

Robert Indelicato

Data Analyst Portfolio

732.915.4785





About Me

My name is Robert, and I am an accomplished Data Analyst skilled in extracting actionable insights from extensive datasets using data-driven approaches to drive growth and propel the business forward.

Introductory Slides - Skillset Analytics



Projects

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

Tools



Microsoft Excel



Tableau



SQL



Python



Github



Jupyter Notebook

Auto Toy Sales

Exploratory Data Analysis of an Auto Toy Company that is selling mini collectables to customers located around the world.

- Explore trends and patterns driving sales
- Geographic heat map
- Product distribution
- Predictive analysis using forecasting
- Summary and recommendations



Auto Toy Sales Company

Data

- Data format - Excel
- Python Libraries - Pandas and Numpy
- Data Wrangling and Subsetting
- Data Consistency Check
- Deriving variables, Grouping & Aggregating data
- Combining, Enriching and Exporting Dataframe
- Automobile Sales data – Via Kaggle

Skills

- Python Libraries - Pandas, Numpy,, Requests, Pylab.
- Machine Learning - Sklearn: Linear Regression, Kmeans, mean_squared_error, r2_score, train_test_split
- Visualization with Python – Folium, Seaborn, Matplotlib & Scipy
- Jupyter Notebook
- Tableau

Challenges

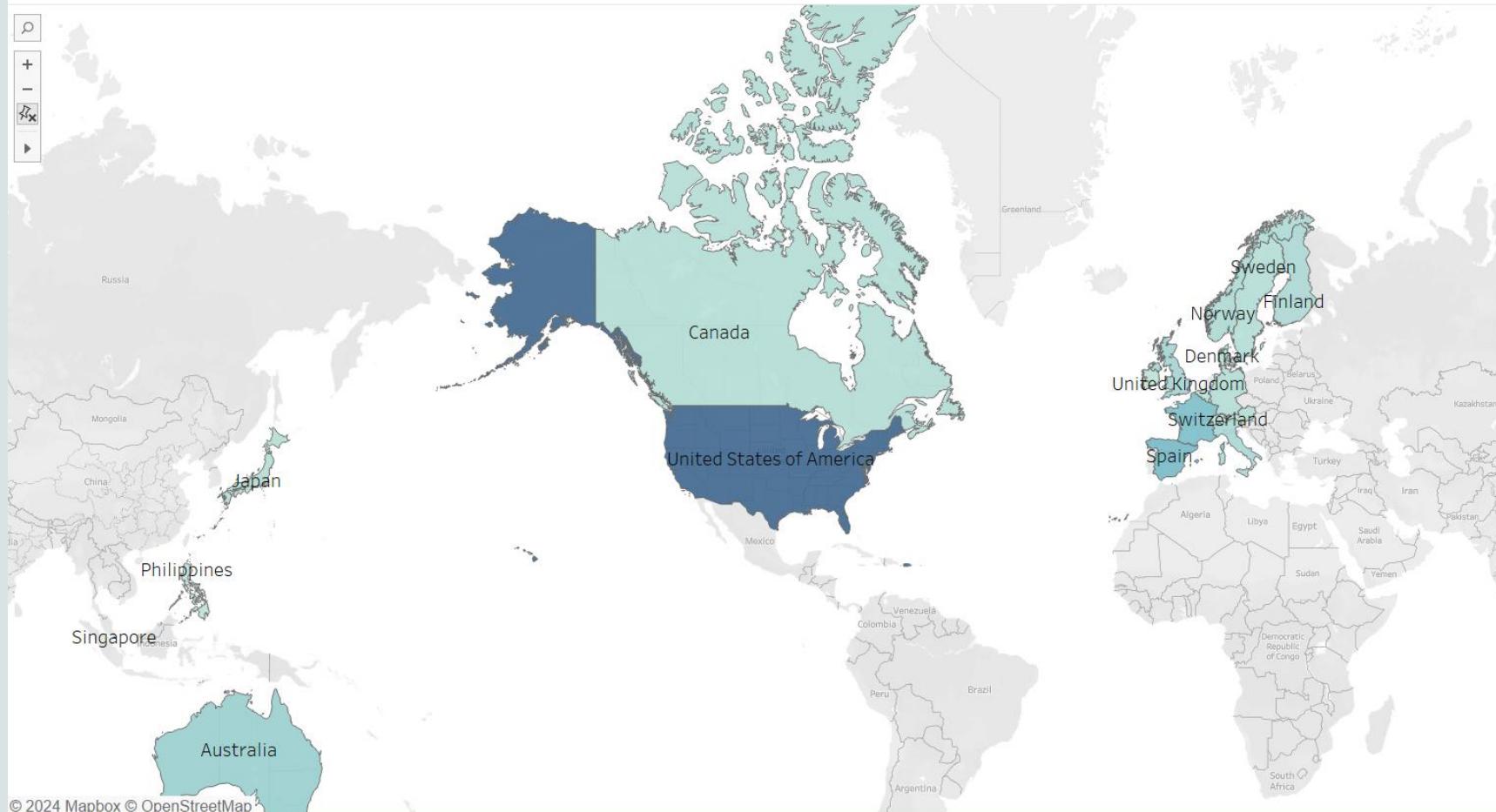
- Data cleaning, missing data, format issues and standardizing the dataset.
- The data has time limitation, only has five months of 2020, not full complete data set.

Insights

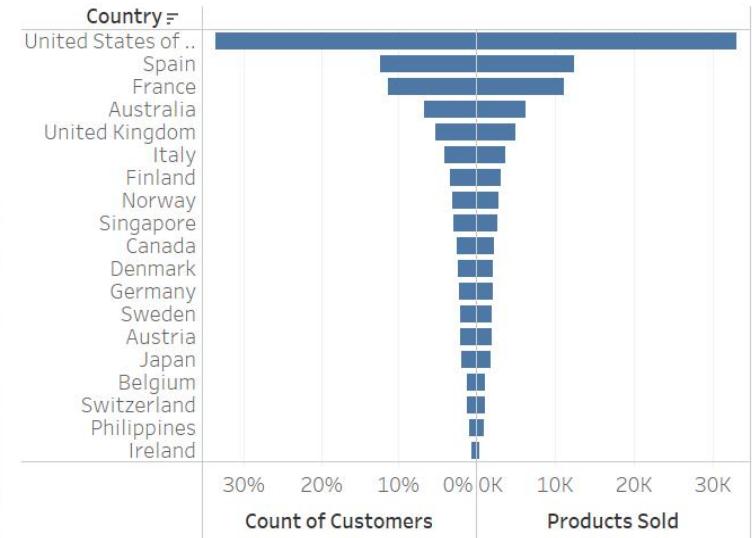
- Few contributing factors on why auto toy sales are healthy
- U.S. and Spain are the 2 highest Distribution countries.
- Classic Cars is the only product with markdown margin, while the rest of the products have high margins.
- Toys cars are highest quantity product and sales.
- In Q1 of 2018 and Q1 2020 Toy sales were surprisingly high.

Geographic Sales: United States is the biggest consumer

Geographic Heat Map

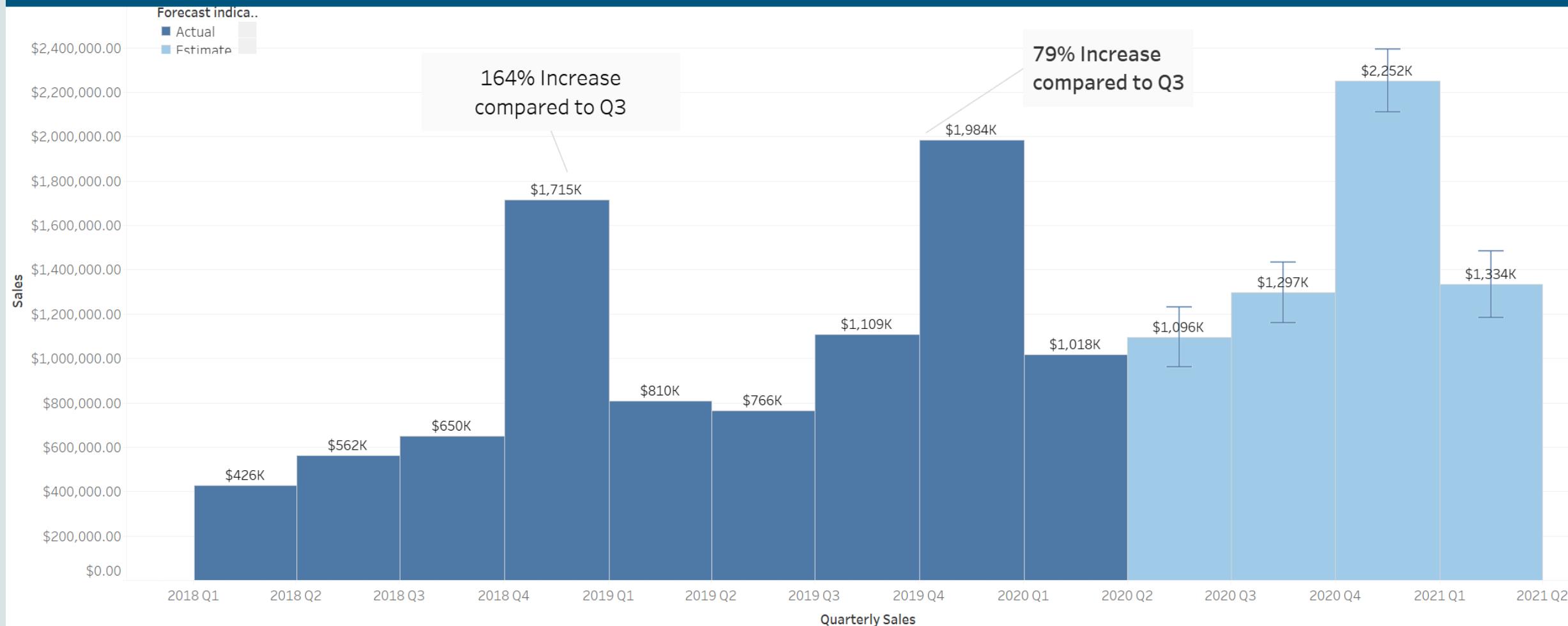


Customers by Country - % of Total



- United States is the largest consumer representing 34% of Sales, followed by Spain and France.

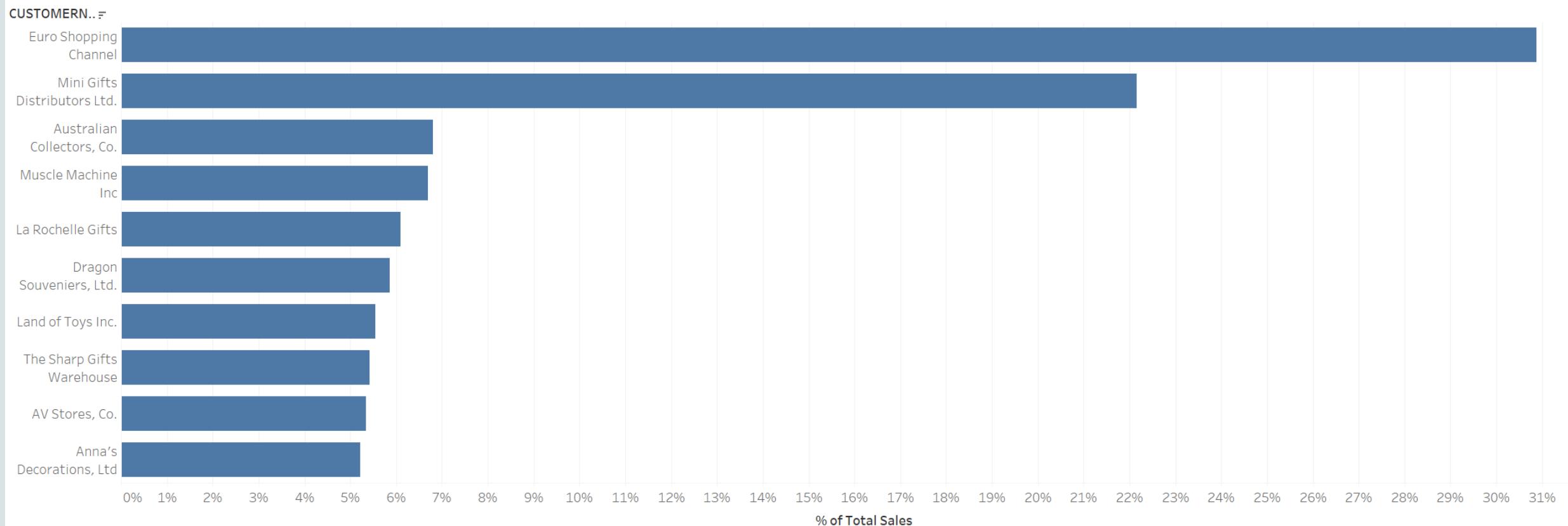
Quarterly Sales Growth: Positive Quarter over Quarter Sales growth.



- Strongest sales are in the 4th Quarter.
- 2020-2021 forecast estimated to be positive.

Top 10 Customers: The top 2 customers represent 53% of Sales.

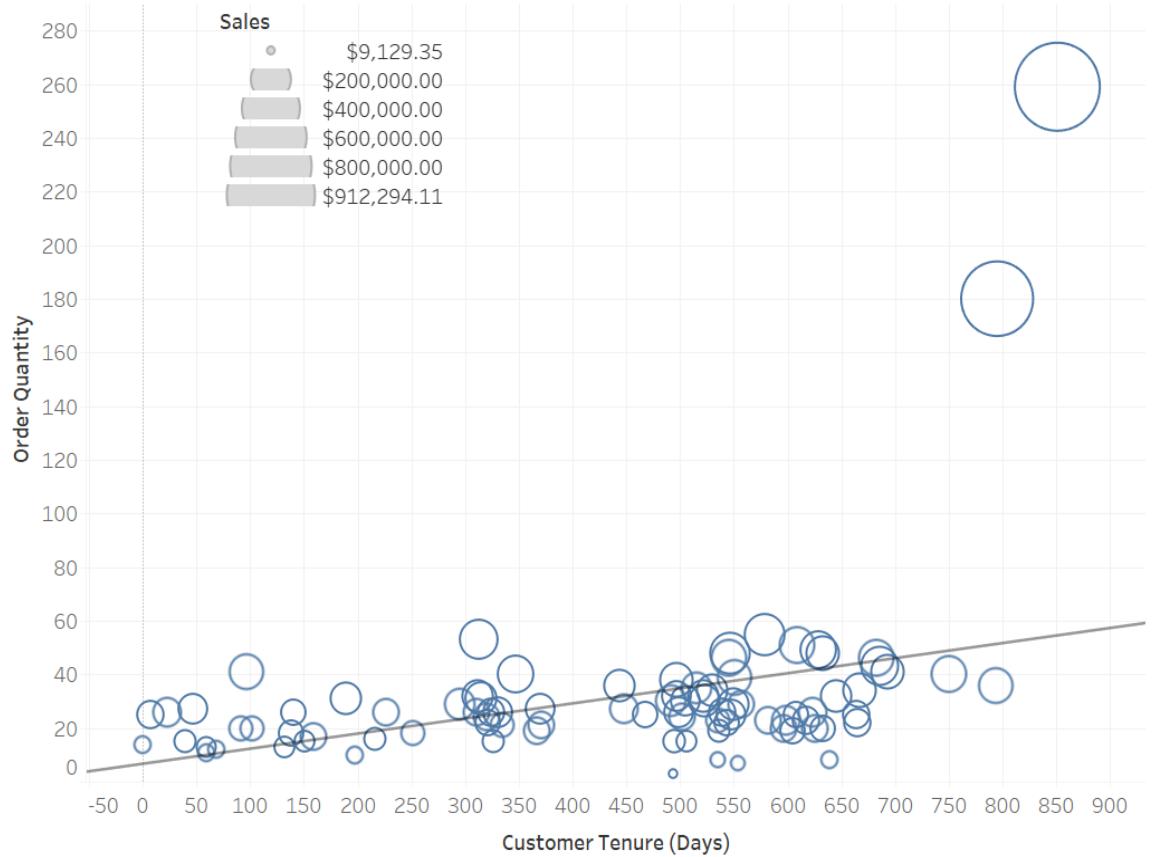
Top 10 Customers as Percentage of Total Sales



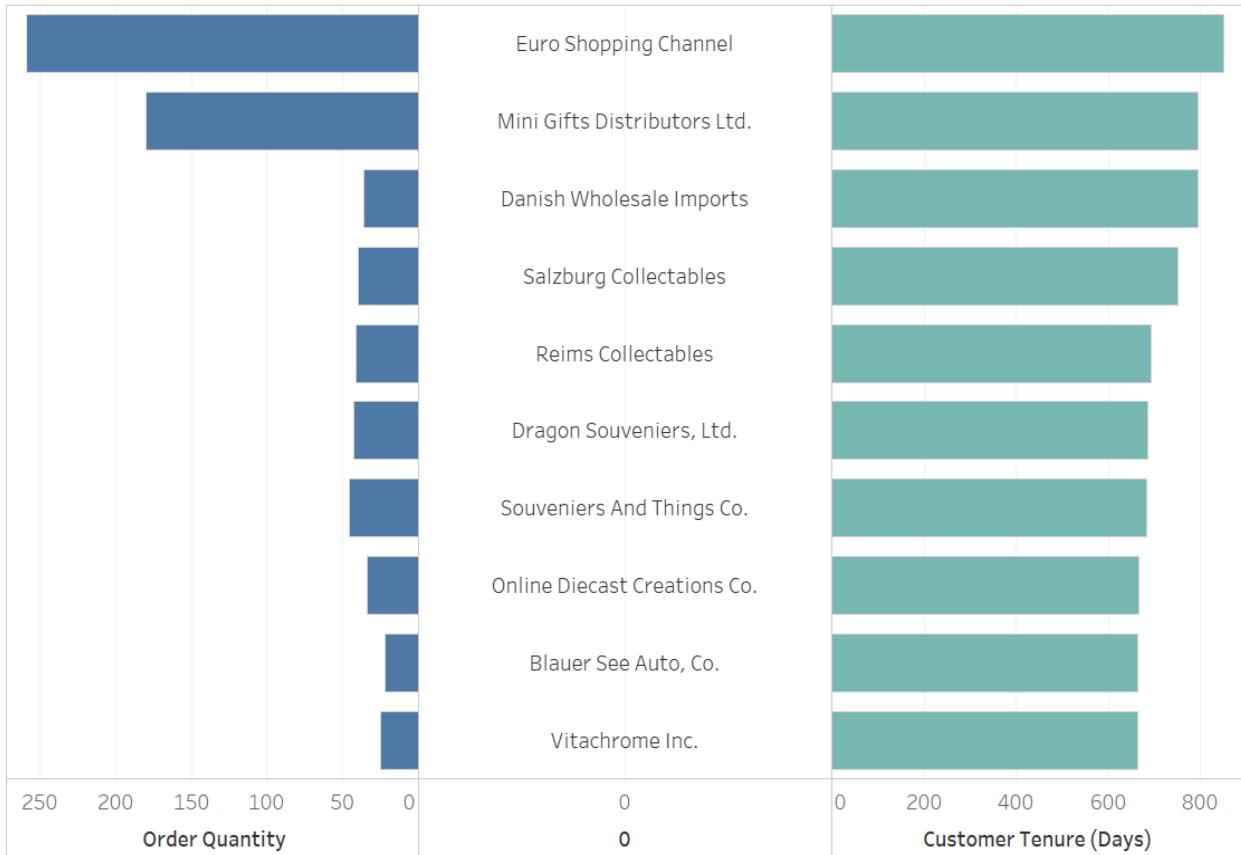
- Euro Shopping Channel followed by Mini Gifts Distributors are the biggest customer by far.
- The top 2 customers represent 53% of sales which can be a risk to the business. Further considerations should be given to expand the larger customer base.

Customer Retention: Customer Loyalty is healthy.

Customer Tenure Correlation



Customer Loyalty

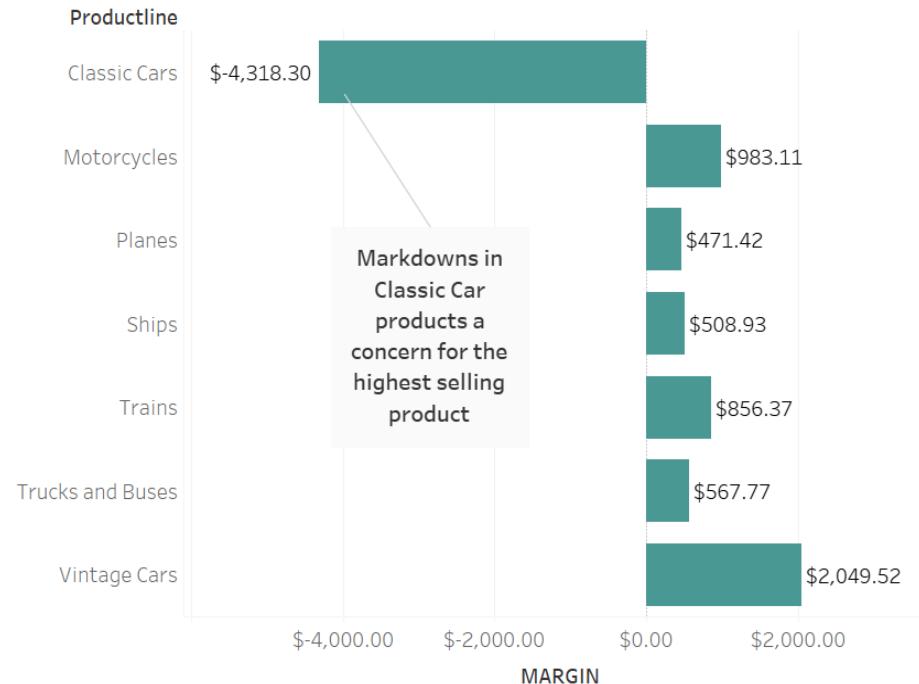


- Positive Coreelation between Customer Tenure and Order Quantity.
- The top 10 customers have stayed with the company for at least 1.8 years and are producing majority of the orders.

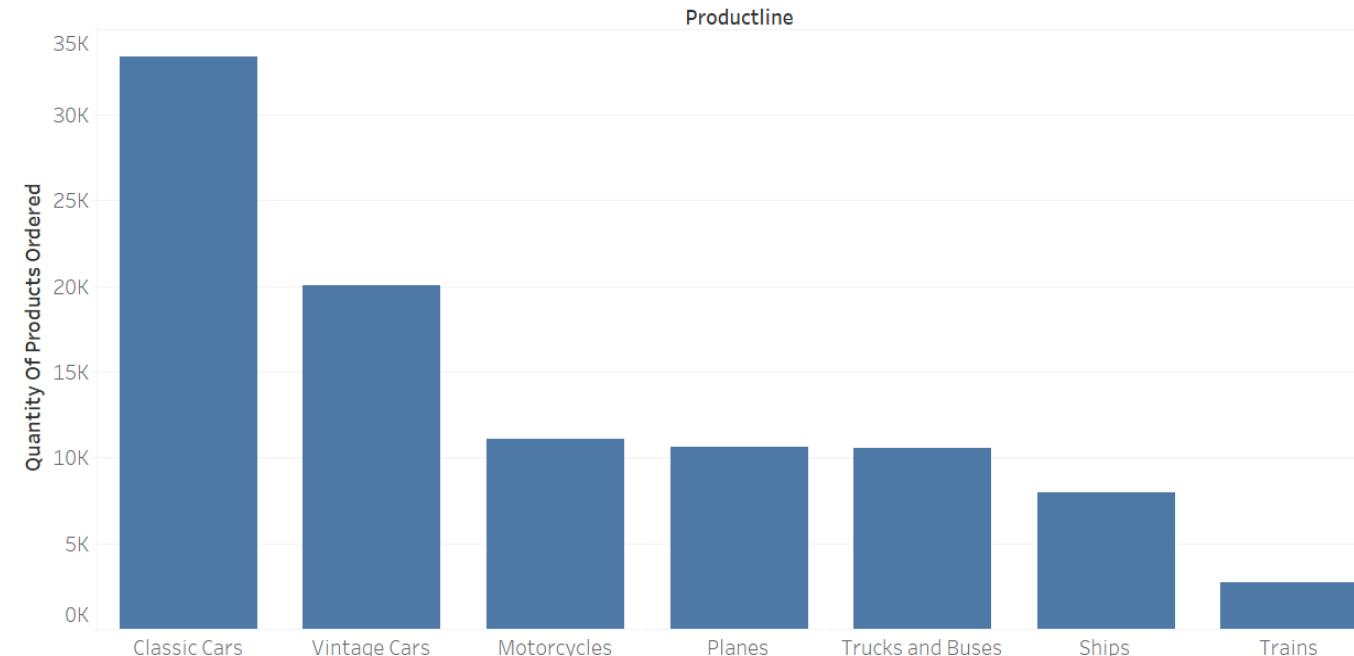
Product Margins:

Markdowns in Classic Cars Product line can be cause of concern.

Markdowns / Markups



Quantities Sold by Product



- The data set does not contain the Cost of Goods Sold (COGS), therefore we applied margins to the MSRP and actual sales price.
- Classic cars are the best selling products however they are marked down from the MSRP price. Reducing COGS or raising retail prices can improve the margins and lead to profitability.
- Margins overall are positive as shown in the Grand Total

Product Distribution: Prioritize distribution based on Cities w..

Quantity ordered

24 3,816



- Prioritize product distribution by stocking best-selling Cities adequately and avoid excess inventory in low-selling Cities.

Recommendations

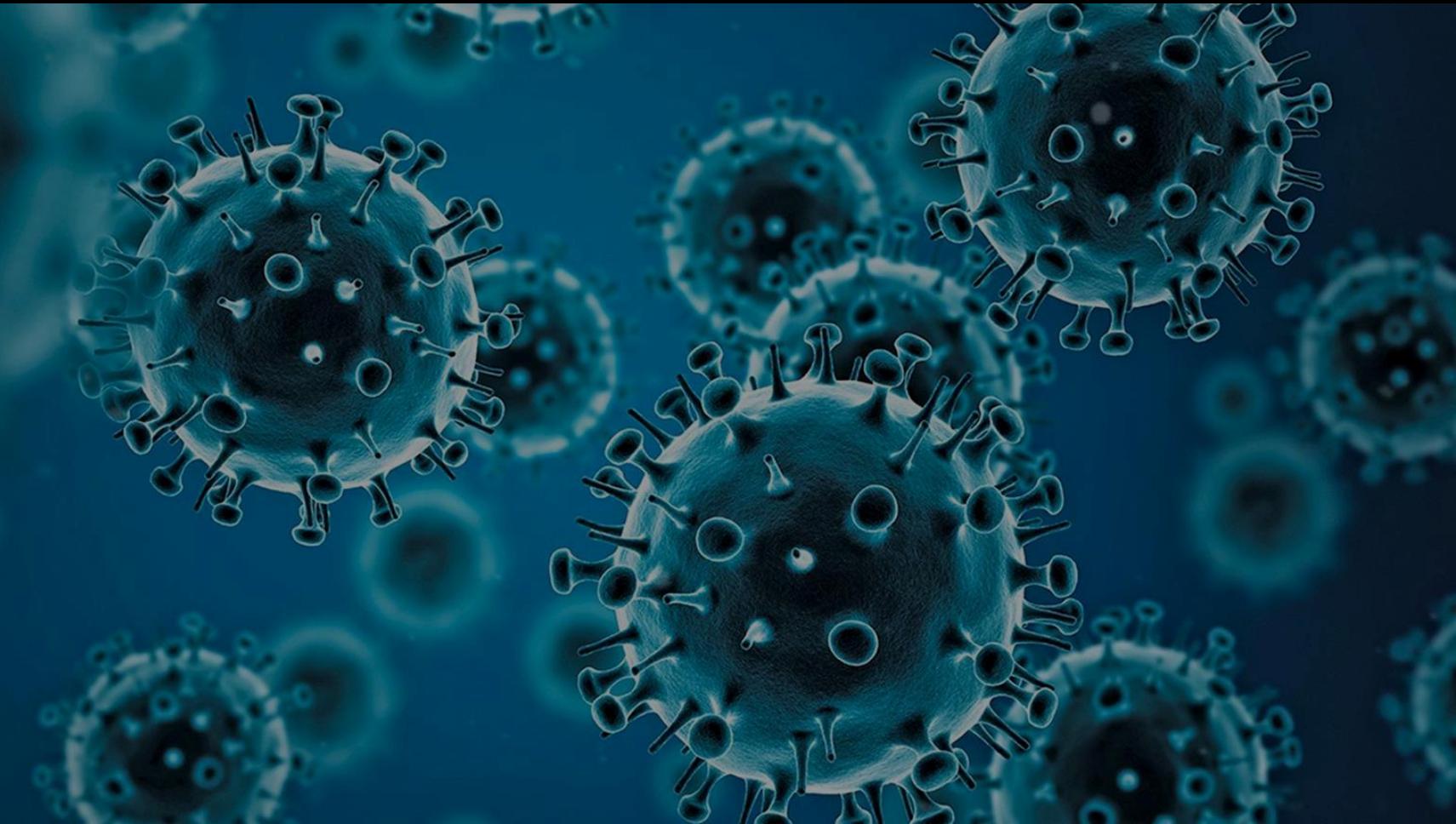
The Auto Toy Sales Company is profitable with strong sale trends as evident in Q/Q sales growth and steady customer loyalty. 2019 produced \$4.7M in sales and there is no reason to believe that 2020 could not have been even better if we had complete data as shown by our forecast model. Margins remain healthy overall for 6 out of 7 products but one was a cause for concern.

The following recommendations should help to further boost profitability and growth:

- Address Classic Car Markdowns which is our highest selling product. Some examples include reducing manufacturing cost and increasing the selling price.
- Capitalize on the seasonality as November and December are the strongest selling Months.
- More efficient product distribution based on the Cities with the highest demand.
- Expansion into new countries and markets.

Influenza Season

Assisting a medical staffing agency in preparing for influenza season by analyzing trends to proactively plan staffing needs nationwide.



Influenza

Data

- Data format - Excel CSV.
- Regions: North America.
- Pivot Tables and Charts.
- Population Data - 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

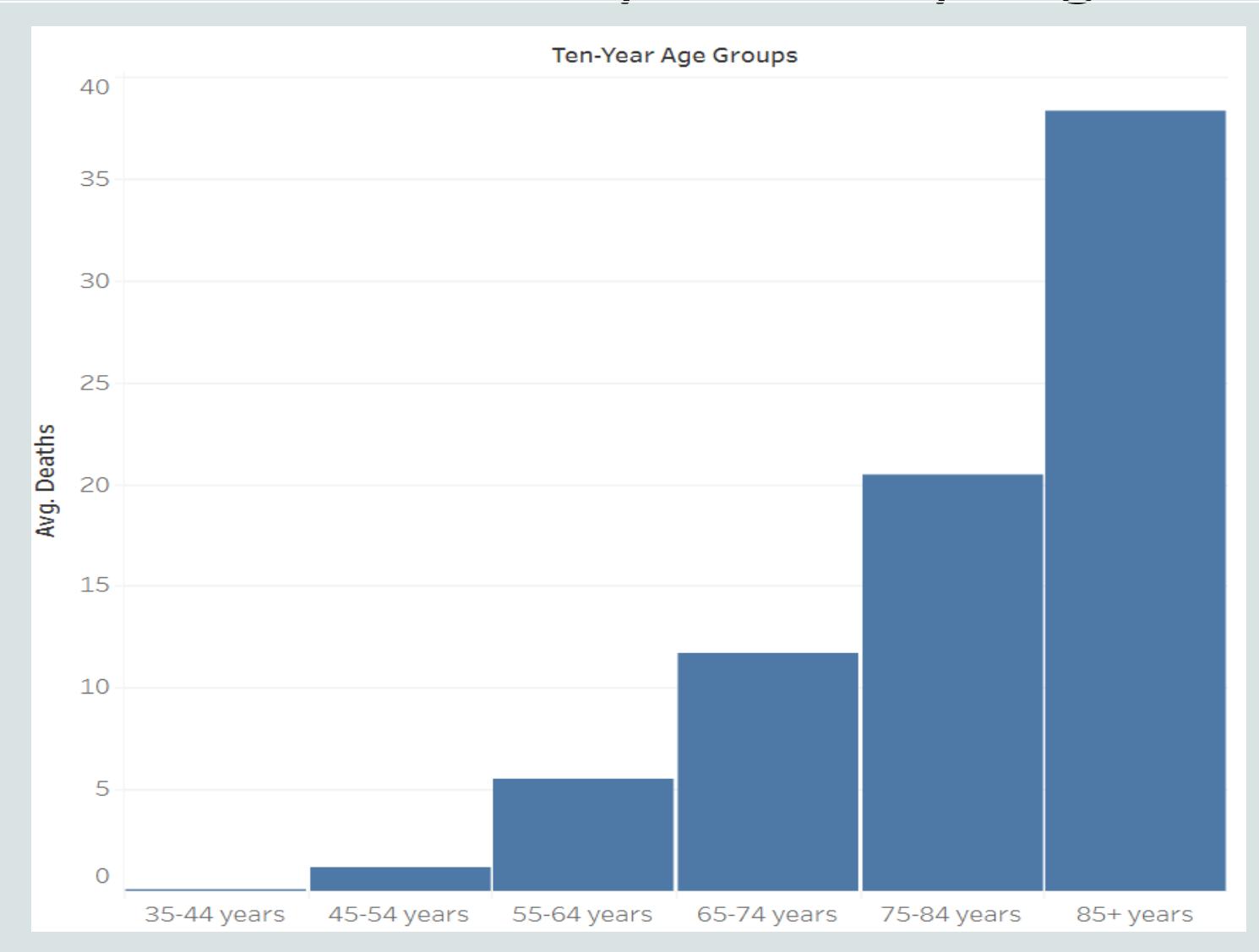
Skills

- Data cleaning, integration and transformation
- Data Profiling and Integrity
- Statistical hypothesis
- Testing
- Tableau/Visual analysis
- Forecasting

Challenges

- Data cleaning, missing data, format issues and standardizing the dataset.
- Possible notice omitted info as raw data has 8% not stated and 81% as suppressed, thus info was limited to elderly, even though vulnerable group involves wider category of people.

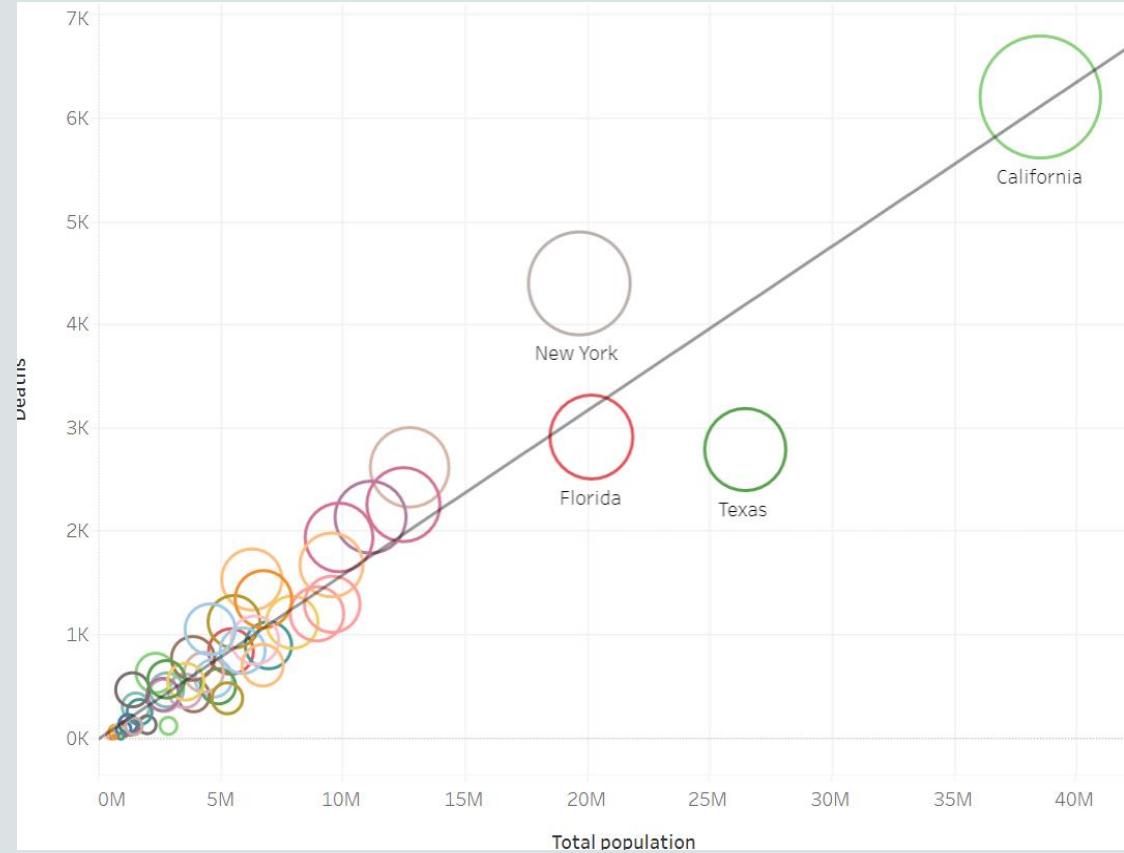
Influenza Mortality Rate by Age-Group



The impact from Influenza significantly increases with age.

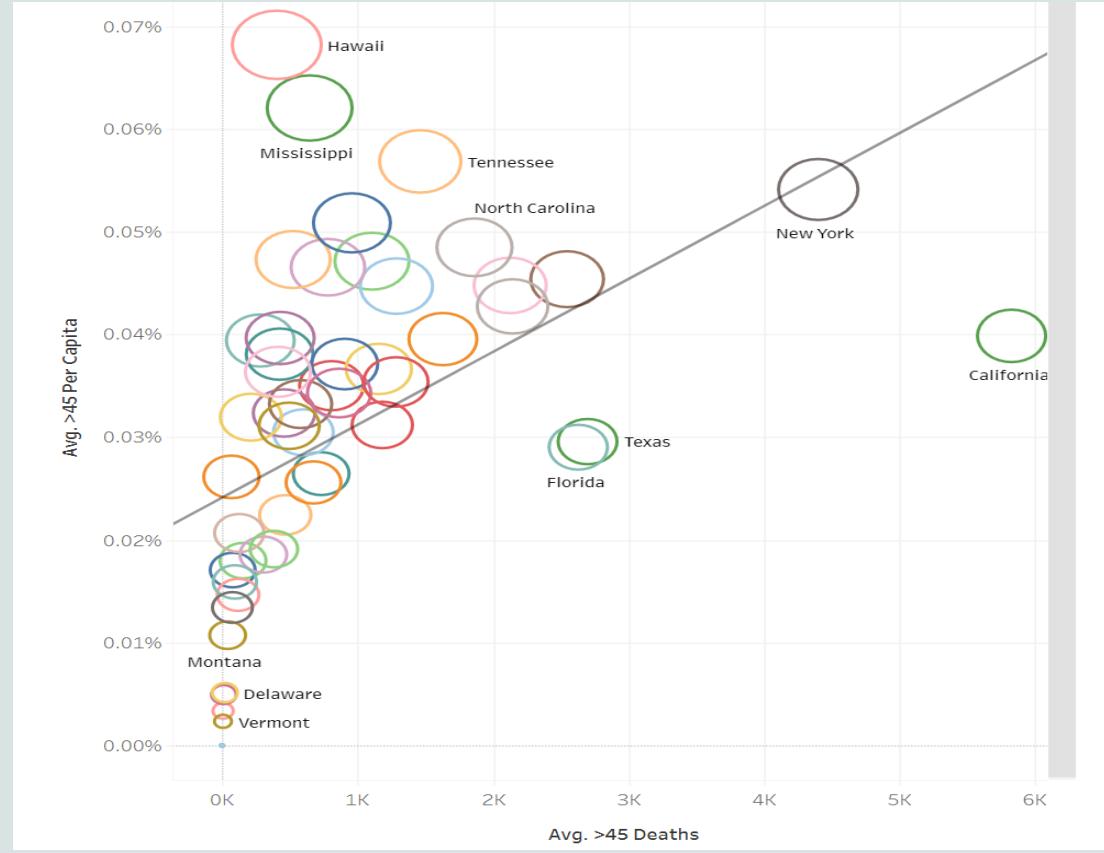
Insights

Influenza Death vs. State Population



California, New York, Florida & Texas were amongst the highest death rate.

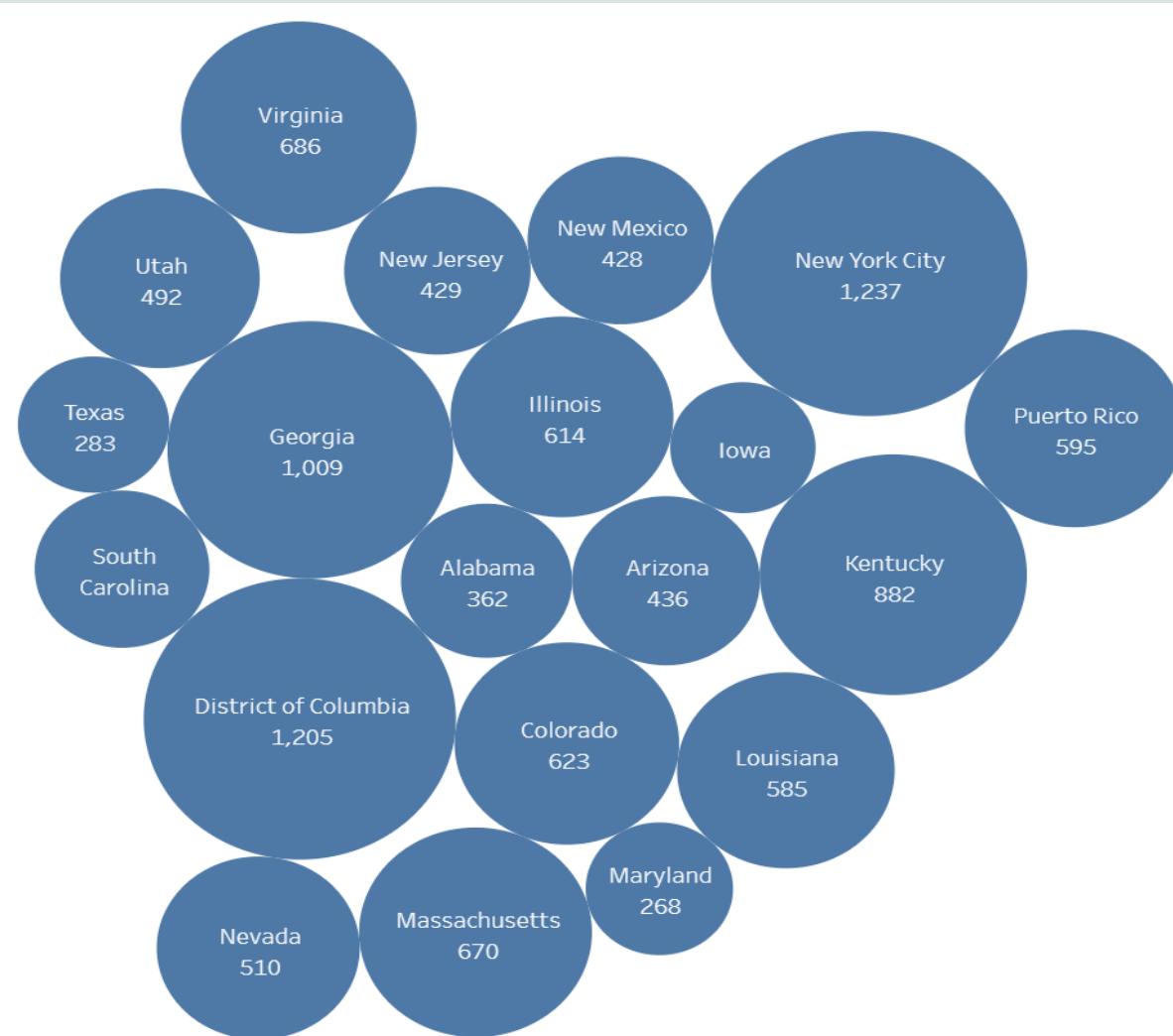
Per Capita Influenza Deaths By State



When viewing States on a per capita death rate, Hawaii & Mississippi rank amongst the highest.

Insights

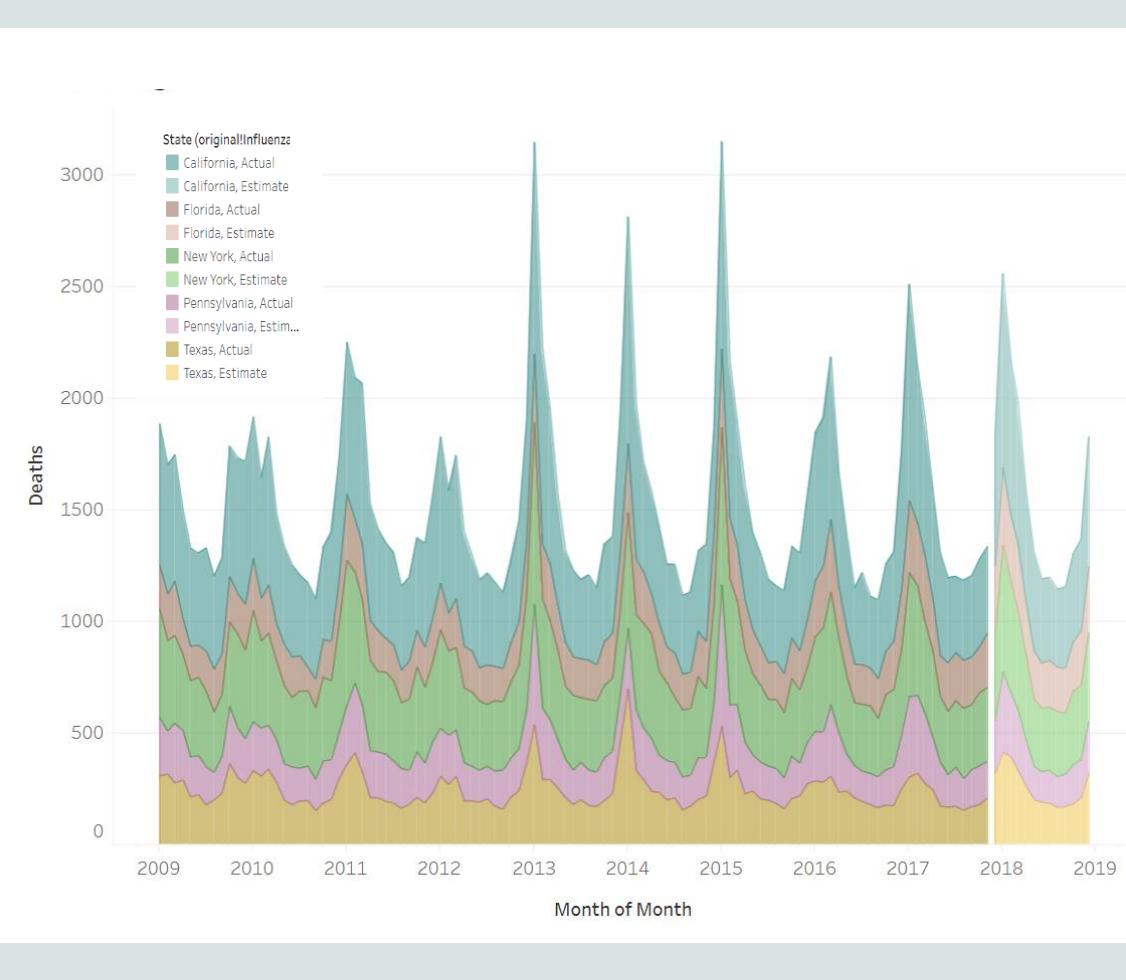
Patient To Provider Ratios



Higher patient to provider ratios in States can result in inadequate medical attention for patients.

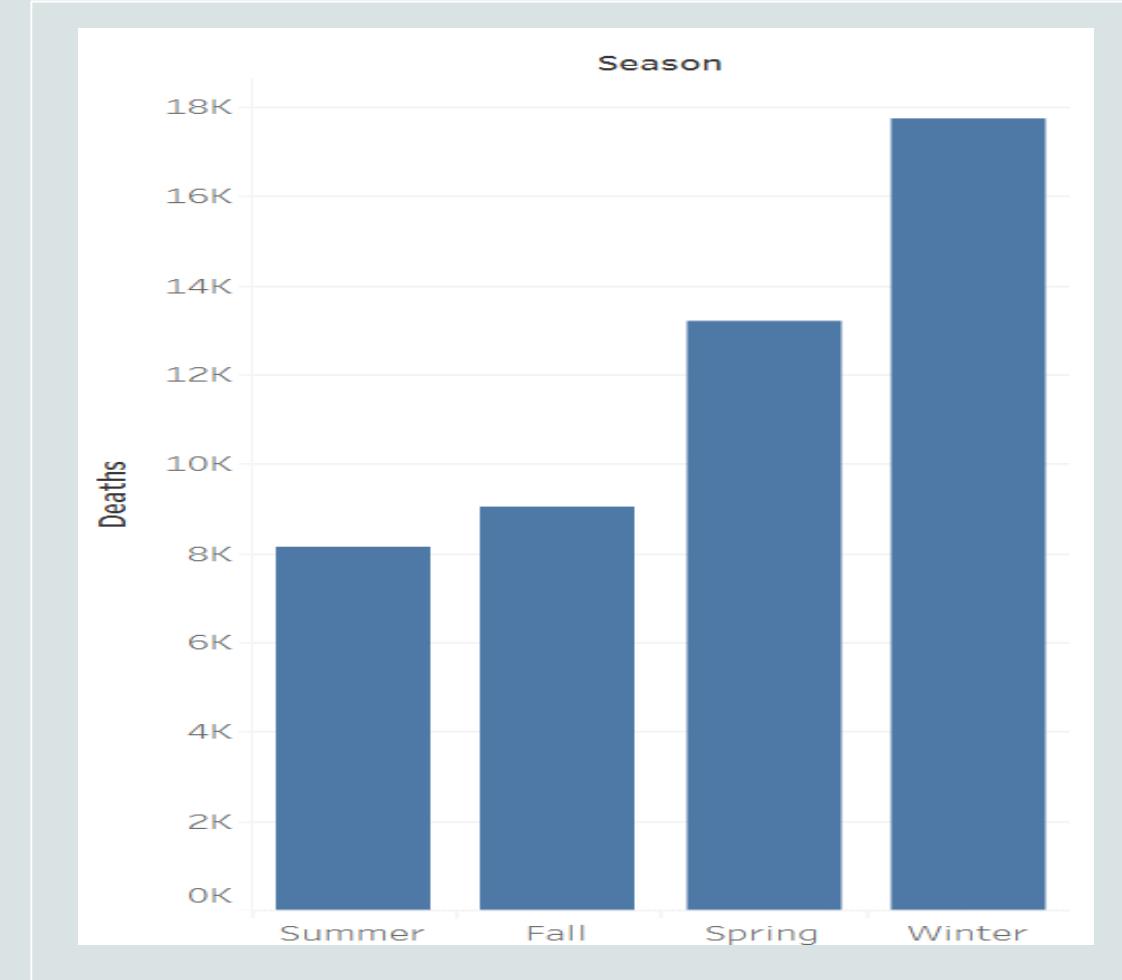
Insights

Forecasting Top 10 States: Influenza Deaths



The trend continues into 2018.

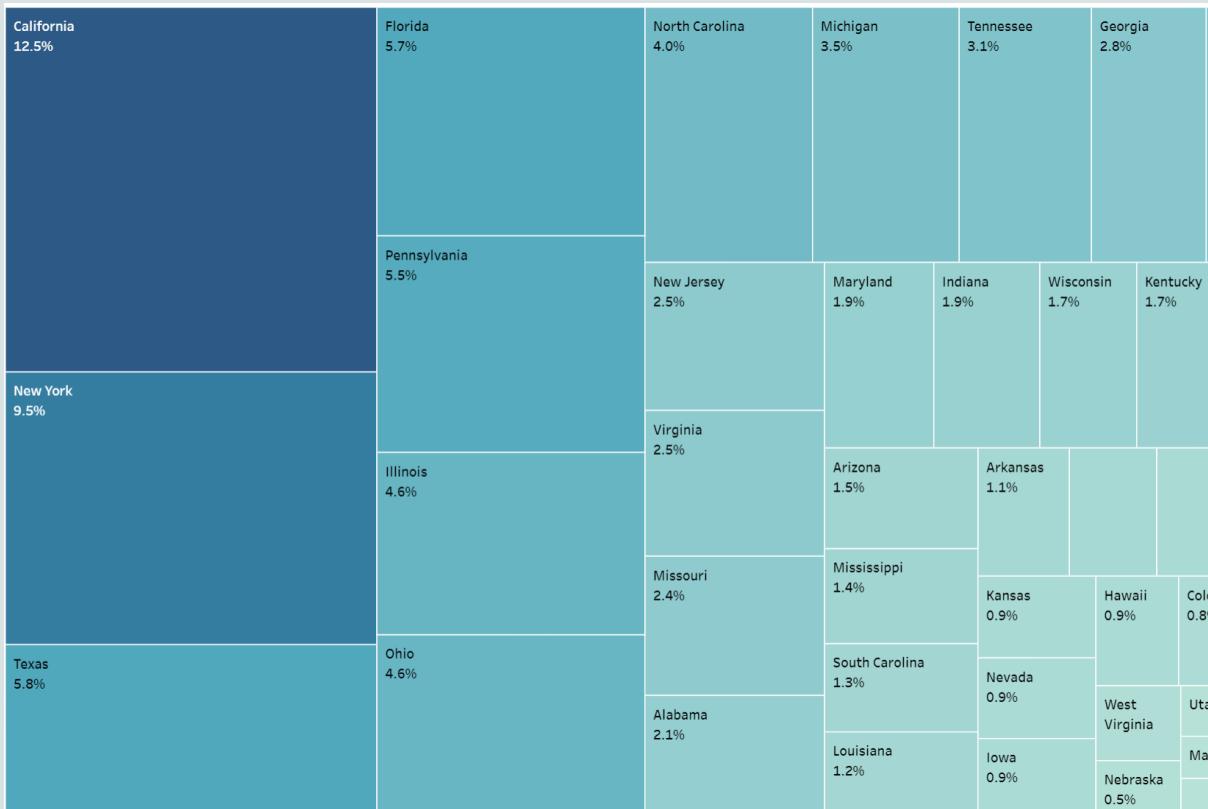
Death Rate Count By Season



Winter has the highest death rates since cold weather affects our bodies.

Mortality Rate by State

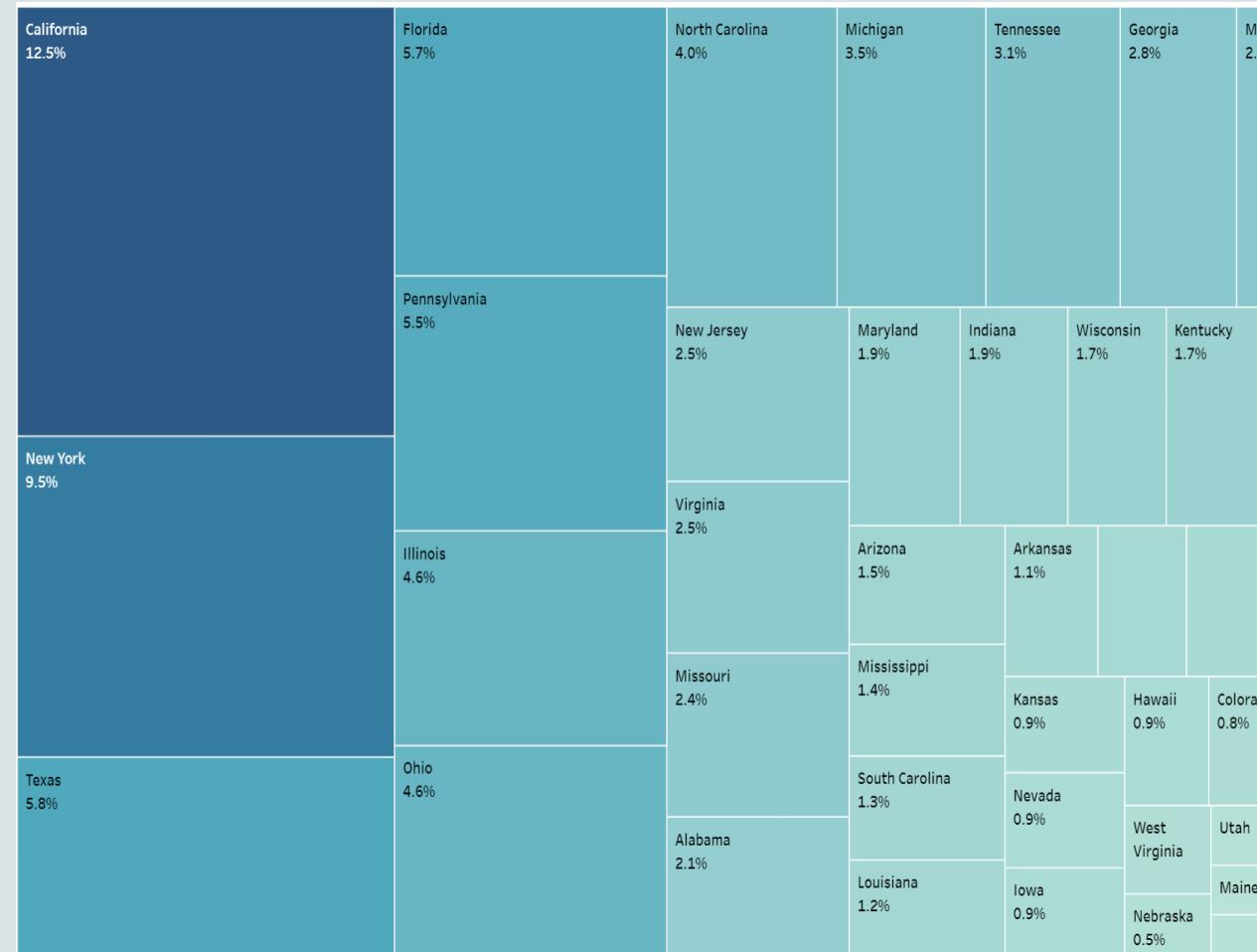
Deployment of medical staff prioritized based on States with highest mortality rate.



Recommendations

Influenza ranks within top 10 for the highest mortality rate and particularly increases with age. Efficient deployment of medical staff & resources can help to flatten the curve and improve patients care. I recommend several proactive measures:

- Deploy medical resources to States based prioritized by Influenzas Death Rate outlined in the tree map.
 - Increase flu shots in States with higher per capita death rate during the CDC recommended months of September and October.
 - Improve patient care by increasing the number of providers to States with a higher patient-to-provider ratio



GameCo

Analyzing global video game sales to help understand how GameCo's new games might fare in the market.



GameCo

Data

- 16.5K titles.
- Data format - Excel CSV.
- Regions: North America, Europe, Japan and others.
- Information: title, platforms, year, genre, publisher.
- Pivot Tables and Charts.
- Data Source: VGChartz

Skills

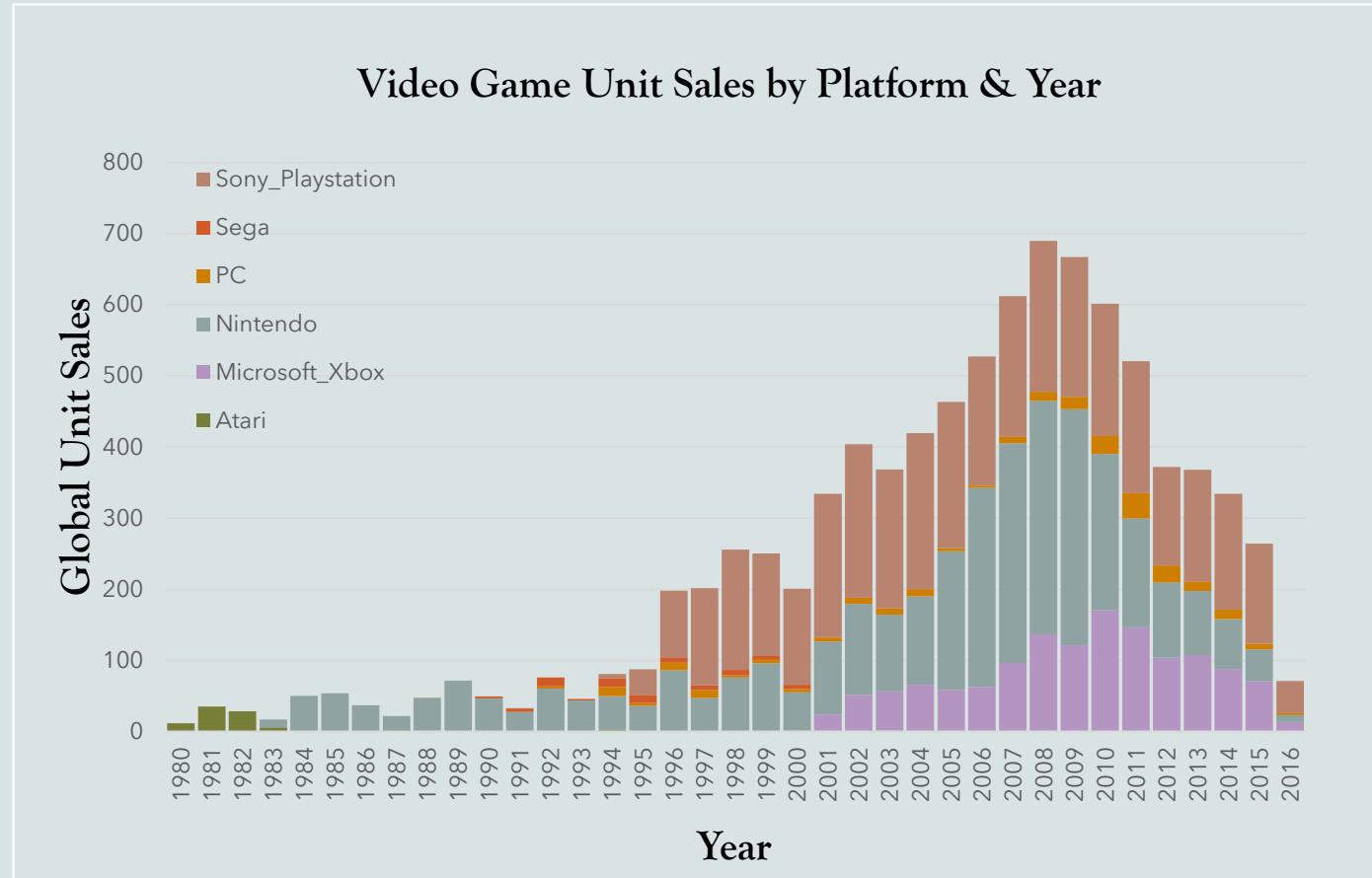
- Data Cleaning
- Data Grouping and Summarizing
- Descriptive Analysis
- Pivot Table
- Visualization Charts in Excel/PowerPoint

Challenges

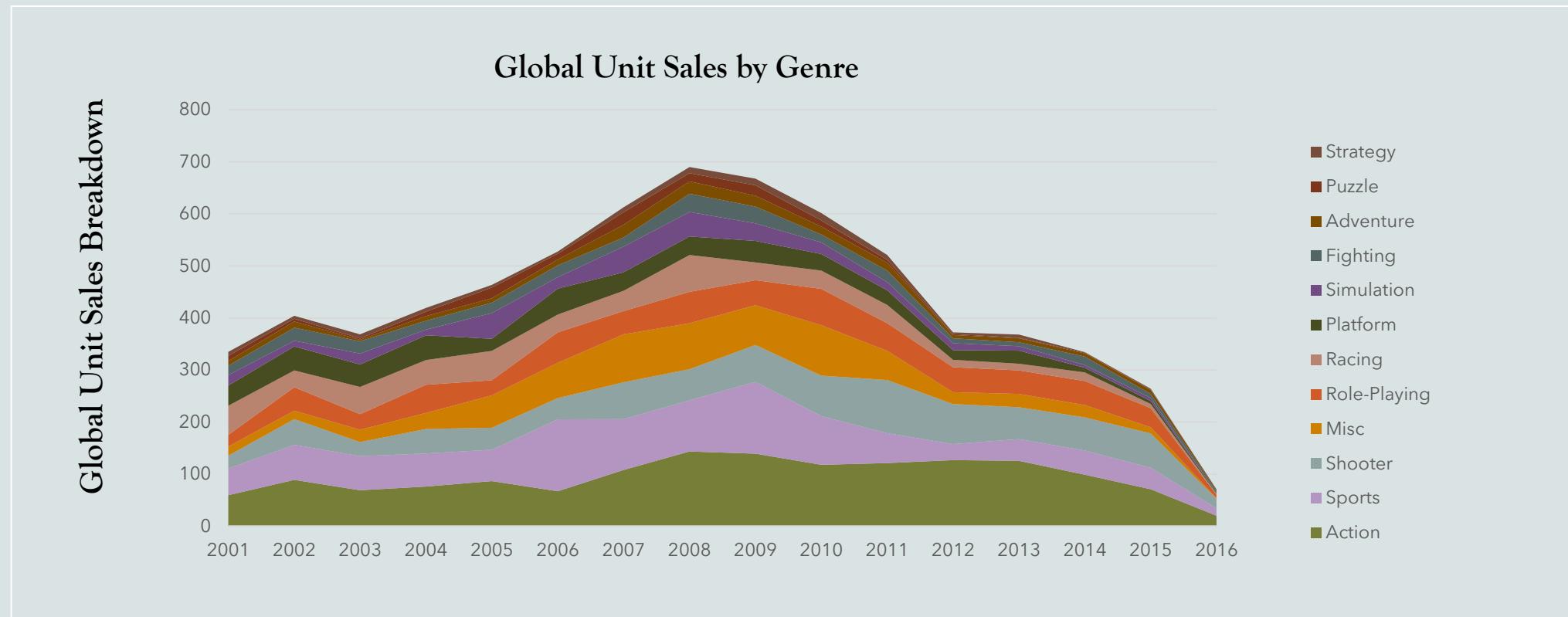
- Data cleaning, missing data, format issues and standardizing the dataset.
- Lack of recent data.
- Decrease of Sales do not reflect the real numbers due to lack of recent data.

History of Video Game Sales

- Atari was the first gaming platform and started in 1980
- In the Mid-90s Video Games started to ramp up with the advent of the Sony PlayStation
- Since the mid 90s, the three Biggest Platform Makers are Sony PlayStation, Microsoft XBox, and Nintendo

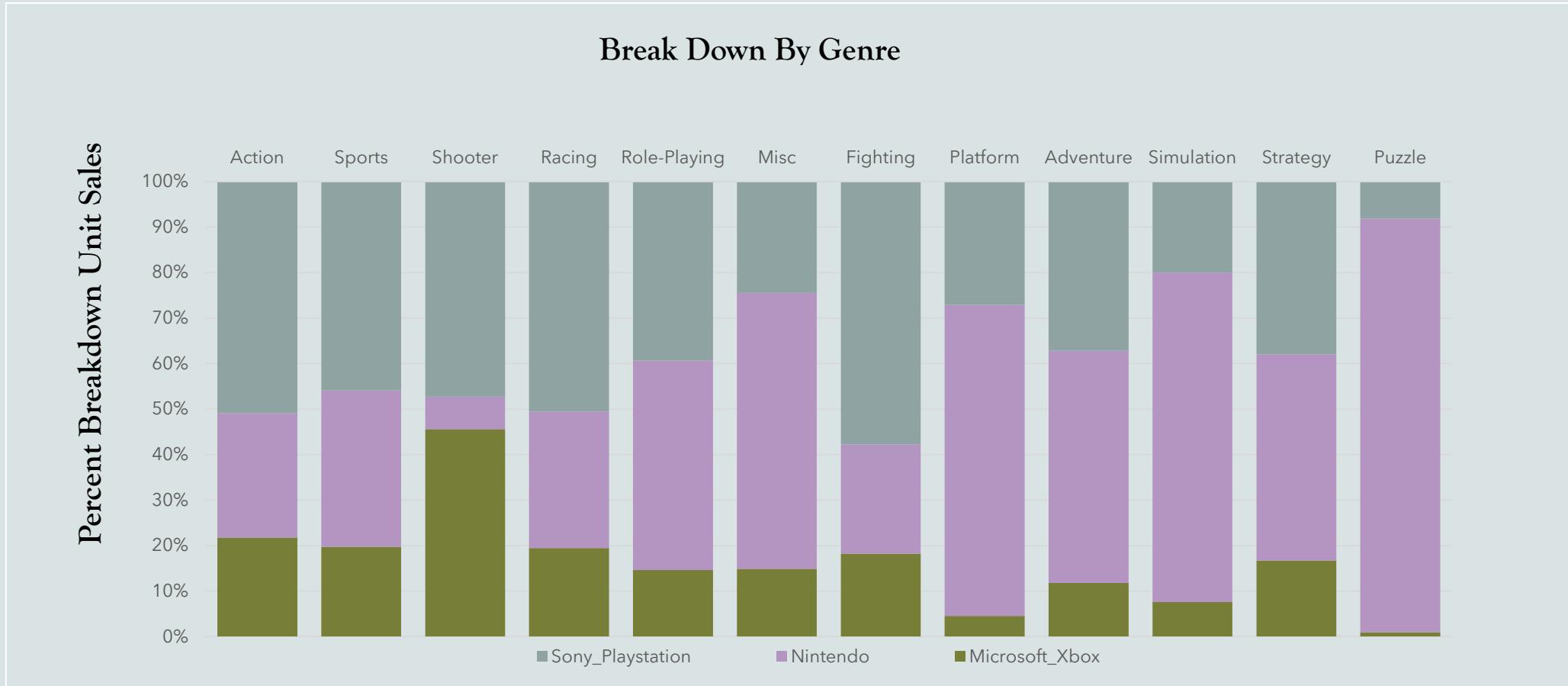


Game Sales by Genre



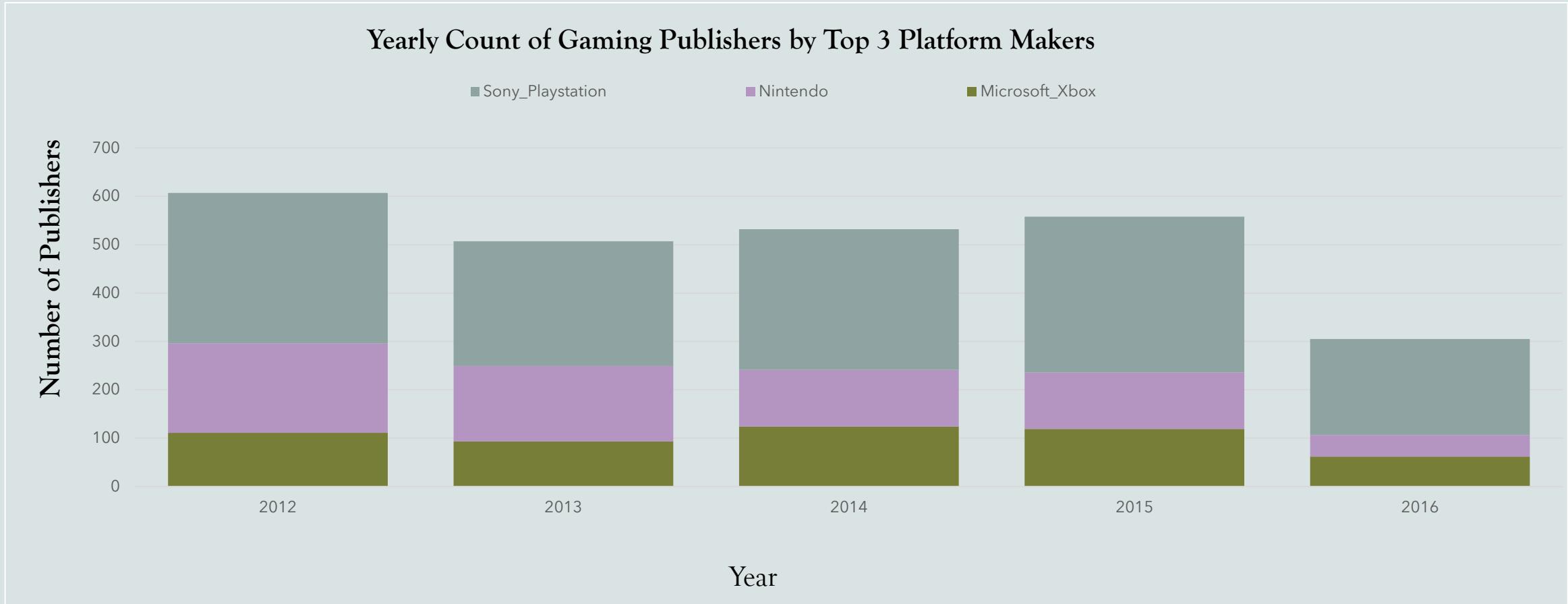
- Top three biggest Genres in terms of global unit sales: Action, Sports and Shooter.

Video Game Sales By Genre



- The Sony PlayStation leads in the Top three biggest Genres in terms of global unit sales which are Action, Sports and Shooter.

Number Of Publishers - Top Three Platforms



- Sony has the most publishers making games for its platforms

Sales Comparison By Platform- Top Three



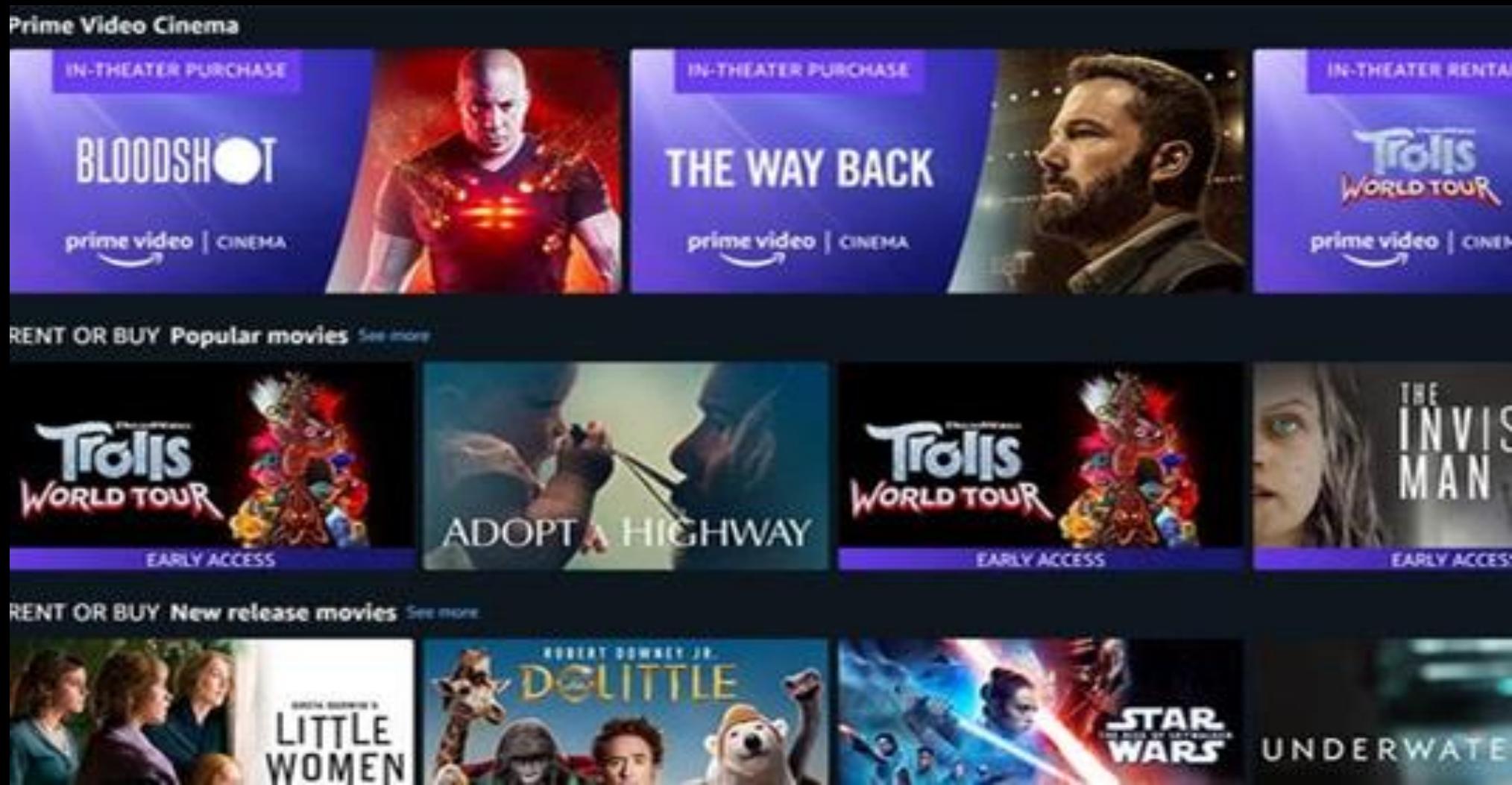
- Sony is the dominant platform as compared to Microsoft and Nintendo
- 2016 has the least sales across the three platforms.
- Sony European sales have been trending upward and have surpassed North American sales with the exception of 2016.
- Microsoft & Nintendo sales have been trending downwards
-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

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.

Rockbusters

Analyze Rockbuster's data through SQL queries for the launch strategy of the new online video service.



Rockbuster

Launch Strategy Analysis

Data

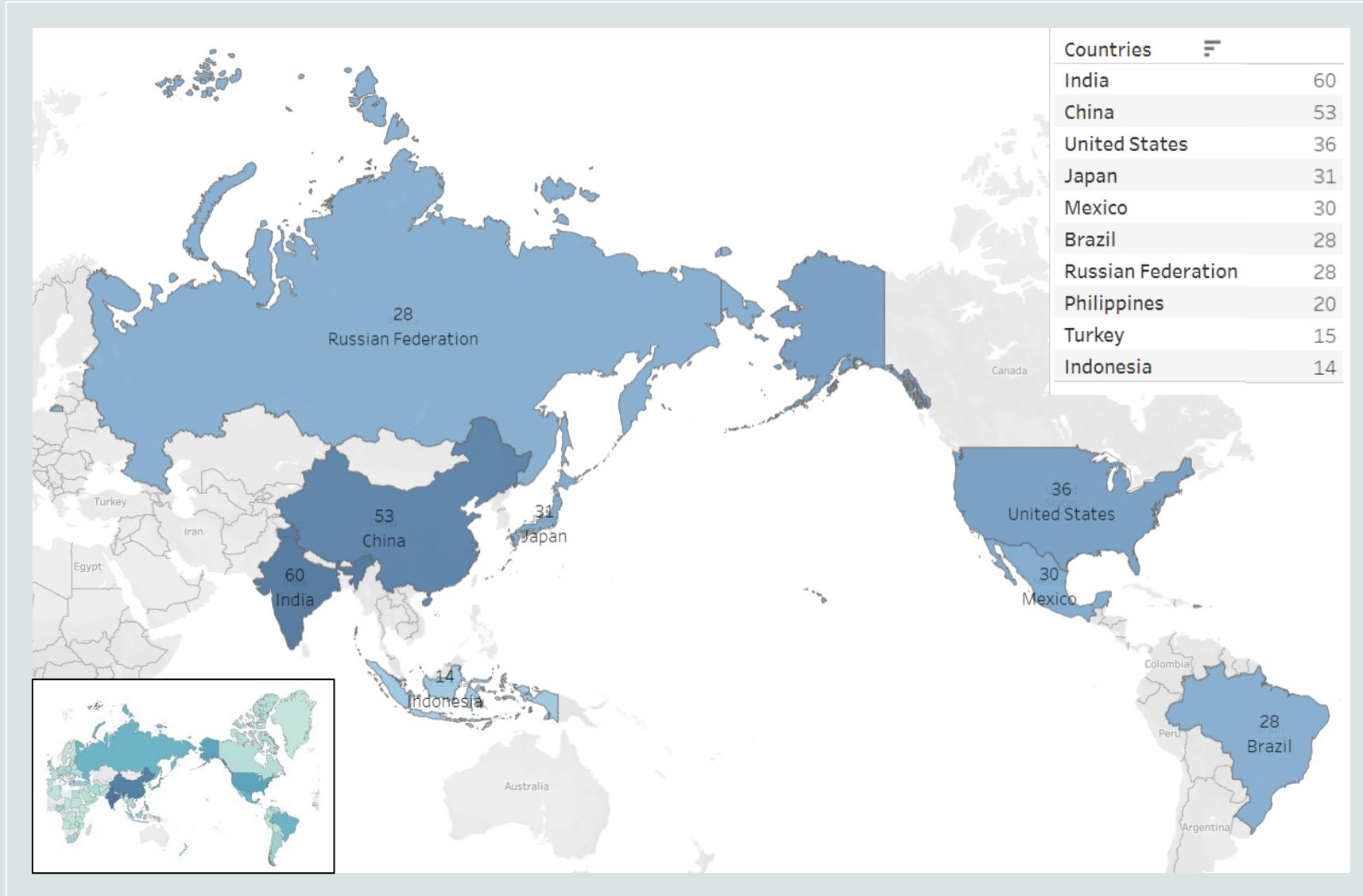
- Data format - SQL.
- Regions: International.
- Develop a relational database.
- Schema
- Performing subqueries
- Data Source: Project Brief

- SQL
- Database querying
- Filtering
- Cleaning and summarizing
- Joining tables
- Subqueries
- Common table expressions

Challenges

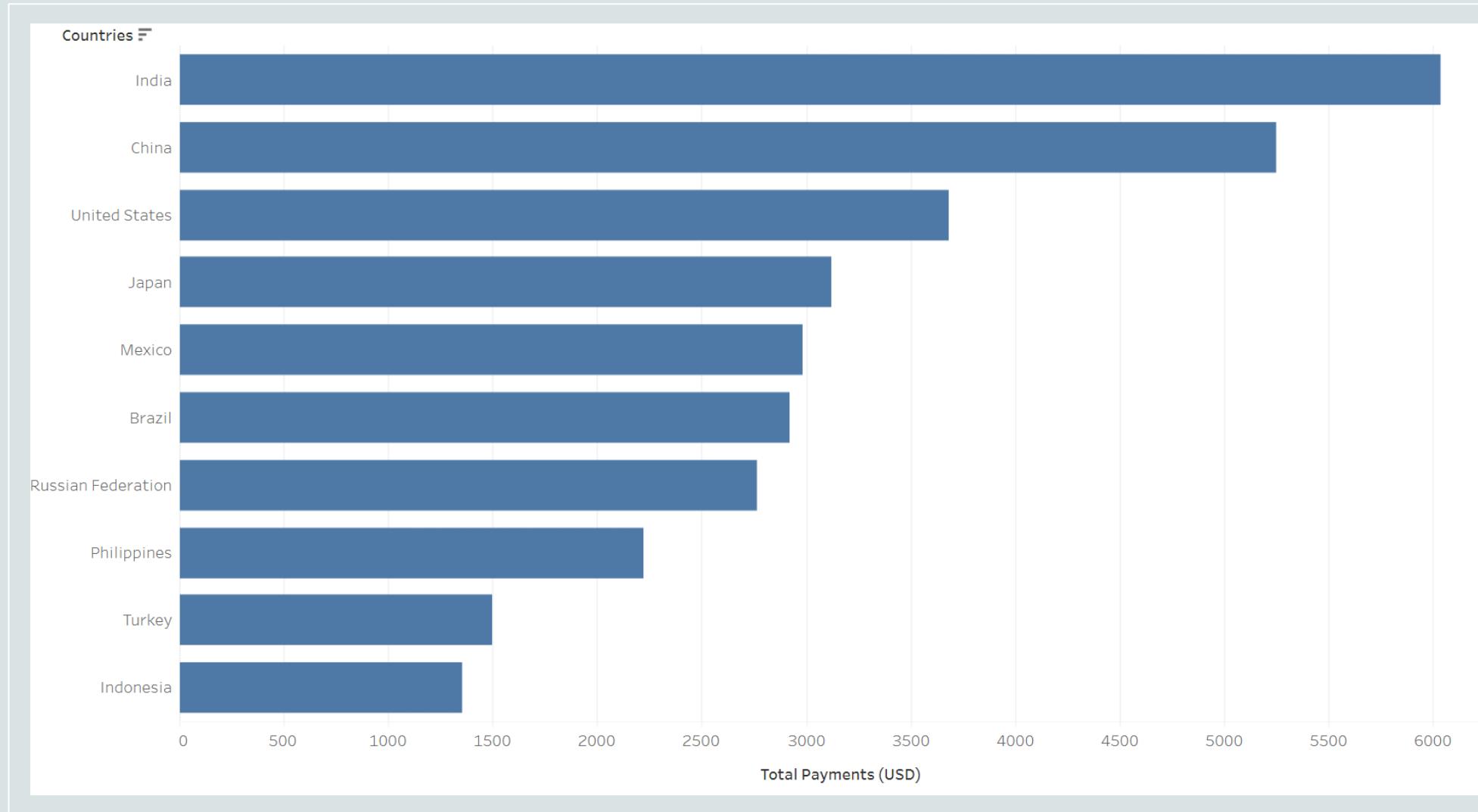
- Data cleaning, missing data, format issues and standardizing the dataset.
- Does not provide info related to historical data or that would allow a market trend analysis or forecast.

Customer Count For The Top 10 Countries



Rockbuster customers are spread around the world in 108 distinct countries. These are the top 10 countries in terms of number of customers.

Top Ten Countries With The Highest Total Payments



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.

Recommendations

- Strategic promotion and engagement: Offer targeted discounts on highest-grossing movies year-round. Especially key markets like India, China, USA to attract new audiences and foster loyalty. Introduce personalized discounts on online accounts sign ups to enhance customer retention.
- Focus on high quality content: Prioritize top performing movies and discontinue underperforming one to streamline the content library. Ensure consistent focus on quality title that resonate with diverse customer bases.

Instacart Basket

Derive insights and suggest strategies for increasing sales based on exploratory analysis



Instacart Basket

Skills

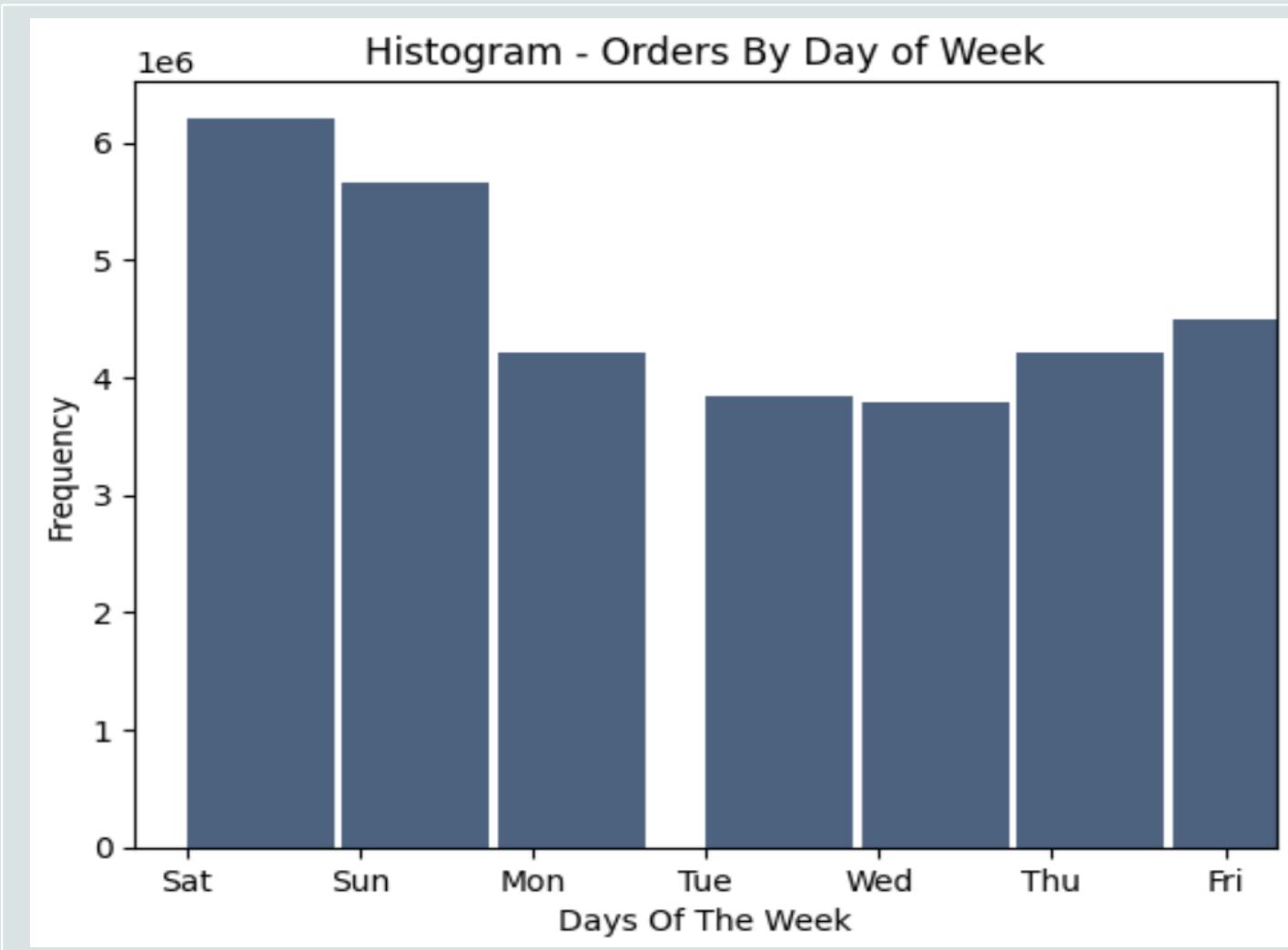
- Python Libraries - Pandas & Numpy
- Visualization with Python - Seaborn, Matplotlib & Scipy
- Jupyter Notebook
- Data wrangling and merging
- Deriving variables
- Grouping & Aggregating data

Data

Challenges

- Data format - Excel
- Python Libraries - Pandas and Numpy
- Data Wrangling and Subsetting
- Data Consistency Check
- Combining and Exporting Dataframe
- Customer Data set - By CareerFoundry
- Instacart Data Set - Via Kaggle
- Data cleaning, missing data, format issues and standardizing the dataset.
- 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.

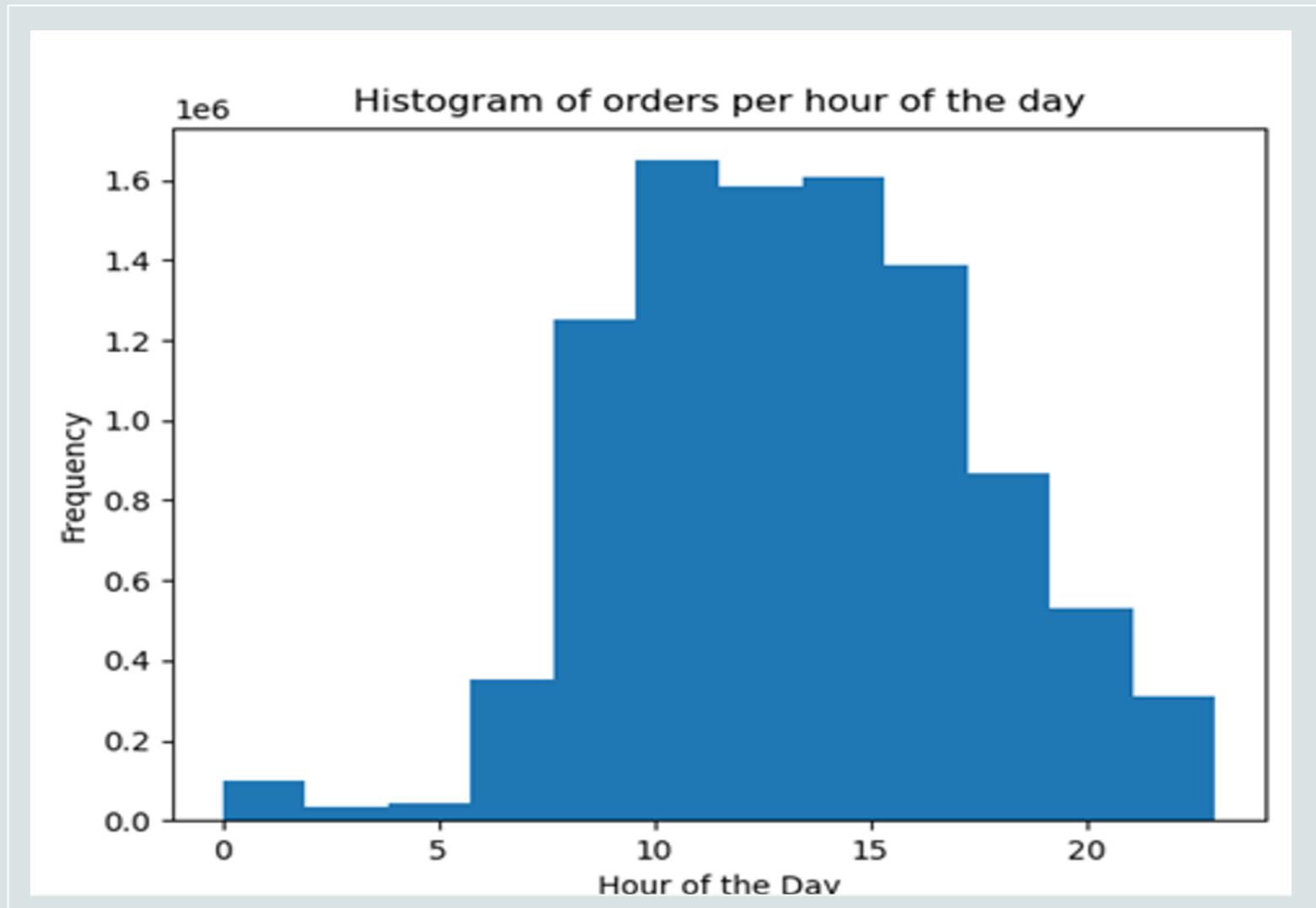
Busiest Days



How busy each day of the week according to number to number of orders..

Saturdays and Sundays are busiest days and Tuesdays and Wednesdays are the least busy days.

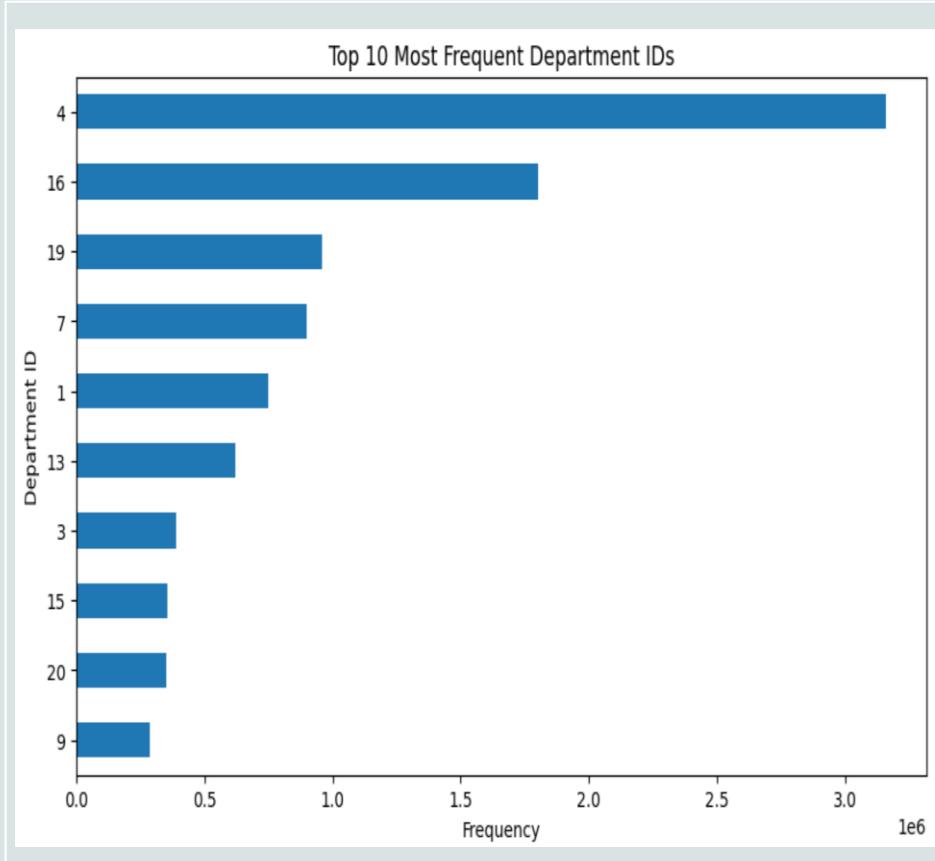
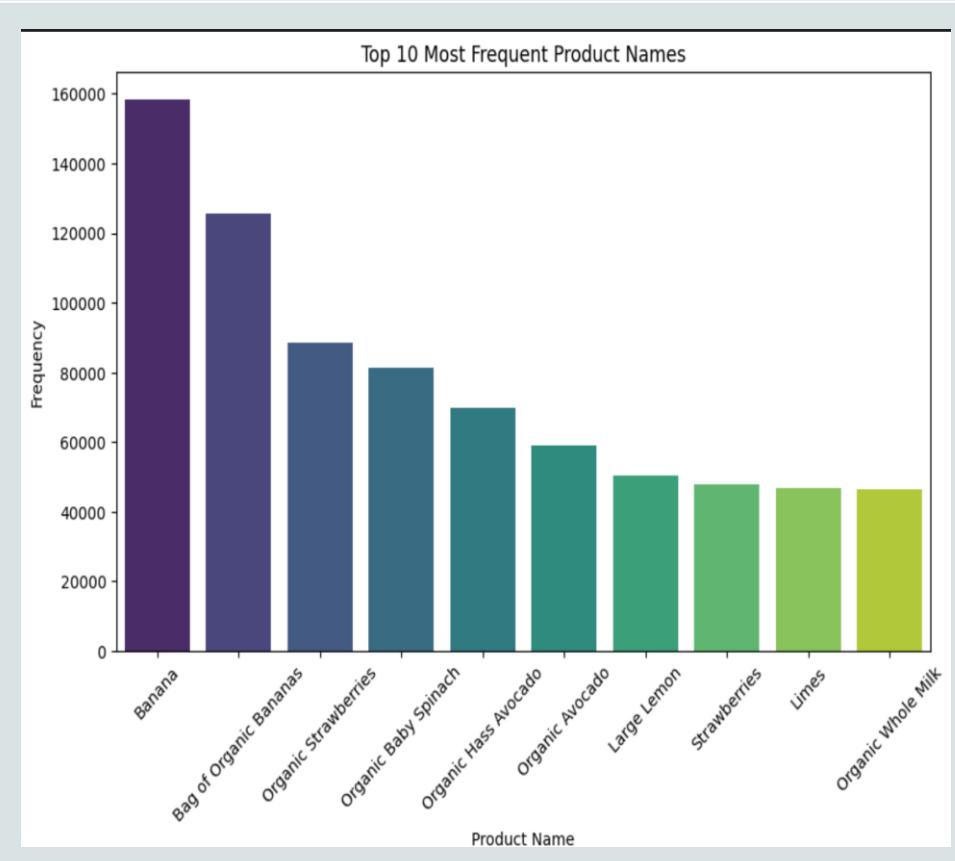
Busiest Hours



The busiest hours, increase at 9 am until 3pm, when it starts to decrease.

Top 10 Products and Departments

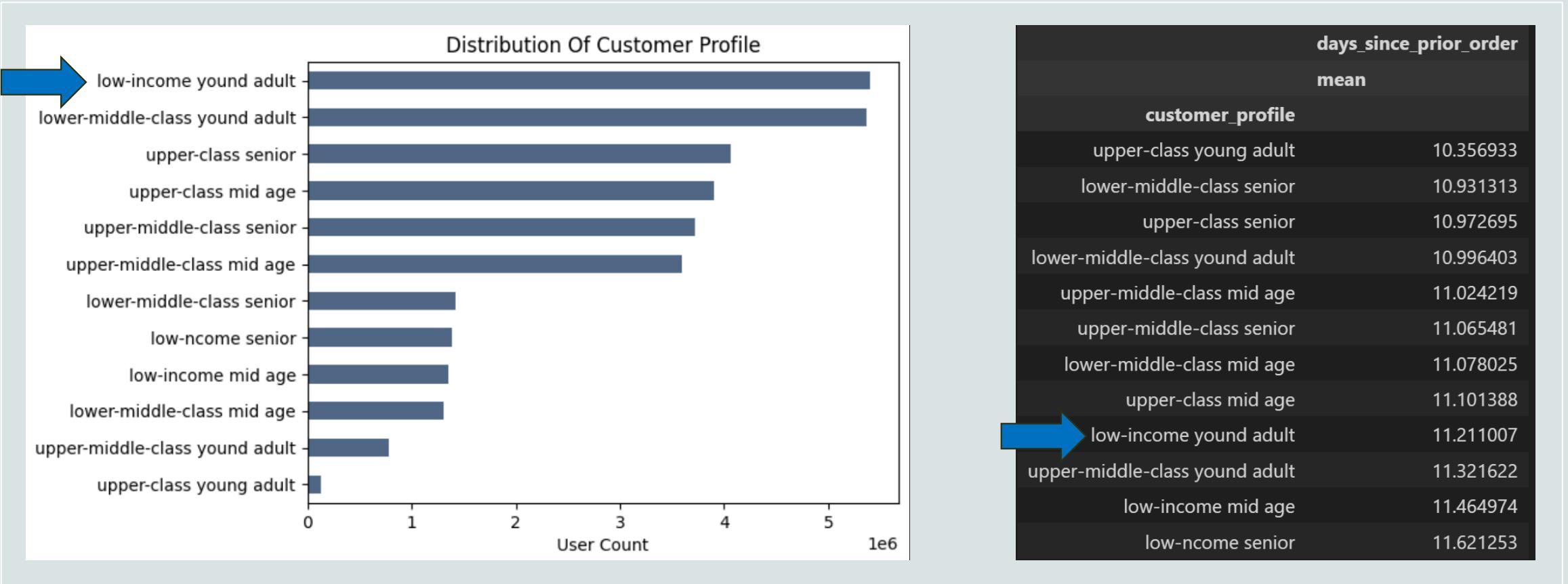
Data Dictionary



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

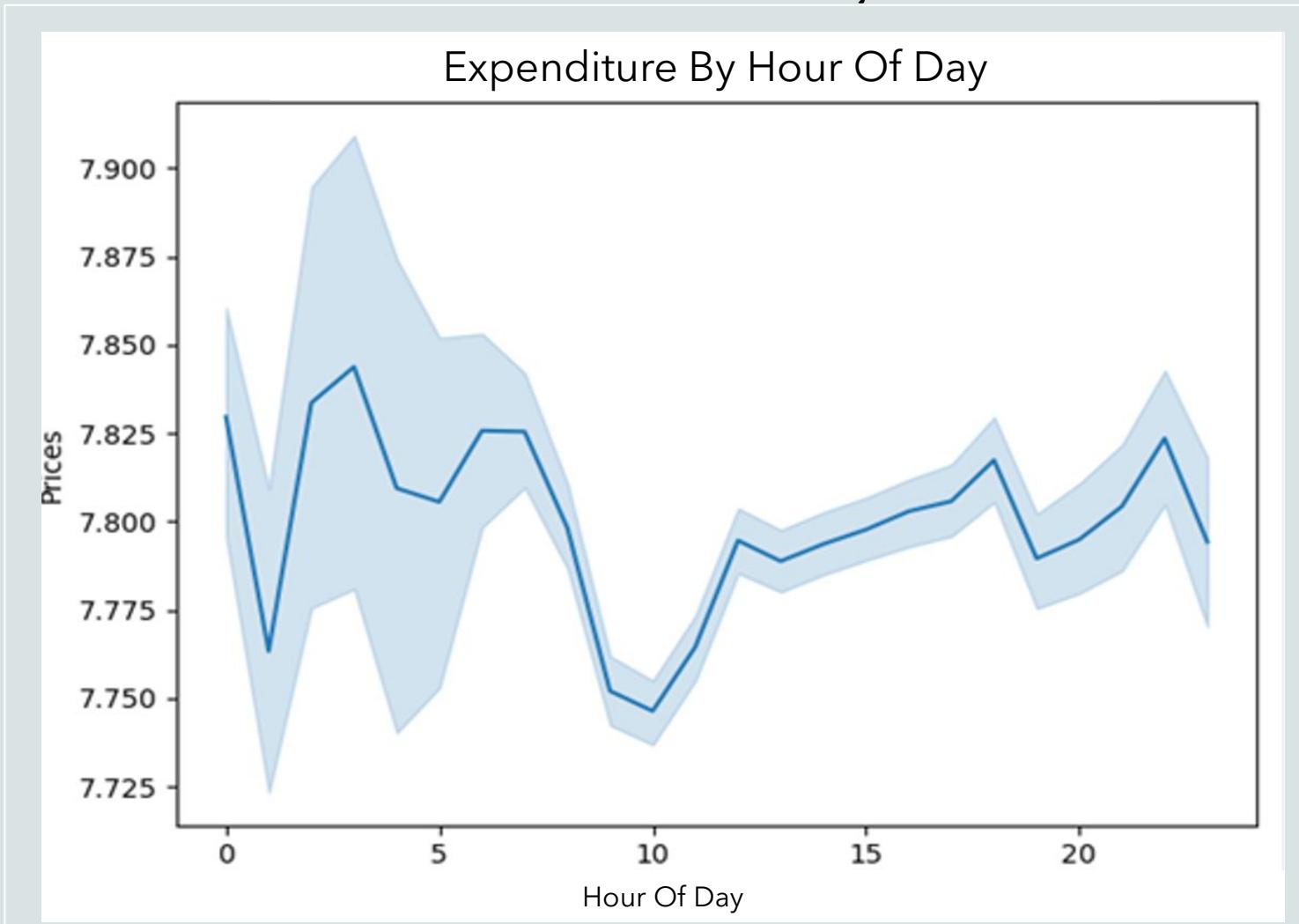
Top 10 Departments consumed by all regions in the Instacart database. Produce is the highest, which could indicate a buying trend by its customers. Within Produce, Bananas are the top purchased Produce.

Customer Profile



Low-income young adults come more often than other demographics, however, they do not return as often as others. This is an opportunity to provide incentives, so they come more often and spend more.

Market Analysis



According to the graph above customers have a higher tendency to spend more between 4 am to 8 am, as well as from 8 pm until 10 pm. This can be an opportunity to sell more expensive products.

Recommendations

- Runs all ads from Monday to Thursday when they are fewer sales to try and boost the sales on these days as well.
- Run the ads in morning for more expensive products
- Give more discounts on the most popular products types (snacks, pantry, and personal care)
- Give more discounts to loyal costumers, so they come back more often and buy more stuff.

Pig E. Bank

Analytical support to the anti-money laundering compliance department to assess client and transactional risk at a global bank.



Pig E. Bank

Data

- Data format - Excel
- Time Series analysis forecasting
- Intro to Data Mining
- Data Ethics: Bias, Security and Privacy
- Data Source: Customer Data Set- By CareerFoundry

Skills

- Intro to Big Data
- Data Ethics:
- Bias Security and Privacy
- Data Mining
- Predictive Analysis
- Time Series Analysis and Forecasting
- Decision Tree

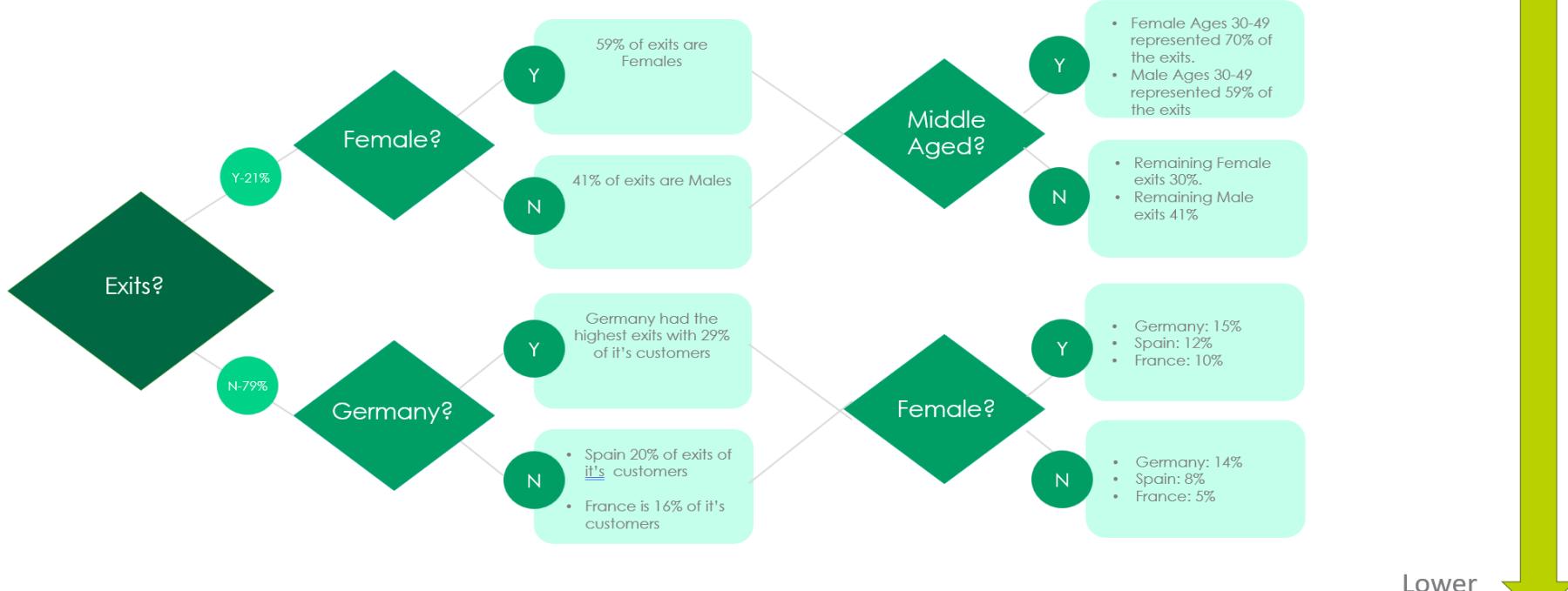
Challenges

- Data cleaning, missing data, format issues and standardizing the dataset.
- Data does not provide date, so it's impossible to observe whether there was a trend of customer loss during specific period or something that occurred instantly

Insights

- Few contributing factors on why clients are leaving the bank can be analyzed by:
- Age
- Gender
- Balance
- Tenure
- Credit Card

Decision Tree



Decision tree based on gender, country & age.

- Females had a higher number of exits as compared to men.
- Within the female population, middle aged females were higher.
- Germany saw the most exits as compared to Spain and France

Recommendations

- Pig E Bank should conduct surveys with their customers regularly, This will help to better understand customers banking needs.
- Provide incentives for females particular the one's who are up in age.
- Offer competitive rates for German customers.
- Since less active customers are more likely to leave, active customers should benefit in some way.

Let's Connect!

Robert Indelicato

Data Analyst

