



A dark blue background featuring an abstract geometric pattern composed of numerous thin, light-colored lines forming overlapping triangles and various polygons.

STAFFORD REECE

Data Analyst

OVERVIEW OF PROJECTS

GAMECO.

Analyzed global video game sales to create a strategy to allocate resources based on its client trends over time.

HOSPITAL STAFFING AGENCY

Evaluated Influenza deaths across the United States of America to better prepare and staff Hospitals to reduce the number of Influenza related deaths.

ROCKBUSTER STEALTH LLC

Examined Rockbuster's top global clients to help launch Rockbusters new online streaming video service.

INSTACART BASKET ANALYSIS

Developed a strong understanding of Instacart's clients and basket of goods to help solve business problems and drive revenue.

NYC YELLOW

Find options to increase revenue and enhance the value proposition to our cab drivers

GAMECO.

Objective:

As such, you've been asked to perform a descriptive analysis of a video game data set to foster a better understanding of how GameCo's new games might fare in the market.

Dataset:

[Video Game Dataset from VGChartz's](#)

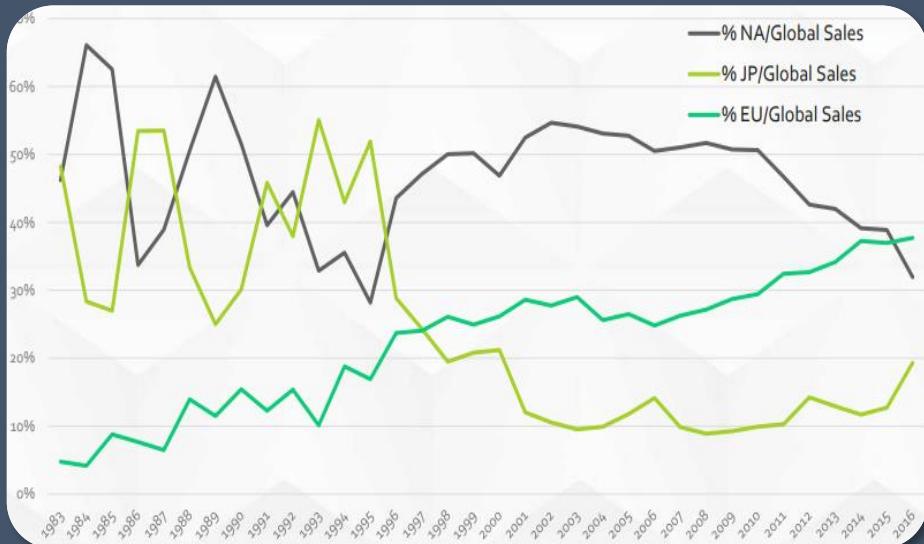
Skills Used:

- Grouping Data
- Summarizing Data
- Descriptive Analysis
- Visualizing Data in Excel
- Presenting Data
- Cleaning and Formatting Data

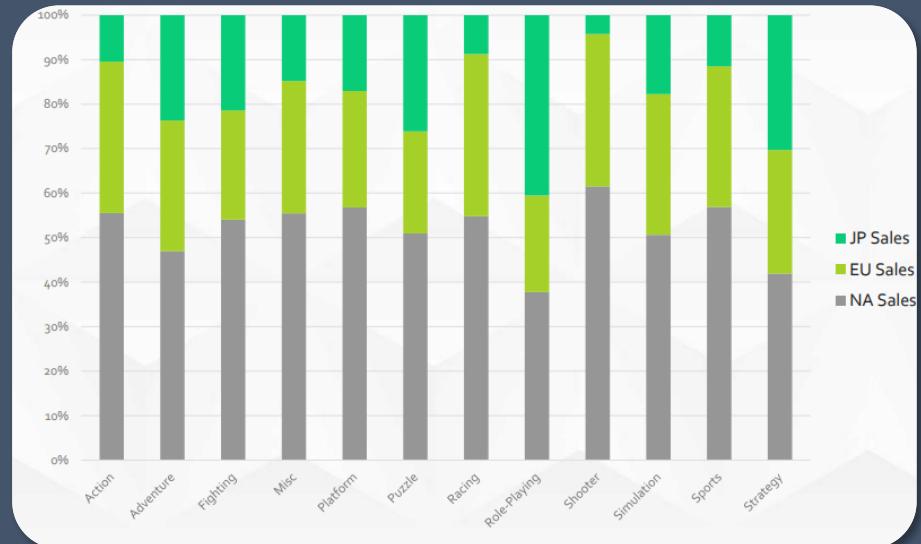
Tools Used:



Regional Revenue Compared Against Global Revenue

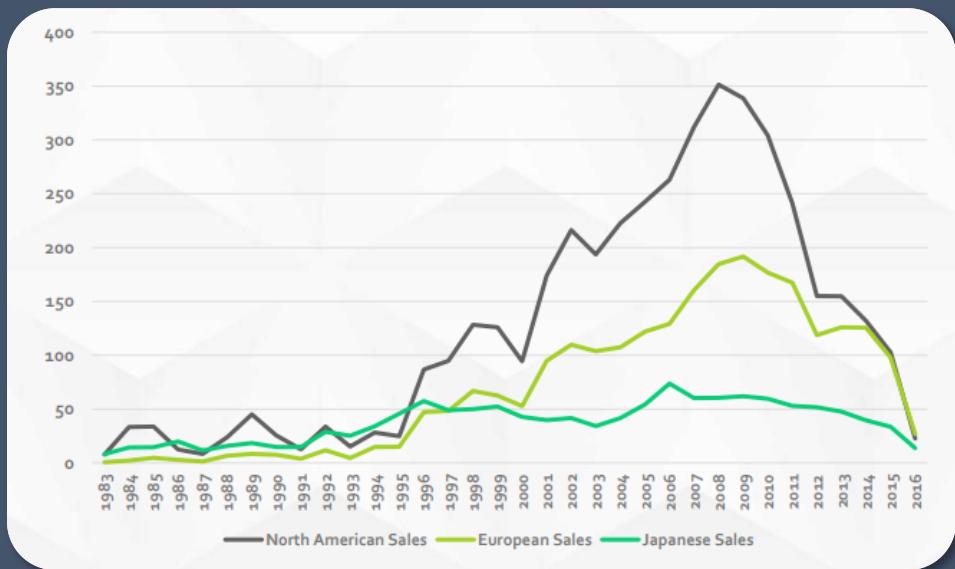


Regional Video Game Genre Preference's in relation to Revenue



- European market share has been steadily gaining without any major dips or long-term corrections
- Japanese market share has declined over the entire period but has steadily gained ground since overall global sales peak in 2008
- It appears that the European and Japanese markets are slowly gaining market share • of global sales
- North America has a strong presence in most categories; but Role-Playing & Strategy are the worst performing genres.
- European Sales has a strong presence in Racing & Shooter games – this is especially true when comparing to Japanese Market share
- Japanese Sales has a strong presence in Role Playing and Strategy Genre's with a weak performance in Shooter, Sports, Action and Racing genres

Regional Sales Data



Regional Market Share of Adventure, Action, Platform, Sports, Strategy, Misc, Fighting Genre's



- We believe a capital allocation change would be most prudent for GameCo's long term viability given the decline in overall sales.

- What we found is that since 2008 Europe has clearly gained a large percentage of market share from North America in North America's most lucrative genre's
- Japan has gained in sales in these genres as well but not nearly as large of gains, but its clear that Europe and Japan are taking market share from North America.
- Based on this information we recommend to allocate less capital towards North America and a higher allocation to Europe for these genres.

Recommendations to Gameco.

- Allocate capital regionally based on the region's market share of global sales.
 - 40% capital allocation to Europe, 35% capital allocation to North America, 15% capital allocation to Japan, the remaining 10% capital allocated to non major regions
 - Review the capital allocation yearly and adjust according to the regions market share of global sales.
- Each region should then allocate the capital its given based on trends, and allocate more capital to genres growing in market share
 - Europe – Increase allocation to Adventure, Action, Platform, Sports, Strategy, Misc, Fighting Genre's
 - Japan – Increase capital allocation to Role Playing and Strategy games – Decrease allocation action, shooter & racing genres
 - North America – Understand that market share is declining but still has strong correlation to overall global sales so keep reinvesting in its strongest genres such as Action, Adventure, platform, Sports, Strategy, Mics, Fighting but reduce allocation to Role Playing as Japan is clearly gaining a larger market share year over year.

HOSPITAL STAFFING AGENCY

Objective:

To help a medical staffing agency that provides temporary workers to clinics and hospitals on an as-needed basis. The analysis will help plan for influenza season, a time when additional staff are in high demand. The final results will examine trends in influenza and how they can be used to proactively plan for staffing needs across the country.

Dataset:

Influenza deaths by geography, time, age, and gender

Source: [CDC](#)

[Download Data Set](#)

2. Population data by geography

Source: US Census Bureau

[Download Data Set](#)

Skills Used:

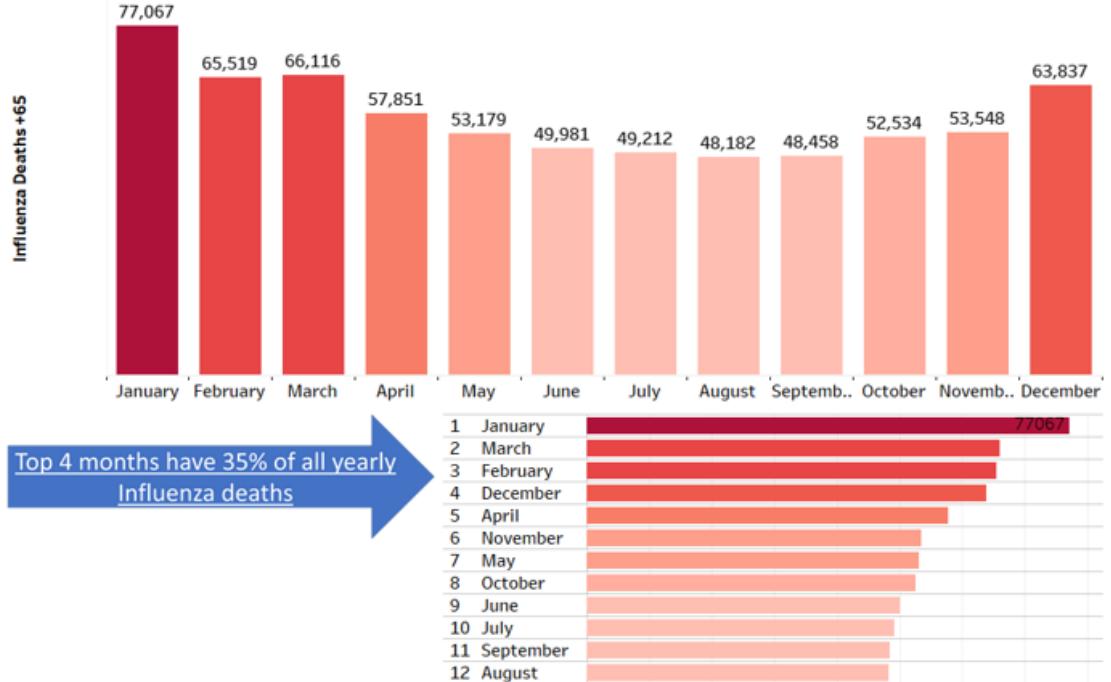
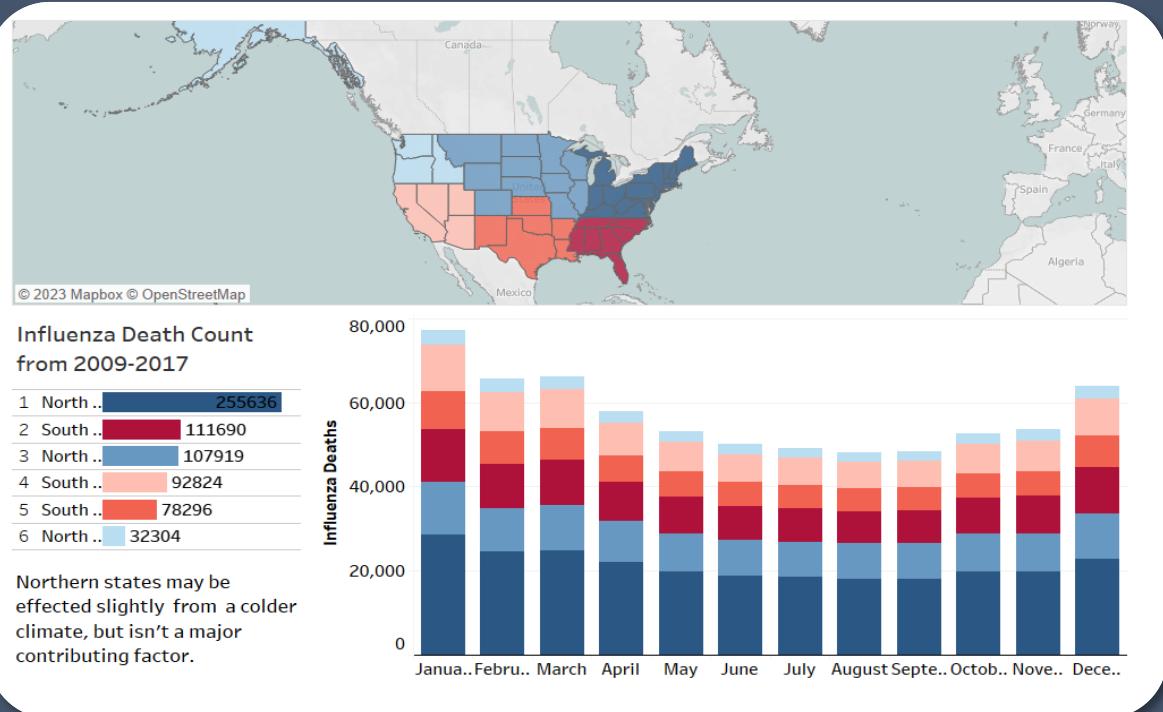
- Translating Business Requirements
- Data Cleaning
- Data Integration
- Statistical Hypothesis Testing
- Visual Analysis
- Forecasting
- Dashboard Design
- Video Presentation

Tools Used:

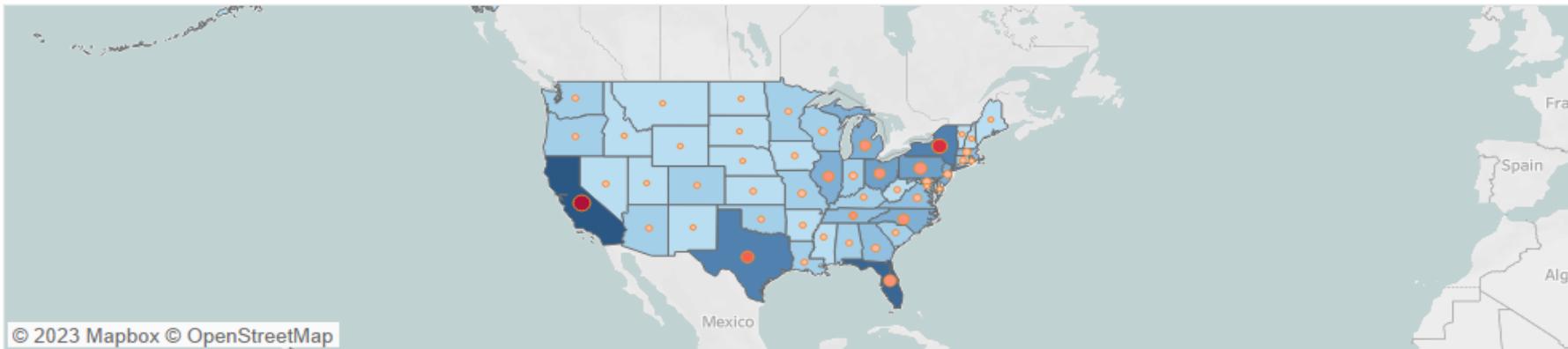


When is the Influenza Season?

Does geography play a role in Influenza deaths?



Does density play a role in Influenza death?



Total Regional Vulnerable Population (+65)

A horizontal bar chart titled 'Ra.. Region' showing the distribution of regions across different Ra.. values. The x-axis represents Ra.. values from 1 to 6, and the y-axis lists the regions. The bars are colored blue, with the length of each bar corresponding to the value of Ra.. for that region.

Ra..	Region	Value
1	North East	~9.5
2	South East	~7.5
3	North Central	~4.5
4	South West	~4.5
5	South Central	~3.5
6	North West	~1.5

Influenza Regional Death Count

A horizontal bar chart titled 'Region' on the y-axis and 'Value' on the x-axis. The y-axis lists six regions: North East, South East, North Central, South West, South Central, and North West. The bars are colored in a gradient from dark red for the North East to light orange for the North West.

Region	Value
1 North East	95
2 South East	75
3 North Central	70
4 South West	65
5 South Central	55
6 North West	35

Based on our preliminary statistical analysis of the CDC Influenza data and US Census Population data we already can surmise..

Prioritize Hot Spot States with high death counts, exclude states with the highest priority region.

The states that are given "Hot Spot" status will be given additional staffing res..

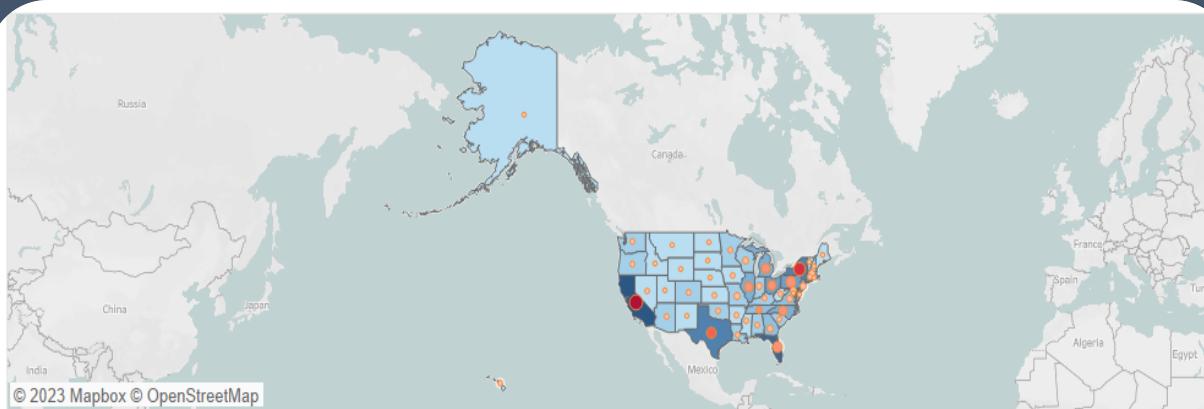
Percent of allocation for HOT SPOT staffing resources

State	Region	Allocation (%)
California	South West	5%
Texas	South Central	4%
Florida	South East	3%
Illinois	North Central	2%
North Carolina	South East	1%

Prioritize additional staffing resources within a given region by the percentage of deaths that state incurred:..

Percent of allocation for Influenza staffing resources

State	Region	Allocation (%)
Texas	South Central	5%
Louisiana	South Central	4%
Arkansas	South Central	3%
Oklahoma	South Central	2%
Kansas	South Central	1%
New Mexico	South Central	0.5%



Percentage of Regional Vulnerable Population(+65)

Rank Region

1	North East	10%
2	South East	8%
3	North Central	6%
4	South West	4%
5	South Central	2%

Percentage of Influenza Regional Death Count

Rank Region

1	North East	5%
2	South East	4%
3	North Central	3%
4	South West	2%
5	South Central	1%
6	North West	0.5%

Recommendations:

1. We should staff hospitals based on the region's percentage of Influenza deaths.
2. Prioritize "Hot Spot" states with additional resources or staffing. We will pull those resources from the region with the lowest density of vulnerable population and lowest overall death count. This region is the Northwest Region. We will allocate the 5% of staffing resources for Influenza in the Northwest Region and allocate that to "Hot Spot" states.

Links:

Full Video Presentation:

[Hospital Staffing Recommendations for Upcoming Influenza Season – YouTube](#)

Full Story Board on Tableau:

[Hospital Staffing Agency Recommendations for Influenza Season | Tableau Public](#)

ROCKBUSTER STEALTH LLC

Objective:

The Rockbuster Stealth Management Board has asked a series of business questions and they expect data-driven answers that they can use for their 2020 company strategy. Here are the main questions they'd like to answer: Which movies contributed the most/least to revenue gain? What was the average rental duration for all videos? Which countries are Rockbuster customers based in? Where are customers with a high lifetime value based? Do sales figures vary between geographic regions?

Dataset:

[Download Rockbuster Data Set](#)

Skills Used:

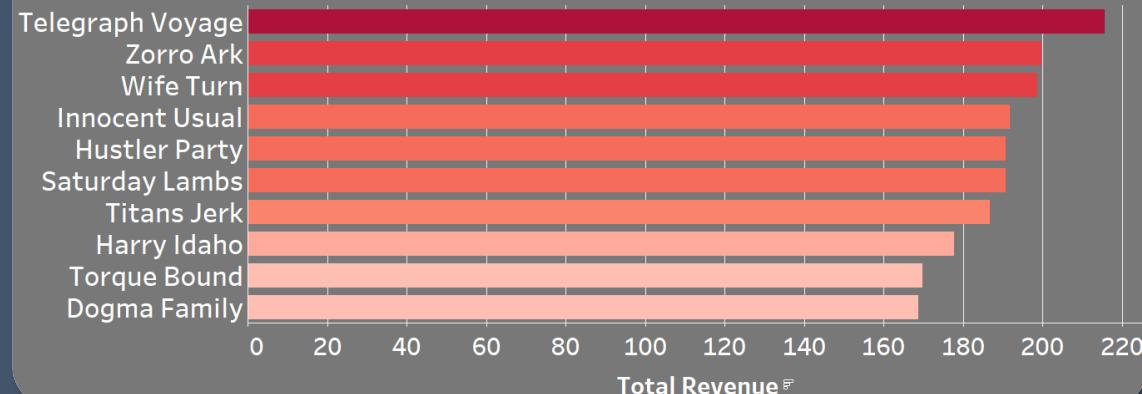
- Relational Databases
- Database Querying
- Filtering
- Cleaning and Summarizing
- Joining Tables
- Subqueries
- Common Table Expressions

Tools Used:

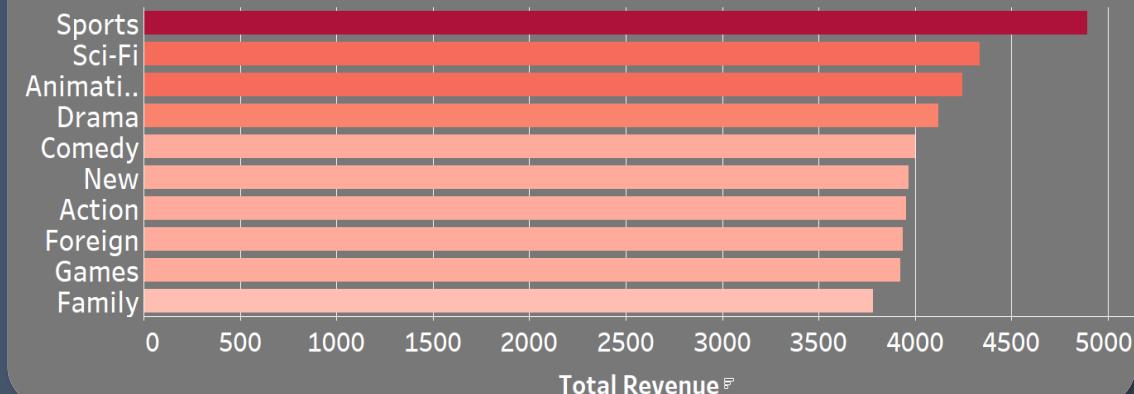


Rockbuster Visualizations

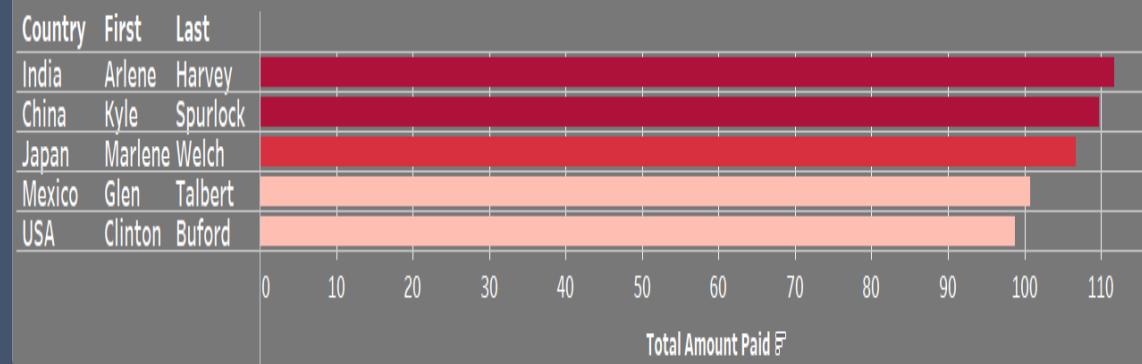
Top 10 Titles by Revenue



Top 10 Genres by Revenue



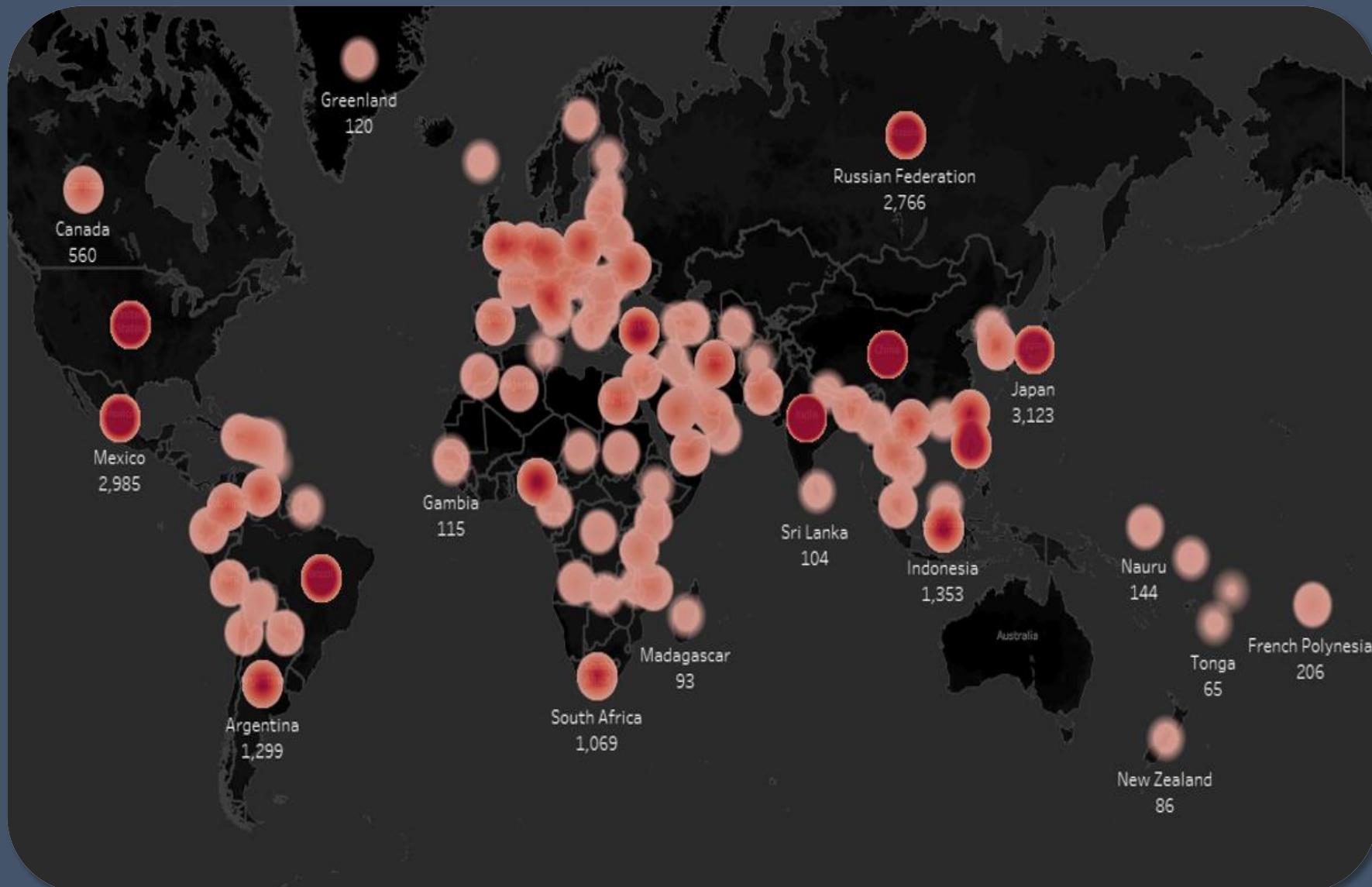
Top 5 Customers Globally per Revenue



Top 10 Countries by Revenue



Rockbuster: Total Revenue Per Country



Recommendations for implementing new business segment

Subscription plan:

- Standard plan of \$9.99
- Premium Plan of \$13.99

Increase spending in largest revenue silos:

- PG-13 Rated Movies
- Movies within Sports Genre
- India, China, & South America

Create Loyalty Program to retain customers in top markets like:

- India, China, USA, Japan, Mexico

- Sports is the largest revenue producing genre comprising 8% of total market share in terms of revenue globally. The top ten titles by revenue command 3% of global sales.
- Based on this information we believe Rockbuster should update the top ten list of movies daily on the new Rockbuster stealth streaming service to constantly drive demand to the new ecosystem.
- Rockbuster should focus a slightly larger number of resources on the Sports genre since it alone compiles 8% of global sales.
- Outside of the sports genre the percentage of sales is similar, and resources should be spread evenly so we can reach the largest amount of the addressable market.

INSTACART BASKET ANALYSIS

Objective:

You're an analyst for an existing company, Instacart, 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. Your task is to perform an initial data and exploratory analysis of some of their data in order to derive insights and suggest strategies for better segmentation based on the provided criteria.

Dataset:

[Download Customer Data Set](#)

Source: www.CareerFoundry.com

[Download Instacart Data Set](#)

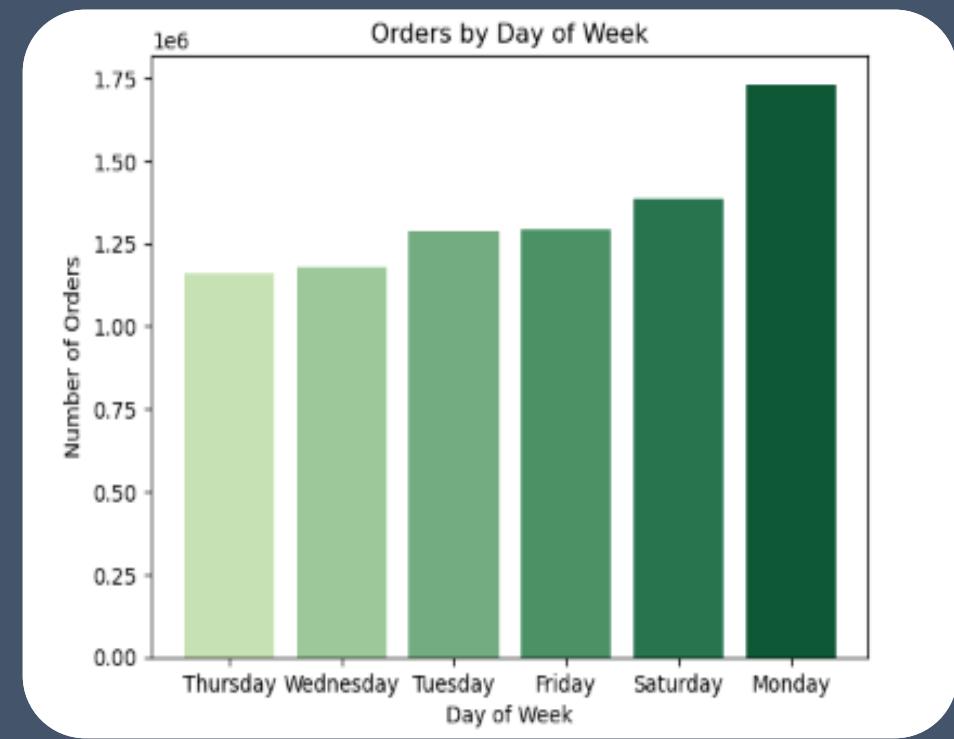
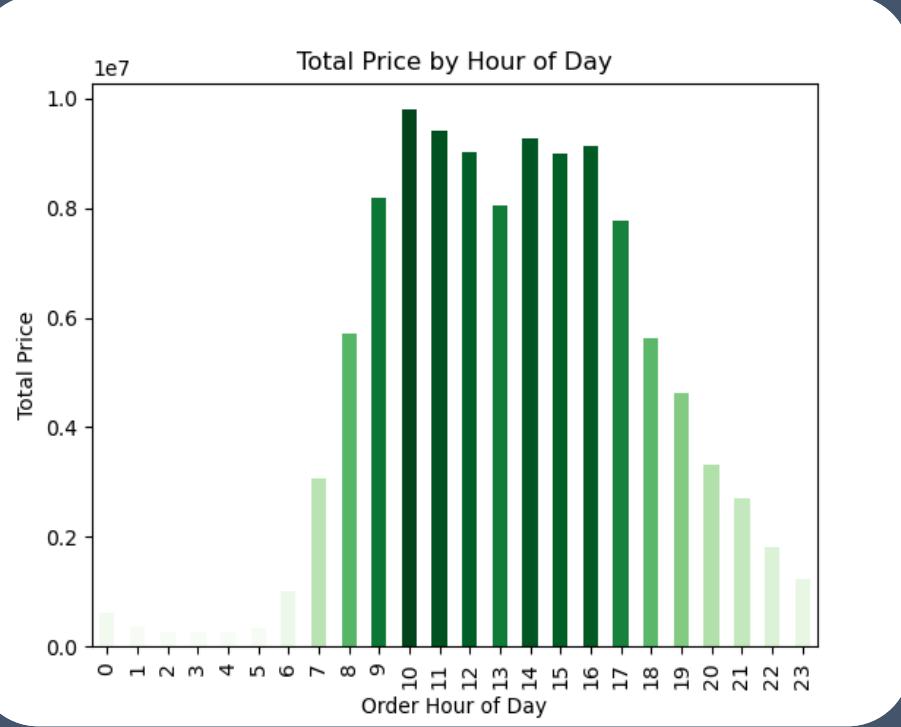
Source: [Kaggle](https://www.kaggle.com)

Skills Used:

- Data Wrangling
- Data Merging
- Deriving Variables
- Grouping Data
- Aggregating Data
- Reporting
- Population Flow

Tools Used:



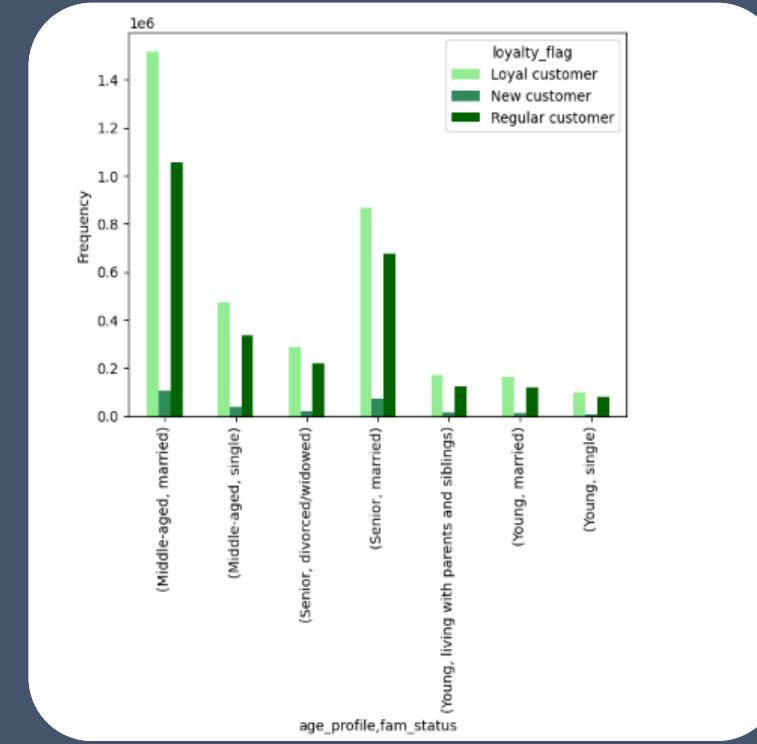
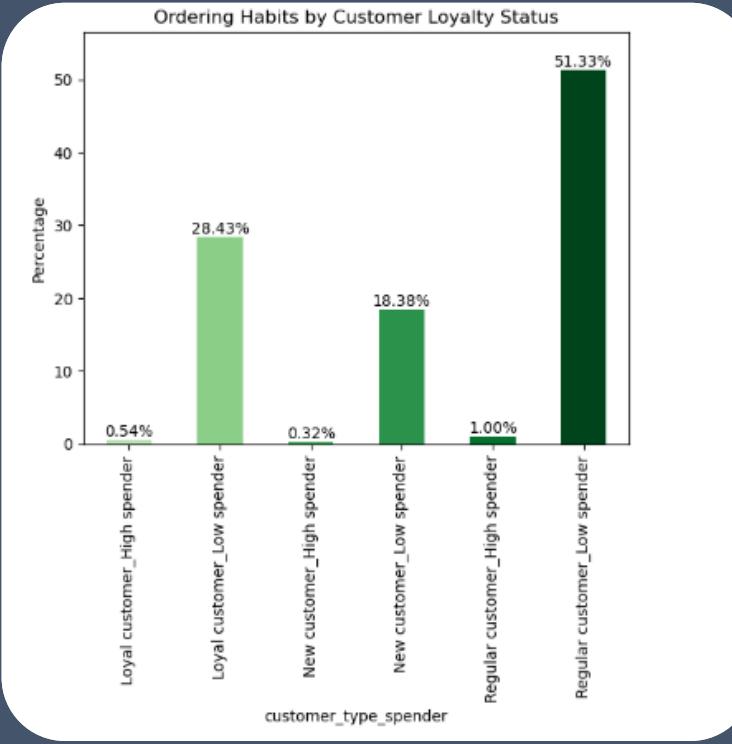
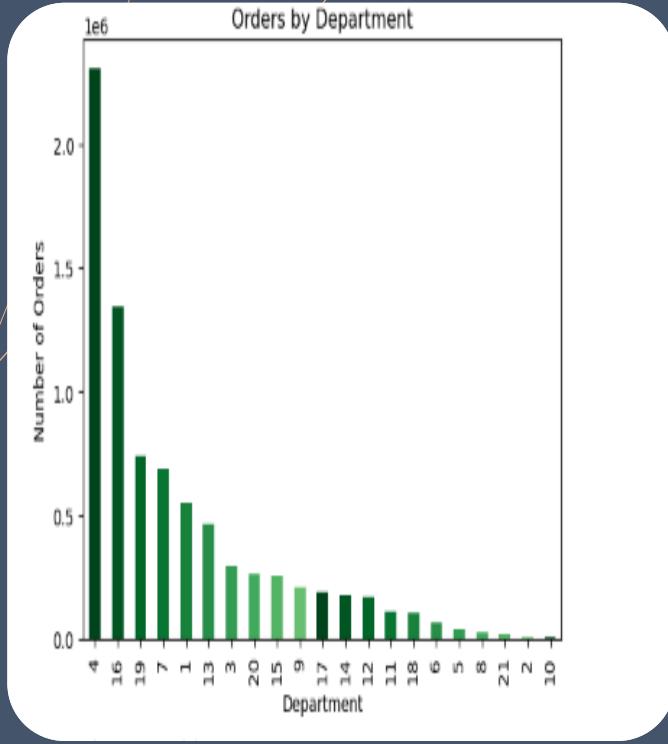


These 2 visuals help to give us a good understanding of the busiest times of the day and the busiest days for Instacart.

Busiest Hours of the Day:
10 AM – 11 AM & 2 PM -3 PM

Busiest Day of the Week:

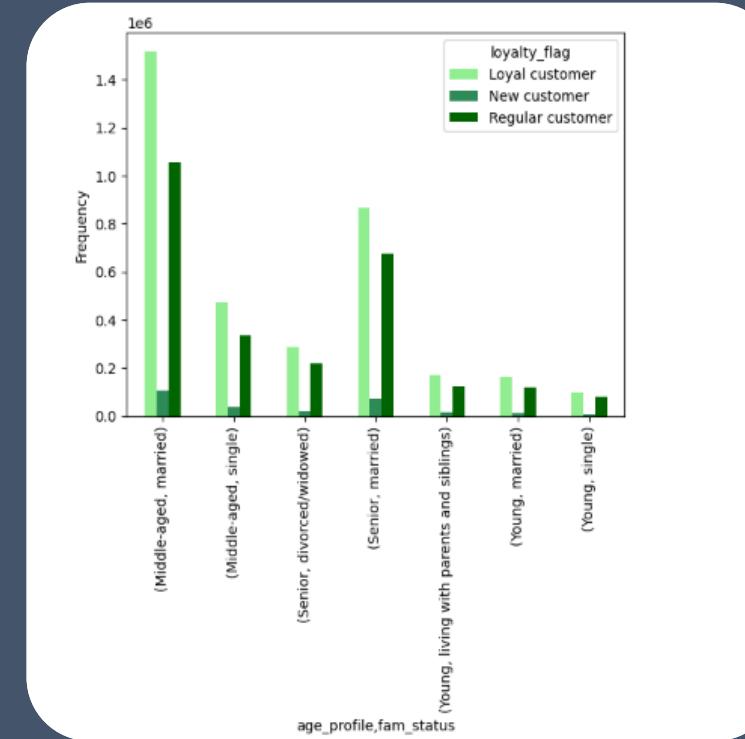
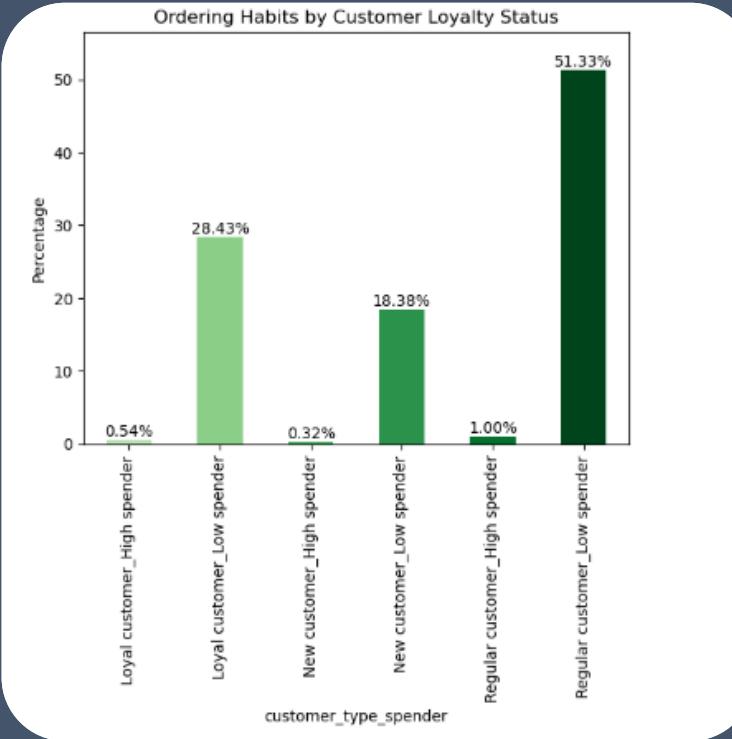
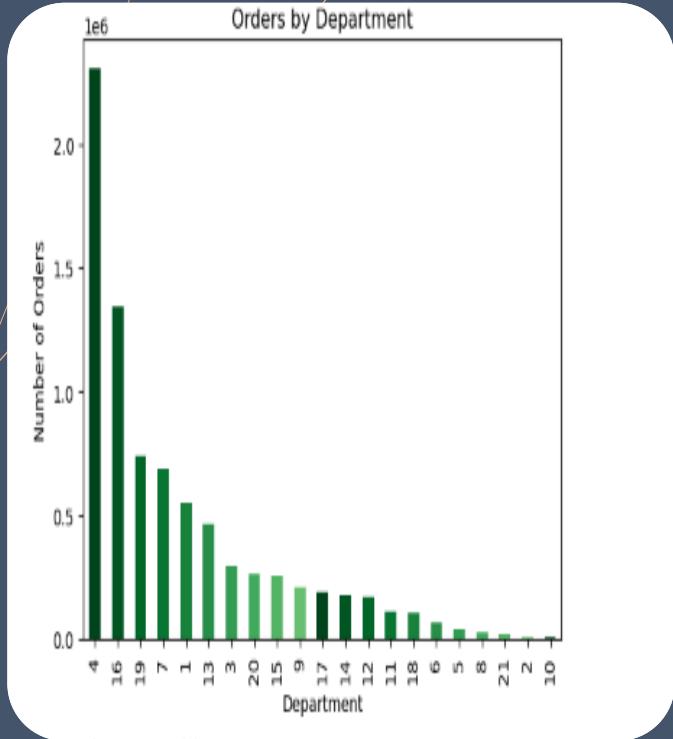
Monday is by far the busiest day for Instacart.



These visuals help us breakdown which departments have the most sales. Produce(4), Dairy/Eggs(16) and Snacks(19).

The next charts break down the ordering habits of the different types of customers that Instacart has Loyal, New and Regular Customer, and if that customer is considered a High or Low Spender with Instacart. From this chart we can see that half or 51% of Instacart's clients are regular customers that spend a low amount. The second highest category is loyal customers that spend low amounts at almost 29% of there customer base.

The 3rd chart to the far right helps to organize customers marital status with how loyal they are to Instacart. Based on this chart the older the client and if they are married the more likely they are to be loyal customers. This makes sense because married couples typically have dual incomes and more need to groceries versus dining out at restaurants, because they have a higher probability of having children.



These visuals help us breakdown which departments have the most sales. Produce(4), Dairy/Eggs(16) and Snacks(19).

The next charts break down the ordering habits of the different types of customers that Instacart has Loyal, New and Regular Customer, and if that customer is considered a High or Low Spender with Instacart. From this chart we can see that half or 51% of Instacart's clients are regular customers that spend a low amount. The second highest category is loyal customers that spend low amounts at almost 29% of there customer base.

The 3rd chart to the far right helps to organize customers marital status with how loyal they are to Instacart. Based on this chart the older the client and if they are married the more likely they are to be loyal customers. This makes sense because married couples typically have dual incomes and more need to groceries versus dining out at restaurants, because they have a higher probability of having children.

NYC YELLOW

Dataset:

NYC Yellow Cab Data

Github:

<https://github.com/MarkSt affordReece/ NYC-Yellow>

Skills Used:

- Grouping Data
- Summarizing Data
- Descriptive Analysis
- Visualizing Data in Excel
- Cleaning and Formatting Data
- Data Wrangling
- Data merging, concatenating,
- Aggregating Data
- Geospatial Data
- Translating Business Requirements
- Data Cleaning
- Data Integration
- Statistical Hypothesis Testing
- Visual Analysis
- Forecasting
- Dashboard Design
- Utilizing Github for project organization and display

Objective:

Utilize the companies' data to increase revenue

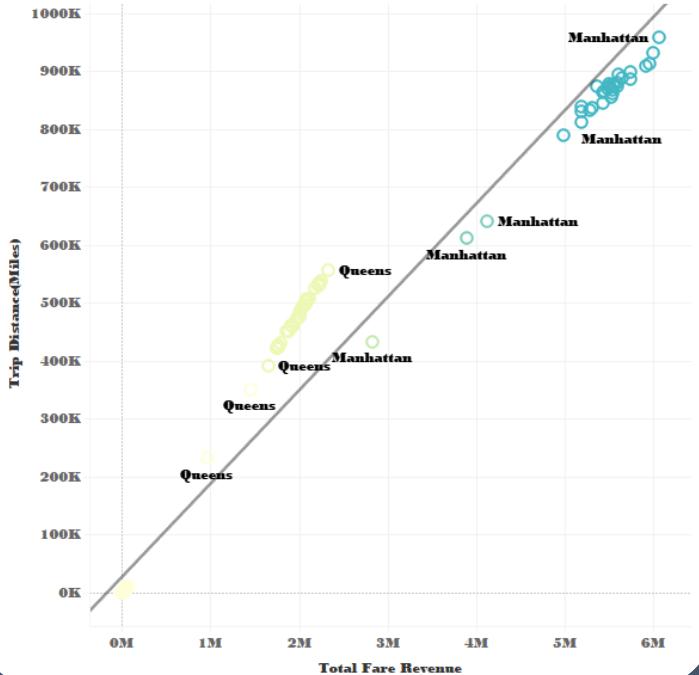
Link to Tableau Visual

[NYC Yellow: Business Report | Tableau Public](#)

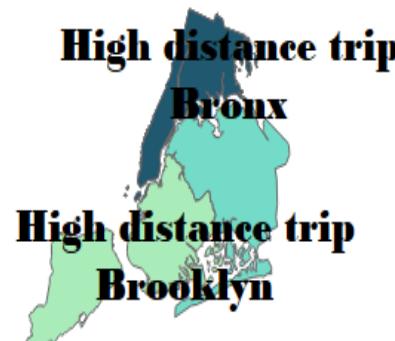
Tools Used:



Trip/Fare/Tips

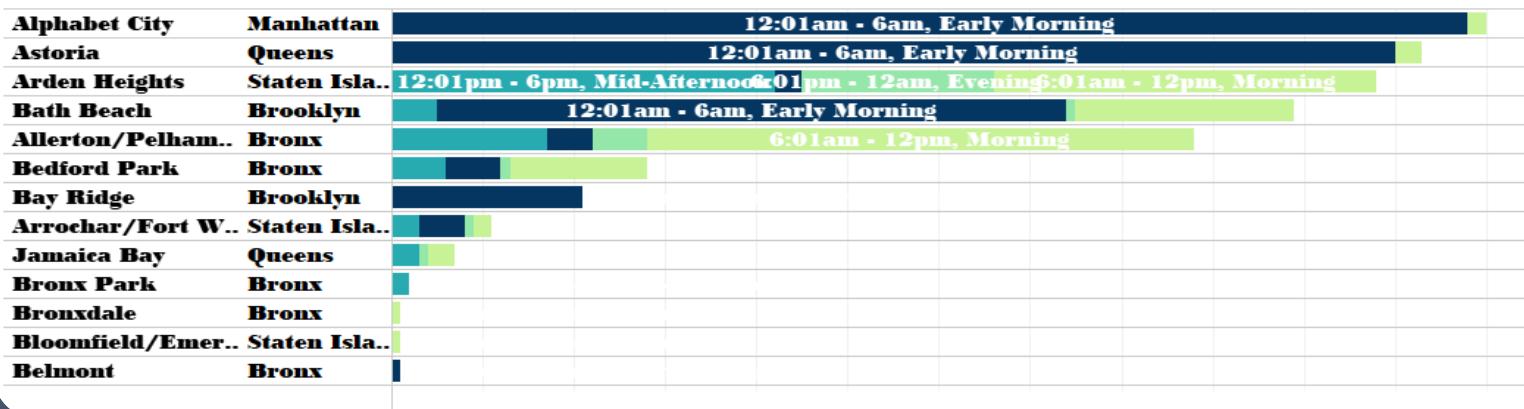


Q1 Time Series: Type of Trip Distance



I created scatterplots that helped visualize the relationship between trip distance and fare amount. I created a geospatial map from and then added informative layers to analyze trip distance and fare amount.

Busiest Borough/Zone for Pick Ups



This graph helps to show the relationship between borough, zone and time of day. These graphs utilized many different flags that I created with python to help create relationships between the data so its easier to see analyze.



THANK YOU

Stafford Reece

Reecems@vcu.edu