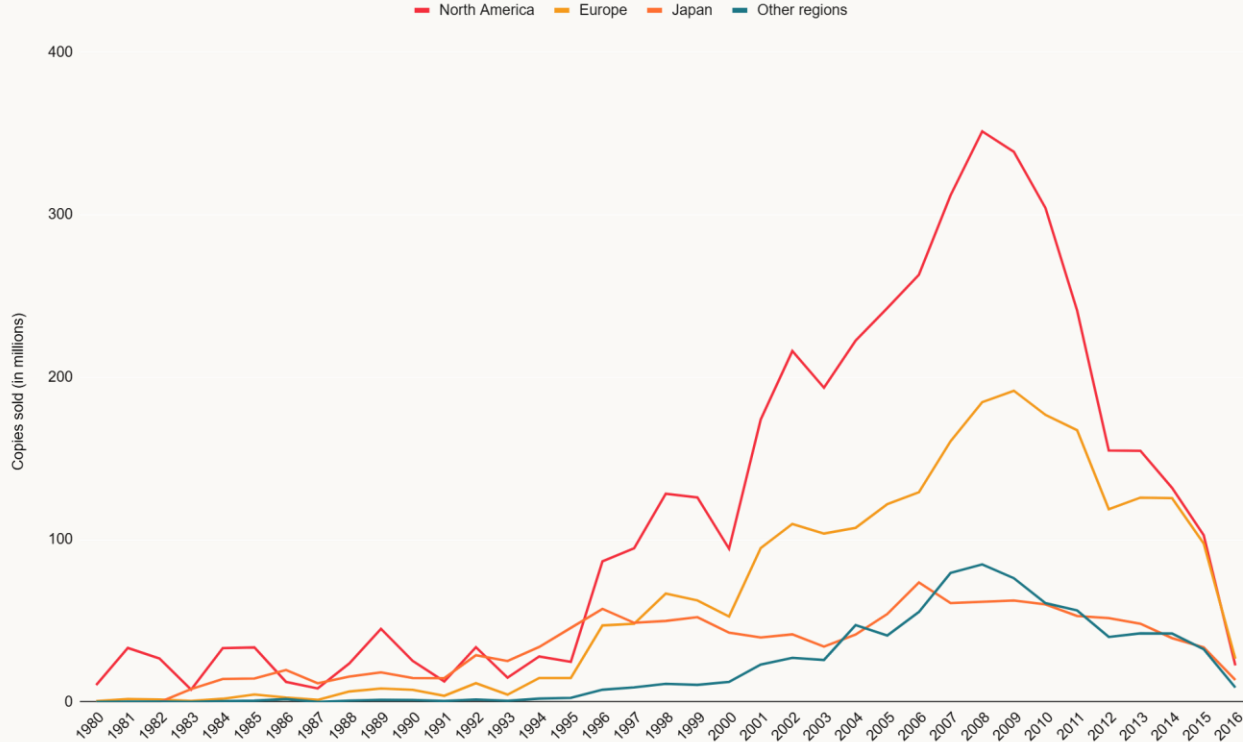
## Other regions

Main competitor: EA | Market share: 18%

Most popular genre: Shooter | 27%



# SALES & BUSINESS UNDERSTANDING



## Current understanding

GameCo's sales for the various geographic regions have stayed the same over time.

## Revised understanding

- Market demand has shifted over time.
- Possible reasons: changes in consumer preferences, like choosing mobile gaming (smartphones) instead of consoles.



# RECOMMENDATIONS



## Budget

Allocating budgets to shooter, action, and sports games. In Japan, besides action, the budget should go to role-playing and adventure.



## Strategy

Investigating what strategies Electronic Arts and Namco (JP) have implemented to outperform their competitors, and what the other publishers may have missed that led to their market share losses.



## Investment

Exploring investment opportunities in the mobile gaming (smartphones) sector.



# Preparing for Influenza Season

Medical Staff Agency in the US

# SUMMARY



Determine when  
and where to  
deploy medical  
staff, and how  
many

GOAL



Data sets covering  
influenza in the United  
States from 2009 to  
2017

**Some limitations:** no data on people's  
socioeconomic status, migration  
status, or number of people they  
share their home with.

DATASETS



Data integration,  
statistical analysis,  
hypothesis testing,  
visualizations and  
storytelling

TECHNIQUES



TOOLS

# THOUGHT PROCESS



## 1. RESOURCES

Given the limited amount of resources (medical staff), we have to optimize them. For that, we answer the following questions.

## 2. WHO?

Is there a population that is more vulnerable to die from the flu? To find the answer, we need to check the flu death rates from the previous years.

## 3. WHEN?

Use visualizations to identify when the flu peaks occur throughout the year.

## 4. WHERE?

Create a heat map to identify the states with higher density of vulnerable population.

## 5. FORECASTING

Use forecasting tools to anticipate the flu season for 2018.

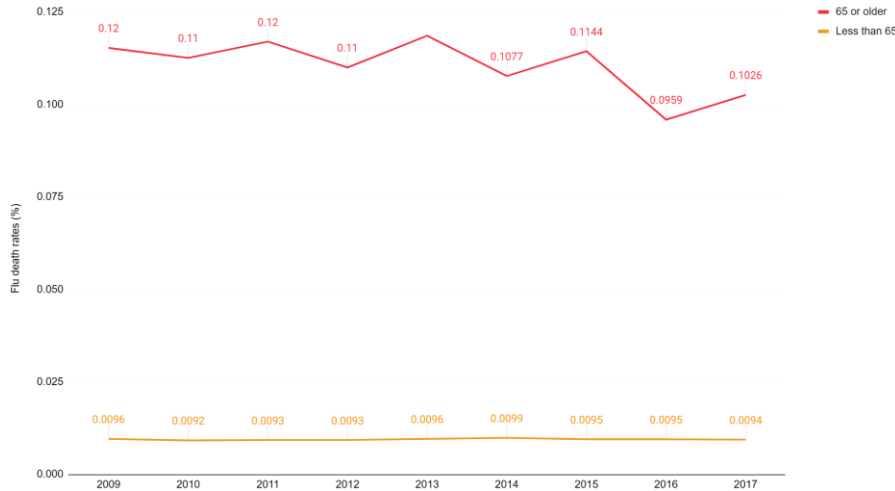
## 6. RECOMMENDATIONS

Create a formula for medical staff allocation prioritizing vulnerable population. Use the data to decide when and where to send the staff.

# VULNERABLE AGE GROUPS | WHEN IS THE FLU SEASON

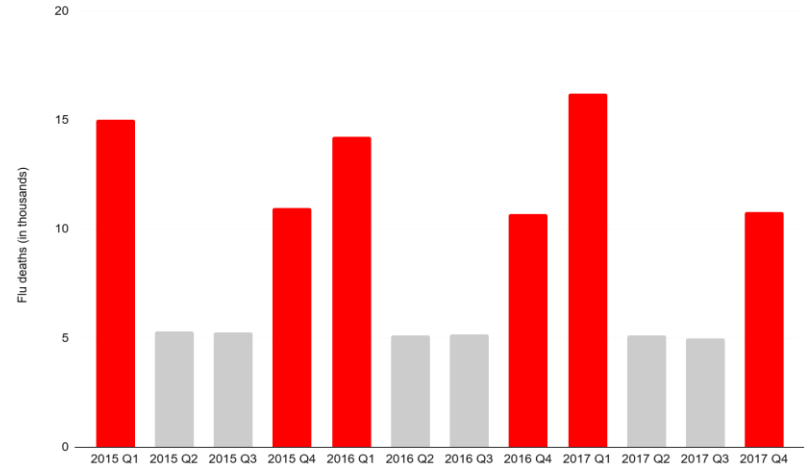
## Flu death rates

Flu death rates among people 65 and older have historically been higher, too



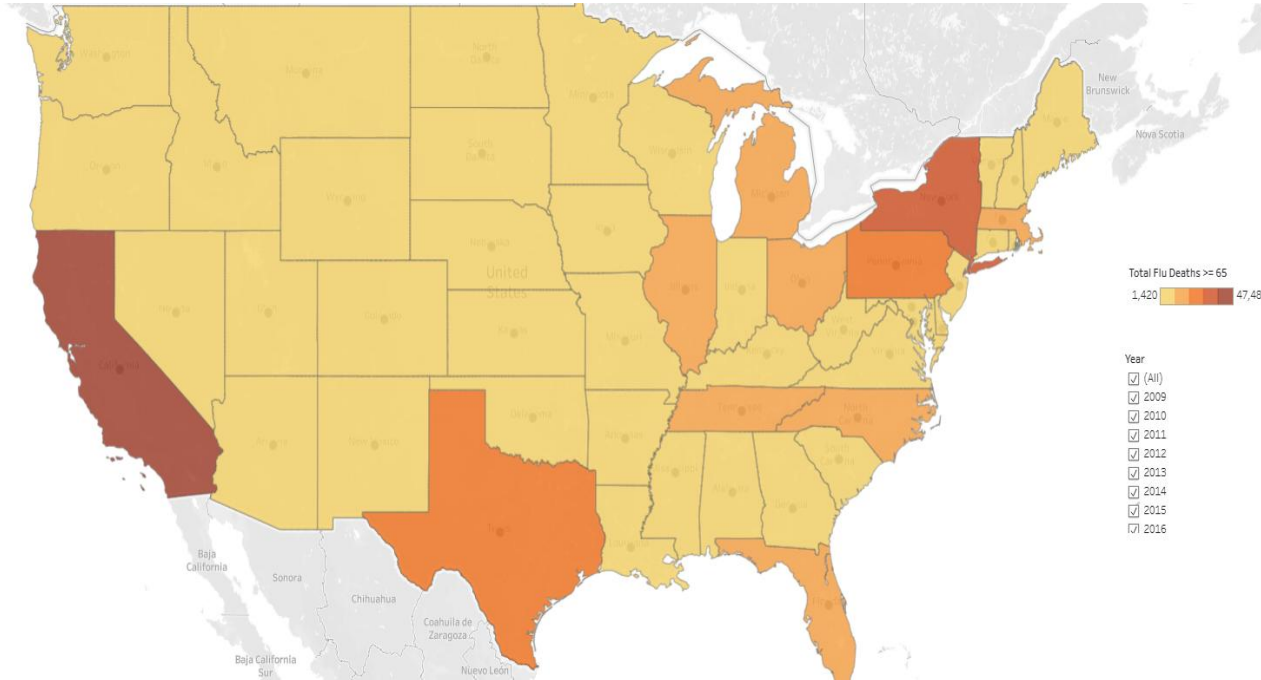
## Flu peaks

Flu death peaks have occurred in the first and last quarters of the year, affecting all regions (2009 - 2017)



\* While the pattern has been consistent from 2009 to 2017, this visualization shows data from 2015 to 2017 only.

# WHERE IS THE MOST VULNERABLE POPULATION?



2009 - 2017

Historically, these states have had the highest flu deaths among the vulnerable population:

- California
- Texas
- Florida
- New York
- Pennsylvania

# Final notes and recommendations

- ❖ The forecast for the next flu season anticipates very similar behaviors than the previous season per state.
- ❖ In the last flu season, the US lost 74,685 people. For the next flu season, we expect 74,200 deaths caused by the flu. That is 1% less than the previous season.
- ❖ The forecast for the next flu season as well as the details for the suggested allocation for the medical staff are available in the table (Annex) at the end of this presentation.
- ❖ I suggest **allocating medical staff based on the expected flu deaths**. Therefore, the total medical staff will be allocated according to the following formula:

$$(\text{Expected flu deaths per state} / \text{Expected flu deaths nationwide}) * 100\%$$

- ❖ Please check the **annex** for more details about both the forecast for 2018 and the medical staff allocation per state.
- ❖ When? The **first and last quarters** of the year will be flu peaks.





# Rockbuster

## Stealth Data Analysis Project

Company: Rockbuster Stealth LLC





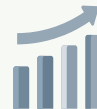
# SUMMARY



## Goal



Use the existing movie licenses to launch an online video rental service to stay competitive



## Datasets

Contains information about Rockbuster's film inventory, customers, payments and other relevant info

## Techniques



SQL, visualizations, storytelling

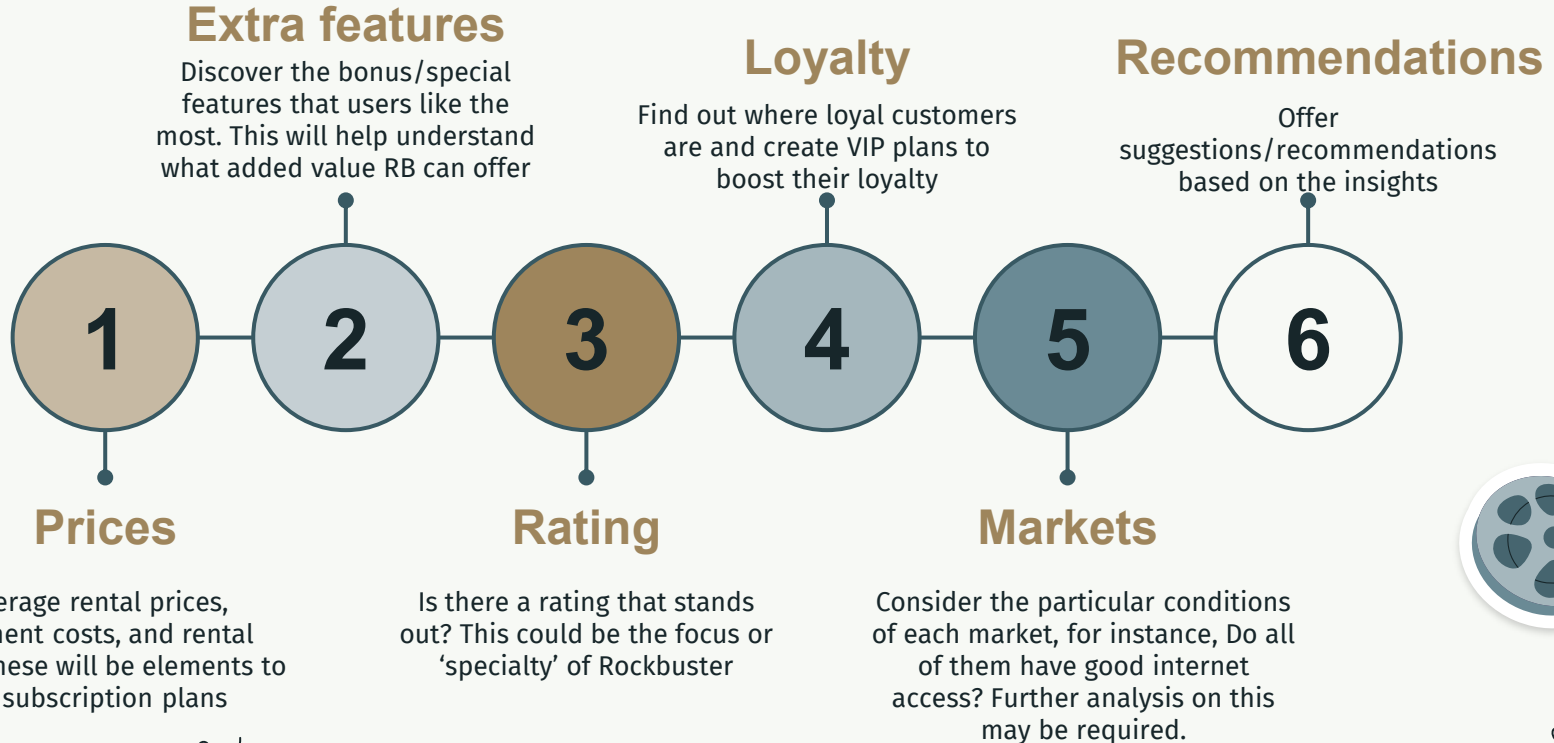
## Tools



Click on the SQL icon to check some of my queries!



# Thought process





4,95 days

Average **rental duration**



USD 2,98

Average **rental rate**



USD 19,98

Average **replacement cost**



PG-13

Most common **rating**



Trailers,  
Commentaries,  
"Behind the Scenes"

Most common **special features**



Academy  
Dinosaur

Movie with **most inventory**



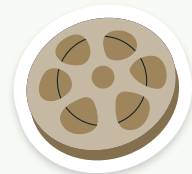
115,27 mins

Average **length** of  
Rockbuster's movies

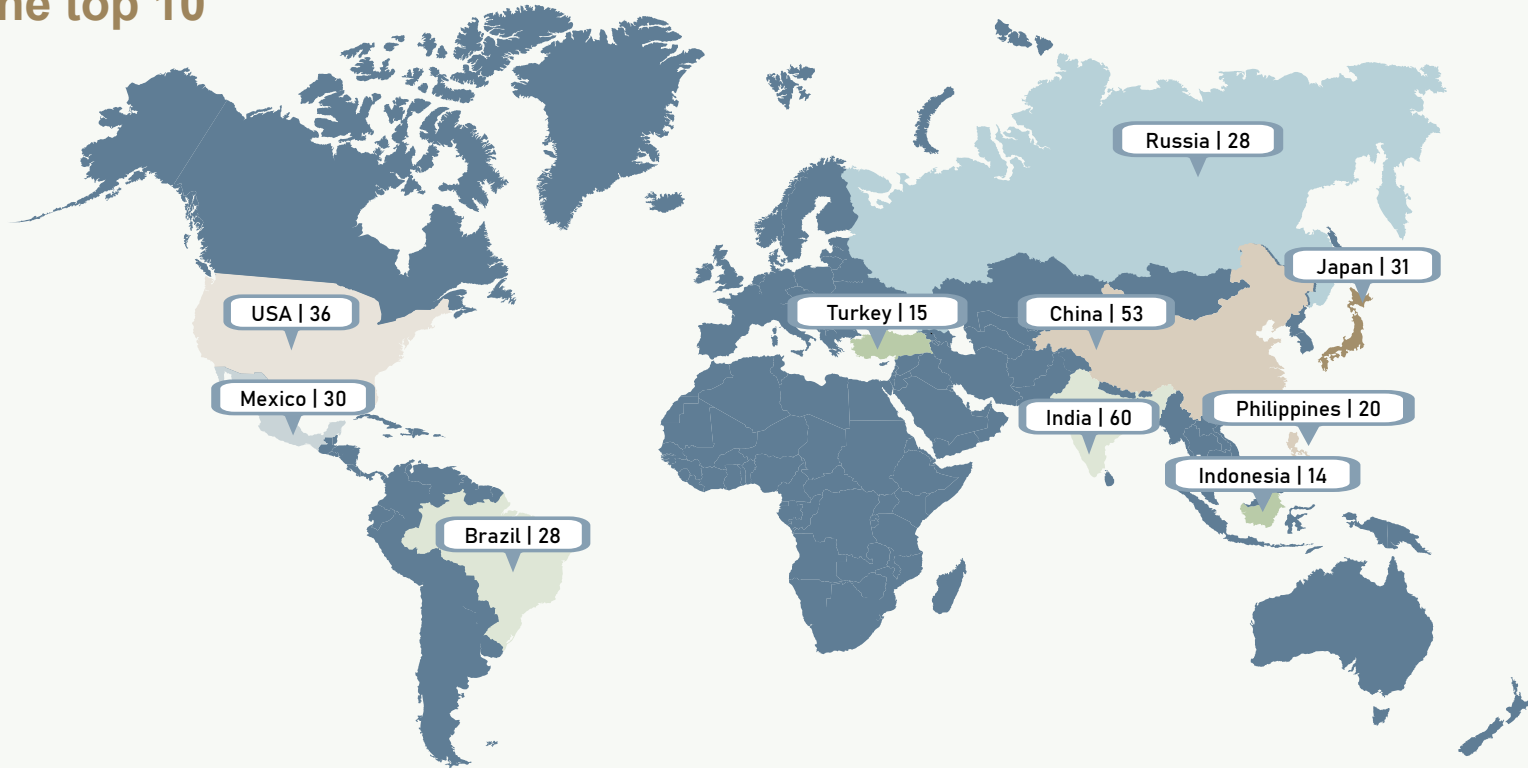


2006

Average **release year**



## Countries with the most customers. Asia leads with 6 countries in the top 10

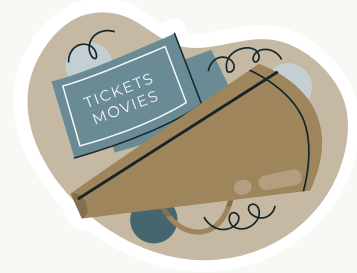


## Top 5 customers who have paid the most to RockBuster

Name	Total paid (USD)	City	Country
Karl Seal	208.58	Cape Coral	USA
Marion Snyder	194.61	Santa Barbara	Brazil
Marcia Dean	166.61	Tanza	Philippines
Louis Leone	156.66	Tanauan	Philippines
Tim Cary	154.66	Bijapur	India



# Recommendations



## Pilot tests

- Do not remove physical stores completely from countries where Internet service is known to be unstable, i.e., some Asian countries.
- Offer online streaming service as an alternative, making sure that it is mobile friendly. Then, evaluate results.
  - Marketing campaigns.



## Subscriptions

Adjust subscription prices per country based on the average cost paid per customer.



## Extra features

Trailers, Commentaries, and Behind the Scenes content can be positioned as exclusive bonuses to stand out from the competitors.



## Ratings

Expand the PG-13 film collection to attract family and teen viewers



## Loyal customers

- Offer early access to premium features or free subscription months as you migrate them into the new model.
- Create "Founders Club" or "Rockbuster Pioneers" badges.



## Asian market

- Develop regional content collections and localized marketing to cater to audiences in India, China, Philippines, and Japan.
  - Consider subtitles or dubbed versions if feasible.



# Grocery basket analysis

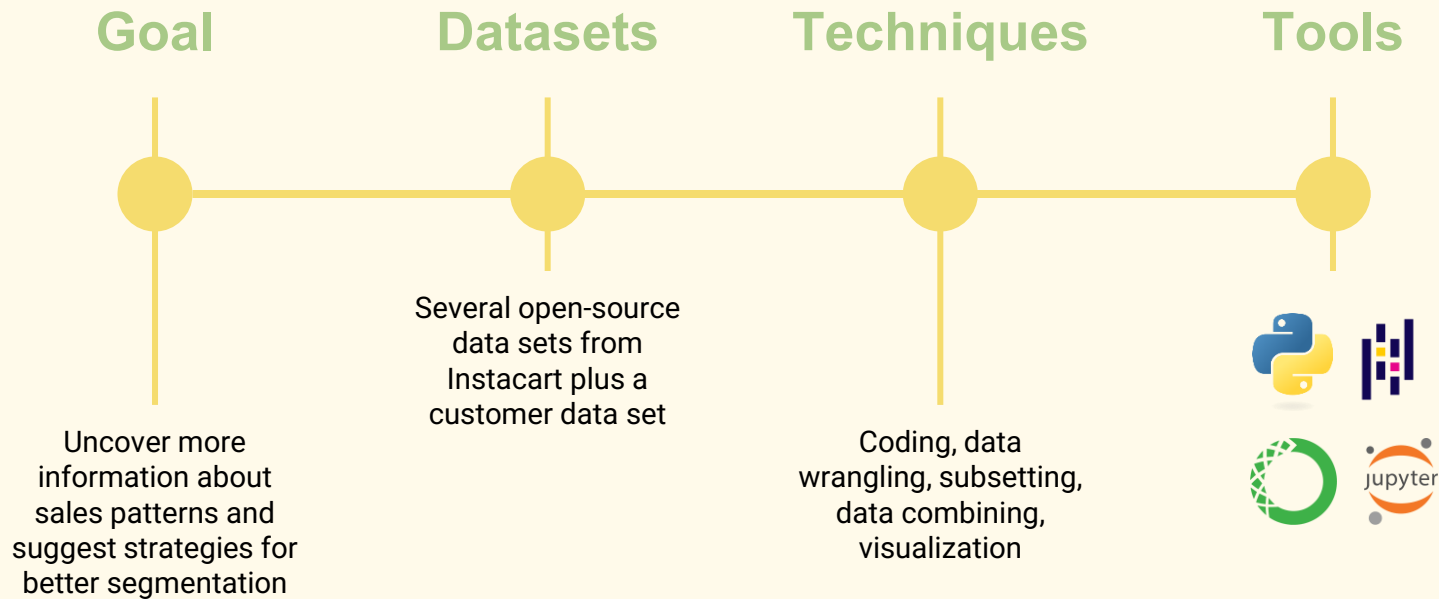
---

**Company: Instacart**



# Summary

---



Click on the jupyter icon to check my code lines and notebooks!



# Thought process

---



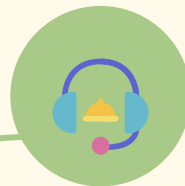
## Questions

This time, the questions also come from depts like the sales one. These are about busy days and hours, customers' loyalty, customers' income, average spending per income group, popularity of products, and possible relationships between income and spending.



## Insights

I used visualizations to better understand the answers to those questions



## Recommendations

Some insights led to recommendations; others can pave the way to further analysis

# Busy days

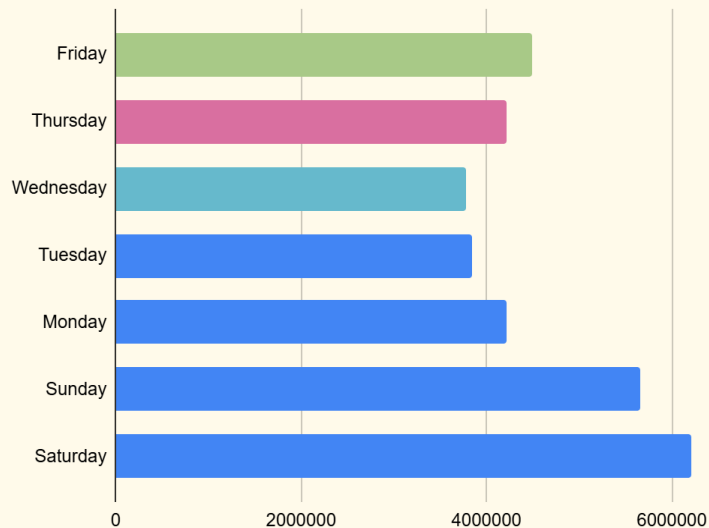
---

## Saturday & Sunday

are the busiest days

## Tuesday and Wednesday

are the least busy days



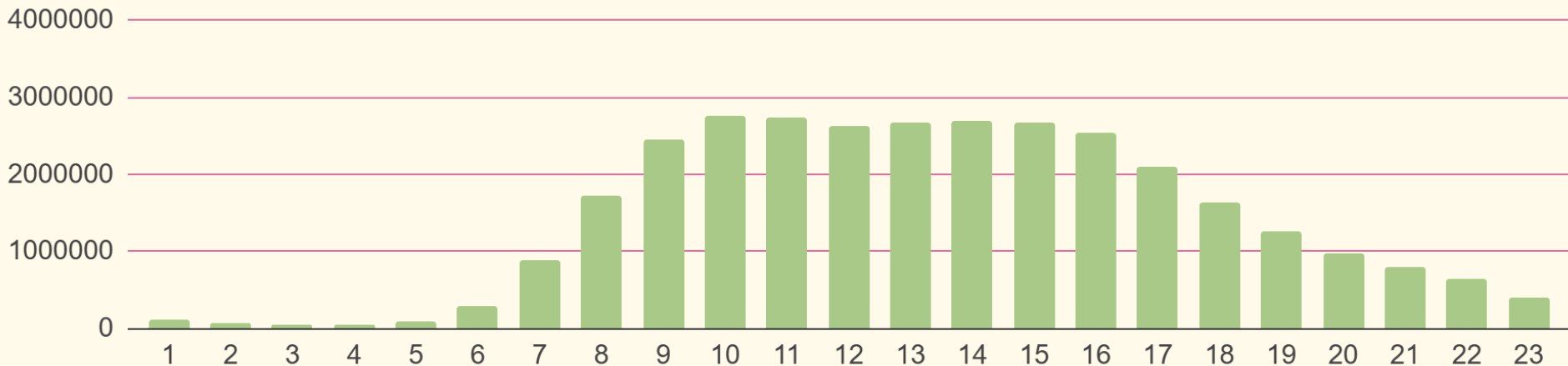
# Busiest hours

9 – 17  
Busiest

0 – 5  
Least busy

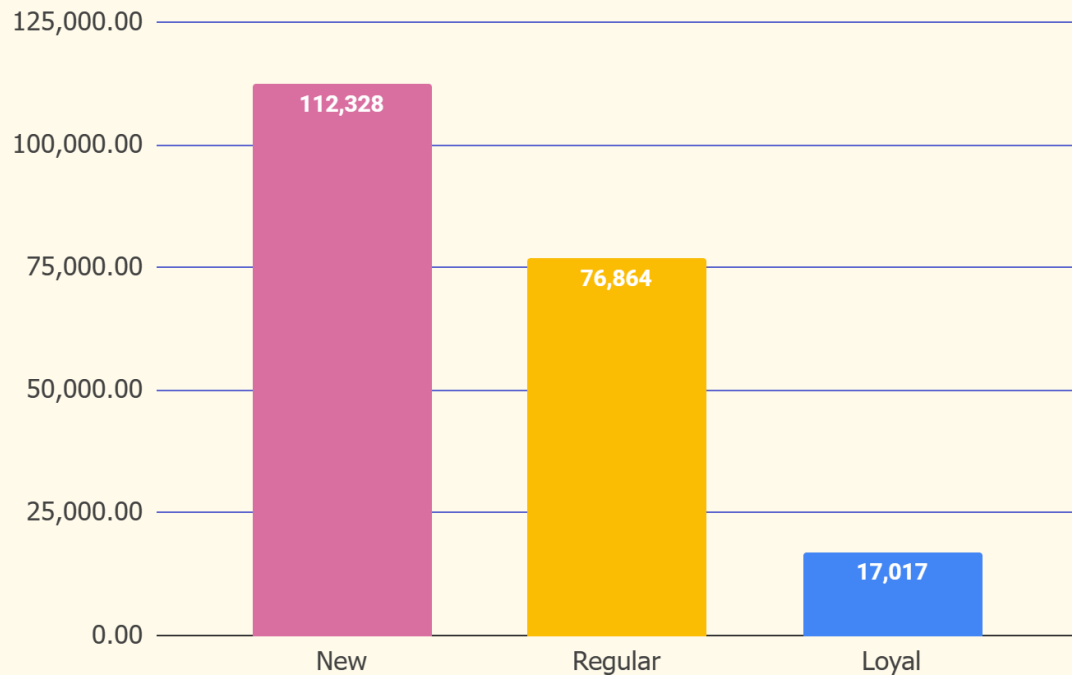
While 0 – 5 is a low-orders timeframe, the average spending per product is slightly higher than during the busiest hours (7,77 – 7,85 vs 7,75).

**Why?** Late-night shoppers may want to make bulk purchases or purchase more expensive products while there's less traffic.



# Customers

---



**206,209**

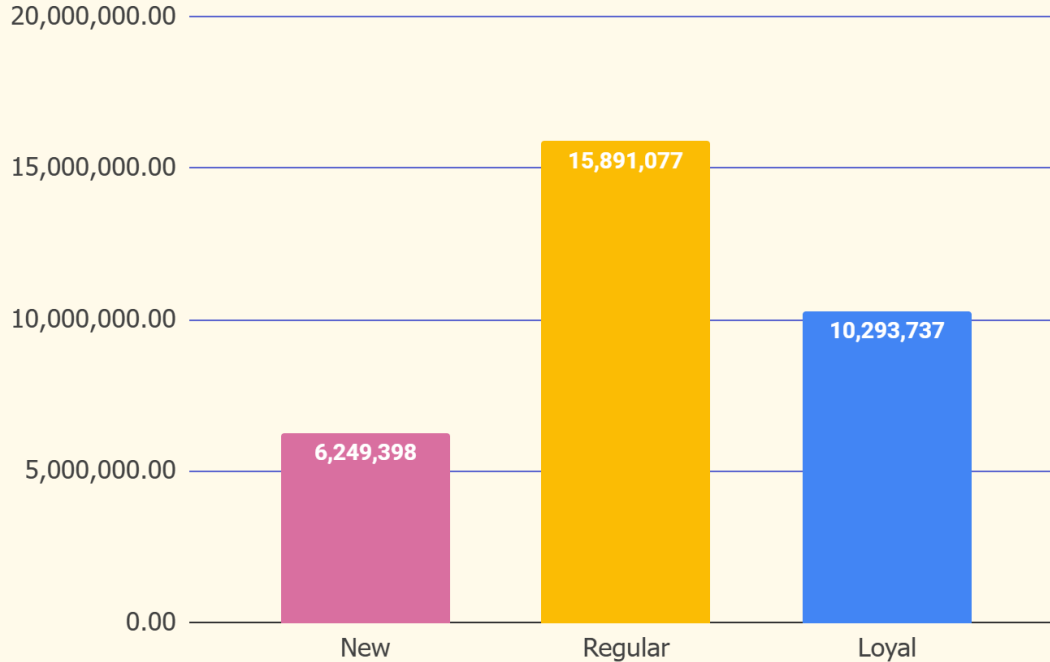
Total customers

## Loyalty

- New customer = less than 10 orders
- Regular customer = between 10 and 40 orders
- Loyal customers = more than 40 orders

# Distribution of orders in terms of loyalty

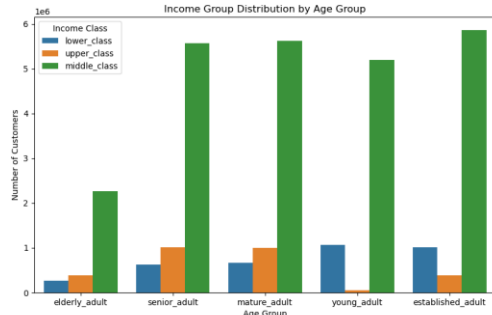
---



**32,434,212**

Total orders

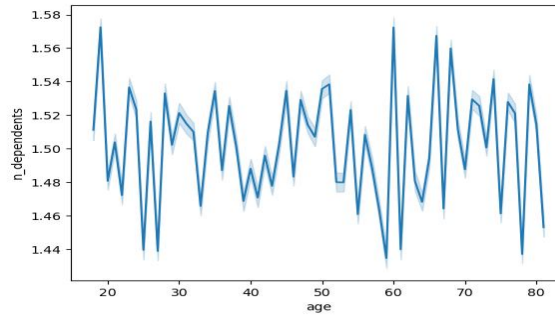
# Other insights (I)



- **Lower** class = < 52,000
- **Middle** class = 52,000 - 156,000
- **Upper** class = > 156,000

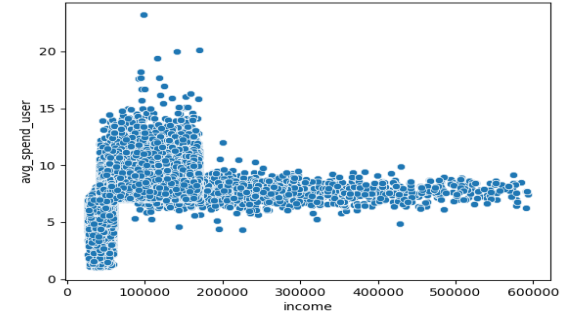
Conclusion: The middle class is the predominant class through all age groups.

Currency: USD  
Ranges taken from the [World Economic Forum](#)



Is there a connection between age and number of dependents?

This visualization suggests that such relationship does not exist.



Is there a connection between income and average spending?

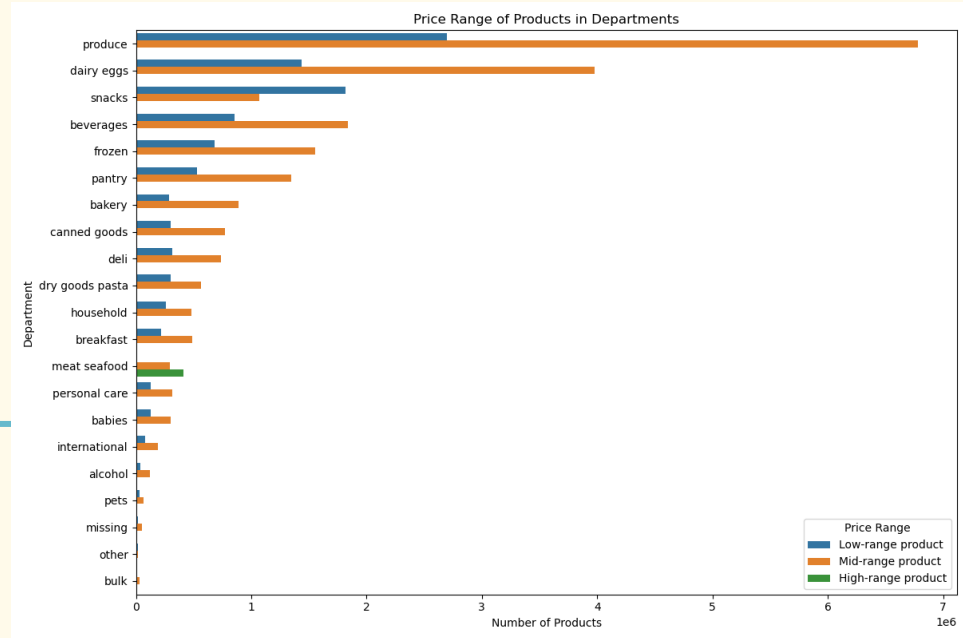
The average spending does not increase as the income increases. In fact, lower- and middle-class individuals tend to spend more than upper class individuals. The latest average spenditure tends to be within a range of 5 - 10.

# Other insights (II)



How's the inventory distributed in terms of price ranges?

The 'Bulk' category has the least inventory. If high-income people use our app to make bulk purchases, this can explain why they're spending less than low- and middle-income customers.



# Product Overview

## Most popular products across the US

- Produce
- Dairy eggs
- Snacks

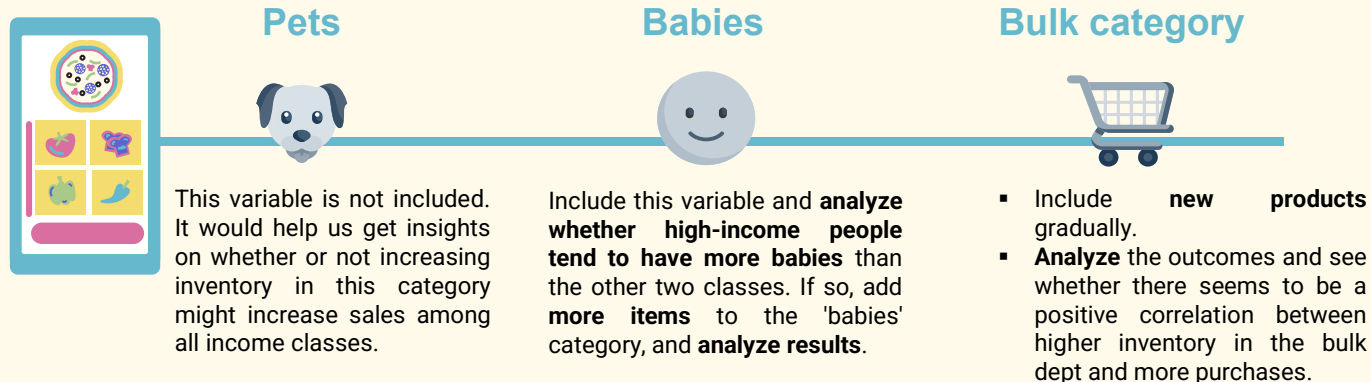
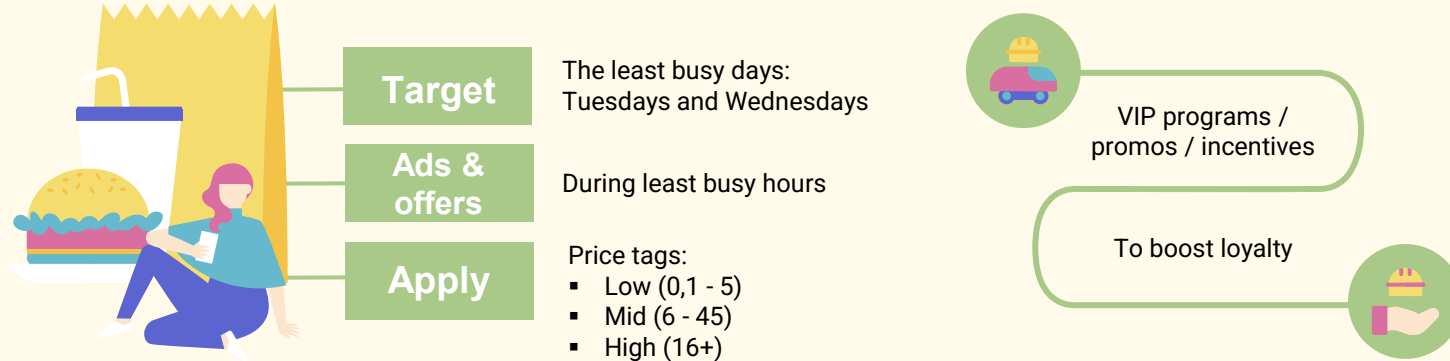
## Price tags

- Low-range (0,1 – 5)
- Mid-range (6 - 15)
- High-range (16+)





# Recommendations and next steps



# Risk factors that have contributed to churn

Pig-E Bank



# Summary

## Goal

Identify the risk factors that have contributed to clients leaving the bank



## Dataset

Pig-E bank client data set



## Techniques

Descriptive statistics, visualizations, decision tree



## Tools



# Thought process

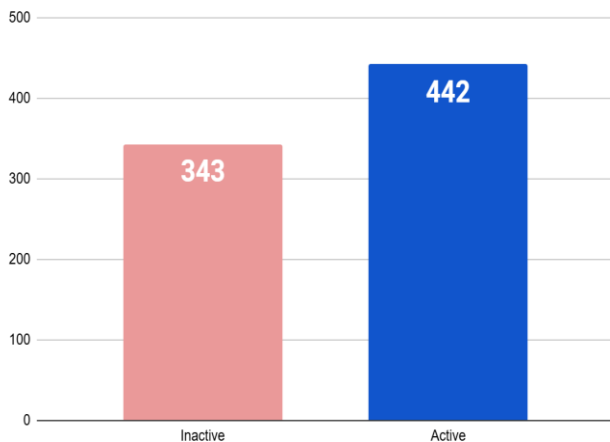
- This is not a linear process, but it comprises variable comparisons.

*Example: compare the behavior of men/women and active/inactive users in the staying vs. leaving alternatives. If the behavior is similar, then I cannot conclude that there's a correlation. Otherwise, if the behavior is the opposite, then I can suggest that there's a correlation between two of the studied variables.*

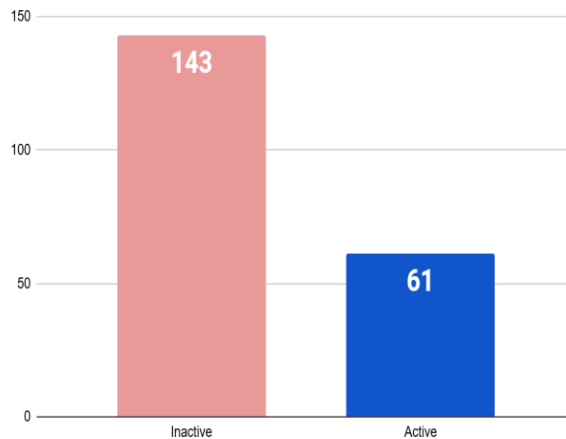
- Another variable of interest is the age variable compared with staying/leaving users.
- Any visualization not included here means it did not produce any relevant relationship.
- Finally, I created a decision tree based on the insights I discovered.

# 1<sup>st</sup> factor: **Inactive** users are more likely to leave

## Staying users



## Leaving users



**70%** of leaving users are **inactive** customers

**56%** of staying users are **active** customers

**Inactive** **29%** of inactive users left

**Active** **12%** of active users left

## 2<sup>nd</sup> factor: **Women** are more likely to leave

### Staying users

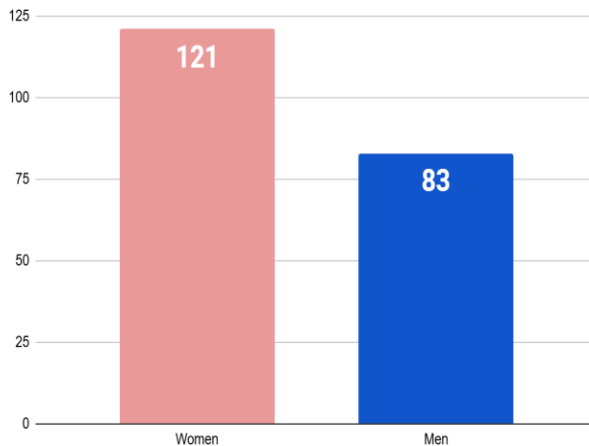
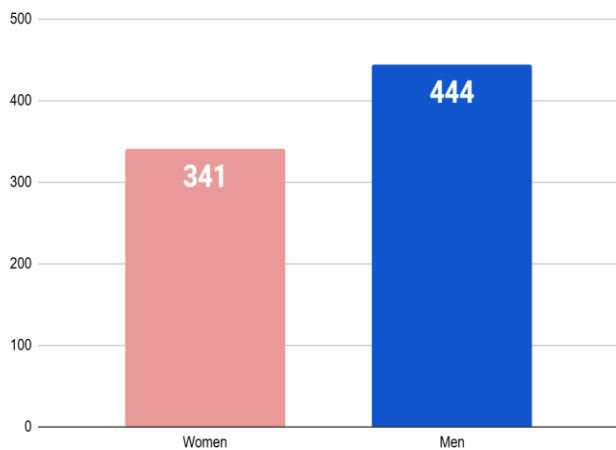
### Leaving users

59% of leaving users are **women**

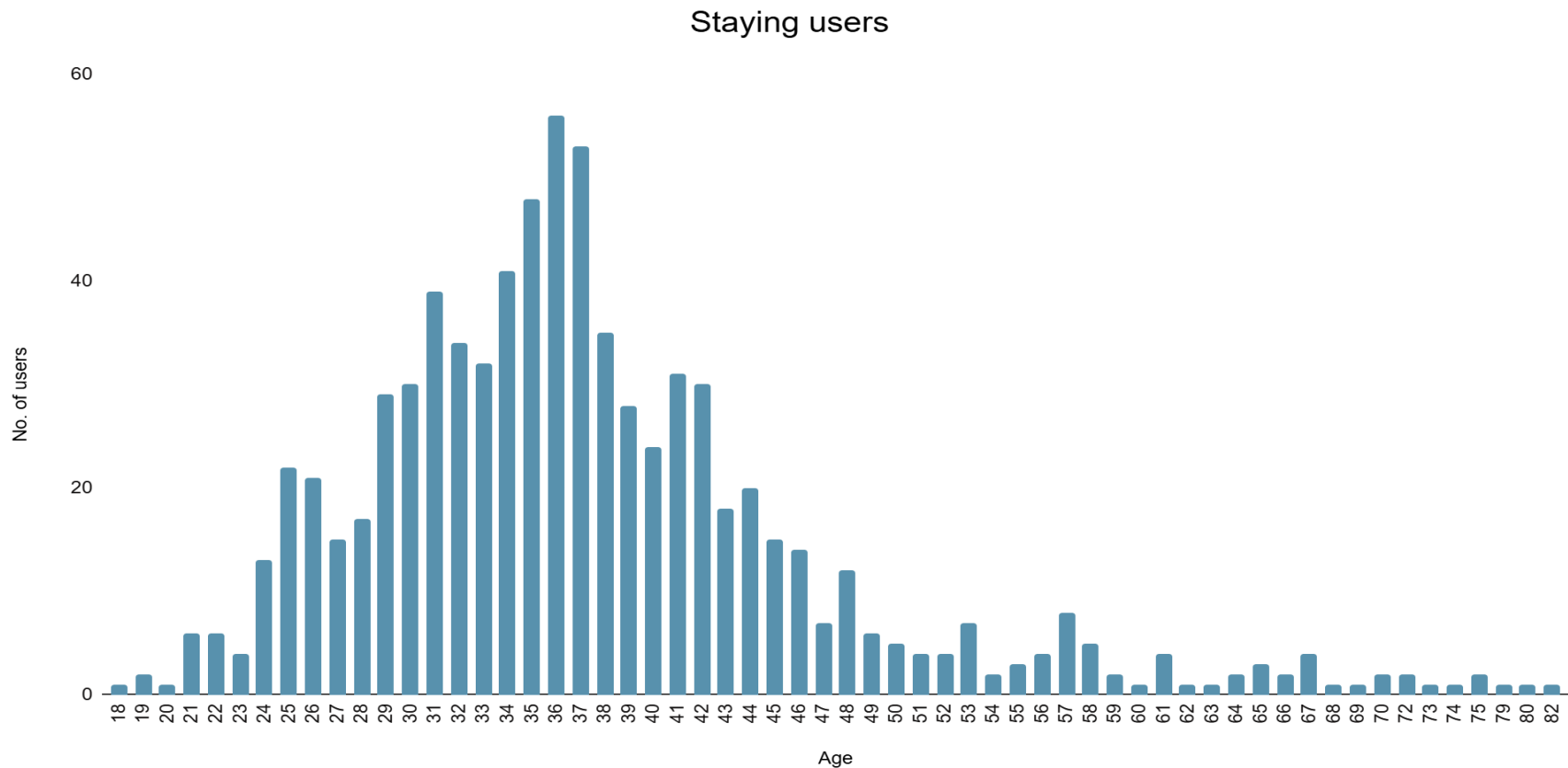
57% of staying users are **men**

**Women** 26% of women left

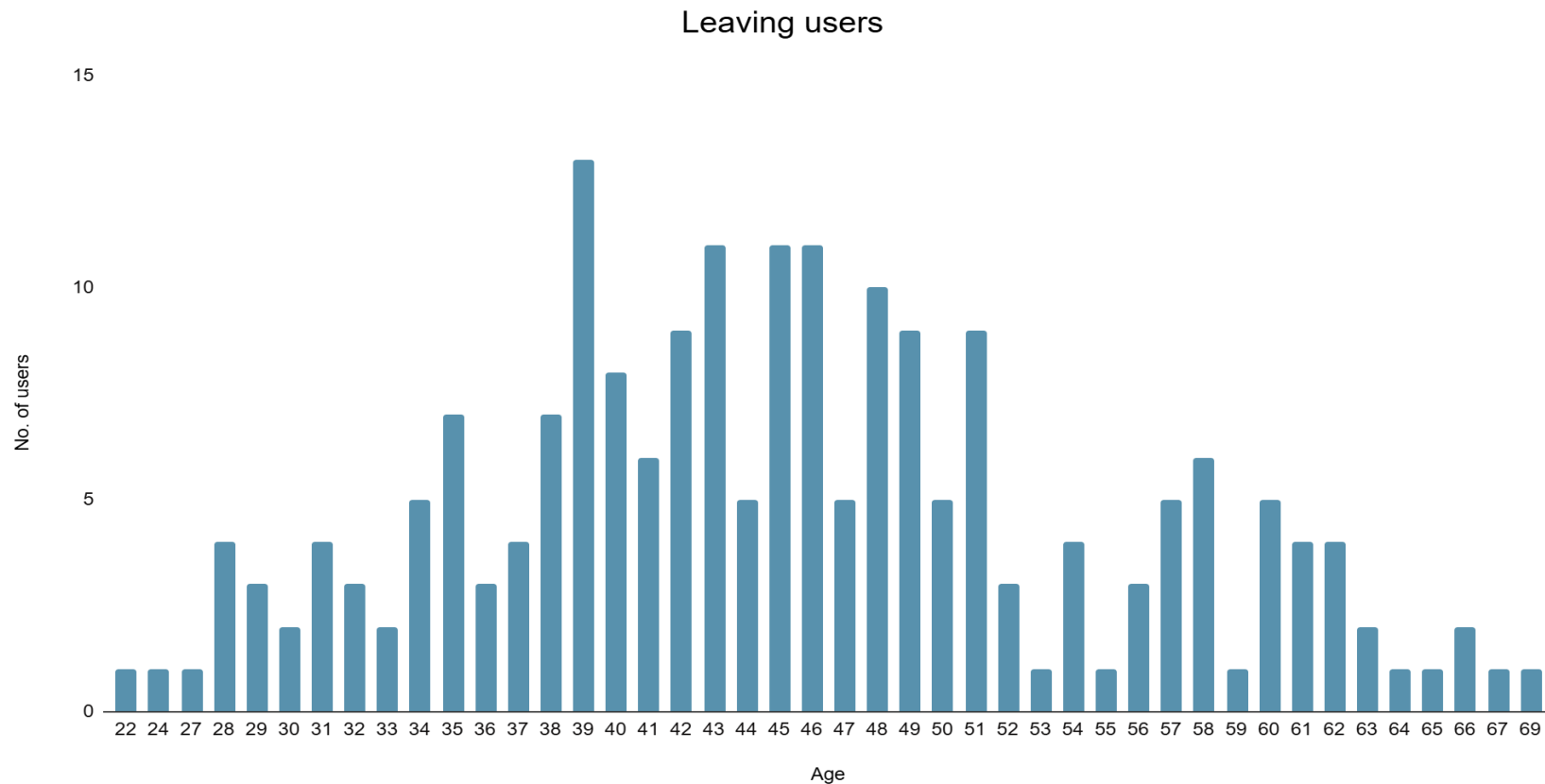
**Men** 16% of men left



3<sup>rd</sup> factor: Users of **54 and older** are (slightly) more likely to leave

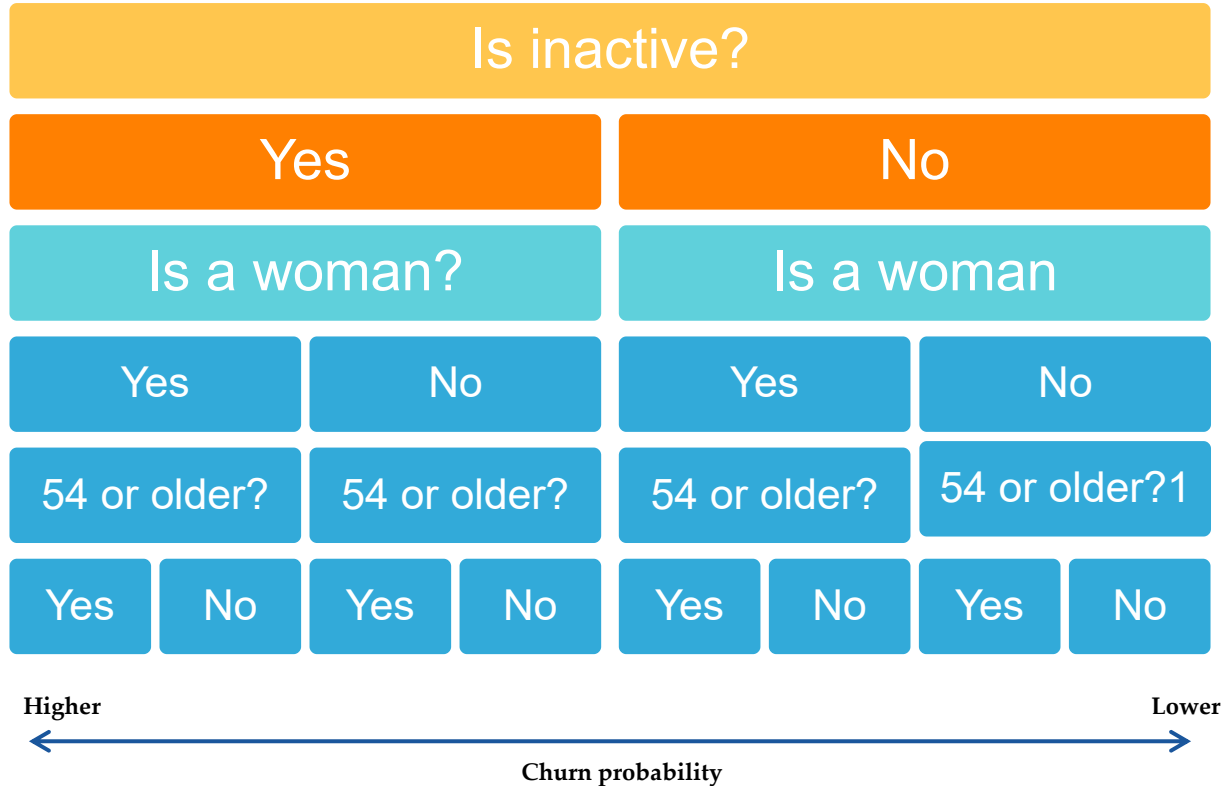


# 3<sup>rd</sup> factor: Users of **54 and older** are (slightly) more likely to leave





# Decision tree



# Suggested next steps

- Collect data about leaving clients, i.e.
  - Why are they leaving?
  - What services they did not find in Pig-E Bank?
  - What would help them use our services more (be active clients).
- Focus on the efforts that can be made to prevent clients' inactivity.
- Further analysis to try to discover why women are clients of 54 and older are more likely to leave.



# 2026 QS World University Rankings

Analysis for the Educational Portal

# CONTENTS



## Summary

Goal, dataset, tools, and techniques used for this project

## Thought process

What's behind this analysis

## Measuring AR vs ER

What these variables mean and their weight in the ranking

## Geographic distribution of AR

Average AR per country

## AR vs ER

Scatterplot

## AR vs ER

Linear Regression

## AR vs ER

Clustering

## AR vs EO

Clustering

## Results





# SUMMARY



## GOAL

Understand whether choosing a university means trading off between academic quality and job opportunities.



## DATASET

2026 QS World University Ranking



## TOOLS



## TECHNIQUES

Coding, data wrangling, subsetting, linear regression, clustering, visualization





# THOUGHT PROCESS

## ASK THE QUESTION

Does choosing a university mean trading off between academic quality and job opportunities?



Academic Reputation  
Score and Employer  
Reputation



Present results, make  
recommendations, and  
suggests next steps

## FIND THE RELEVANT VARIABLES

## RECOMMENDATIONS AND NEXT STEPS

## HYPOTHESIS AND VARIABLES TESTING

Use linear regression, clustering,  
and other tests to analyze the  
relationship between vars



# MEASURING Academic Reputation vs Employer Reputation



## Academic Reputation

Measures: reputation of **institutions and their programs.**

Weight: **30%** in the flagship QS World University Rankings.



## Employer Reputation

Measures: reputation of **institutions and their programs among employers.**

Weight: **15%** in the flagship QS World University Rankings.



ABC





# Geographic distribution of the academic reputation

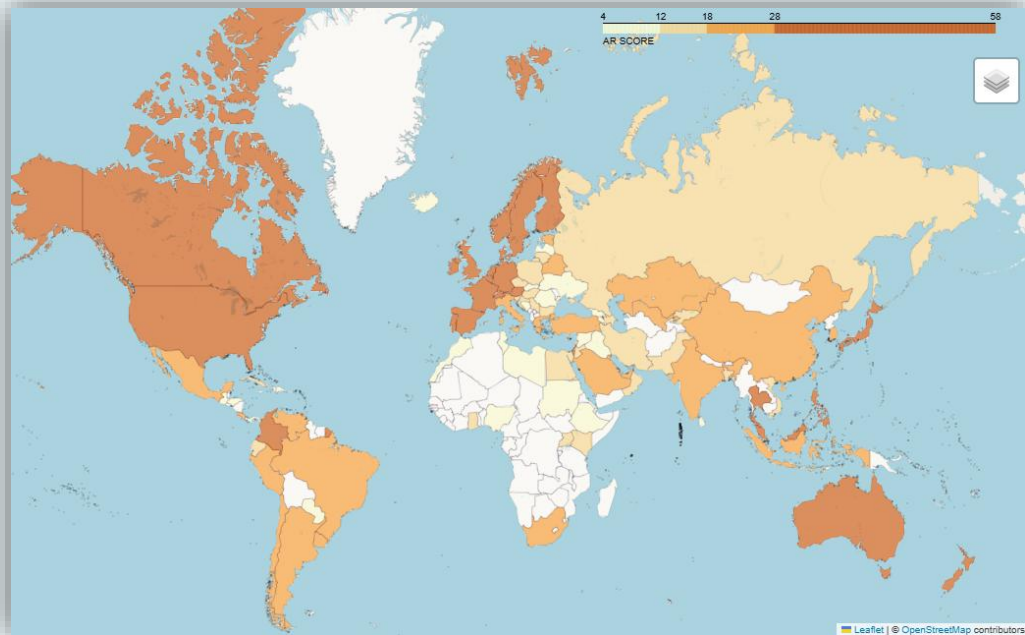
## Pros and cons of this approach

This map provides a useful visual overview of **average academic reputation** by country, based on ranked universities.

## Cons

**Significant variation** in the number of ranked universities per country.

Example, China has 72 ranked universities and Chile only 20. The average AR score for Chile is still higher—suggesting that sample size and university selection can heavily influence country-level averages.







# Academic Reputation vs Employer Reputation



The data supports the hypothesis, ***the higher the Employer Reputation, the higher the Academic Reputation.***





# Linear regression



## Meaning

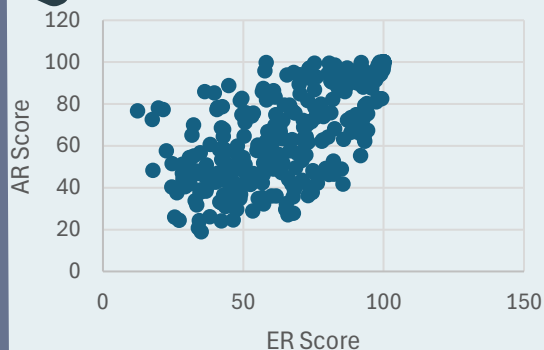
The regression line reinforces this relationship. Still, not every university perfectly fits the trend.

## Analogy

If AR scores were like exam grades, this model is typically off by about **13 points (12.87)** — so if it predicts a score of 80, the real score might be closer to 67 or 93.

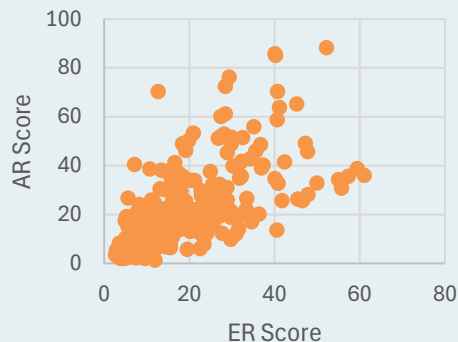


# Clustering: Academic Reputation vs Employer Reputation



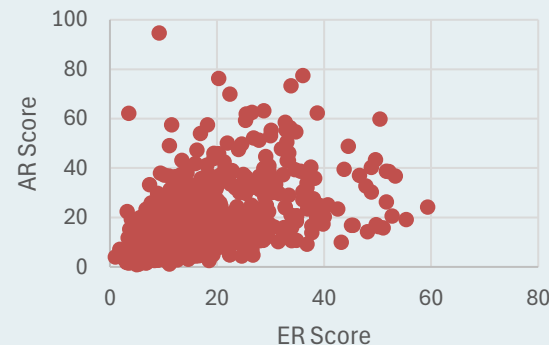
Cluster 1

Strong link: high AR and high ER go hand in hand



Cluster 2

Some schools are stronger academically but less recognized by employers, or vice versa.



Cluster 3

Some schools are stronger academically but less recognized by employers, or vice versa.



# 300

Universities in the cluster No. 1 | Top cluster


# 317

Universities in the cluster No. 2 | Middle cluster

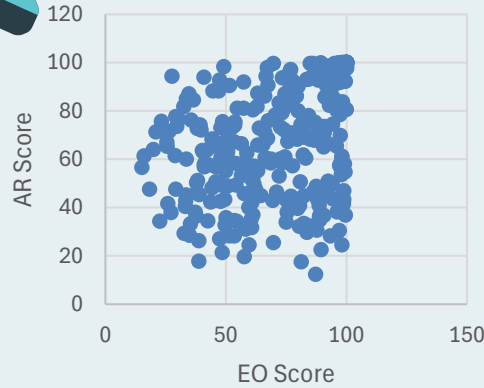


# 884

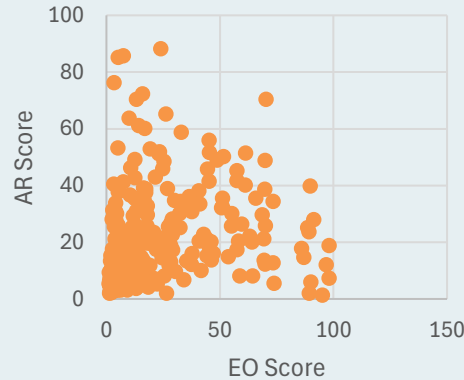
Universities in the cluster No. 3 | Bottom cluster



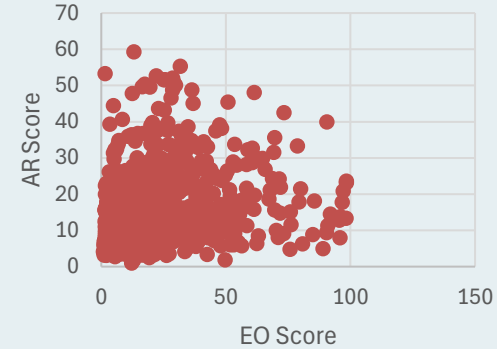
# Clustering: Academic Reputation vs Employment Outcomes



Cluster 1



Cluster 2



Cluster 3

A strong employer reputation **doesn't guarantee** strong employment results.

**Meaning:** reputation opens doors, but actual job performance depends on more than reputation alone.



# RESULTS AND RECOMMENDATIONS



## No trade-off



Students **don't face a strict trade-off**.  
Academic and employer reputation usually reinforce each other.

## Employment predictor



Neither reputation metric is a perfect predictor of employment success. **Other factors matter.**

## Thresholds



Universities with **AR above 66.3** and **ER above 64.5** are most likely to offer both strong academic and employer recognition.

## Suggestion



When it comes **to job outcomes**, I suggest looking beyond reputation.



# THANKS!

**CREDITS:** These presentation templates were created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik.

For more information about editing slides, please read our FAQs or visit Slidesgo School:

<https://slidesgo.com/faqs> and <https://slidesgo.com/slidesgo-school>