

Data Analyst Portfolio

ROBERT INDELICATO
<https://www.linkedin.com/in/robert-i-b0714b122/>

<https://public.tableau.com/app/profile/Robert.Indelicato/vizzes>

<https://github.com/rindelicato>

About Me

- My name is Robert, and I am an aspiring data analyst with 4 years of experience as a Research Analyst. Results-oriented Business Analyst skillful in managing and breaking down large volumes of information. Proactive at heading off issues in operations, workflow, and production by uncovering trends affecting business success. An efficient team member with a collaborative approach to completing projects and mentoring/training new employees to become professional analysts. Passionate about increasing knowledge to drive career growth.



Projects

1. GameCo - Market Analysis
 2. Influenza - Forecasting
 3. Rockbuster Stealth - SQL
 4. Instacart - Python Analysis
 5. Pig E. Bank - Big Data Analysis
-

Tools

- Microsoft Excel



- Tableau



- SQL



python™

- Python

- Github



GameCo

- Perform descriptive analysis & recommendations to improve company sales based on their historical data.



Goal, Data limitation, and Skills

Perform Descriptive Analysis of Video Game Market Data from 1980 to 2016 and provide insights about how GameCo's new game will fare in the market

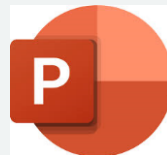
It tracks the total number of units sold, but it does not represent financial figures. Data has time limitation, since shows information from 1980 to 2016 only, which probably does not represent the current market. Data from 2016 is incomplete. Some top rank games sales data show inconsistencies, as they show zero sales in certain market, i.e. Gta V in Japan, even though sold millions of copies.

- Data Cleaning
- Data Grouping and Summarizing
- Descriptive Analysis
- Pivot Table
- Visualization Charts in MS Excel/PowerPoint

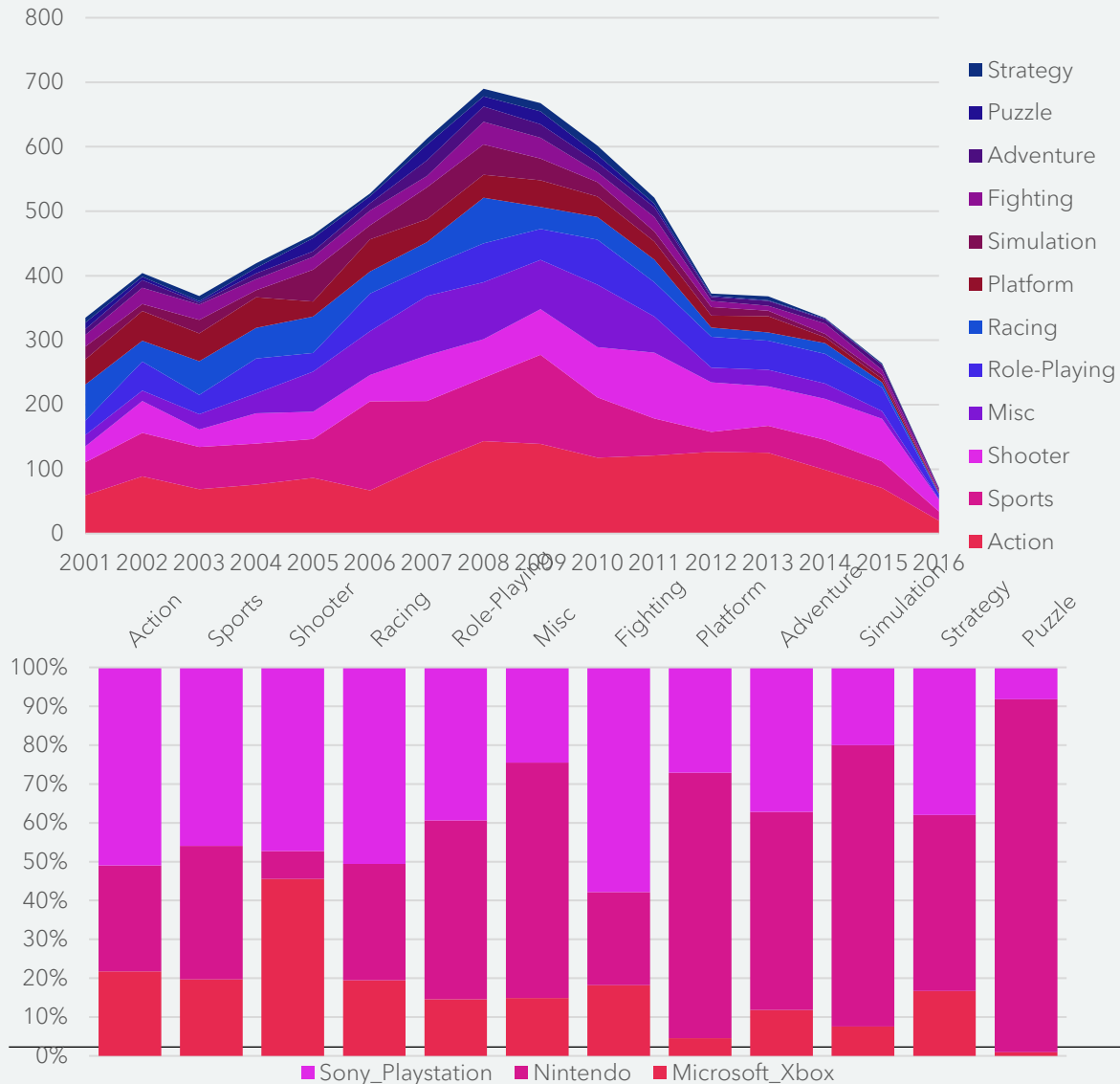
Data Source

[https://images.careerfoundry.com/public/courses/intro-to-data/Project Brief Intro to Data Analytics.pdf](https://images.careerfoundry.com/public/courses/intro-to-data/Project%20Brief%20Intro%20to%20Data%20Analytics.pdf)

Tools:



Overview



A CLOSER LOOK AT 2001-2016 SINCE THIS WAS THE FIRST YEAR SONY, NINTENDO & MICROSOFT PLATFORMS MANUFACTURERS HAD THE MARKET SHARE OF GAMES.

THE SONY PLAYSTATION LEADS IN THE TOP THREE BIGGEST GENRES IN TERMS OF GLOBAL UNIT SALES WHICH ARE ACTION, SPORTS AND SHOOTER.

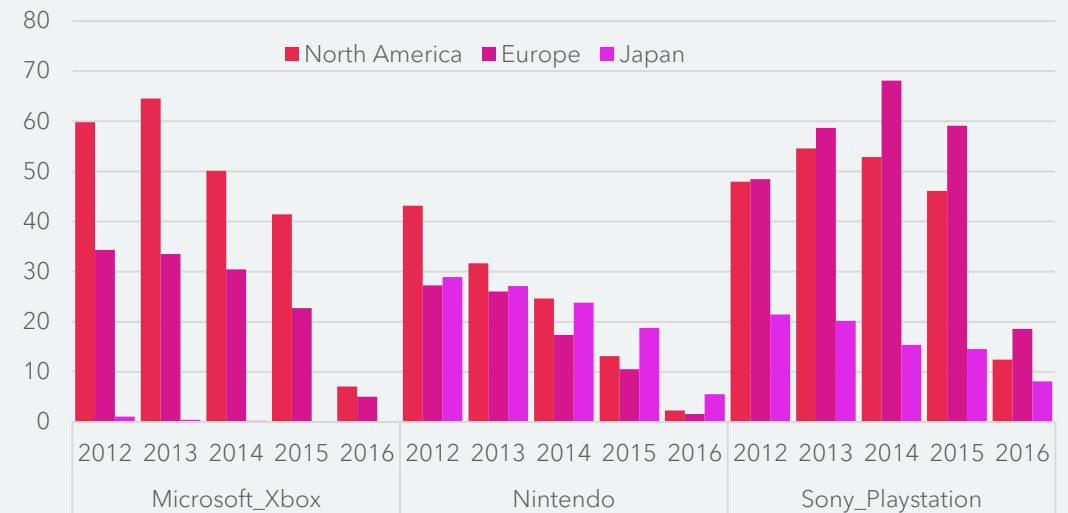
ALTHOUGH NINTENDO CLEARLY LEADS IN PUZZLES, THIS GENRE IS THE SMALLEST CATEGORY.

Market Analysis

DESCRIPTION OF GRAPHS

- Sony is the dominant platform as compared to Microsoft and Nintendo
- 2016 has the least sales across the three platforms.
- Sony European sales have been trending upward and have surpassed North American sales with the exception of 2016.
- Microsoft & Nintendo sales have been trending downwards
- Japan sales are mostly Nintendo and Sony although sales for both have been trending downward.

FIVE-YEAR REGIONAL SALES TREND BY PLATFORM



Summary

- Sony has the most publishers making games for its Platforms.
 - Nintendo is losing market share rapidly.
 - Although Sales have been on decline across the three major Gaming Platforms since 2008, Sony is the dominant Gaming Platform leading in Sales globally.
 - Historically North America has been the largest sales region for Sony however, its European sales have surpassed North American sales in the last several years.
 - Action, Sports and Shooter are the top three genres and Sony leads in these categories.
 - Sony should continue to focus on European Sales and look into why North American sales are slowly in decline.
 - Further investments in the top three genres (Action, Shooter & Sports) may improve sales.
-

The background of the slide is a dark teal color with a pattern of stylized, glowing influenza virus particles. These particles are spherical with a textured surface and numerous small, protruding spikes or hemagglutinins. They are scattered across the frame, with some appearing more prominent than others, creating a sense of depth and movement.

Influenza Season

- Perform descriptive analysis based on influenza historical data to provide recommendations based on trends in influenza and how the medical staffing agency should plan for staffing needs across the country

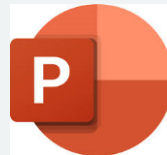
Goal, Data limitation, and Skills

- Analyze US Influenza Season data and support the medical staffing agency to develop an action plan and determine when and where to send the additional staff, to treat the extra patients .
- Population Data: Time Lag it is present, since the US census population data is carried each 10 years, which means it may lead to misinterpretation about the US population depending when it is consulted. In this case, the data in this project is from 2009 to 2017, thus it does not represent the current reality of the US population.
- Influenza Deaths: it was possible to noticed omitted information as the raw data has 8% of their information as "Not Stated" and 81% as "Surpressed". Thus, the analysis and recommendations were limited to the elderly population (65+ years), even though the vulnerable group involves a wider category of people, such as pregnant or children under 59 months.
- Data Profiling and Integrity
Data Quality Measures
Data Transformation and Integration
Hypothesis Testing
Data visualization and Storytelling with Tableau

Data Source

Project Brief Population Data - By US Census Bureau Influenza Deaths - By CDC Influenza Visits Data Set - By CDC Influenza Lab Test Data Set - By CDC Survey of Flu Shots in Children - By CDC

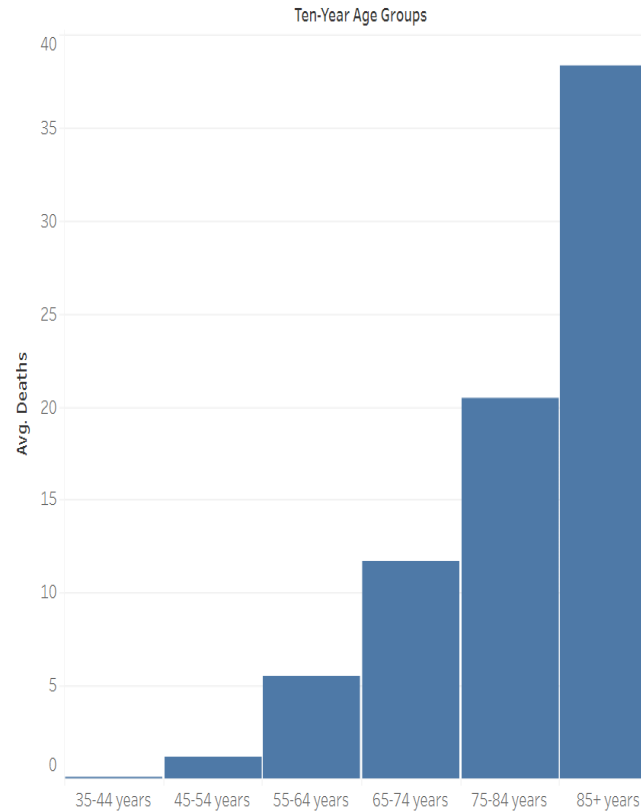
Tools:



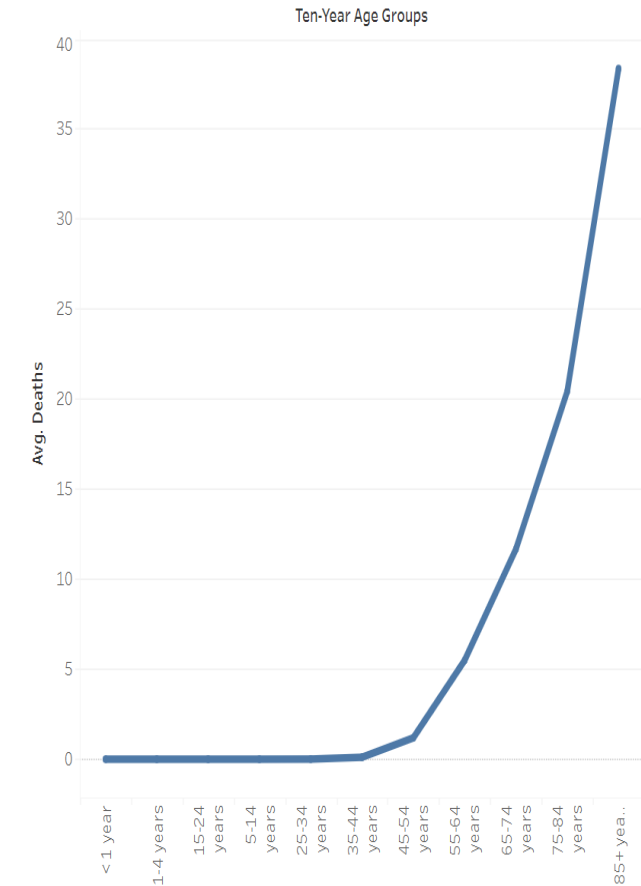
Influenza Mortality Rate by age-group

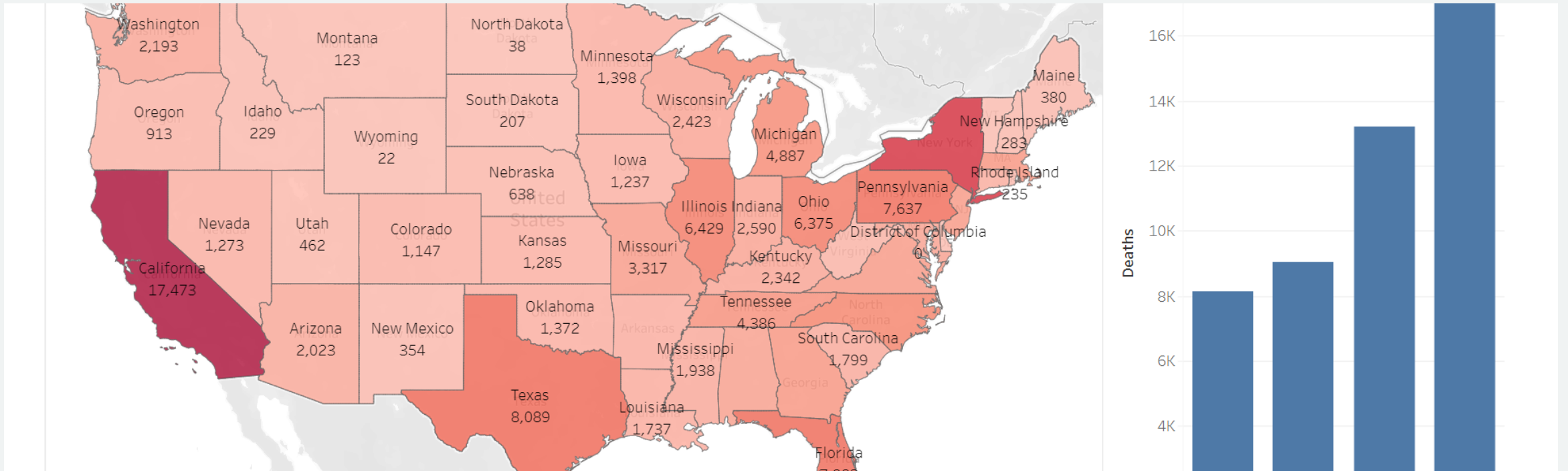
- I intergraded the age groups into 10-year age groups
- Ages 35-44 years don't have the highest mortality rate.
- If these people are not helped, they will likely pass away.

Influenza Mortality Rate: Age-Adjusted Analysis for Individuals Aged 45 and Above



Age-Group-Curve





States with highest influenza mortality rates by individuals aged 45 and above

- The Heat Map shows California, New York and Texas are the highest States with total deaths.
- The Bar Graph shows the deaths in seasonality: Summer, Fall, Spring, and Winter. Winter has the highest death rates since cold weather affects our bodies.

Recommendations

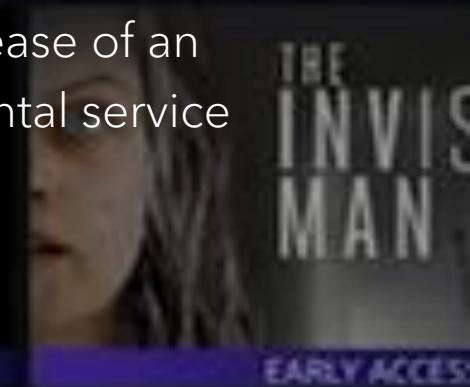
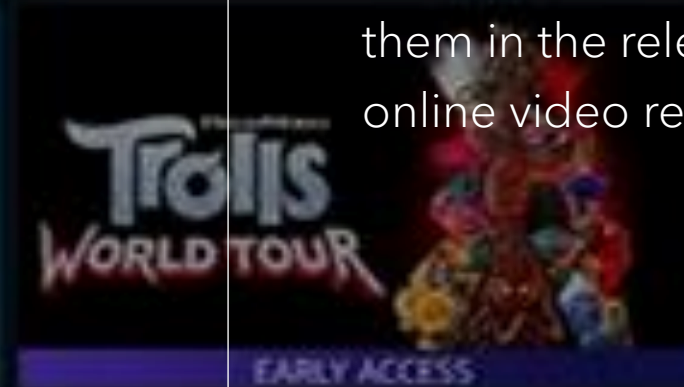


- As previously slated, Influenzas ranks within top 10 for the highest mortality rate and particularly amongst the elderly's ages 45 and above. With more efficient deployment of medical resources, we can help to flatten the curve and improve patients care. I recommend several proactive measures.:
- Deploy medical resources to Staes based on the Influenzas Death Rate proportions outlined in the treemap.
- Increase providing flue Shits in Staes with higher per capita death rate during the CDC recommended months by September and October.
- Improve patient care by increasing the number of providers to Staes with a higher patient-to-provider ratio

Rockbusters

- Analyze Rockbusters data through SQL queries and provide answer to all ad-hoc business questions from various departments inside the company to support them in the release of an online video rental service

RENT OR BUY Popular movies [See more](#)



RENT OR BUY New release movies [See more](#)



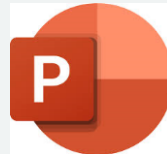
Goal, Data limitation, and Skills

- Provide support to Rockbuster Stealth's Business Intelligence department to help with the launch strategy for their new online video service.
- The data provided serves to show the customer base, films, categories, employees, among other information related to the company's operation. However, it does not provide information related to historical data or that would allow a market trend analysis or forecast..
- Filtering Data in SQL
- Summarizing and Cleaning Data in SQL
- Joining Tables
- Performing Subqueries
- Common Tables Expression
- Presenting SQL Results
- Visualization and Storytelling in Tableau

Data Source

https://images.careerfoundry.com/public/courses/data-immersion/A3/A3_Data_Project_Brief%20.pdf

Tools:



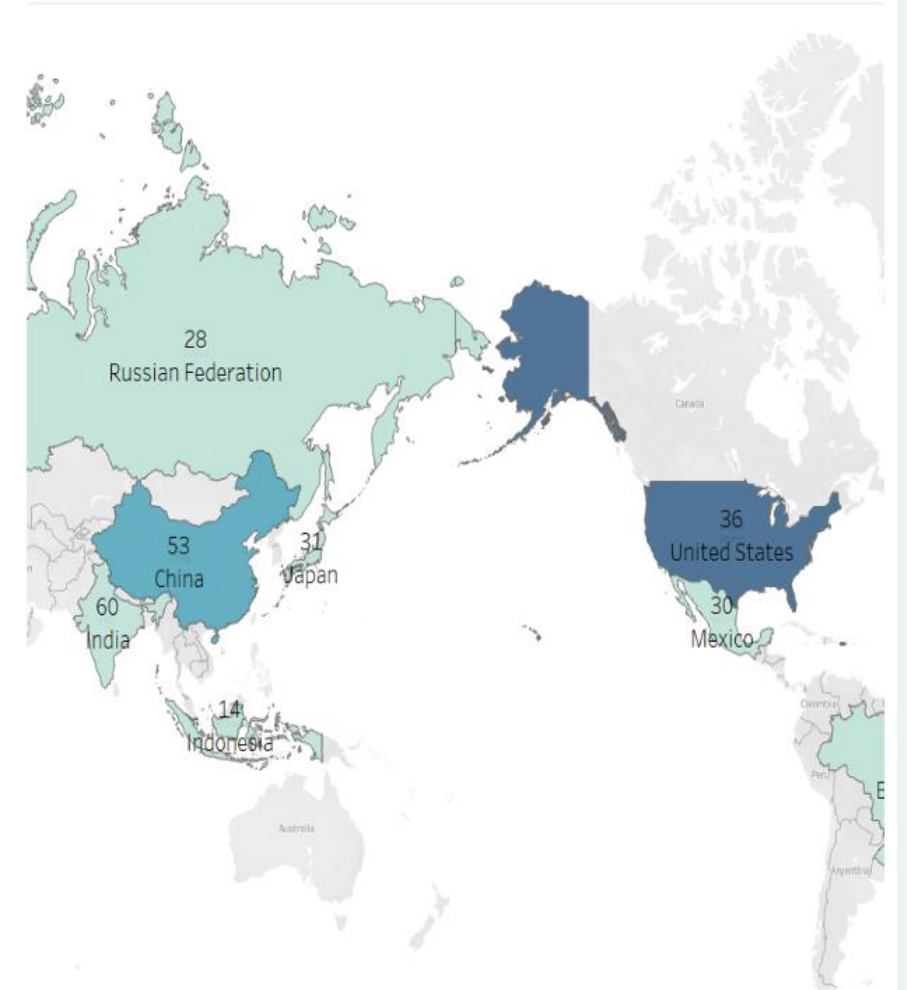
Overall Analysis

- 1,000 unique films
 - 17 Genres
 - Only one language: English
 - 599 customers
 - Length: Average 115 minutes Minimum 46 minutes Maximum 185 minutes
 - Rental Days: Average 5 days Minimum 3 days Maximum 7 days
 - Replacement Cost Average \$ 19,98 Minimum \$ 9,99 Maximum \$ 29,99
-

Overall Analysis

- Rockbuster customers are spread around the world in 108 distinct countries. The top 10 countries in terms of number of customers are India, China, United States, Japan, Mexico, Brazil, Russian Federation, Philippines, Turkey, Indonesia

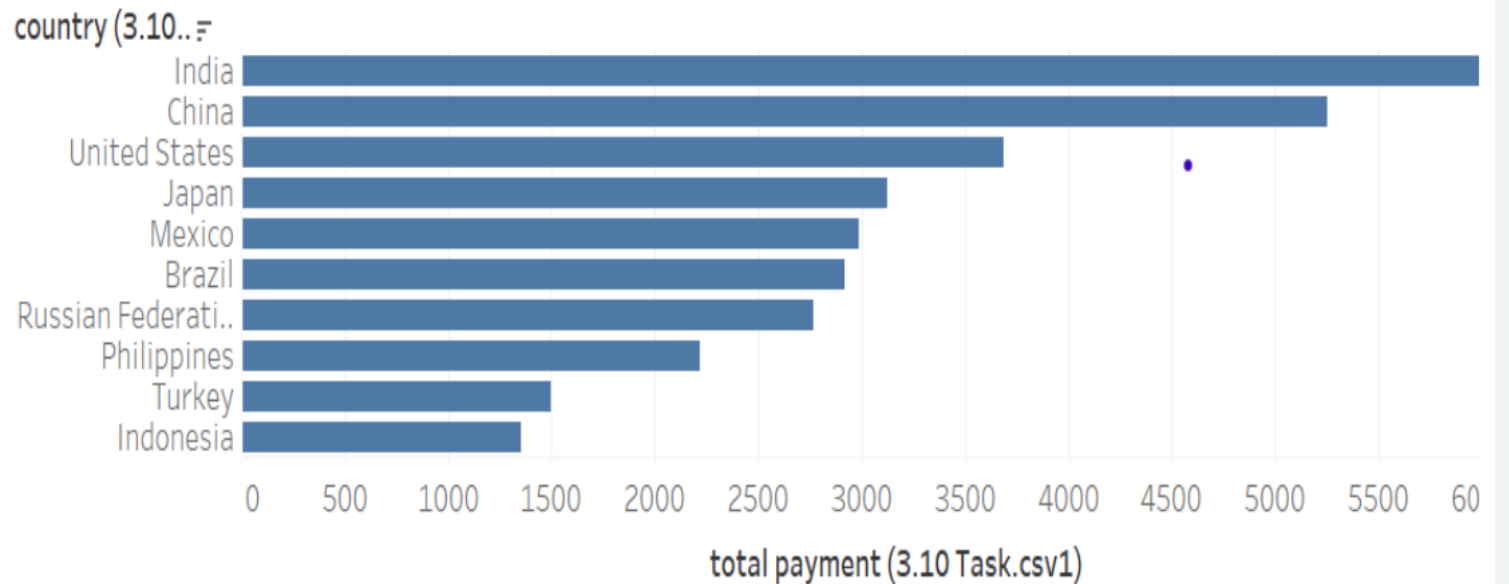
	country character varying (50)	customer_count bigint
1	India	60
2	China	53
3	United States	36
4	Japan	31
5	Mexico	30
6	Brazil	28
7	Russian Federation	28
8	Philippines	20
9	Turkey	15
10	Indonesia	14



Overall Analysis

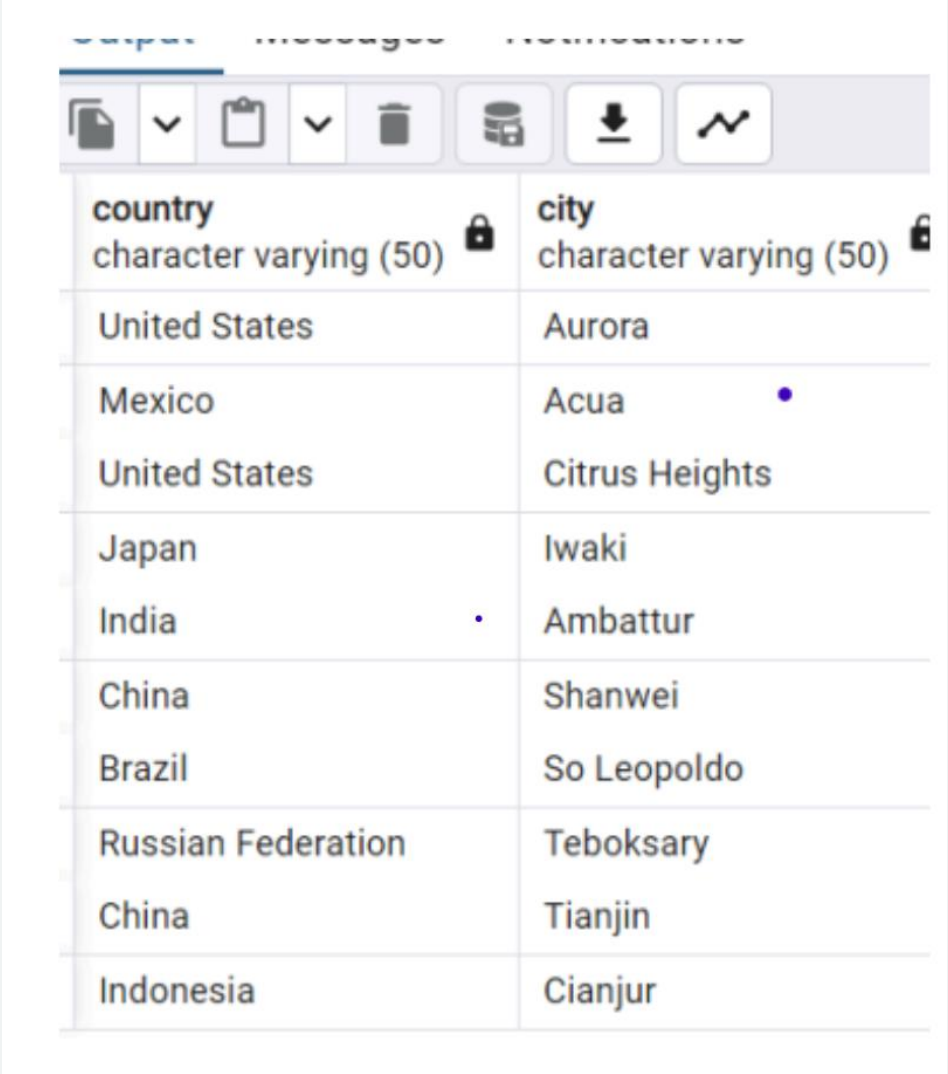
- 10 countries have revenue greater than 1.000. The top 10 countries in terms of customers are the same top 10 countries with higher revenue.

Countries with total payment



IN WHICH CITIES DO THE CUSTOMERS LIVE

- Within the Top 10 countries the top cities are:
- Aurora, United States Acua, Mexico Cireus Heights, United States Iwaki, Japan Ambattur, India Shanwei, China So Leopoldo, Brazil Teboksary, Russian Federation Tianjin, China Cianjur, Indonesia



The image shows a screenshot of a data table with two columns: 'country' and 'city'. Both columns are labeled 'character varying (50)' and have a lock icon. The table lists the top 10 countries and their corresponding top cities. The interface includes a toolbar with icons for file operations, a search bar, and a chart icon.

country character varying (50)	city character varying (50)
United States	Aurora
Mexico	Acua
United States	Citrus Heights
Japan	Iwaki
India	Ambattur
China	Shanwei
Brazil	So Leopoldo
Russian Federation	Teboksary
China	Tianjin
Indonesia	Cianjur

WHO ARE THE TOP 5 CUSTOMERS?

- 1. Arlene Harvey From Ambattur, India \$ 111,76
 - 2. Kyle Spurlock From Shanwei, China \$ 109,71
 - 3. Marlene Welch From Iwaki, Japan \$ 106,77
 - 4. Glen Talbert From Acua, Mexico \$ 100,77
 - 5. Clinton Buford From Aurora, United States \$ 98,76
-

Recommendations

- Increase offer in terms of:
 - Number of films
 - Genre
 - Language
 - Pricing
 - Differentiate the price according to the duration of the rental
 - Create a loyalty scheme
 - Working together with the top clients and promoting them as Rockbuster Ambassadors, the goal is to increase customers and rentals
 - After implementing the new strategy, continue analyzing insights and communicating new recommendations based on what data says
-

Instacart Basket

- Derive insights and suggest strategies for increase sales based on initial data and exploratory analysis of Instacart data by python libraries, such as Pandas, Numpy and Matplotlib.

Goal, Data limitation, and Skills

- Support Instacart to uncover more information about their sales patterns. Also, perform an initial data and exploratory analysis to derive insights and suggest strategies for better segmentation based on the provided criteria.
- The data has time limitation, since comprise information from 2017 only. The customer data provided by CareerFoundry was fabricated and used for the solo proposes of study.
- Python Libraries - Pandas and Numpy
- Data Wrangling and Subsetting
- Data Consistency Check
- Combining and Exporting Dataframe
- Deriving, Grouping and Aggregating Data
- Data Visualization with Python - Seaborn, Matplotlib and Scipy
- Coding Etiquette

Data Source

https://images.careerfoundry.com/public/courses/data-immersion/A4/A4_Data_Immersion_Project_Brief.pdf

Customer Data Set - By CareerFoundry

Tools:



Busiest days

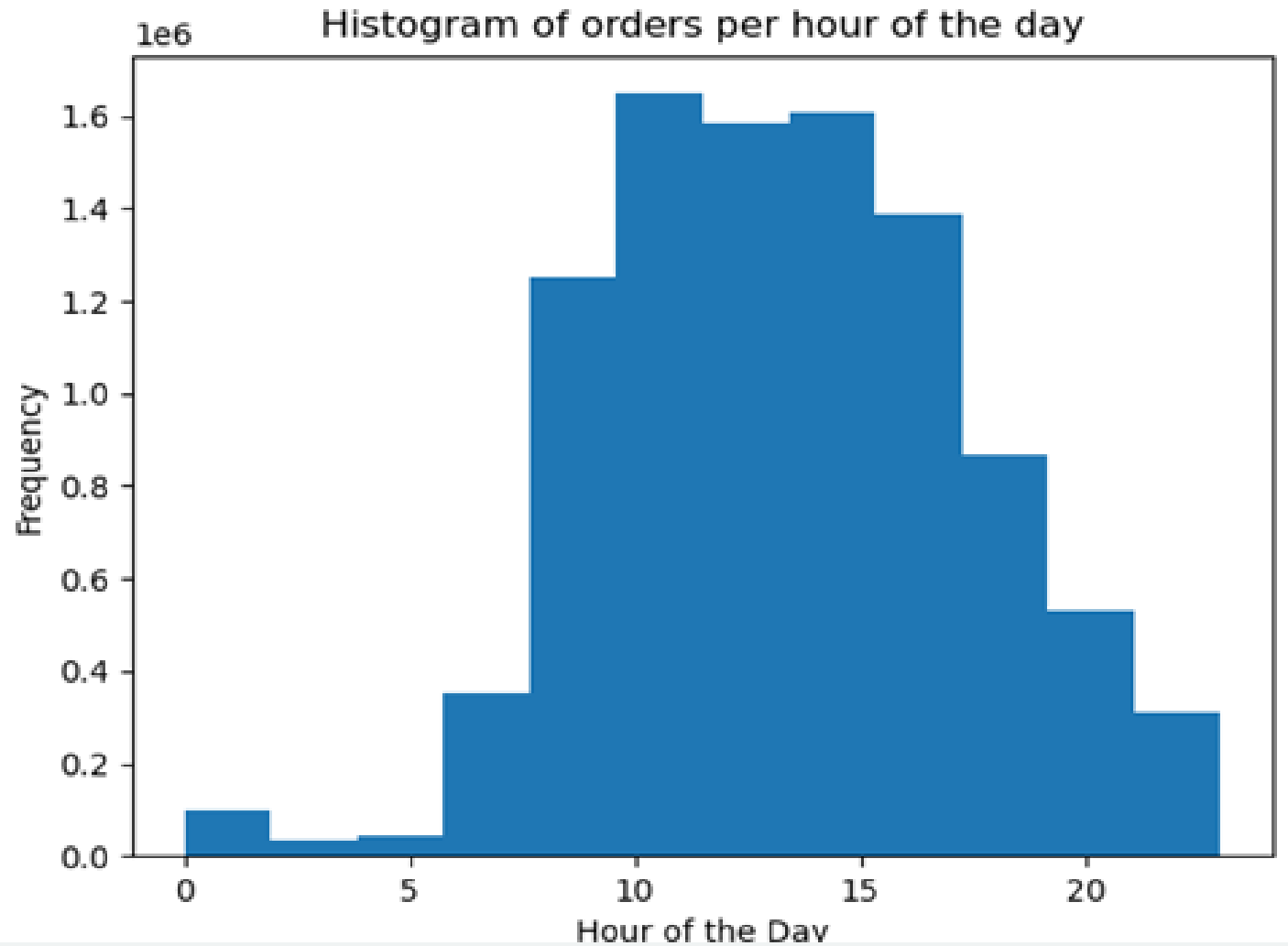
Saturday (0) and Sunday (1) are the busiest days



How busy is each day of the week according to number of orders. Saturday and Sunday are the busiest days and Tuesday and Wednesday are the least busiest days.

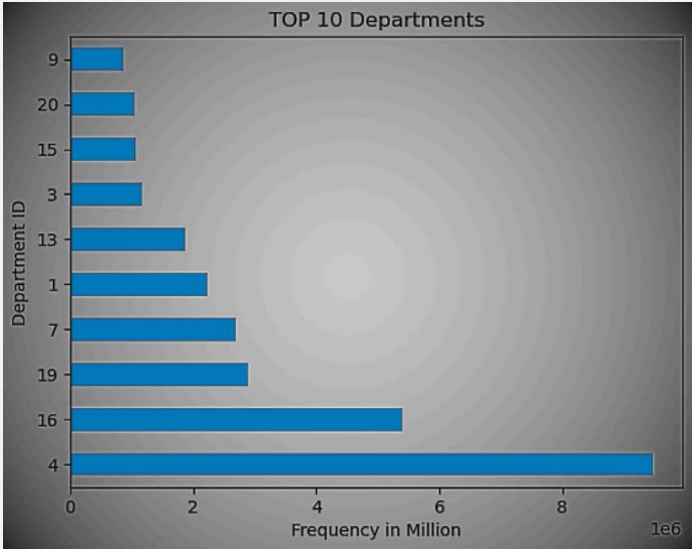
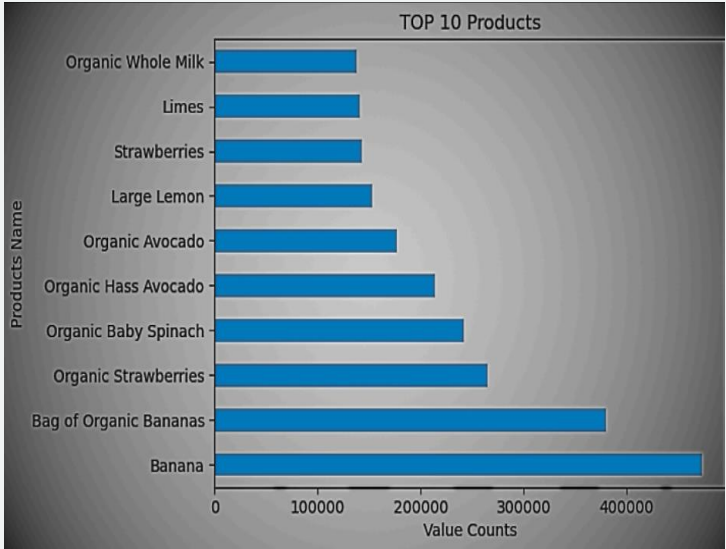
Busiest Hours

The busiest hours, it is possible to notice that the access volume increase at 9 am until 3 pm, when start to decrease.



Top Products and Departments

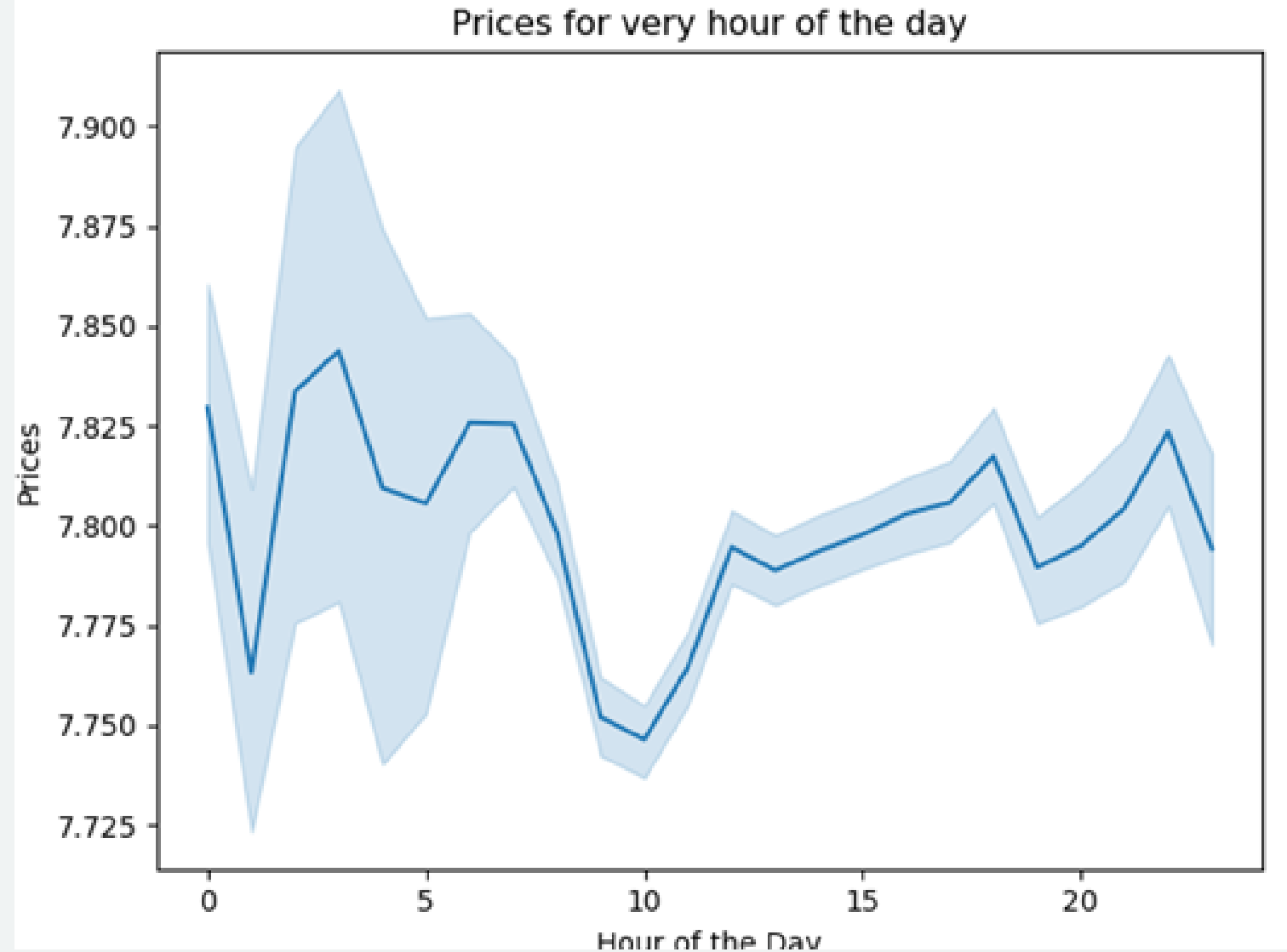
As can be observed, the top 10 products consumed by all regions in the Instacart database are inside the top 1 department in sales frequency, that is "Produce", which could indicate a buying trend by its customers.



Department_id	Department
9	Dry Goods Pasta
20	Deli
15	Canned Goods
3	Bakery
13	Pantry
1	Frozen
7	Beverages
19	Snacks
16	Dairy Eggs
4	Produce

Market Analysis

According to the graph above customers have a higher tendency to spend more between 04 am to 08 am, as well as from 8 pm until 10 pm.



Recommendations

- **Busiest Day**

Saturday (0) and Sunday (1) are the days that have a higher frequency of requests. Thus, these days would be the most recommended for an advertising, as it would reach a higher number of people. As for the best period of time to run these ads, considering that from 09 am to 3 pm Instacart has a higher volume of order, this shows a major activity of customers, compared to the rest of the day, increasing the reach of an ad

- **Top Products**

Market research could be carried out to assess why these products and departments are so popular and why others are not. From this, the responsible department could study measures to be adopted to increase its sales, such as adding a variety of popular products (i.e. banana) or making combos between products that are popular and others that are not , to assess public acceptance for certain products, studying their viability in the market or replacement with another that is more successful.

- **Expenditure Habits**

Considering expenditure habits, it is possible to notice that more expensive products are consumed during the night (8 pm to 10 pm) and early morning (04 am to 08 am). Thus, considering the information regarding the busiest period of time mentioned above, it is possible to use this information to evaluate what type of product should be advertised, depending on the time of day. Therefore, during the night period (8pm - 10pm) and early morning hours (04am - 08am), more expensive products could be highlighted, while during the busiest period (09 am - 3pm), more affordable products could be featured.



FIG E. B A N K

- Analytical support to the anti-money-laundering compliance department through data-related projects that assess client risk and transaction risk, as well as reporting on metrics to support the bank in running their compliance program more efficiently.

Goal, Data limitation, and Skills

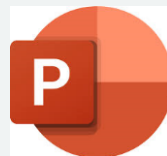
- Provide analytical support to company's anti-money-laundering compliance department, through a variety of data-related projects that help the bank assess client risk and transaction risk, as well as reporting on metrics. Also provide support to help build and optimize models that assist the bank in running their compliance program more efficiently
- Data does not provide date, so it is not possible to observe whether there was a trend of customer loss during a specific period or whether it is something that occurs constantly.
- Intro to Big Data
- Data Ethics: Bias, Security and Privacy
- Intro to Data Mining
- Intro to Predictive Analysis
- Time Series Analysis and Forecasting

Data Source

https://images.careerfoundry.com/public/courses/data-immersion/A5/A5_Data_Project_Brief.pdf

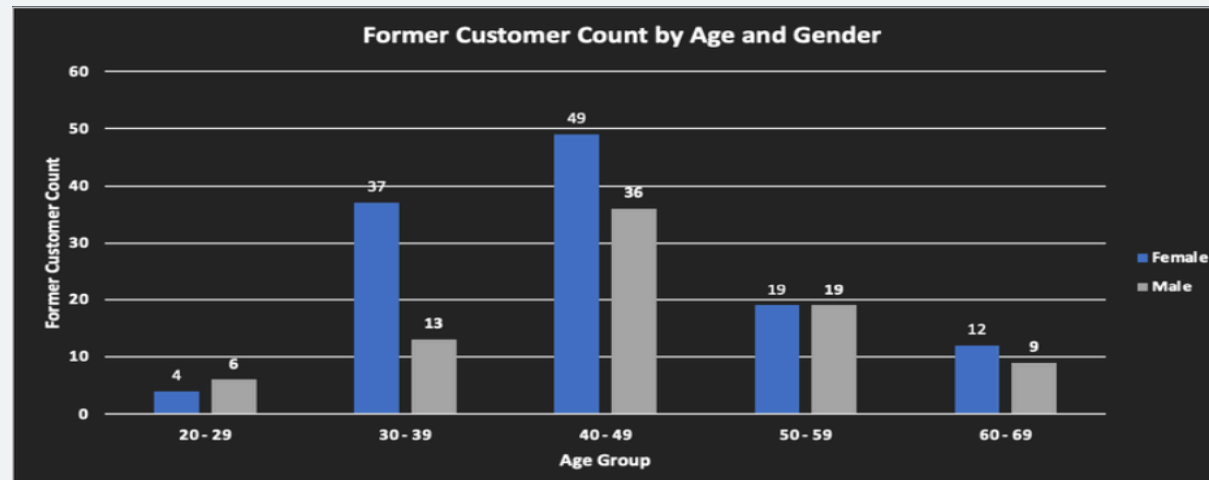
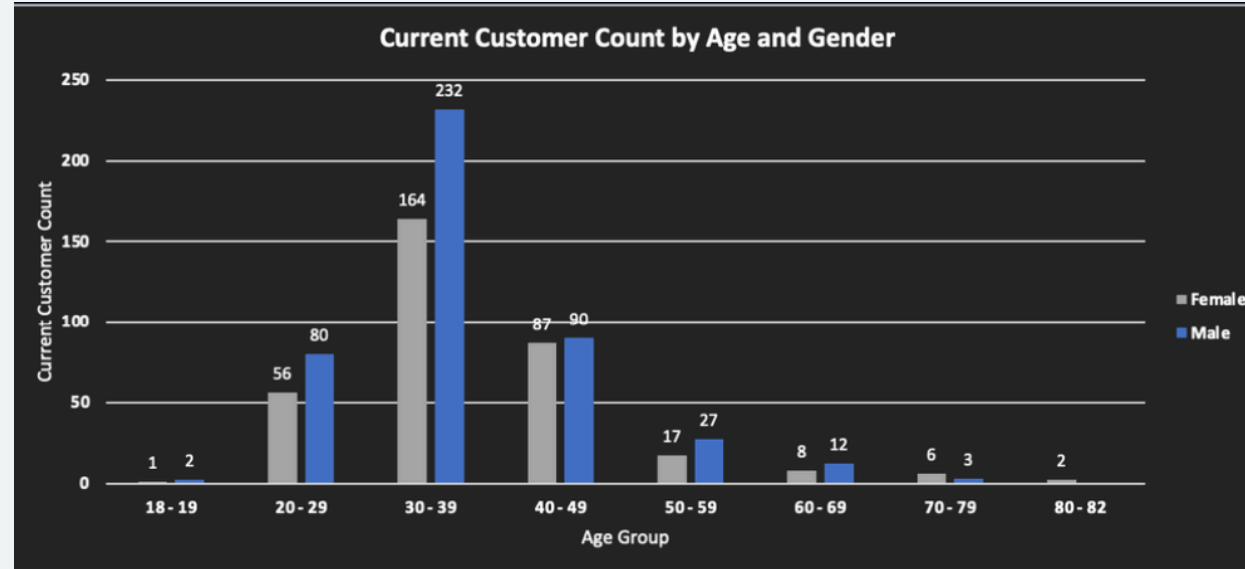
Customer Data Set - By CareerFoundry

Tools:



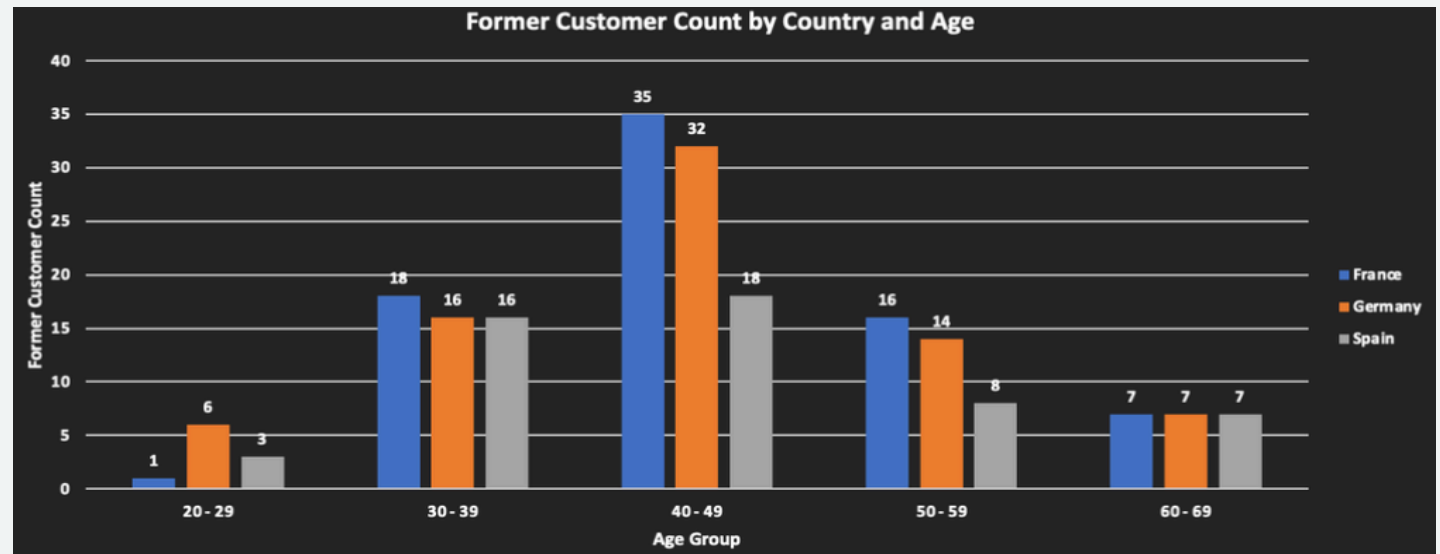
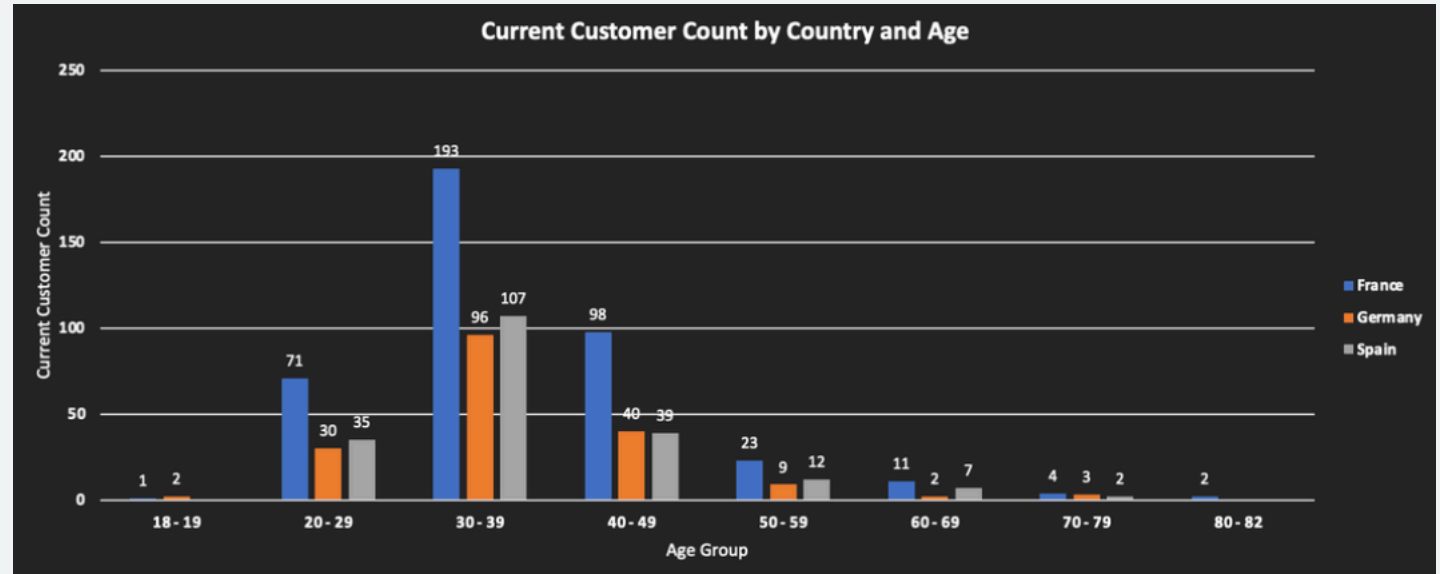
High Risk – Age and Gender

The slides above show that age and gender are risk factors, as: Most of the former customers are in the group of 40-49 years old; Most of the former customers are women.



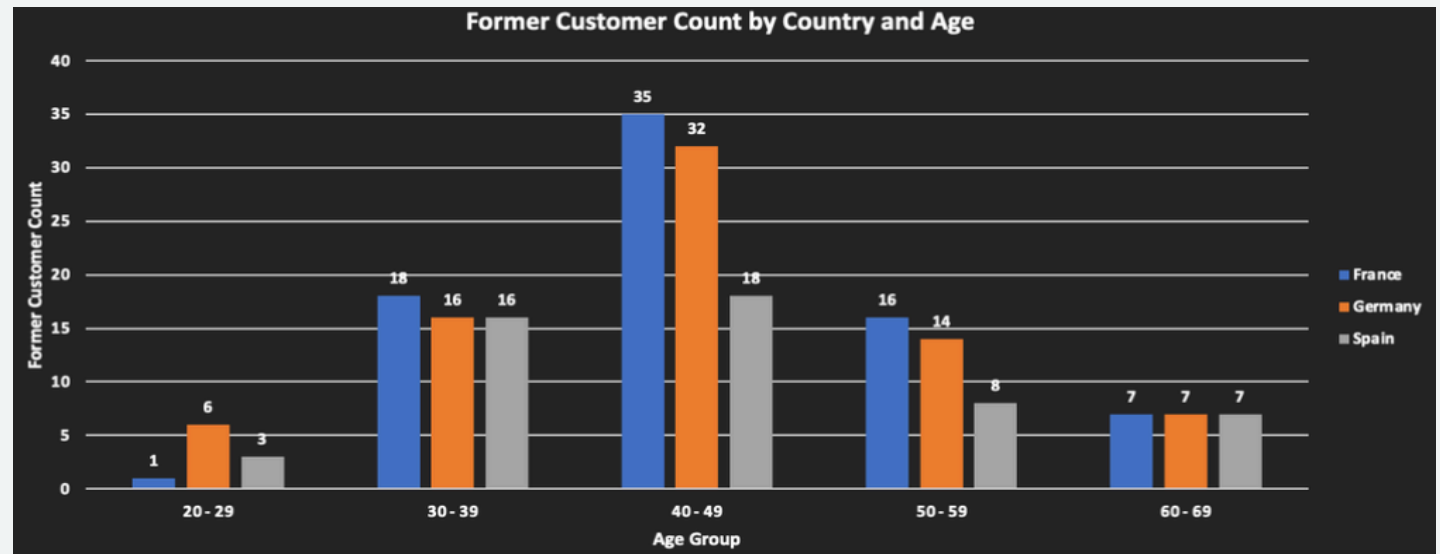
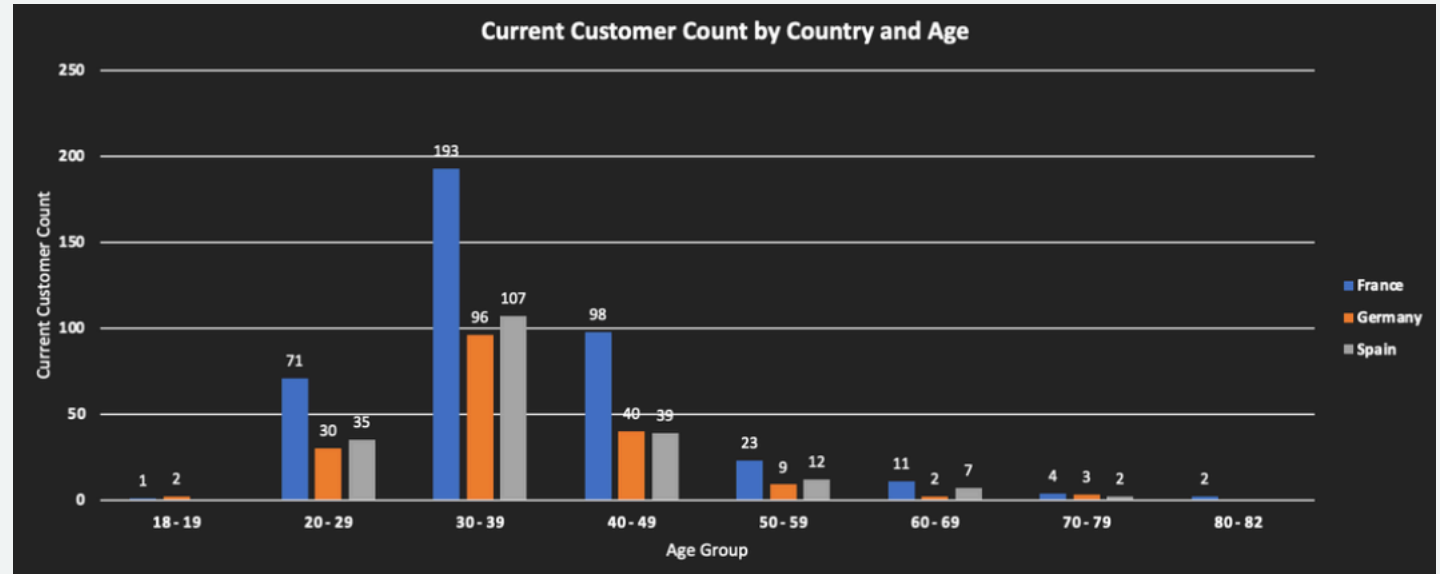
High Risk – Country

Continuing the results shown in the previous slide, Country can also be a high-risk factor as we can notice that there is a high number of people from Germany that leaves the company.



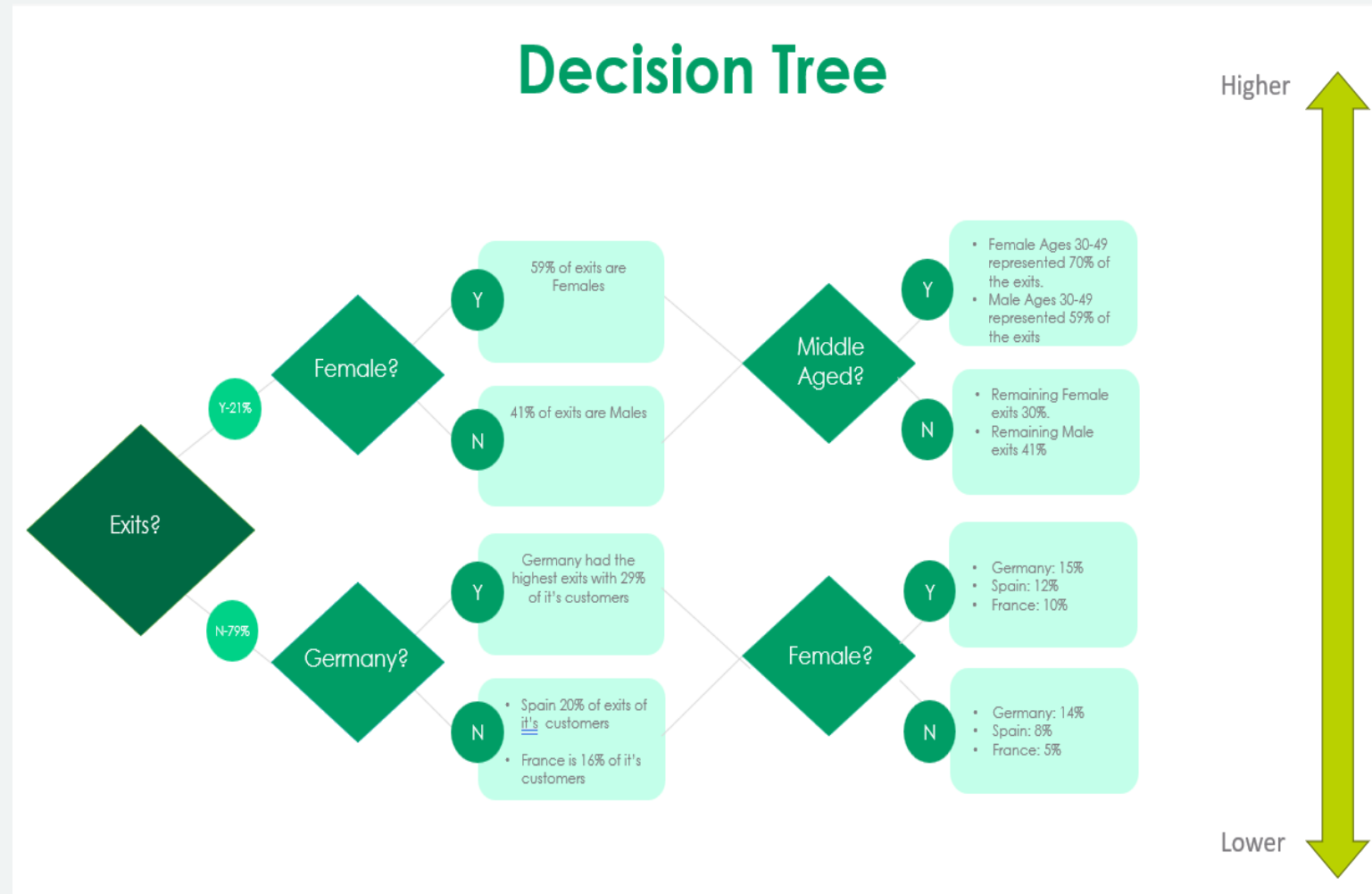
High Risk – Country

Continuing the results shown in the previous slide, Country can also be a high-risk factor as we can notice that there is a high number of people from Germany that leaves the company.



Decision Tree

- The Decision tree shows more Females are leaving banks In Germany than France and Spain combined.
- Spain 20% of exits customers and France has 16# of exits customers.
- The age groups for male and females are 30-49. With 70% females leaving and 59# males leaving.



Recommendations

- Focus on young people demographic Considering that the population under 39 years of age constitutes the majority of customers, Women it would be advisable to create programs or products focused on this category of customers, so that the losses caused by older customers who leave the bank could be mitigated.
 - Therefore, considering that the majority of young customers do not have a consolidated credit score or cannot afford high bank fees, the following services may be attractive:
 - Creation of an account or special products for people up to an age to be determined by the bank, in which they will not be charged special fees or offer services such as access to credit or mortgage at a reduced rate, even without a credit score, so they can start their life. –
 - Sign-Up bonuses or Reward Programs: depending on the age of the person opening the account, any benefit can be taken into account when contracting services, which, combined with benefits such as No-fee banking mentioned above, could result in a customer who would remain with the account for longer
-