

# Data Analytics Project Portfolio

BY  
COURTNEY  
SHATLEY

# Table of Contents

---

GameCo

Slide 3

---

Influenza Preparation

Slide 9

---

Rockbuster

Slide 15

---

Instacart

Slide 20

---

Global Bank

Slide 26

---

New York Citi Bike

Slide 29

# GameCo

---

# GameCo Summary

---

## Project Goal

- Analyze global videogame sales data to gain insights into future opportunities and development.

## Tools

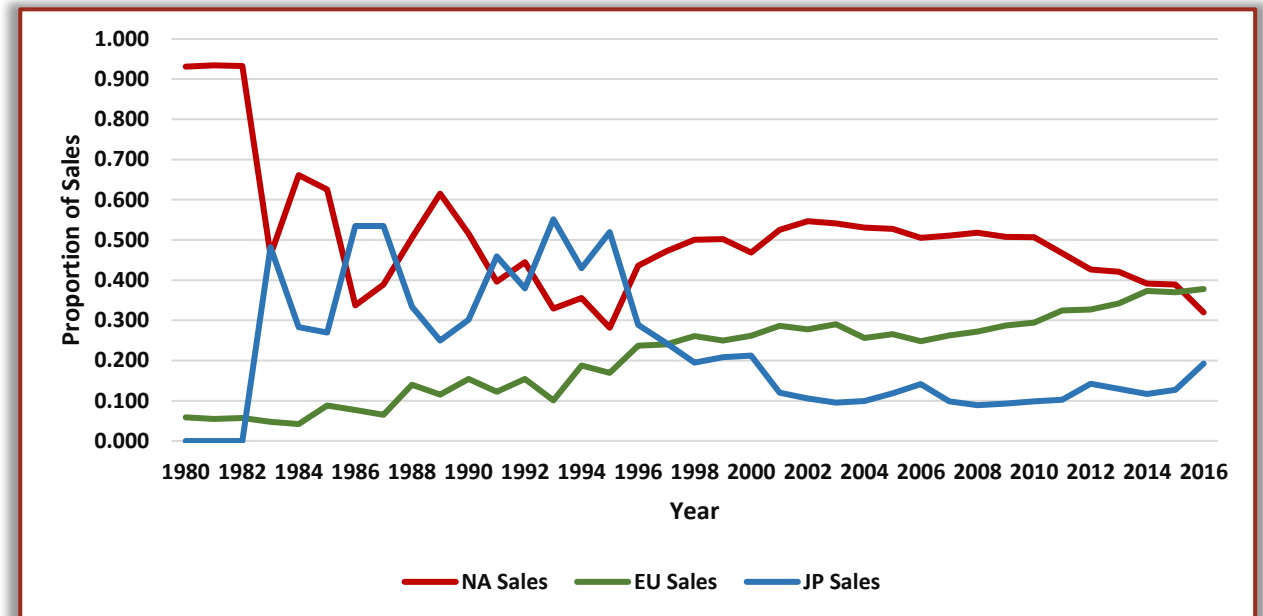
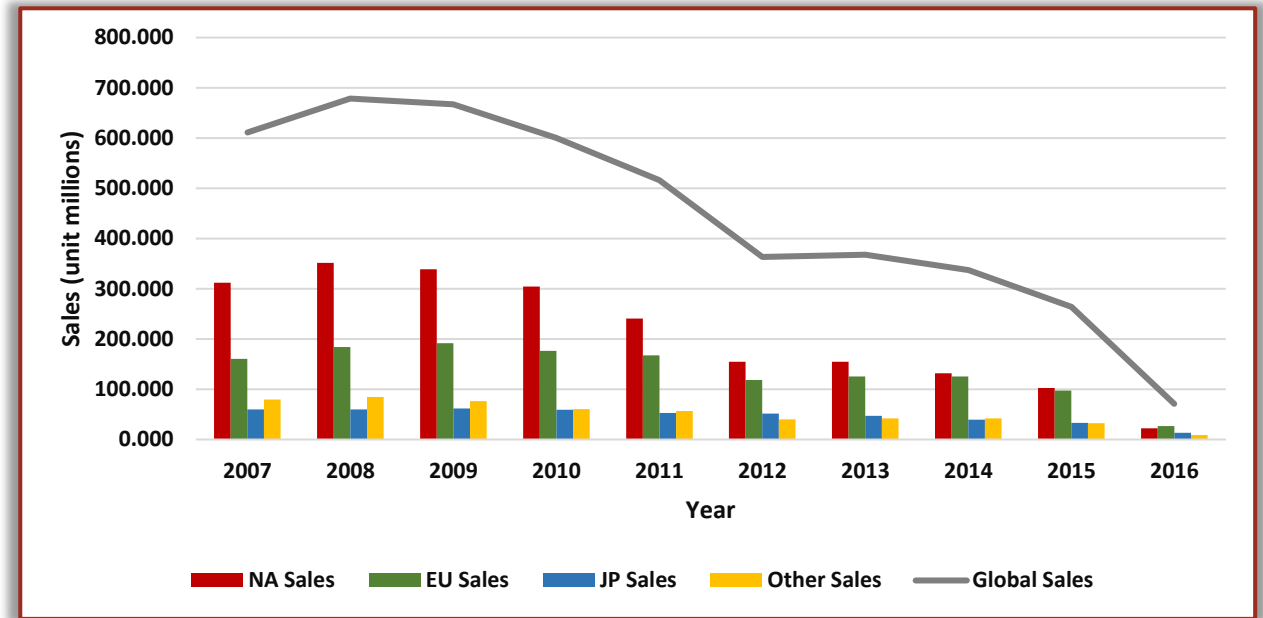
- Microsoft Excel
- Microsoft Powerpoint

## Skills

- Data Cleaning
- Data Grouping & Summarizing
- Descriptive Analysis
- Developing Insights
- Visualization
- Storytelling

# Sales Overview

- Overall sales have been decreasing with significant drops in 2012 and 2016.
- European sales have been steadily increasing and make up the highest proportion of sales in 2016.



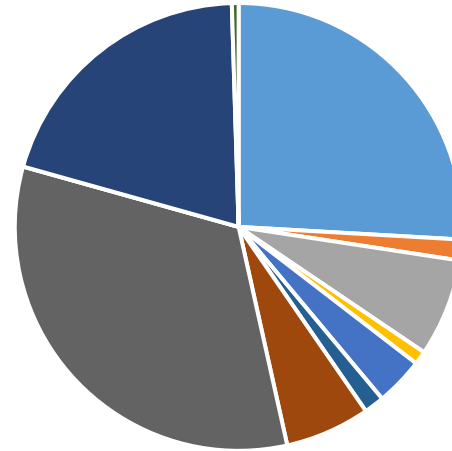
# Genre Popularity

➤ Four genres dominate sales across all regions.

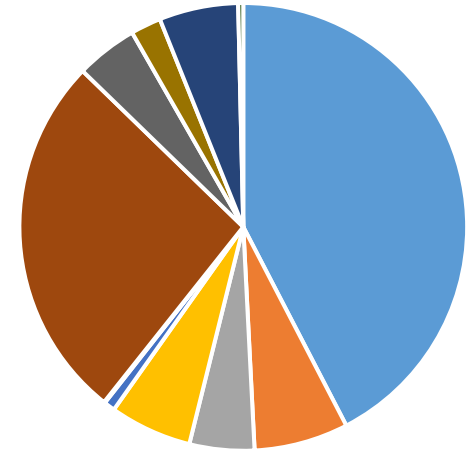
NA	JP	EU	Global
Shooter	Action	Shooter	Action
Action	Role-Playing	Sports	Shooter
Sports	N/A*	Action	Sports

\*There is no third genre in Japan representing >20% of sales.

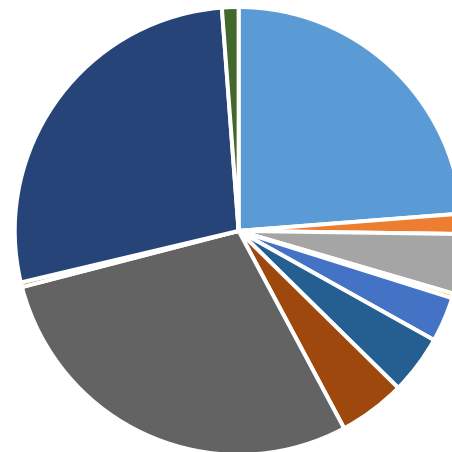
NA Sales 2016



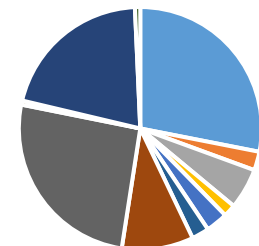
JP Sales 2016



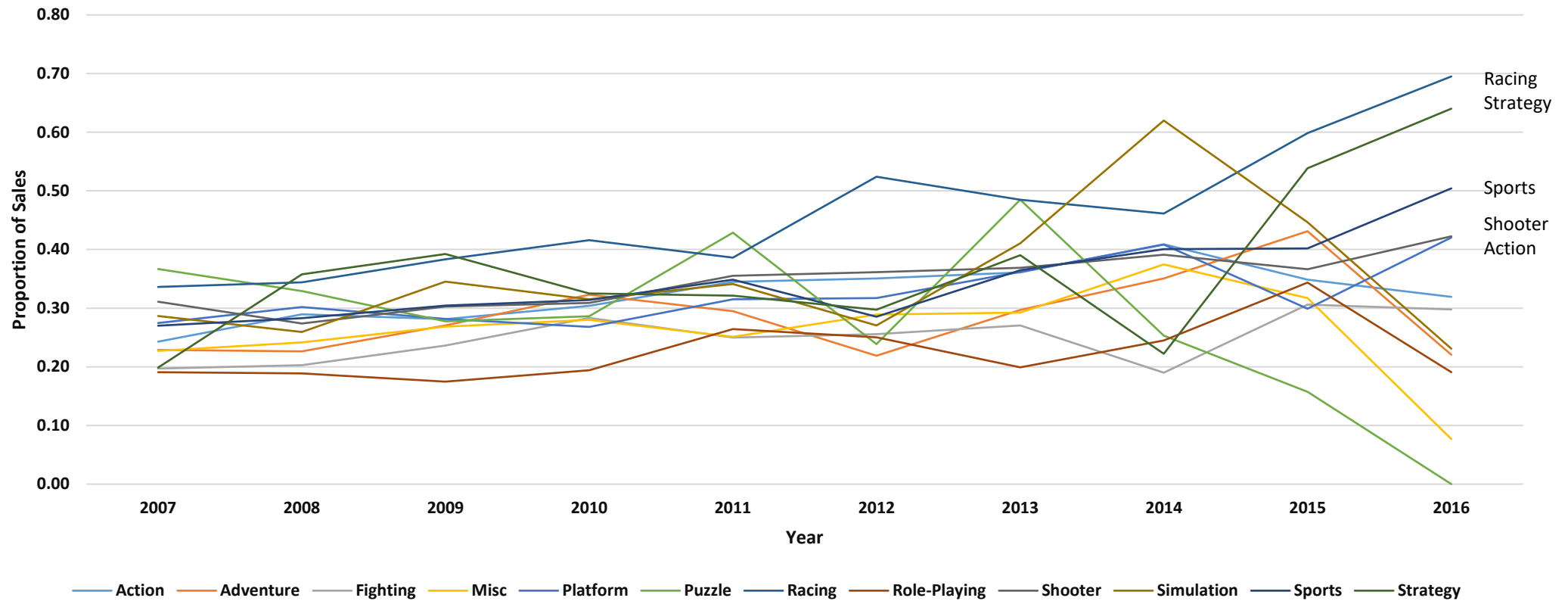
EU Sales 2016



Global Sales 2016



EU Proportion of Global Sales by Genre  
Past 10 Years



- Racing and strategy games lead the increase observed in the European market.
- Sports, shooter, and action games have continued in popularity.

# Key Takeaways

---

- The percentage of global sales from Europe has been increasing.
  - Racing and Strategy games contribute most to this trend despite not being the most popular genres.
  - Investment in the European market may be a profitable step in 2017.
- Global and regional sales in 2016 comprise mainly of four popular genres.
  - Action, Shooter, Sports, and Role-Playing genres are the most popular.
  - Action games are significantly popular across all regions and contribute the most to global sales.
  - Games of these genres will likely continue to be popular in 2017.



# Influenza Preparation

---

# Influenza Preparation Summary

---

## Project Goal

- Define flu season and assess impact on vulnerable population.
- Determine where to send medical staff for the coming influenza season.

## Tools

- Microsoft Excel
- Microsoft Powerpoint
- Tableau

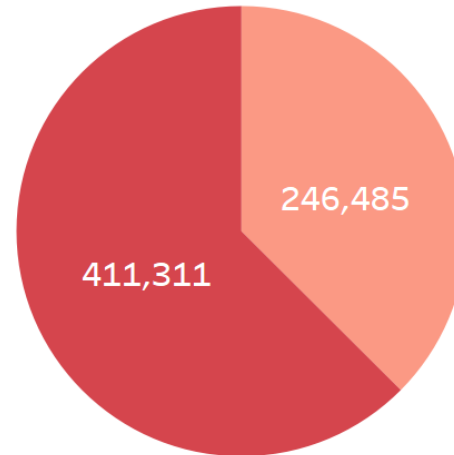
## Skills

- Data Cleaning
- Data Integration
- Data Transformation
- Forecasting
- Visualization
- Storytelling with Tableau

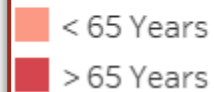
# Vulnerable Population

- Individuals over 65 years have a significantly higher influenza mortality rate.

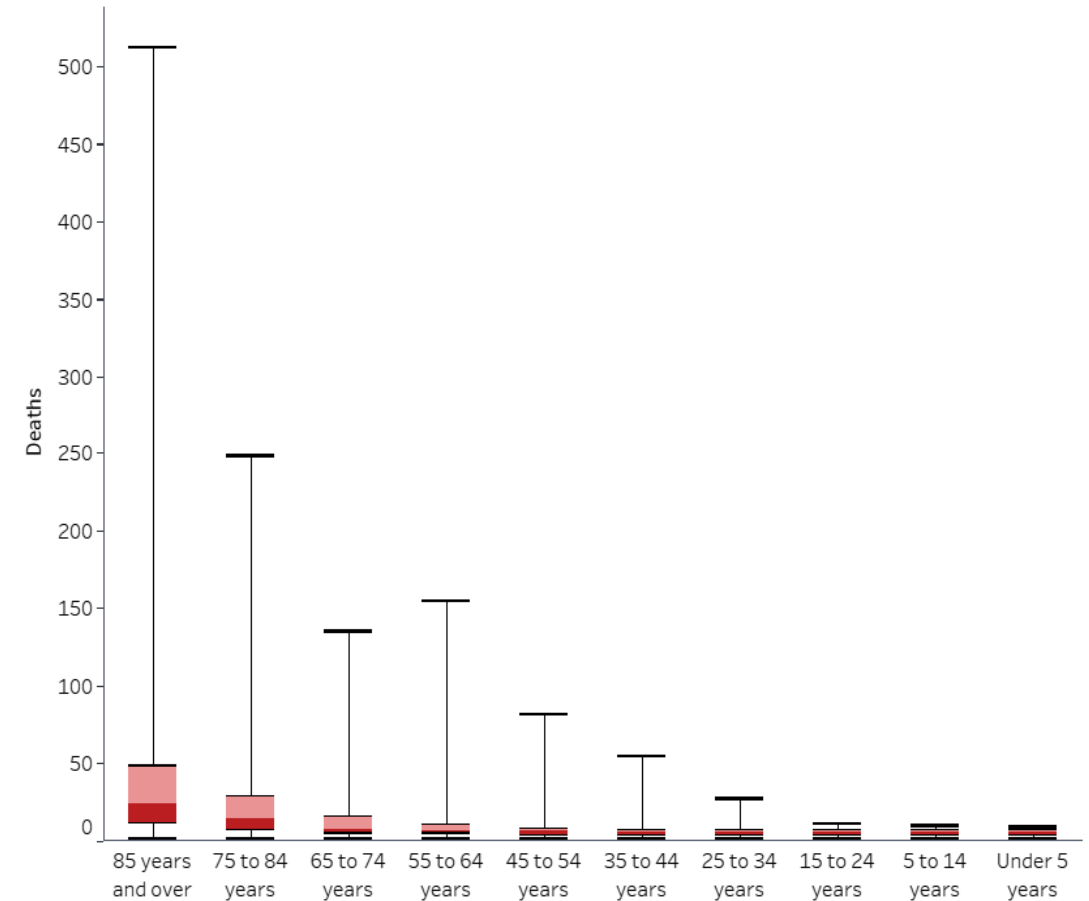
Total Influenza Deaths By Age Group  
(2009-2017)



Measure Names

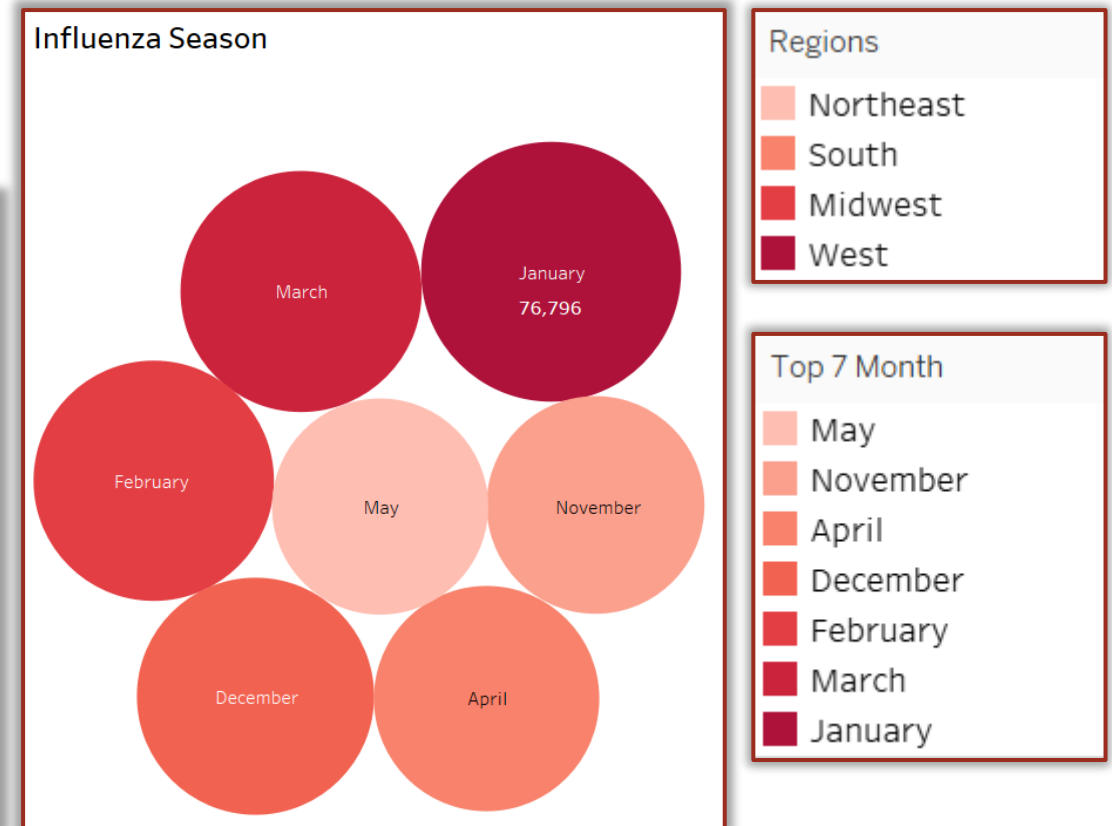
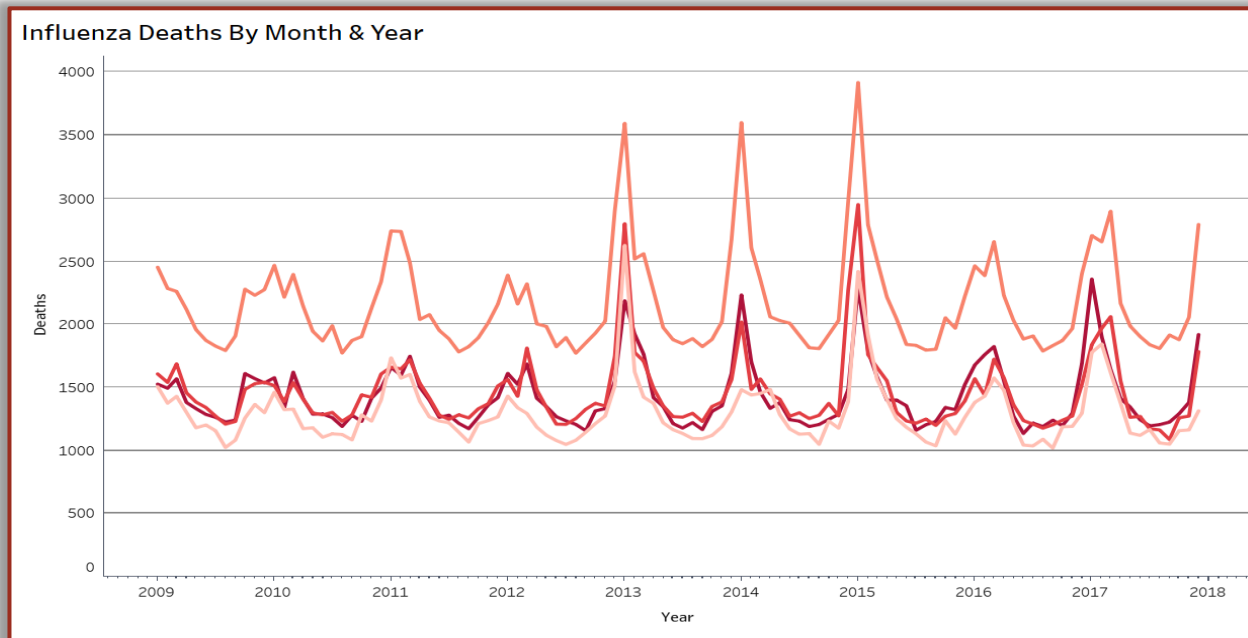


Total Influenza Deaths By Age Group  
(2009-2017)



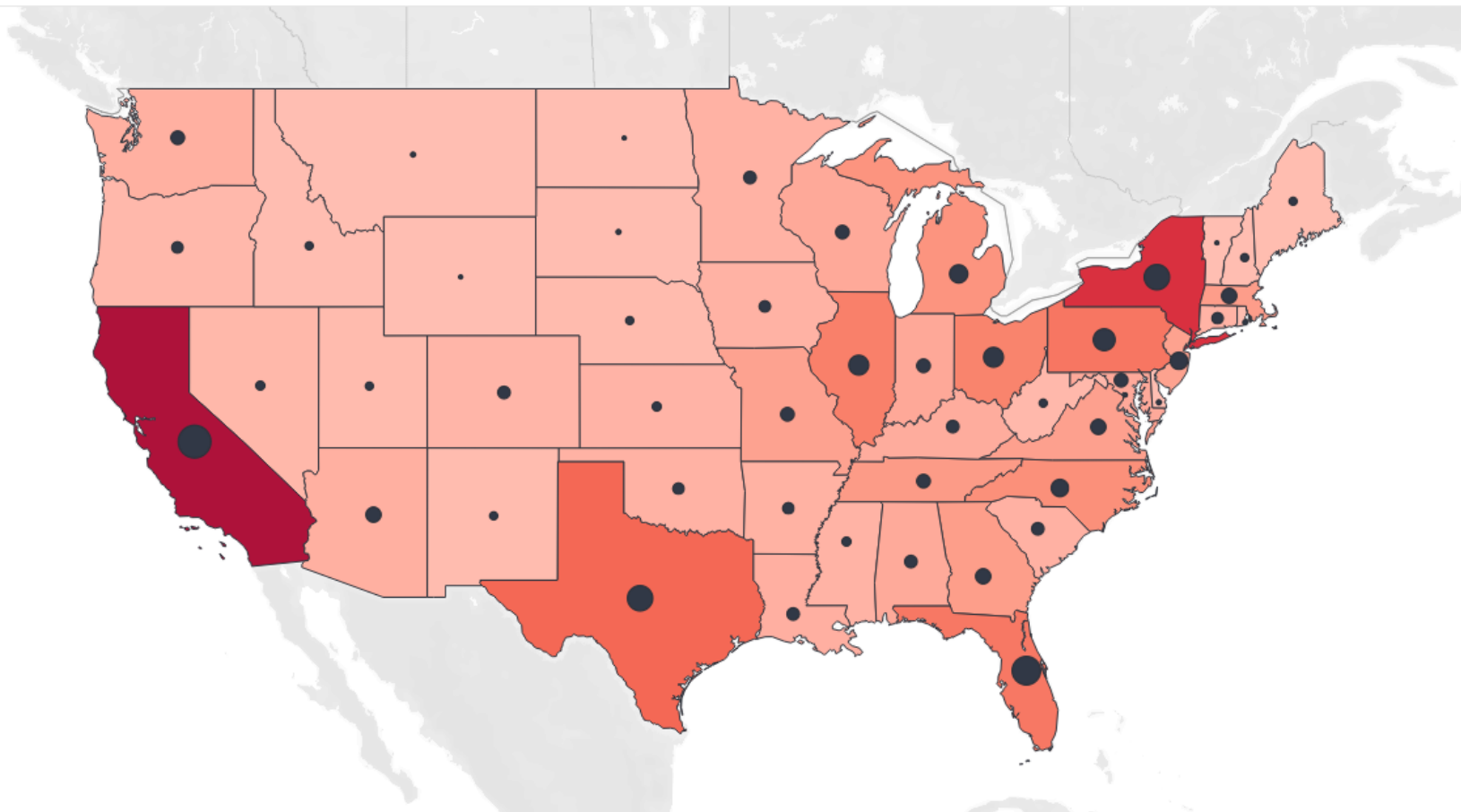
# Influenza Season

- The pattern of influenza deaths is seasonal and peaks in January.



## Population and Influenza Deaths

Ages > 65 Years



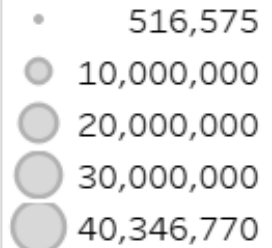
Influenza Deaths



Population > 65 Years



Population



- California and New York feature the highest influenza death count.
- There is a positive correlation between state population and influenza deaths.

# Key Takeaways

---

- Individuals over 65 years of age are significantly more vulnerable to influenza related death than their younger counterparts.
- The seasonal fluctuation of influenza deaths is consistent over the years and is likely to continue into 2018. Flu Season may be approximated as the months December-March with the peak of influenza deaths occurring in January.
- There are particular states where the influenza death count is significantly high throughout the years 2009-2017 and these states will require additional medical staffing in 2018.

# Rockbuster

---

# Rockbuster Summary

---

## Project Goal

- Assist in the launch of an online video rental service by analyzing historical sales data.

## Tools

- Microsoft Excel
- Microsoft Powerpoint
- SQL
- Tableau
- DBVisualizer

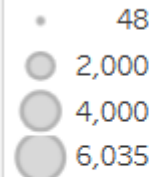
## Skills

- Relational Databases
- Database Querying
- Filtering
- Summarizing Data in SQL
- Subqueries
- Joining Tables
- Presenting

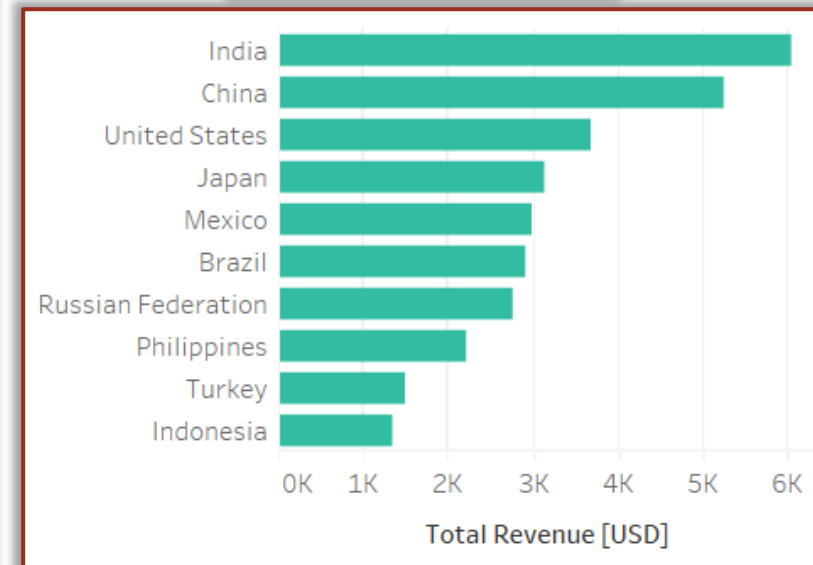




#### Total Revenue



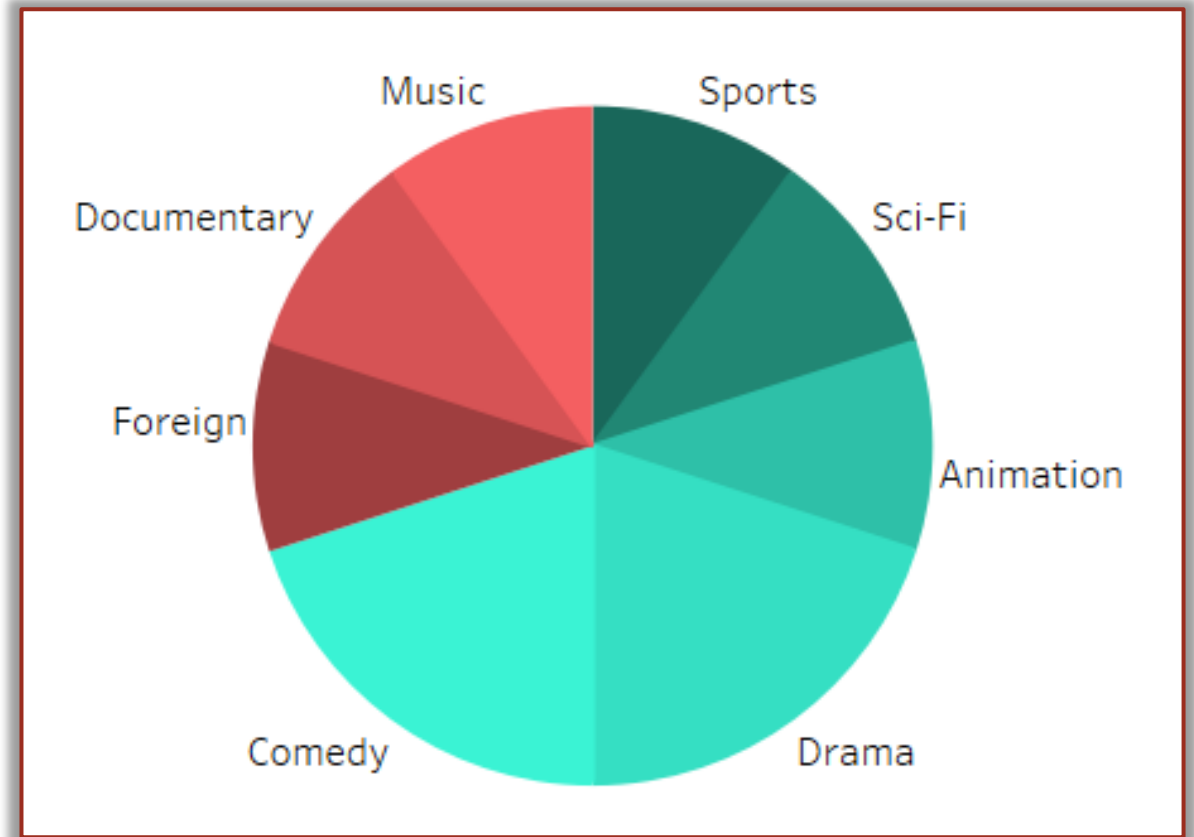
#### Count Of Customers



- Total revenue across all 109 countries was \$61,312.
- India, China, and United States provided the highest revenue.

# Genre Popularity

Film	Genre
Telegraph Voyage	Music
Zorro Ark	Comedy
Wife Turn	Documentary
Innocent Usual	Foreign
Hustler Party	Comedy
Saturday Lambs	Sports
Titans Jerk	Sci-Fi
Harry Idaho	Drama
Torque Bound	Drama
Dogma Family	Animation



- Most popular genres are Comedy, Drama, and Animation.
- 50% of the top 10 films belong to those genres.

# Key Takeaways

---

- The highest revenue countries are India, China, and United States.
  - Target markets in high revenue countries.
  - Consider loyalty program to hold onto lifetime customers.
  - Consider referral discounts to increase customer numbers in existing markets.
- Sports, Sci-Fi, and Animation are the most popular genres.
  - Continue to supply high revenue genres.
  - Boost inventory of other genres belonging to popular movies such as Foreign, Documentary, and Music.
- There are opportunities to expand film inventory.
  - Include films spanning years outside of 2006 to interest customers within various demographics.
  - Increase variety of languages especially those common in the high revenue countries identified.

# Instacart

---

# Instacart Summary

---

## Project Goal

- Use exploratory analysis to uncover patterns in customer orders from Instacart data.

## Tools

- Microsoft Excel
- Microsoft Powerpoint
- Python
- Panda
- Jupyter

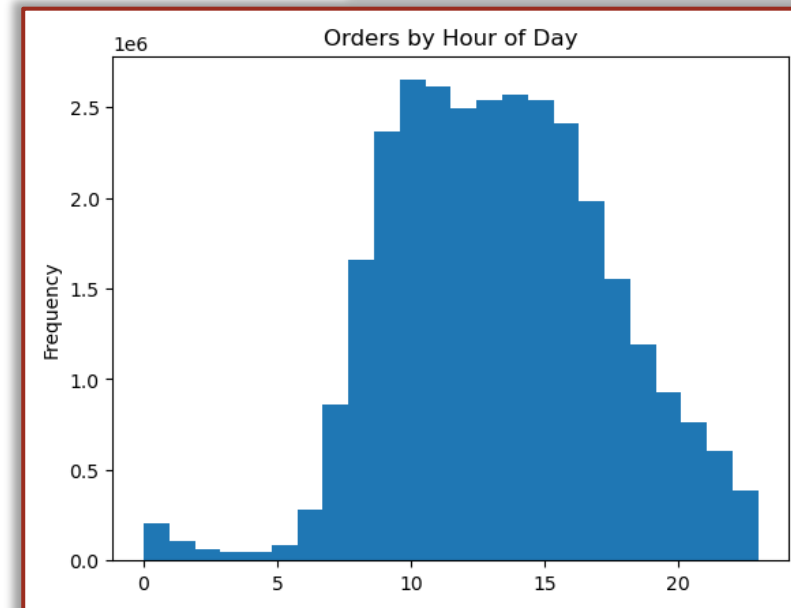
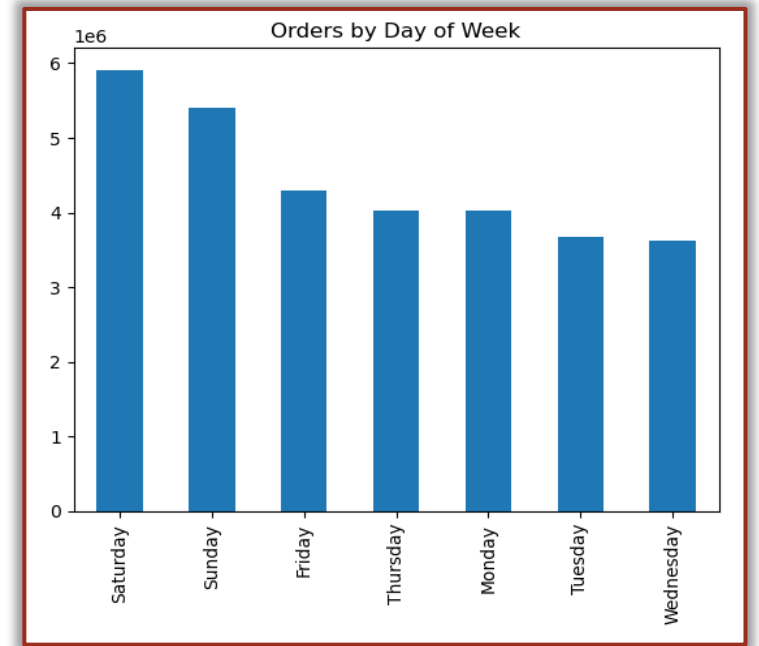
## Skills

- Data Cleaning, Wrangling, & Subsetting
- Data Consistency Checks
- Deriving New Variables
- Grouping & Aggregating Data
- Data Visualization

# Customer Activity

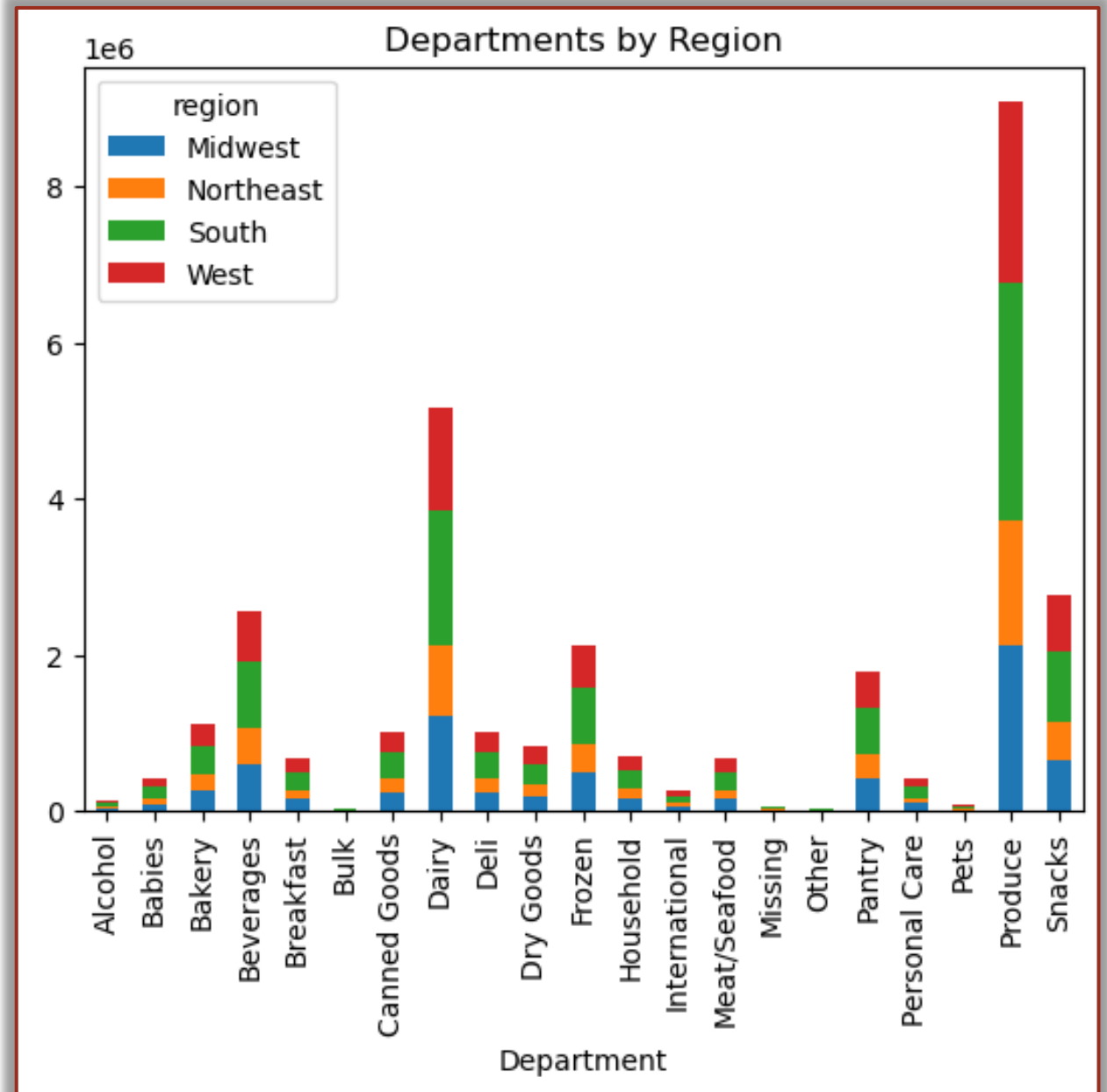
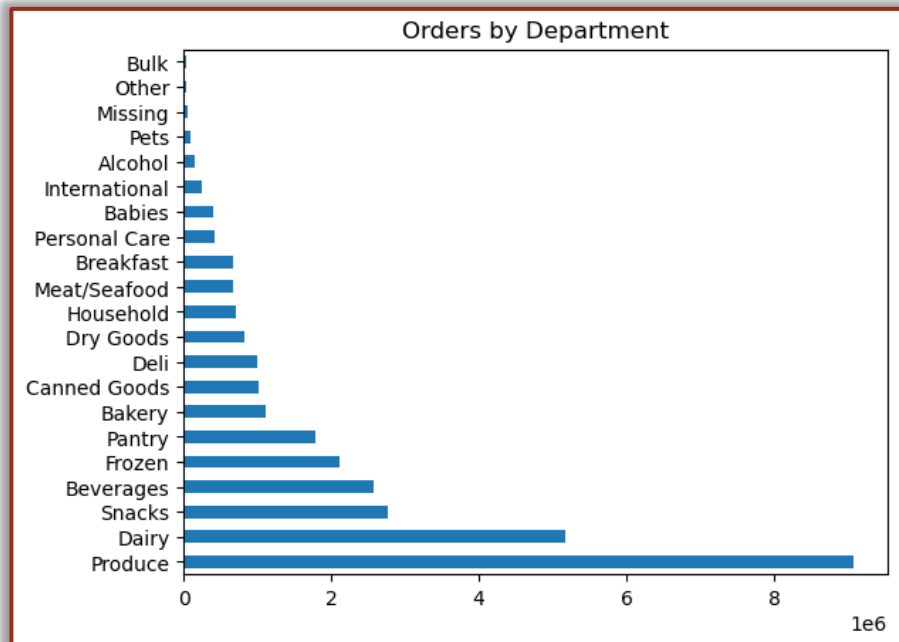
---

- Saturday and Sunday are the most popular days for customers to use Instacart.
- The busiest hours are between 9AM and 5PM.



# Popular Products

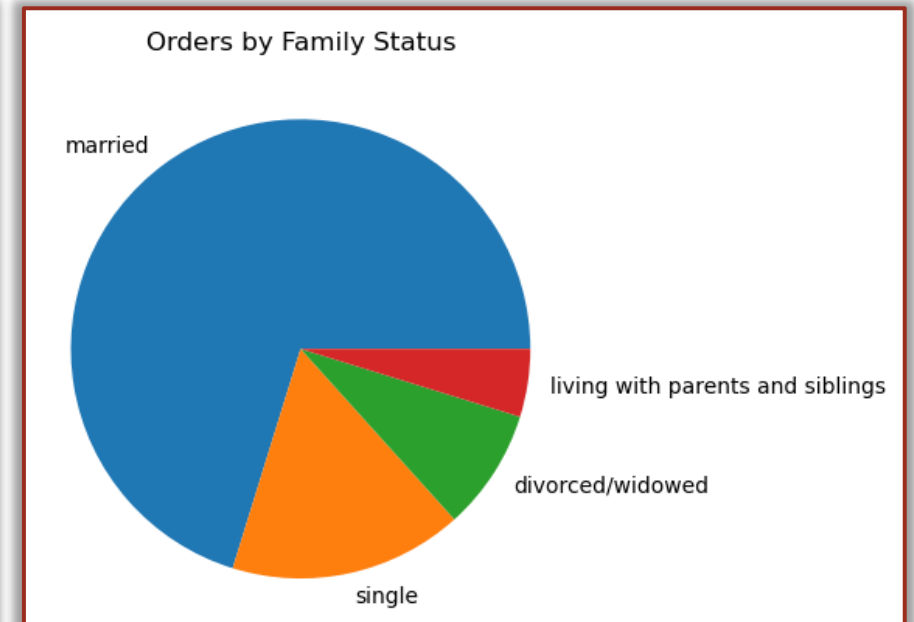
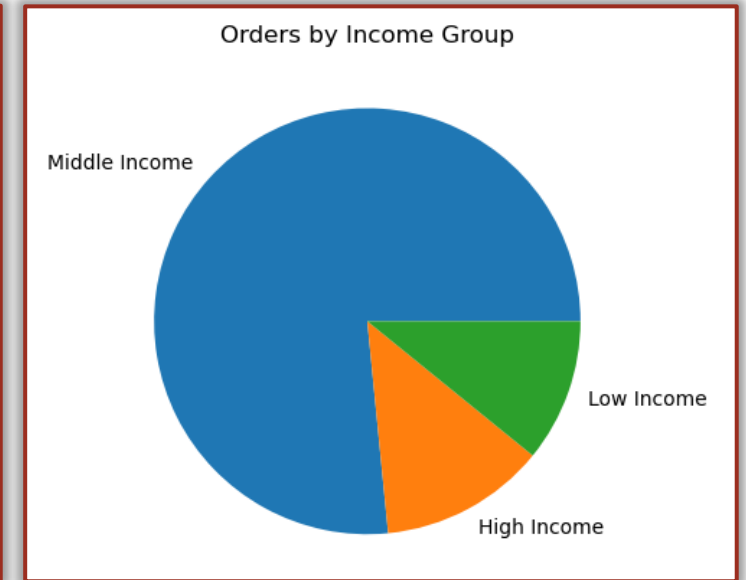
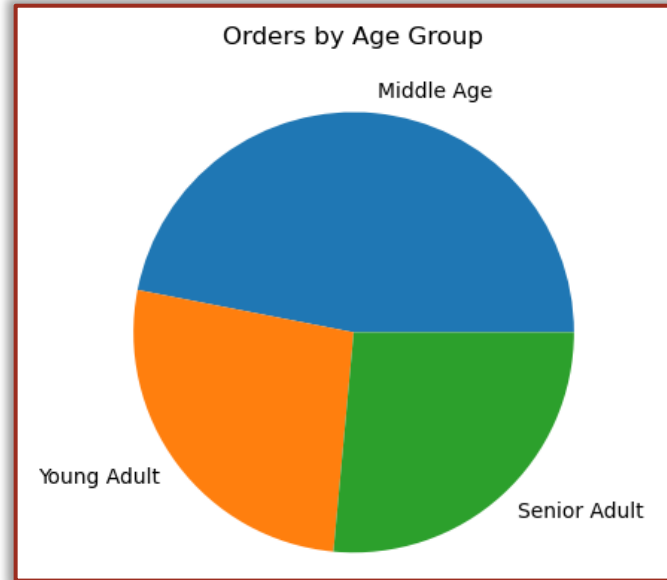
➤ Produce and dairy are the most popular products across all regions.



# Customer Population

➤ Majority of orders come from customers who are:

- Middle income
- Middle aged
- Married
- With dependents





# Key Takeaways

---

- Most orders are made on the weekends between the hours 9AM and 5PM.
  - Consider busy days/times when scheduling ads.
  - Ensure that products are stocked during these periods of high order volume.
- Produce and dairy are the most popular products.
  - Boost marketing ads that emphasis the freshness of these products.
- Middle income, middle aged, married, and customers with dependents make up most of the customer base.
  - Consider these demographics in marketing by catering ads to average family households.

# Global Bank

---

# Global Bank Summary

---

## Project Goal

- Identify leading indicators for customers leaving the bank.

## Tools

- Microsoft Excel
- Microsoft Powerpoint
- Github

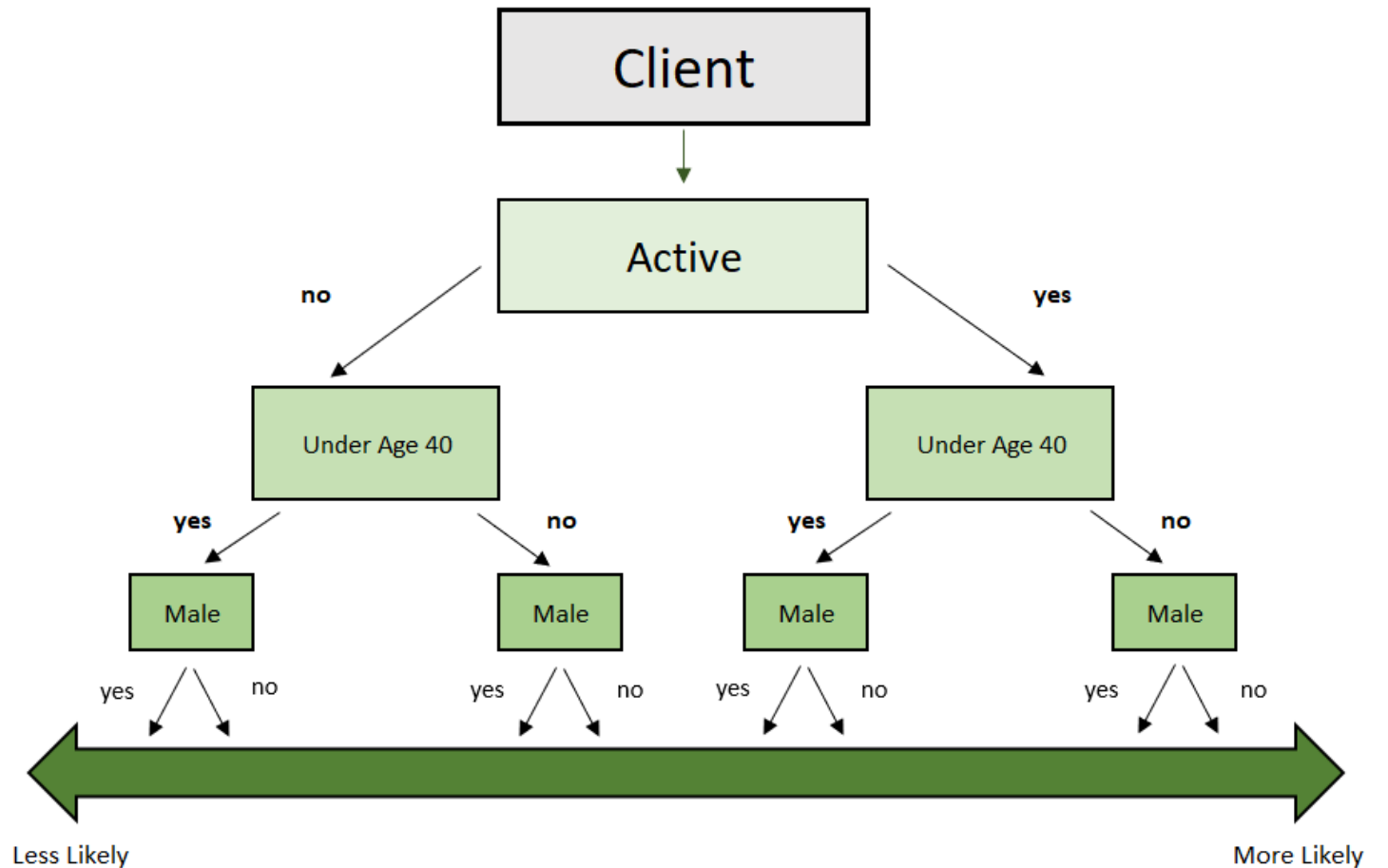
## Skills

- Big Data
- Data Ethics
- Data Mining
- Predictive Analysis
- Time Series
- Forecasting

## Decision Tree for Likelihood of Clients Leaving Pig E. Bank

➤ Leading indicators for customers leaving the bank include account activity, age, and gender.

➤ Additional factors include country and account balance.



# New York Citi Bike

---

# New York Citi Bike Summary

---

## Project Goal

- Analyze rental data from September 2013 to better understand who is using NYCB, what times and days show more rental activity, and which stations are the most popular.

## Tools

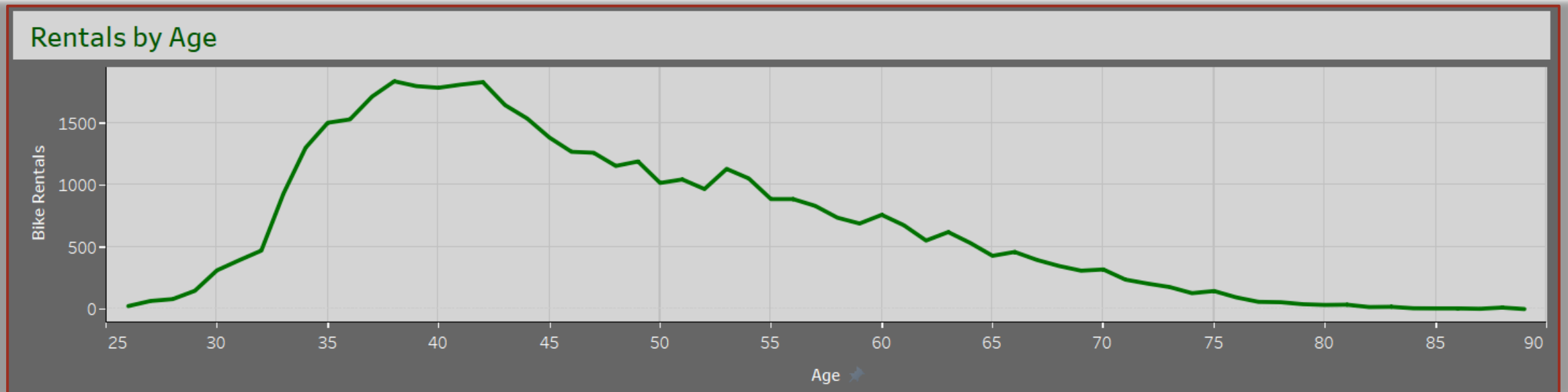
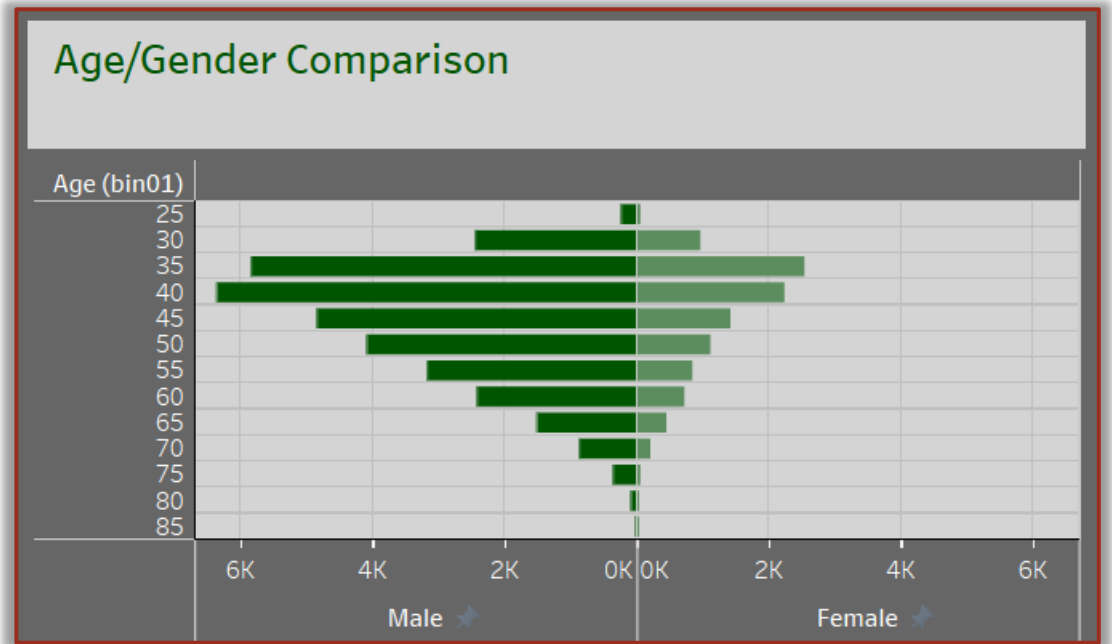
- Microsoft Excel
- Microsoft Powerpoint
- Python
- Panda
- Jupyter
- Tableau

## Skills

- Big Data
- Data Ethics
- Data Mining
- Correlation
- Predictive Analysis
- Time Series Analysis and Forecasting
- Visualization
- Storytelling with Tableau

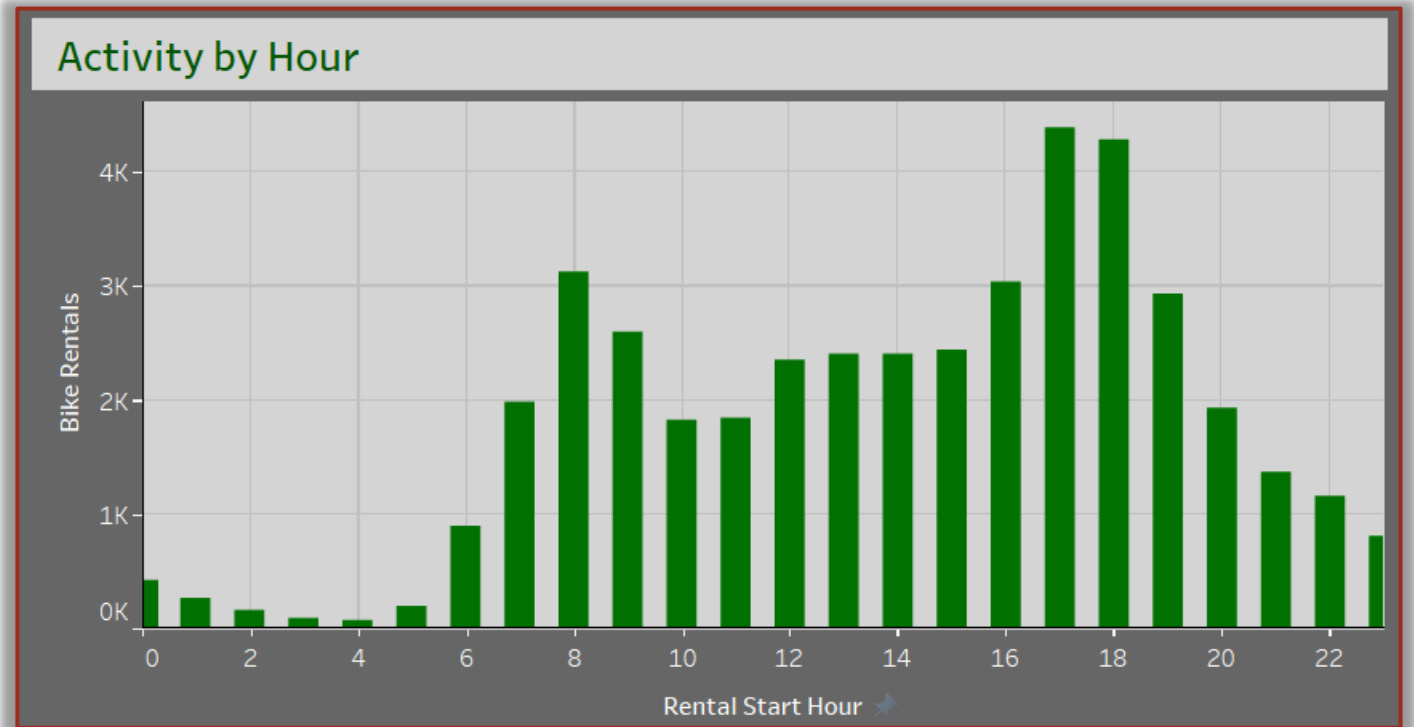
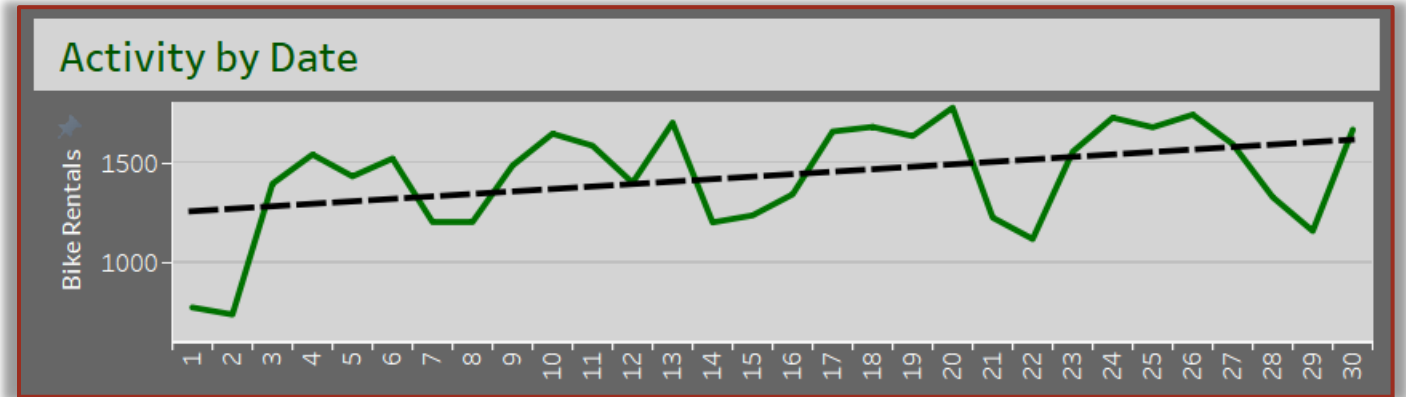
# Customer Demographic

- Majority of customers are 33 to 50.
- There are over 3x more male than female customers.



# Temporal Analysis

- Rental activity increased throughout the month of September.
- Peak hours for bike rentals are 8AM and the hours between 4PM and 6PM.

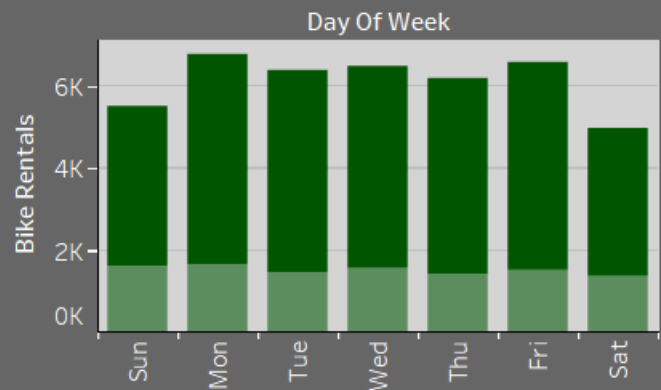




# Demographic Habits

- Young adults rent more during nighttime hours.
- Male customers rent more during weekdays.

Rental Day by Gender

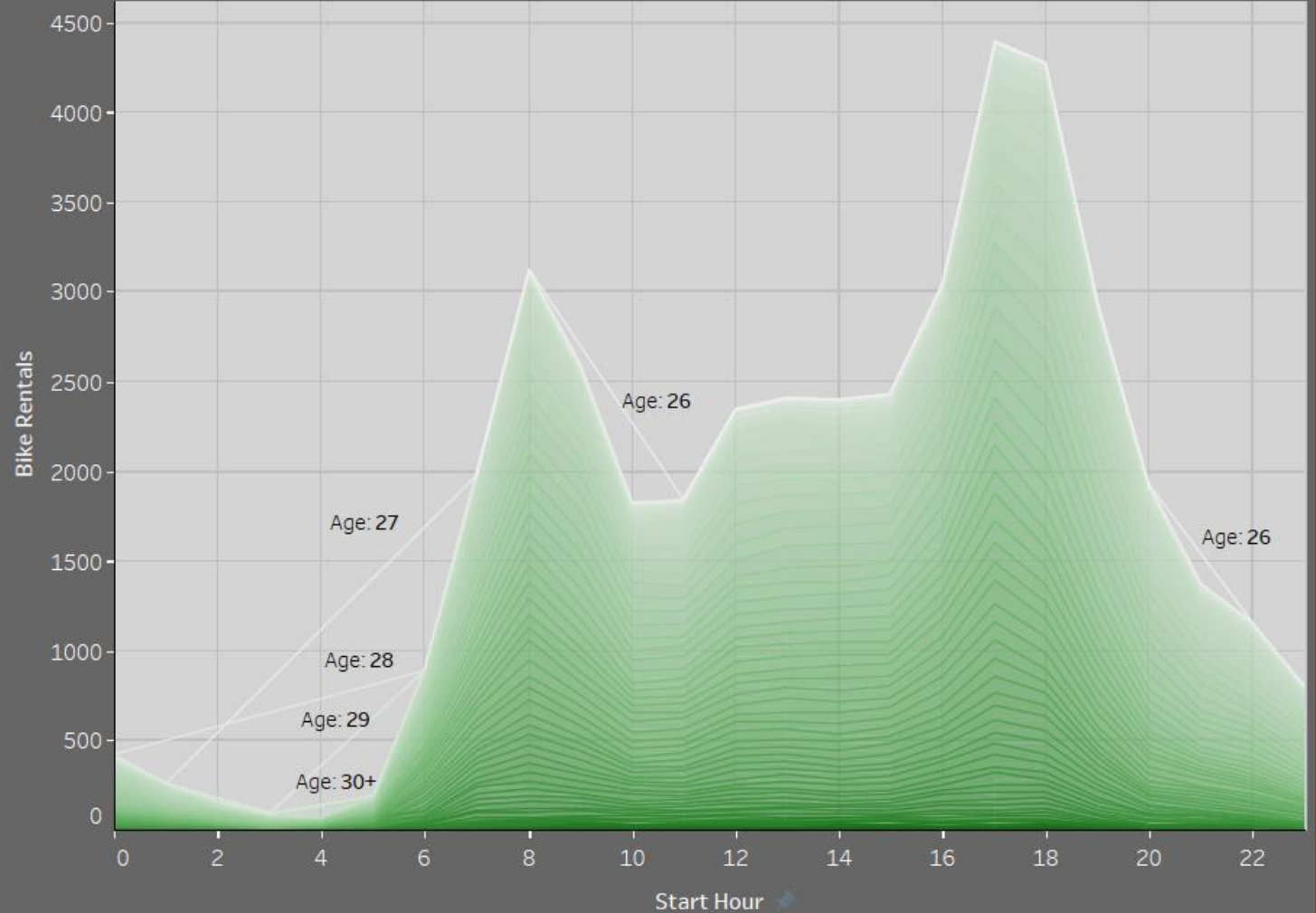


Gender

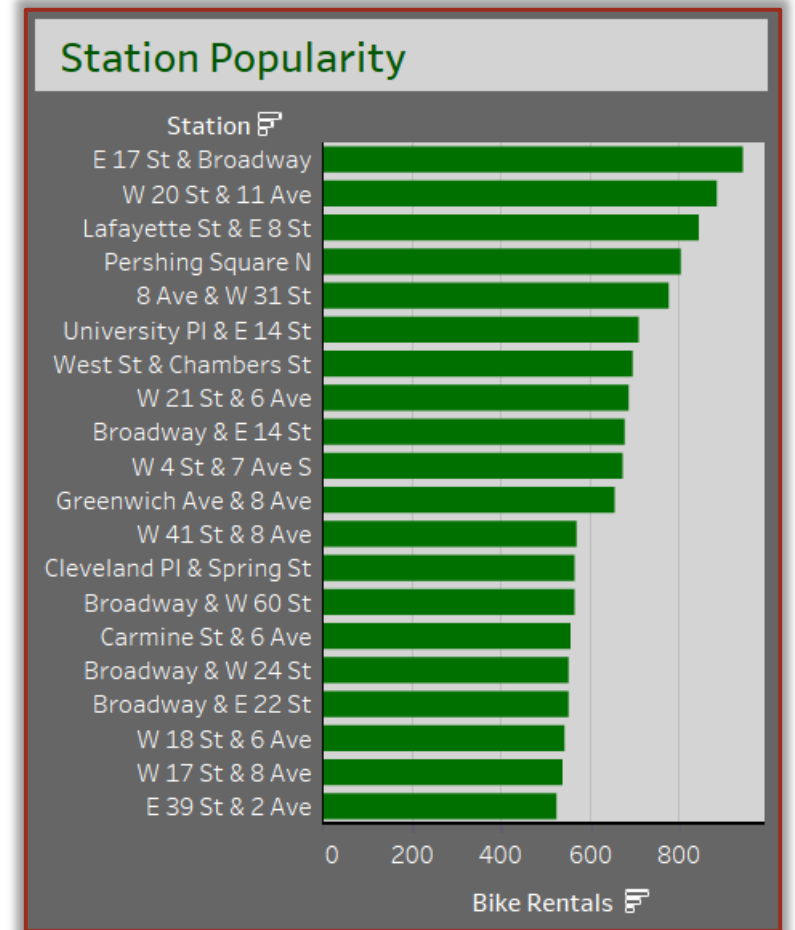
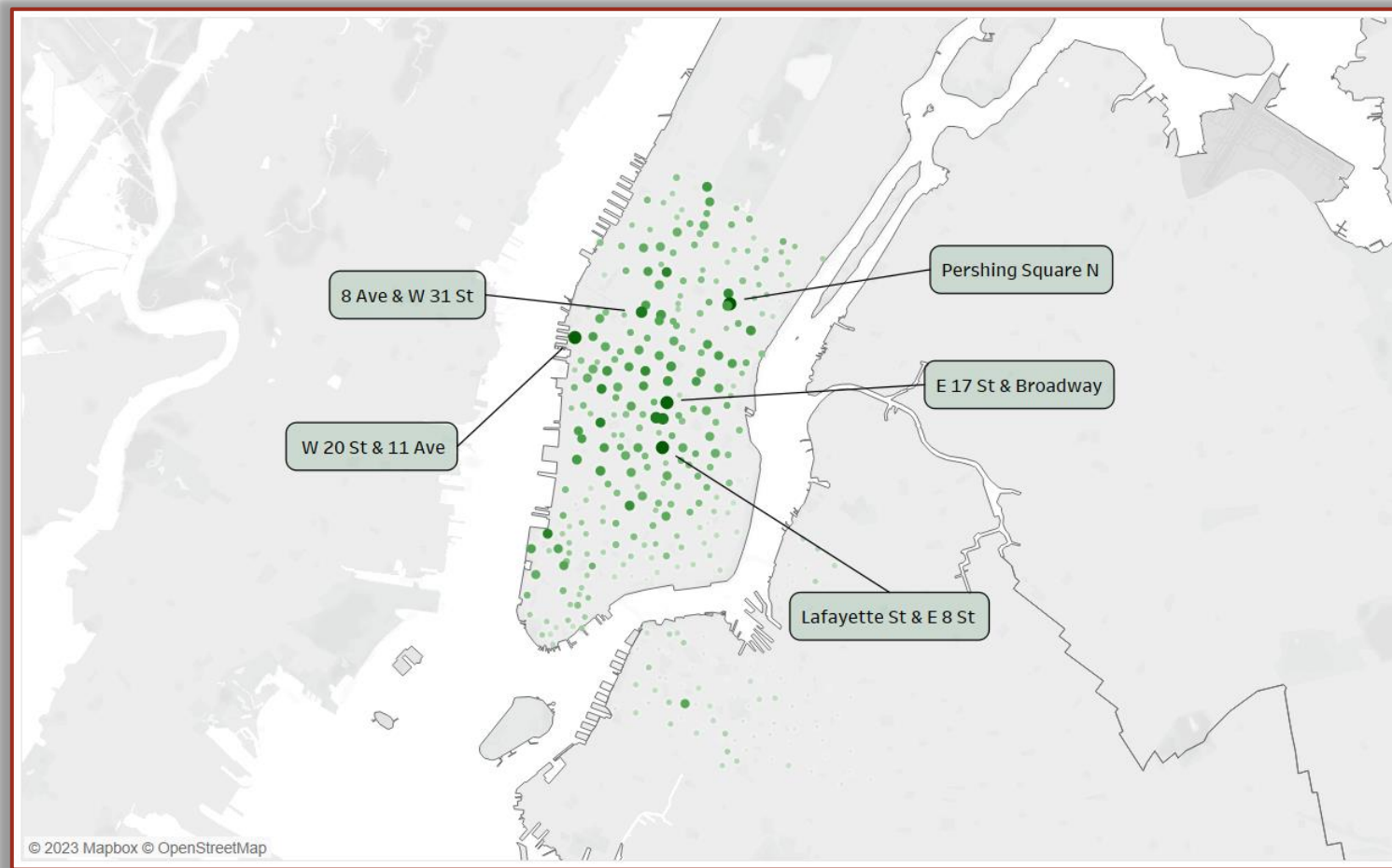
Female

Male

Rental Hour by Age



# Popular Station Locations



➤ Stations in the borough of Manhattan are more popular than those in Brooklyn.

# Key Takeaways

---

- Majority of customers are 33 to 50 with the peak age around 40.
- There are over 3x the number of male customers than there are female.
- Rental activity increased throughout the month of September.
- Weekdays show higher count of bike rentals for male customers.
- Peak hours for bike rentals are 8AM and the hours between 4PM and 6PM.
- Stations in the borough of Manhattan are more popular than those in Brooklyn.

# References

---

## **GameCo**

- [Github: GameCo Project](#)

## **Influenza Preparation**

- [Github: Influenza Project](#)
- [Tableau Dashboard](#)
- [Video Presentation](#)

## **Rockbuster**

- [Github: Rockbuster Project](#)

## **Instacart**

- [Github: Instacart Project](#)

## **Global Bank**

- [Github: Bank Project](#)

## **New York Citi Bike**

- [Tableau Dashboard](#)
- [Github: NYCB Project](#)