



MARIA MASKE

Data Analyst Portfolio



<https://github.com/mmaske/>



<https://www.linkedin.com/in/mmaske/>

PROJECT OVERVIEW

1. Global Happiness (2015-2023)



- Exploratory Analysis of the happiness Data from the Gallup World Poll

2. Instacart



- Python – based project on Consumer Behavior Trends and Marketing Analysis

3. Rockbuster



- International Business Analysis of Online Video Rental Services through SQL

4. Influenza season in the USA



- National Medical Staffing Distribution based on Historical trends

5. GameCo



- Global Marketing Analysis of the Gaming Console Industry

6. PIG E. Bank



- Predictive Analysis of customer retention risk for a Global Finance Service Company

WORLD HAPPINESS (2015-2023)



OBJECTIVE

Each year, the **Gallup World Poll** surveys adult citizens in nearly **160 countries** and consists of over 100 global and region-specific questions to calculate the Happiness score of each country. A lot has changed globally since 2020 and led to a lot of uncertainty. Therefore we **analyze** the data to find out if the **happiness score has changed between 2015 and 2023**.

DATA

- Project reflection
- World Happiness Reports 2015 -2023 published by the Sustainable Development Solutions Network

TECHNIQUES APPLIED

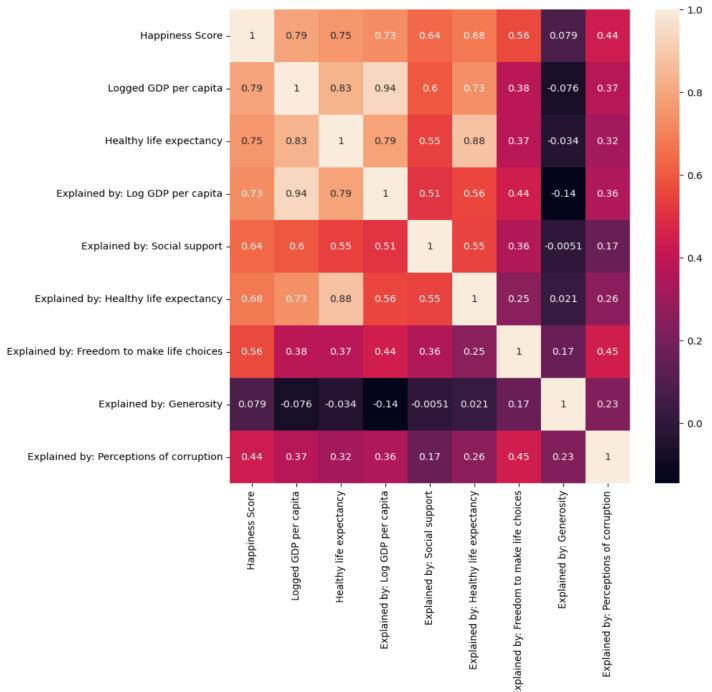
- Linear Regression
- Cluster Analysis
- Spatial Analysis
- Python & Tableau Visualization

TOOLS



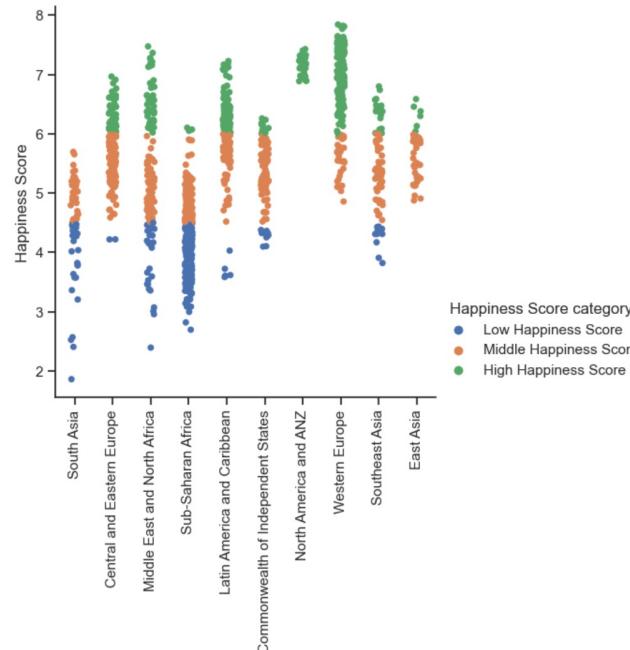
ANALYSIS

Correlation Heat Map of the measures of the Happiness score



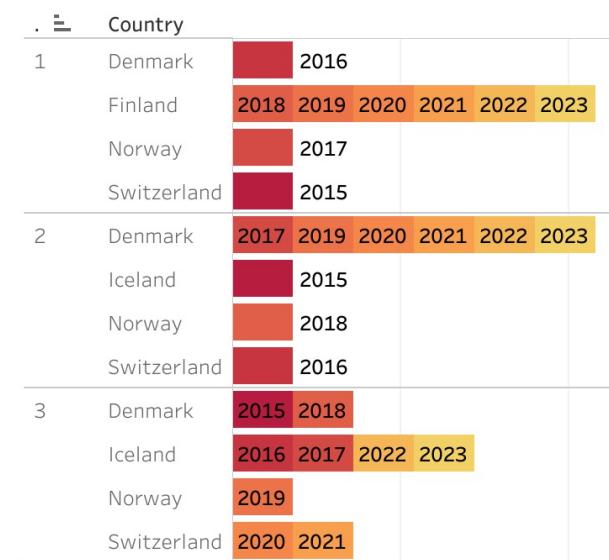
The correlation heat map shows that the logged GPD per capita has the highest impact on the happiness score of a country.

Happiness score categories by region



The happiest region is North America and ANZ, while the countries with the highest Happiness scores are located in Western Europe.

Happiness Ranks by Country and Year



The country with the highest Happiness score is Finland since 2018.

SOME RECOMMENDATIONS & FURTHER INSIGHT

Recommendations for further analysis:

- Collection of further data, especially for the Sub Sahara and Middle East and North Africa regions, in order to close the data gaps here.
- Investigate more deeply why the overall happiness score for South Asia has fallen so much in the last three years.
- And it would also be interesting to see to what extent the happiness score per country also differs by gender or age.

Limitations

- We don't have data for every year for every country since many countries don't do the survey annually. We have especially for the year 2022 many missing values.
- Therefore our analysis may not be as accurate and could be slightly biased due to the circumstance that most missing values stem from the region Sub Sahara and Middle east and North Africa.

Further Insights:

Project Overview



Python Scripts:



Tableau Dashboard:





OBJECTIVE

Instacart is an online grocery store that operates through an app that wanted to uncover more information about their sales patterns. Therefore we analysed the existing data in order to derive insights and suggest strategies for better segmentation based on the provided criteria.

DATA

- [Project Brief](#)
- Instacart provided an Data Set featuring customer information and purchase history

TECHNIQUES APPLIED

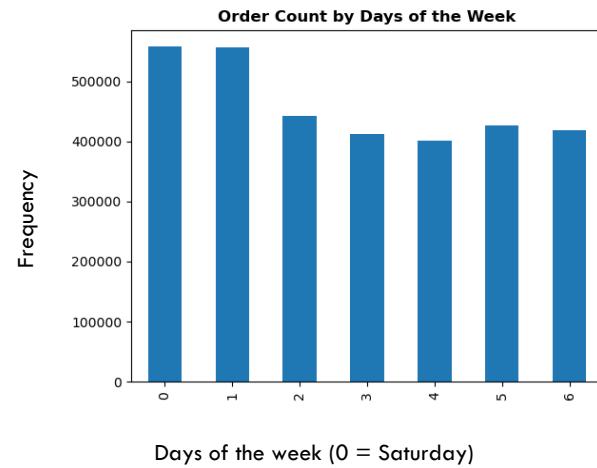
- Data Cleaning: Wrangling
- Combining and Exporting Data
- Grouping Data and Aggregating Variables
- Python Visualization
- Excel Report

TOOLS



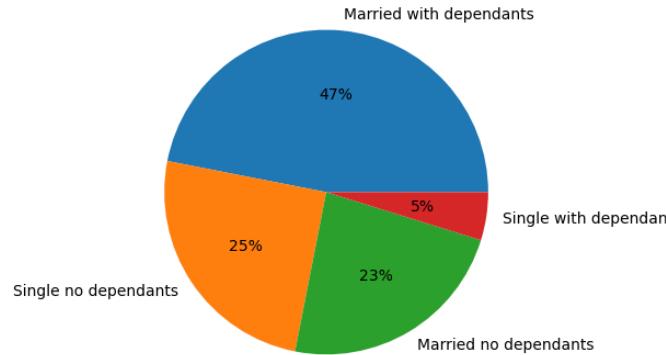
ANALYSIS

Busiest Day of the week



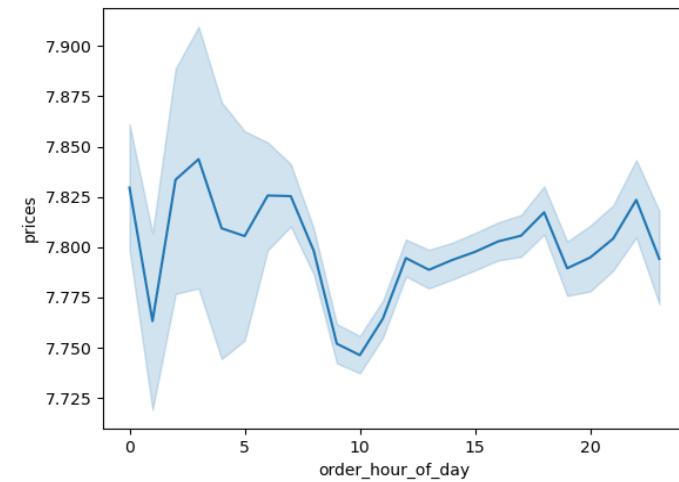
The busiest days of the week are the weekends, with Saturdays being the busiest day.

Customer Profile



The majority of customers are parents with a partner followed by single adults.

Times of the day people spend the most money



Instacart's customers spend the most money on Friday and Saturday with expenditure peaking from 2am to 7am.

SOME RECOMMENDATIONS & FURTHER INSIGHT

Recommendations:

- **Customer profile marketing**

- Advertise affordable products within specific departments such as produce, babies and pantry items to middle age group and families who generate significant revenue. Additionally ,targeting the low income profile customers with cost saving promotions would be crucial.

- **Time-based marketing**

- To increase revenue, schedule ads on Wednesday and Tuesday after 15, as these are the least busiest days and sales begins to decrease after 15 up to morning 6.

- **high prices products Marketing**

- Instacart should place more ads for high-priced products in the early morning hours from 4 am until 7 am.

Further Insights:

Python Scripts:



Detailed final report:





ROCKBUSTER

OBJECTIVE

Rockbuster Stealth LLC is a movie rental company that wanted to launch an online video rental service in order to stay competitive. analysed the existing data to help with the launch strategy for the new online video services for 2020.

DATA

- Project Brief
- This Data Set contains important information about Rockbuster's.

TECHNIQUES APPLIED

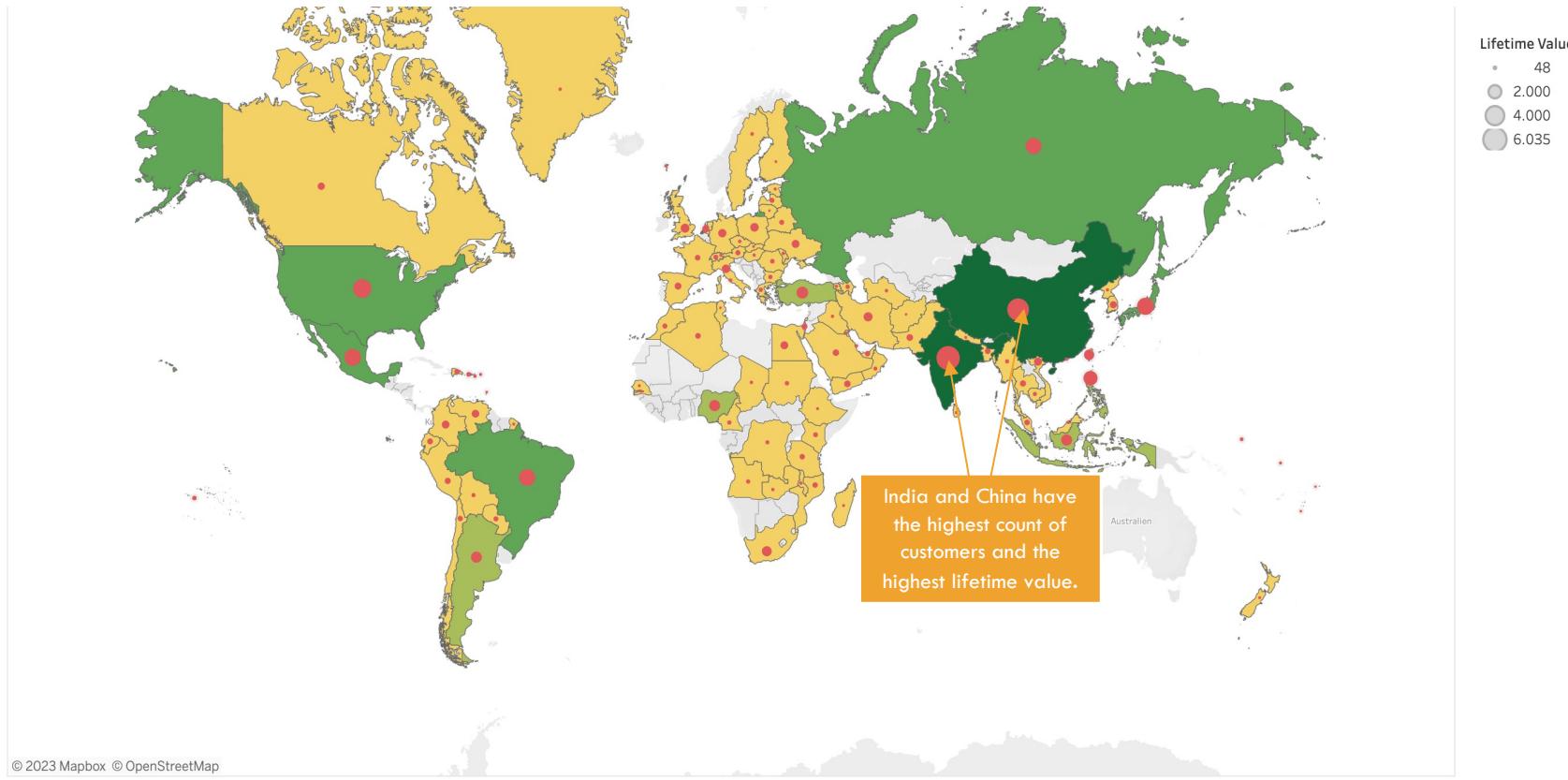
- Data Cleaning and Summarizing
- Relational Databases
- Creating a Data Dictionary
- Entity Relationship Diagram (ERD)
- Filtering Data
- Joining Tables
- Subqueries

TOOLS



ANALYSIS

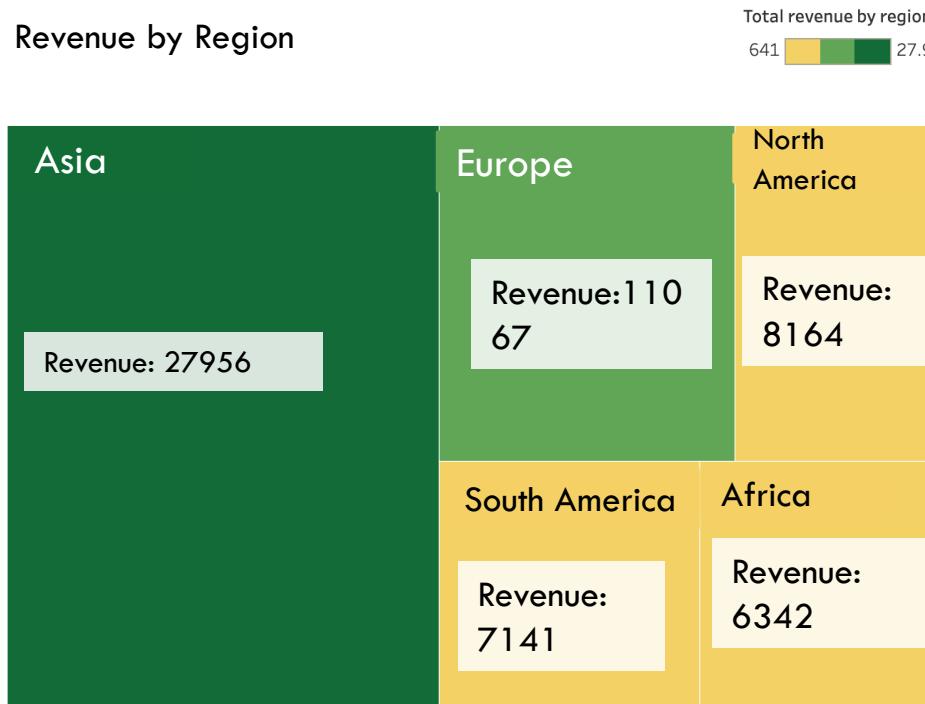
Total Customers per Country



- The Philippines, the United States and Brazil have a high **lifetime value** too.
- The **Top 10 Cities** are
 - Aurora,
 - Iwaki
 - Cirtus Heights,
 - Ambattur,
 - Shanwei,
 - Acua,
 - So Leopoldo,
 - Tianjin,
 - Cianjur
 - Teboksary

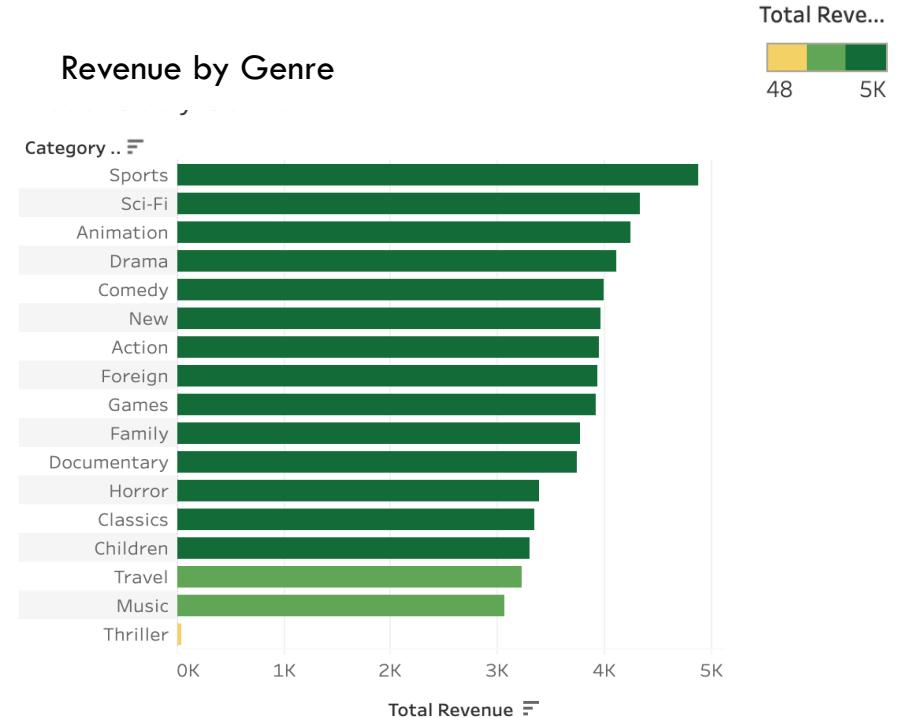
ANALYSIS

Revenue by Region



- **Asia, Europe and North America** are the top regions in terms of total revenue.
- The **lowest** revenue has **Ozeania**.

Revenue by Genre



- The **top** genre by revenue is **sports**.
- It's followed by **Sci-Fi, Animation and Drama**.
- **Thriller** is the **bottom** genre by revenue

SOME RECOMMENDATIONS & FURTHER INSIGHT

Recommendations:

- **Regional marketing**

- Focus on advertising in the regions **Asia, Europe and North America** since these regions have a high revenue.

- **Genre-based marketing**

- Focus marketing on movies in the genres with the highest revenue: **Sports, Sci-Fi, Animation**

- **Inventory update**

- Offer movies in **other languages** than English, e.g. especially **Hindi and Mandarin** since India and China have the largest customer base

- **Loyalty program**

- Implement a **loyalty & recommendation scheme** to increase customer retention and to attract new customers.

Further Insights:

- **SQL code:**



- **Final presentation:**



- **Tableau dashboard:**



- **Data dictionary:**



INFLUENZA SEASON IN THE USA



OBJECTIVE

A medical staffing agency wanted to provide temporary workers to clinics and hospitals on an as-needed basis. Therefore I analyzed the flu trends in the USA in order to provide more accurate recommendations on how many workers to send to each state.

DATA

- Project Brief
- This Influenza Deaths | Source – CDC
- U.S Population | Source – U.S Census Bureau
- Influenza Visit | Source – CDC

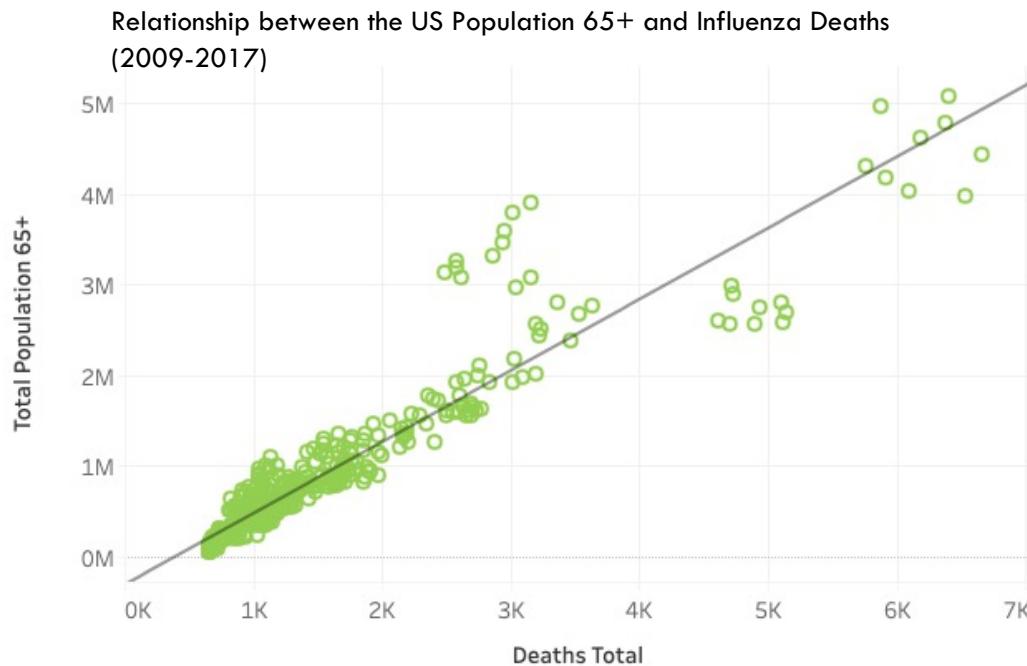
TECHNIQUES APPLIED

- Data Cleaning
- Data Transformation and Integration (* Excel: PivotTables, VLOOKUP, Power Point)
- Data Profiling and Integrity
- Designing a Data Research Project
- Data Visualization and Storytelling (Tableau)

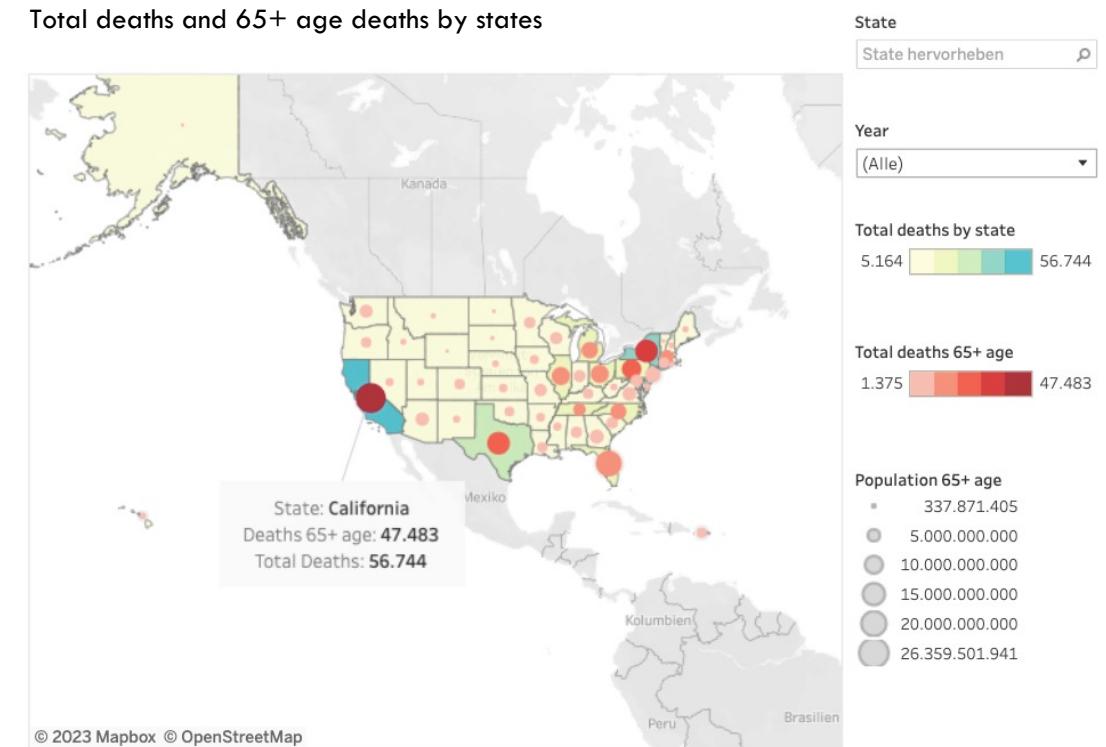
TOOLS



ANALYSIS



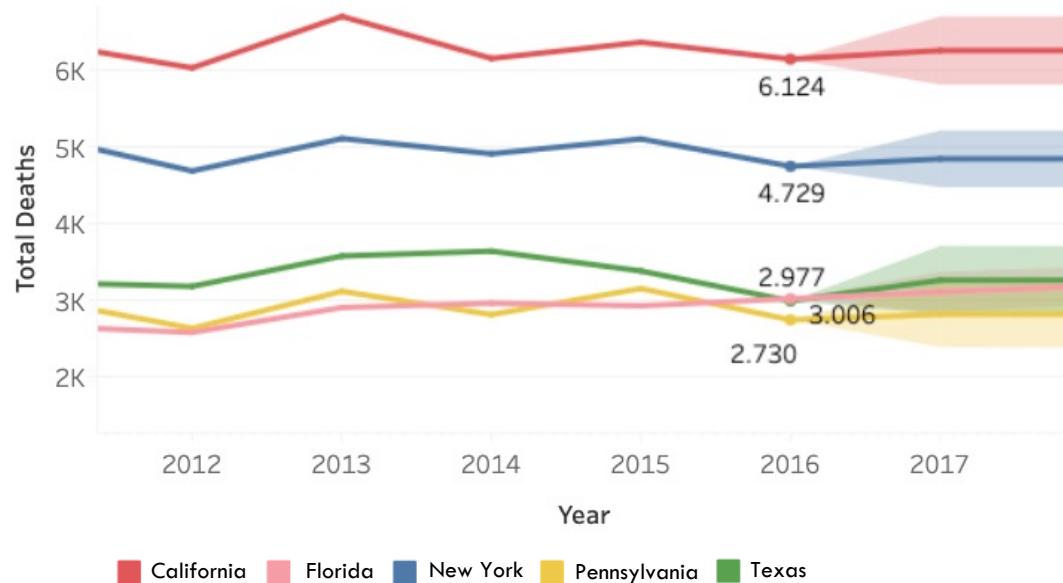
- The **65+ age-groups** have the highest flu death proportion per age group.
- They are **more likely to suffer fatalities** from the flu.
- The **correlation** between the number of 65+ citizens per state and the number of influenza deaths is **significant**.



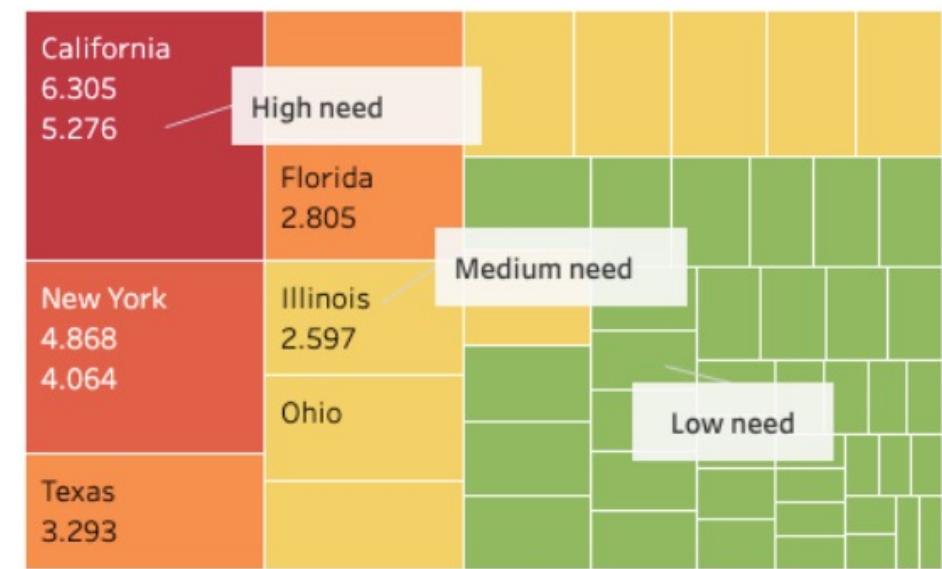
- California** has the highest population proportion of 65+ years olds and therefore the **highest average death rate**.
- It is followed by **New York, Texas, Florida and Pennsylvania**.

ANALYSIS

Flu mortality forecast for the 5 largest 65+ state populations



Flu deaths per state in population 65+ (2009-2017)



- States with a high population proportion of 65+ years olds have a **higher influenza death rate regardless of the year**.
- Five of them are expected to suffer from **at least 2,4K related flu deaths** in 2018.

- Therefore California, New York, Texas, Florida and Pennsylvania are in **high need of additional medical staff**.
- States like Illinois and Ohio have a medium need concerning additional medical staff.

SOME RECOMMENDATIONS & FURTHER INSIGHT

Recommendations:

- Deployment seasons
 - Additional medical staff should be deployed during the **flu season** (**November – April**).
- Allocation staffing resources
 - The staffing resources should be allocated **based on the tiers „High-need“, „Medium-need“ and „Low-need“**, which is visualized with the tree-map.
 - The states with the highest need of additional staff are **California, Florida, New York, Texas and Pennsylvania**.

Further Insights:

Interim report:



Detailed data analysis:



Tableau storyboard:



Video presentation:





GAMECO

OBJECTIVE

The video game company GameCo wanted to get more insight into the development and marketing of new games. I performed descriptive analysis to gain insights into the current video game landscape for a marketing and sales' team planning for 2017.

DATA

- Project Brief
- Data sourced from VGChartz

TECHNIQUES APPLIED

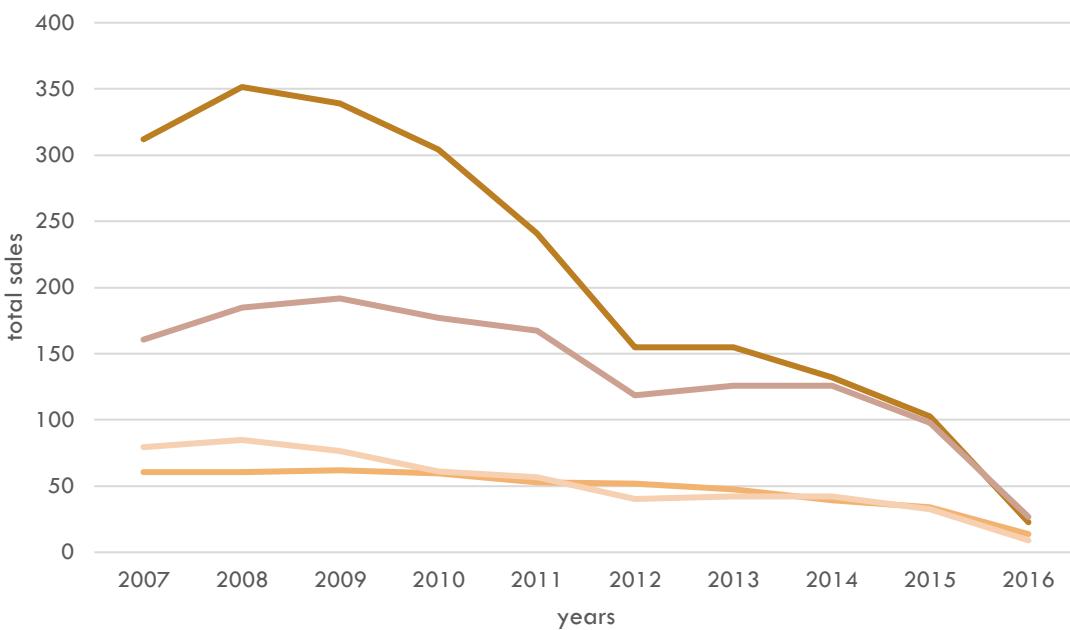
- Data Cleaning
- Data Integrity and Quality (PivotTables)
- Data Grouping and Summarizing (Calculated Fields)
- Descriptive Analysis

TOOLS



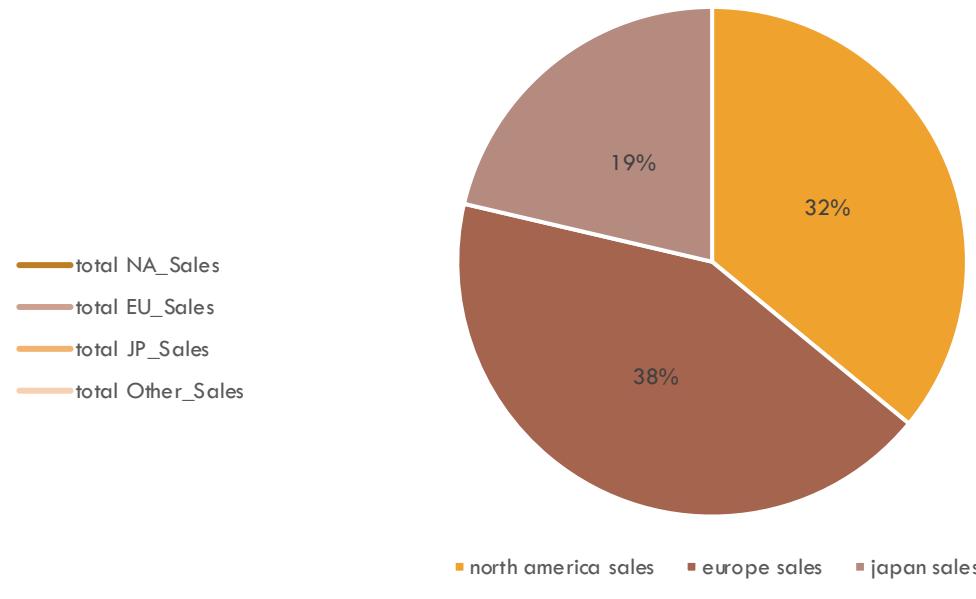
ANALYSIS

Total sales of the various regions between 2007 and 2016



- Sales have dropped across regions in recent years for the retail sector.

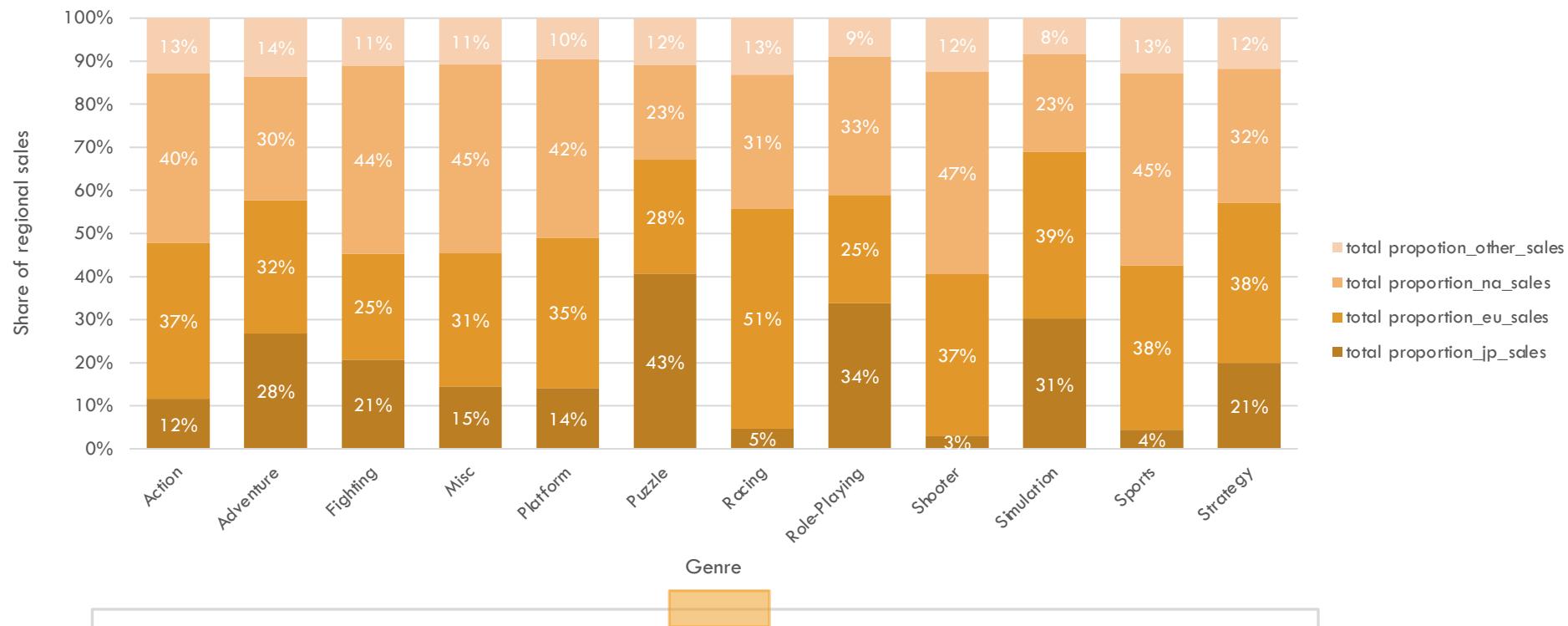
Regional proportion of total sales for 2016



- There has been a **clear change in regional data in recent years as the especially the share of NA-sales in global sales decreased**, while the share of JP-sales and EU-sales in global sales increased.

ANALYSIS

Share of regional sales by genre between 2012 and 2016



- Certain genres sell better in the different regions.
- E.g. shooter are the highest selling genre in north america, while puzzle is the highest selling genre in japan.

SOME RECOMMENDATIONS & FURTHER INSIGHT

Recommendations:

• North America market

- As the competitors are losing their retail sales in the NA market,

GameCo could gain sales by **doing targeted marketing for shooter games.**

• Japan market

- GameCo should focus their marketing on **puzzles and role-playing** games in japan.

• European market

- The mayor focus should be on the development of the EU market since has the higehst sales propotion of the total sales at the moment. Genres like action, shooter and sports should be marketed.

Further Insights:

Detailed data analysis:



Final presentation:



Project reflections:





PIG E. BANK

OBJECTIVE

The **PIG E. Bank** wanted to understand the factors behind them losing clients. I used a data mining mechanism and build a decision tree to predict which clients might leave the bank.

DATA

- Project Brief
- Client Data Set

TECHNIQUES APPLIED

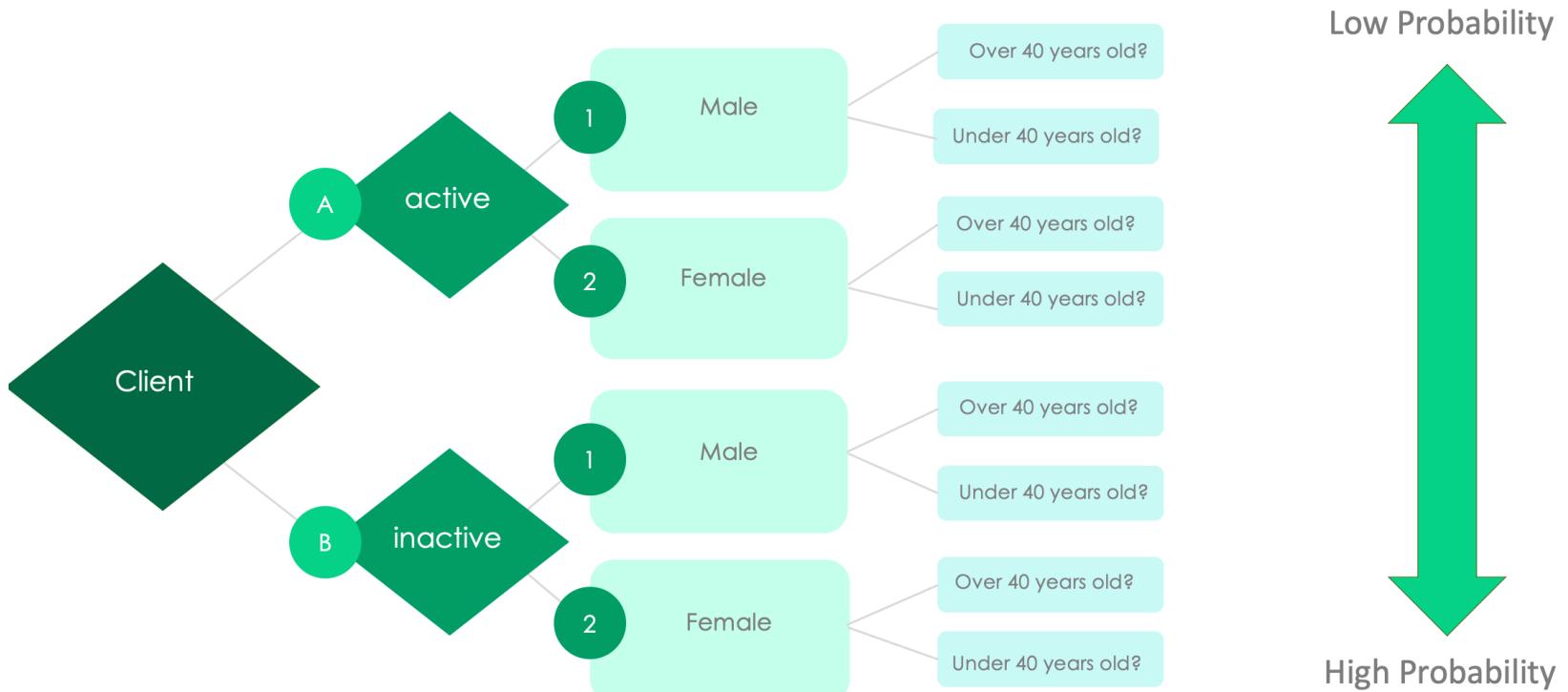
- Data sorting, filtering and cleaning
- Grouping & summarizing data
- Descriptive analysis
- Data Mining
- Building a decision tree as a data mining algorithm

TOOLS



ANALYSIS

The following decision tree demonstrates the risk factors that contributes to a client's likelihood to leave Pig E. Bank.



SOME RECOMMENDATIONS & FURTHER INSIGHT

Recommendations:

- Top three factors that lead to clients leaving

- The top three factors are active membership, gender and age.

- Therefore PIG E. Bank should focus on clients, who fullfill these factors

- Activity:

- They could create a **bonus or loyalty programm** where clients get points for being active that they can redeem for bonuses.

- Age:

- They could **give people under 40 special discounts** or interest rates to retain these clients.

- Gender

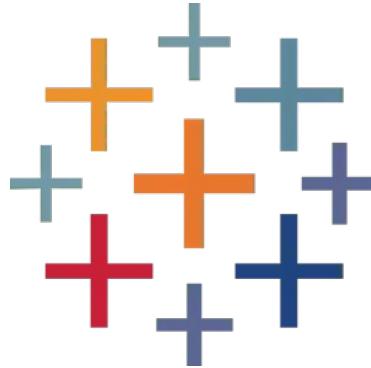
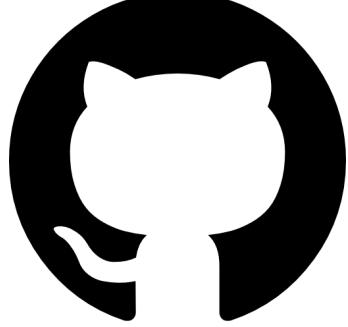
- They can find out which **financial topics are particularly interesting for women** and build a specific program based on them in order to lose fewer women and gain more clients.

Further Insights:

Detailed data analysis

& Documentation:





DO YOU HAVE ANY QUESTIONS?

mariamaske2063@gmail.com