



SALTANAT TULEGABYLOVA

Data Analyst Portfolio

CONTENT

About me

- My Technical Skills

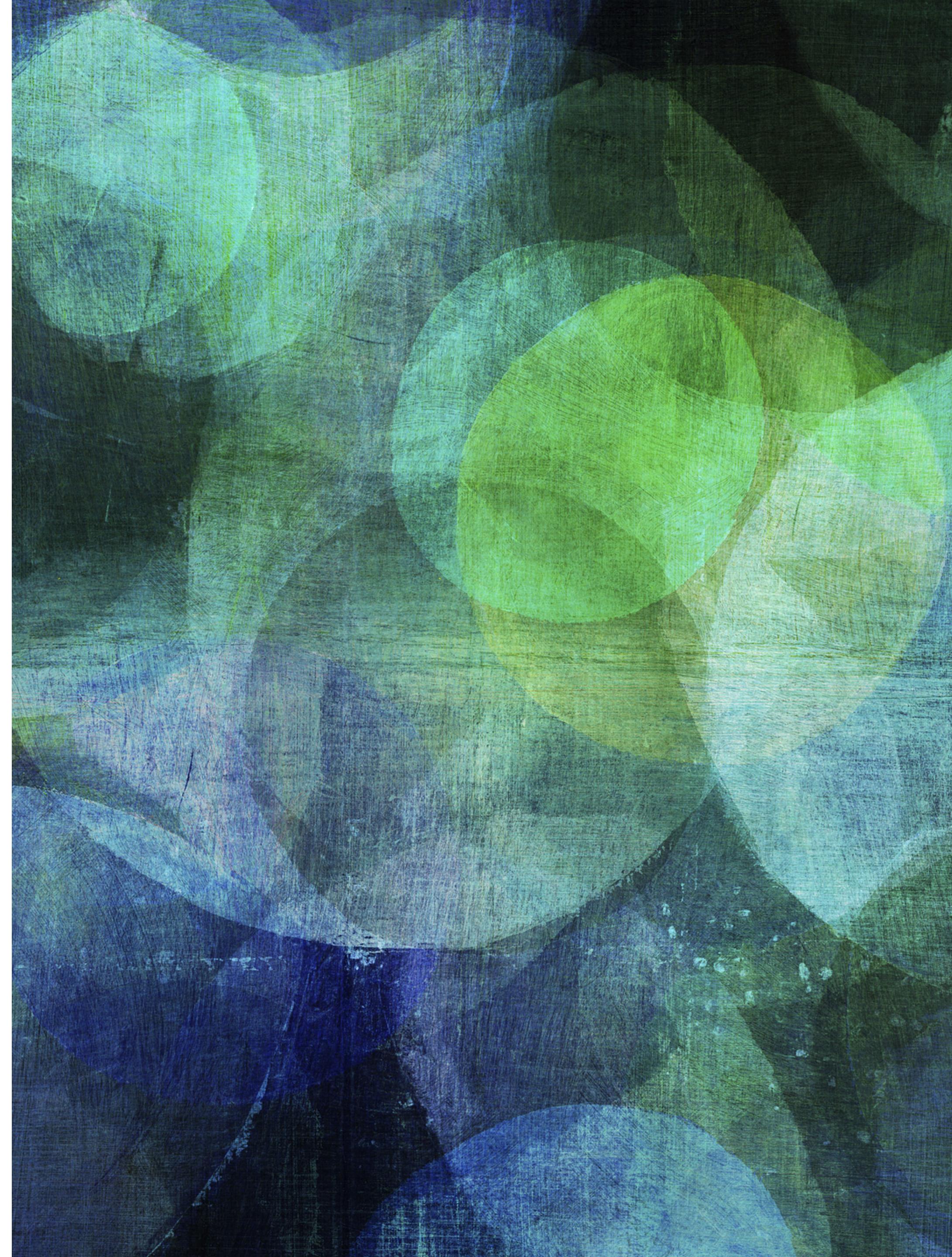
- My Analytical Skills

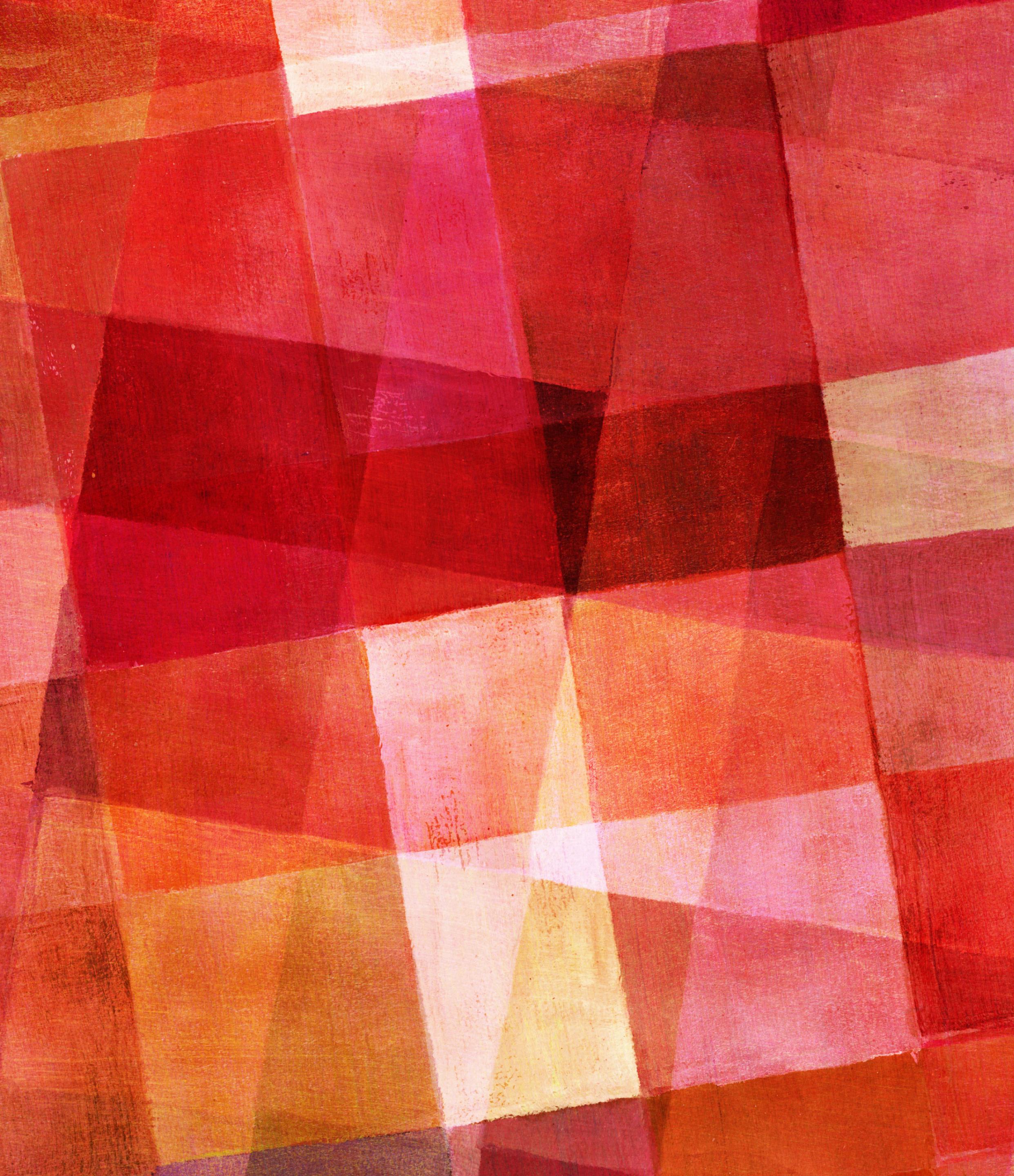
GameCo

Influenza Season

Rockbuster Stealth

Instacart





ABOUT ME

I am a data analyst with a sociology background, driven by a passion for leveraging data to solve problems and optimize processes. My experience includes social research, project management, and statistical analysis. I excel in attention to detail, data mining, and report writing, and am eager to apply these skills to contribute effectively in data-driven environments.

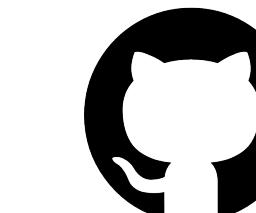
Contact:



saltatulegabylova@gmail.com

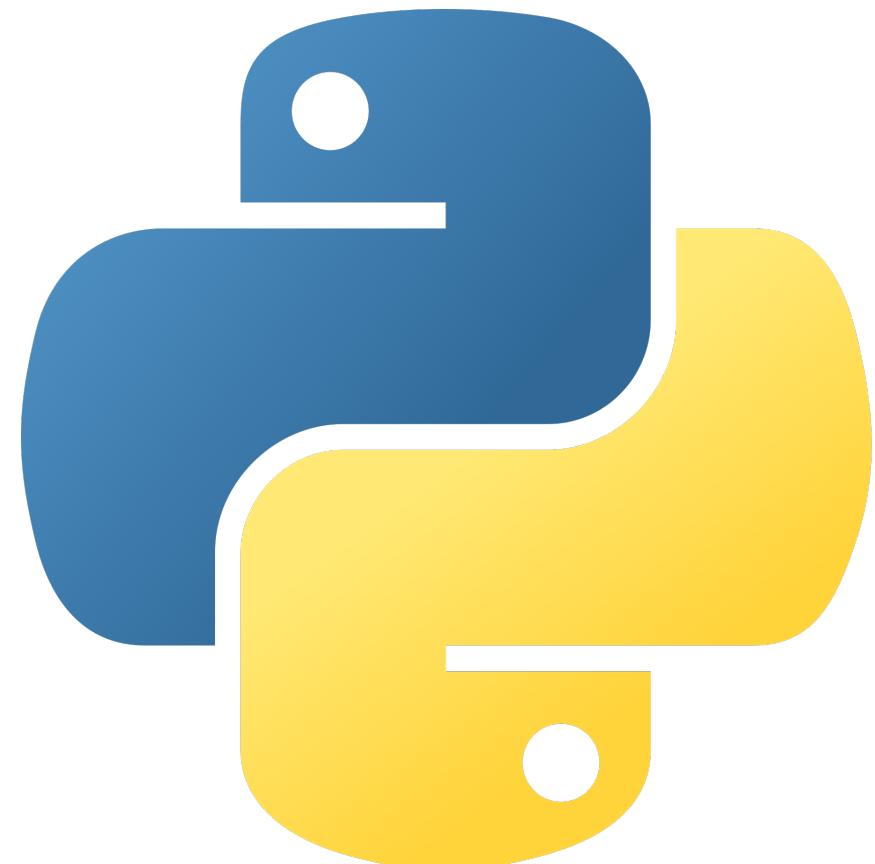


linkedin.com/in/saltanat-tulegabylova-a6aa1b234

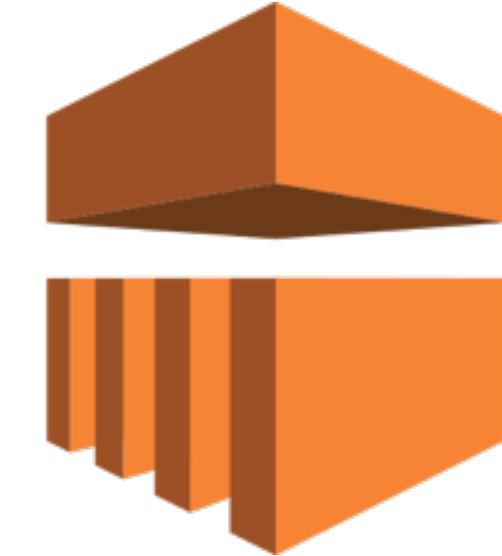


<https://github.com/salta89>

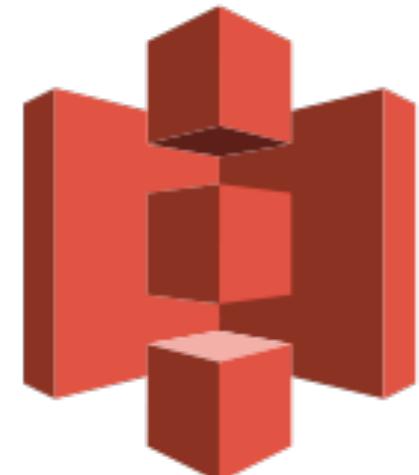
MY TECHNICAL SKILLSET



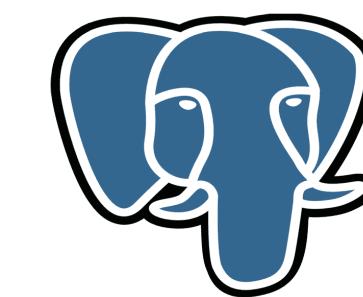
amazon
KINESIS



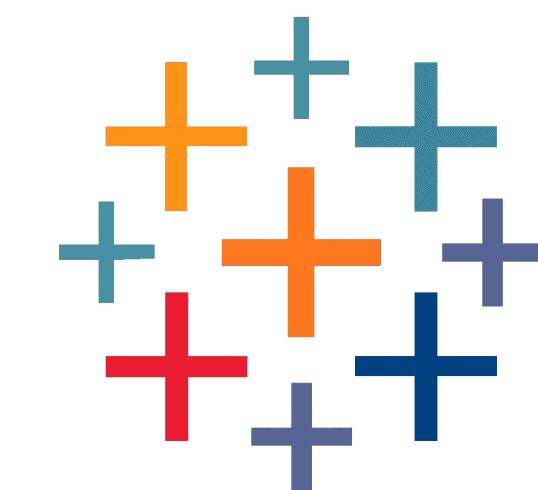
amazon
EMR



amazon
S3

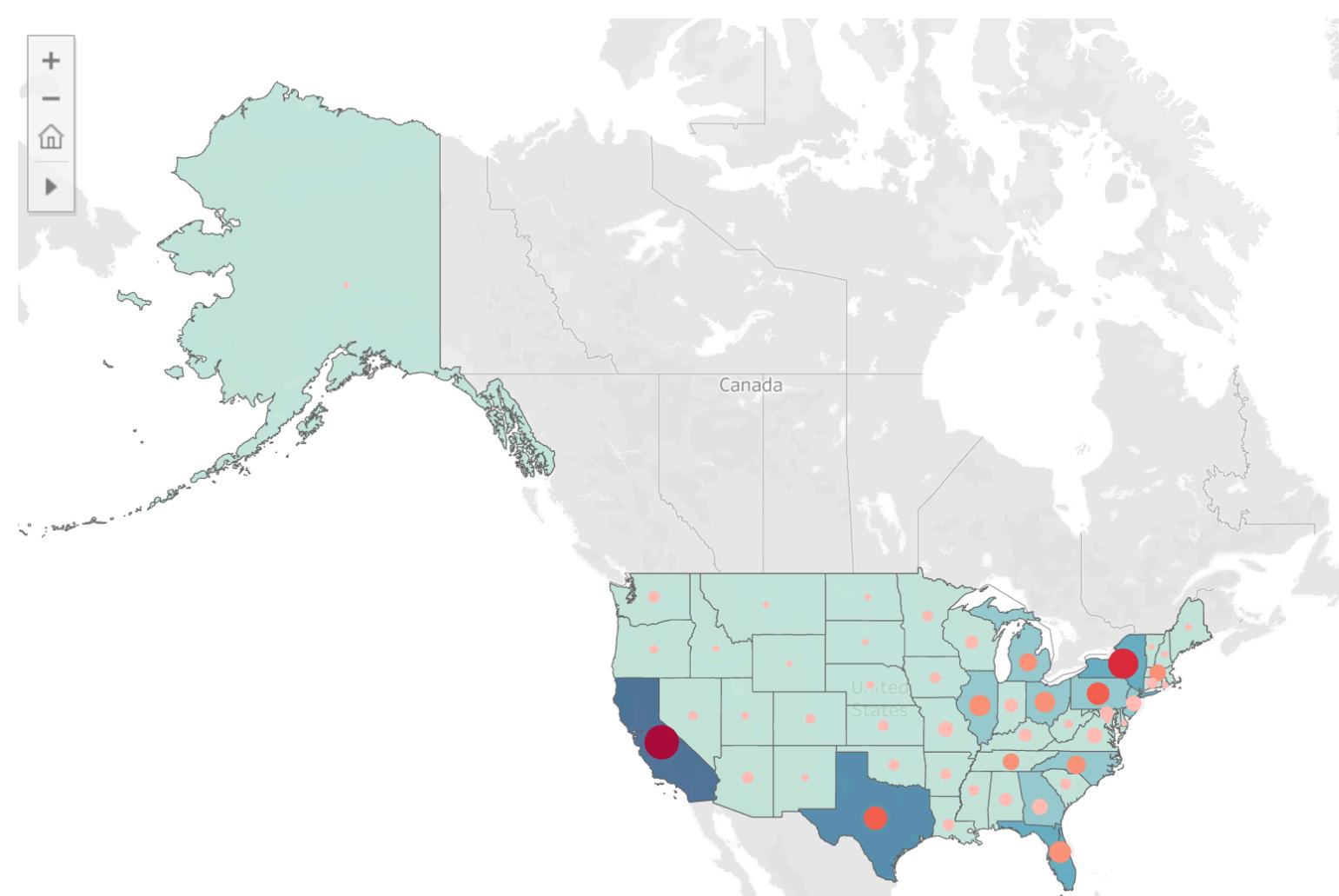
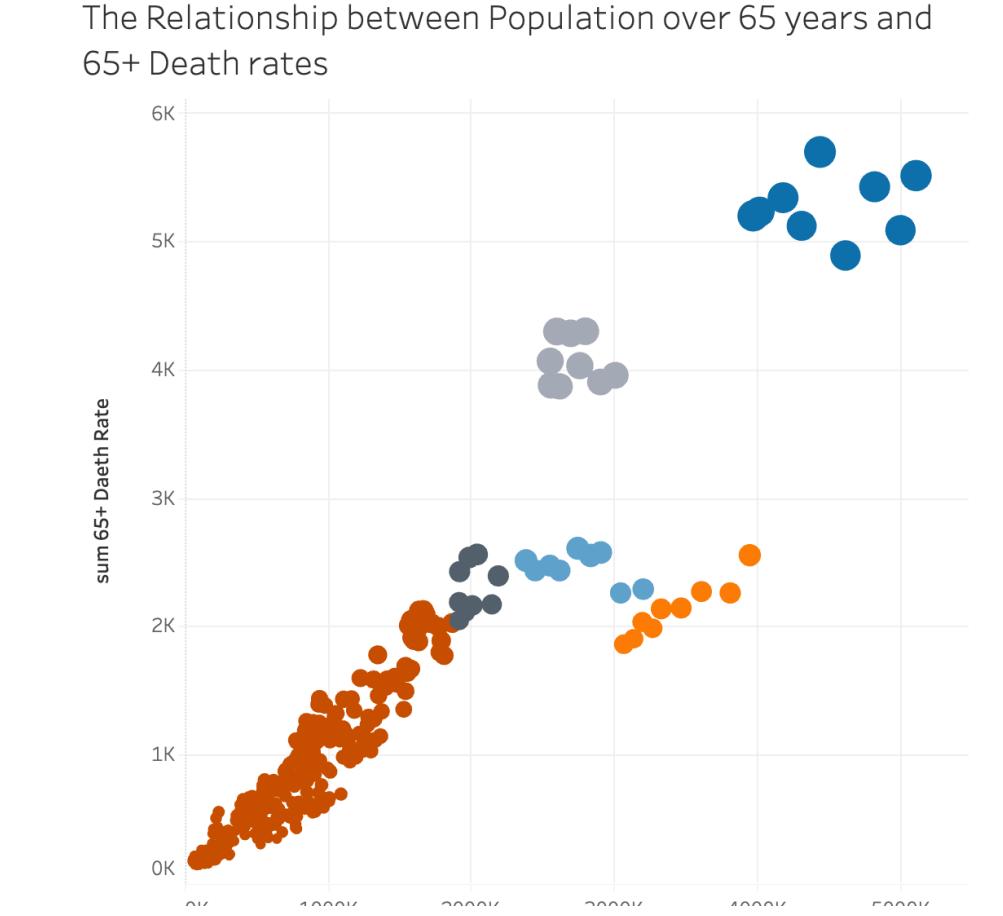
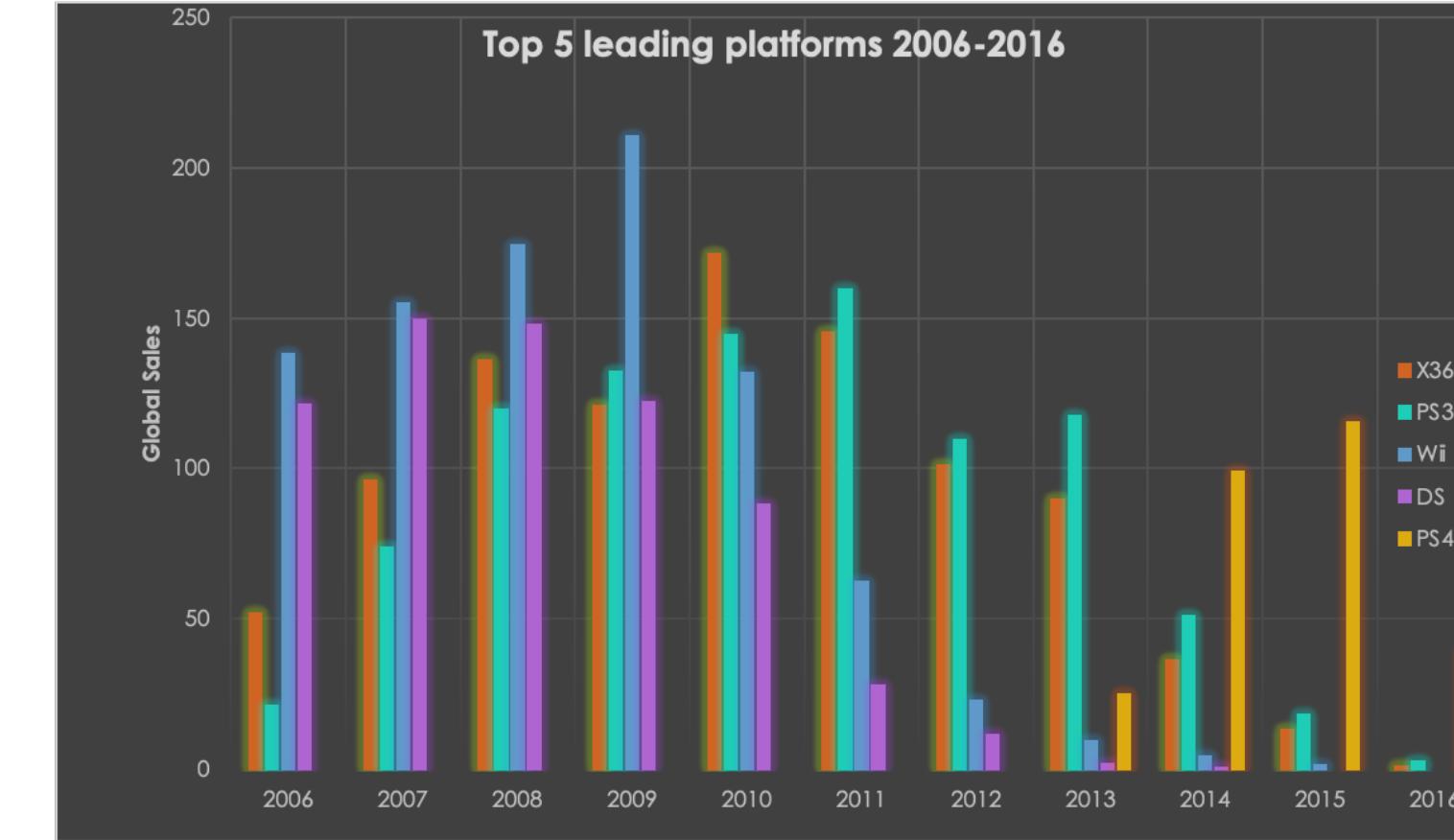
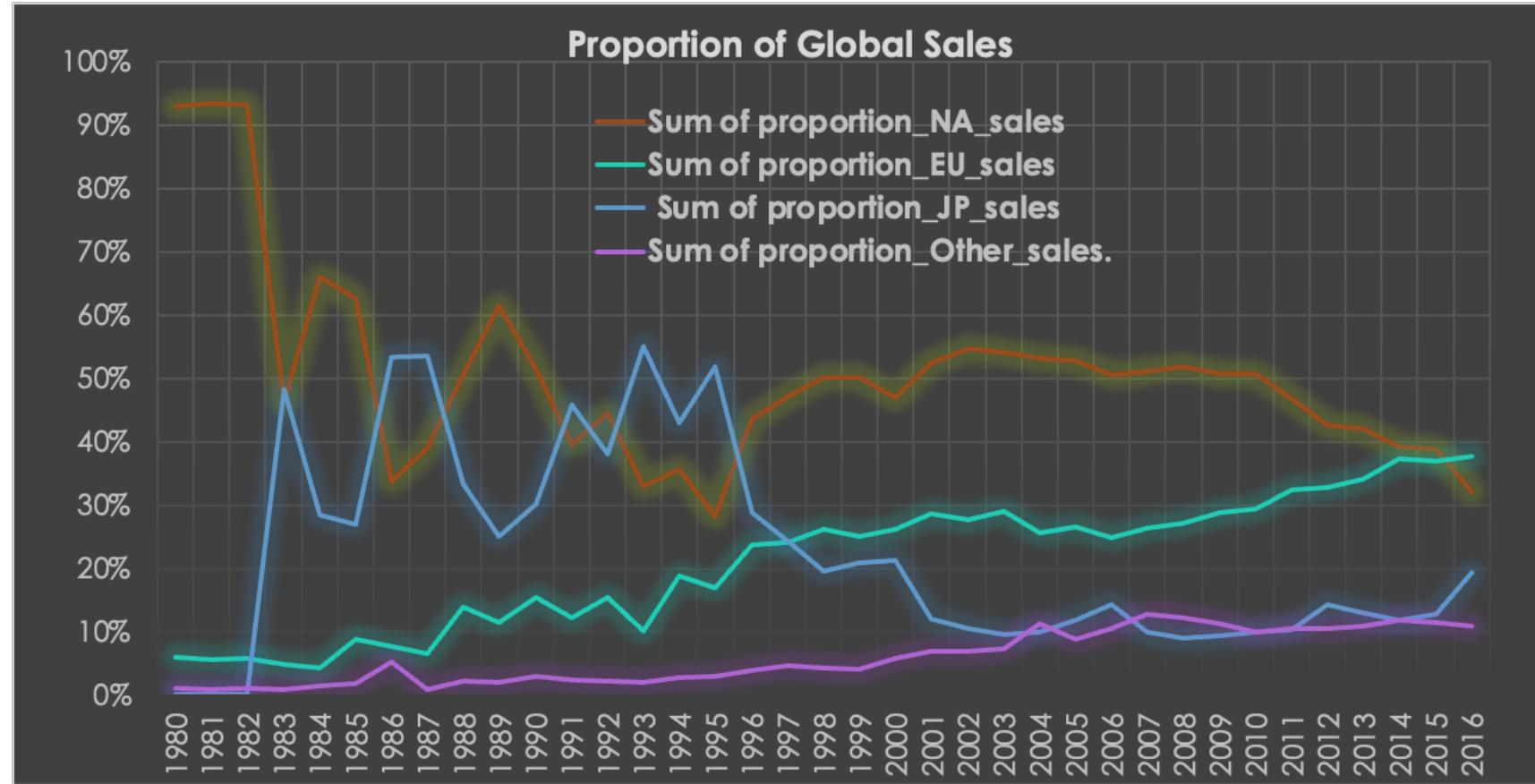


PostgreSQL



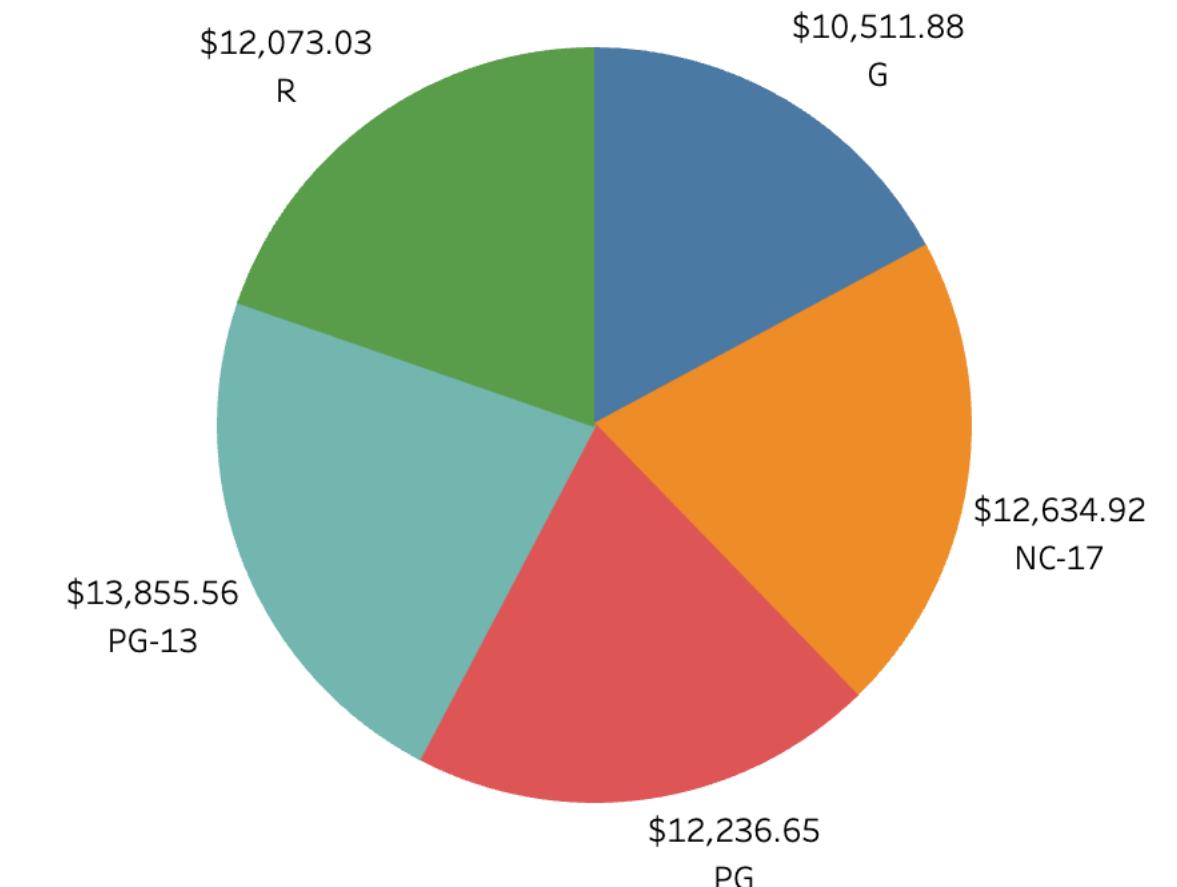
+ tableau

MY ANALYTICAL SKILLSET



No idea
Human Papiloma Virus
Not sure
Human something virus
Human papillomavirus
Dont know Something virus Virus
Don't know
I do not know

Human papilloma virus



GAMECO

- **Objective:** GameCo is a fictional video game company aspiring to develop new games. They want to use data to better understand of how new games might fare in the market.
- **Data:**
 - The data was sourced from [VGChartzV](#)
 - Covers historical sales of video games (for games that sold more than 10,000 copies) spanning different platforms, genres, and publishing studios.

PROJECT OVERVIEW

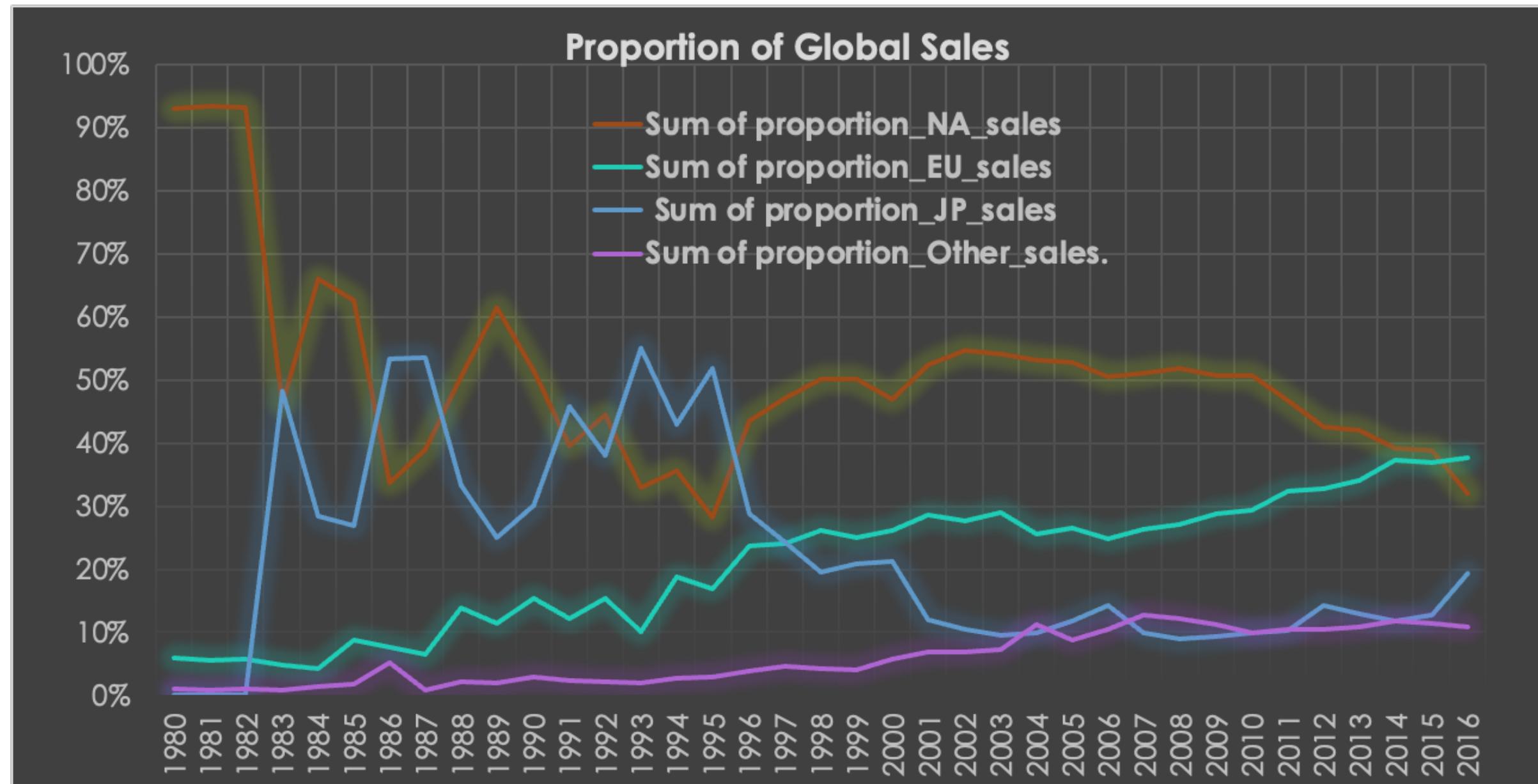
KEY QUESTIONS:

SKILLSET

- Are certain types of games more popular than others?
- What other publishers will likely be the main competitors in certain markets?
- Have any games decreased or increased in popularity over time?
- How have their sales figures varied between geographic regions over time?
- Excel
- Grouping data
- Summarizing data
- Descriptive analysis
- Visualizing results in Excel
- Presenting results

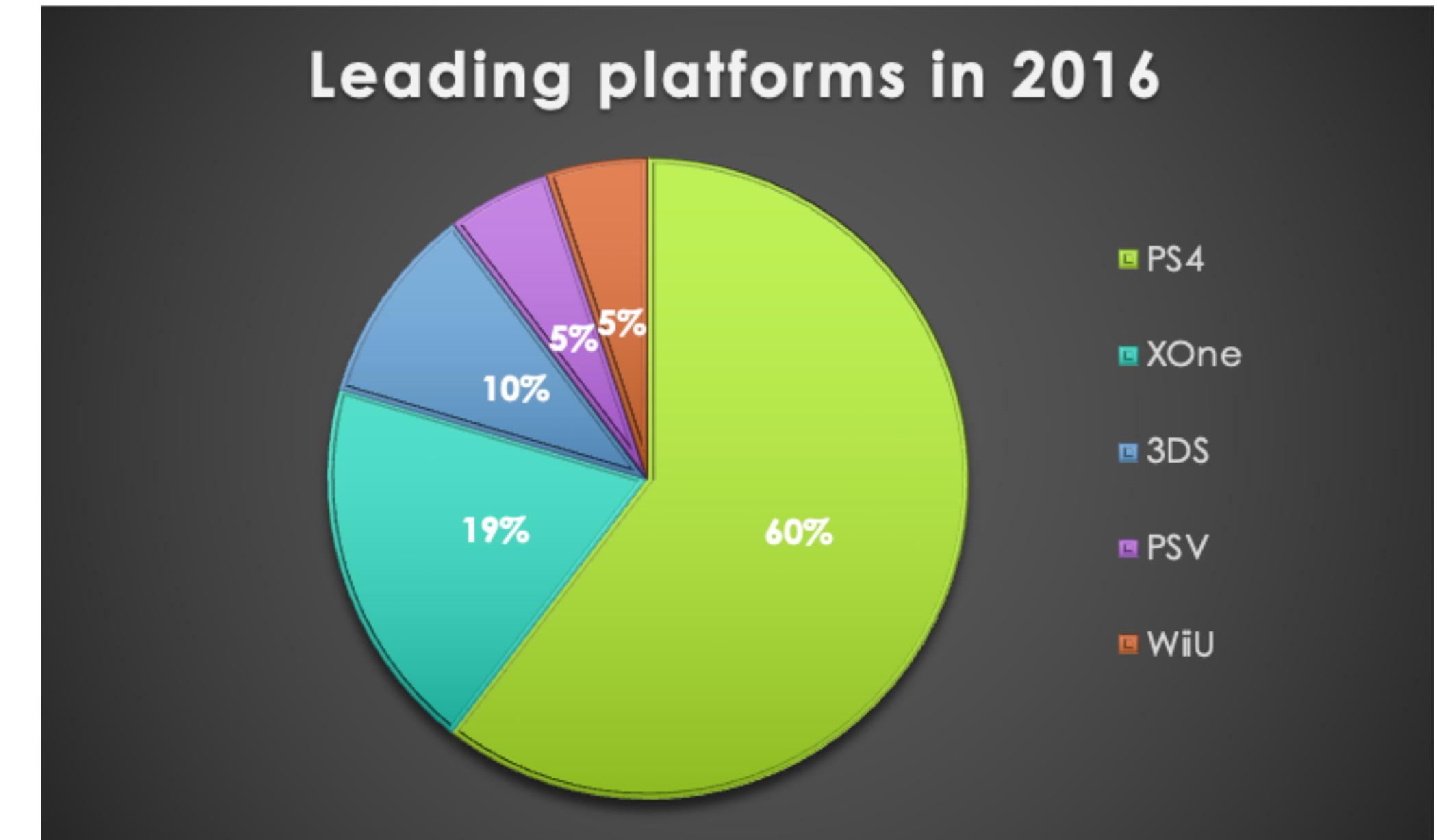


VISUALIZATIONS



Sales in each region in proportion to global sales.

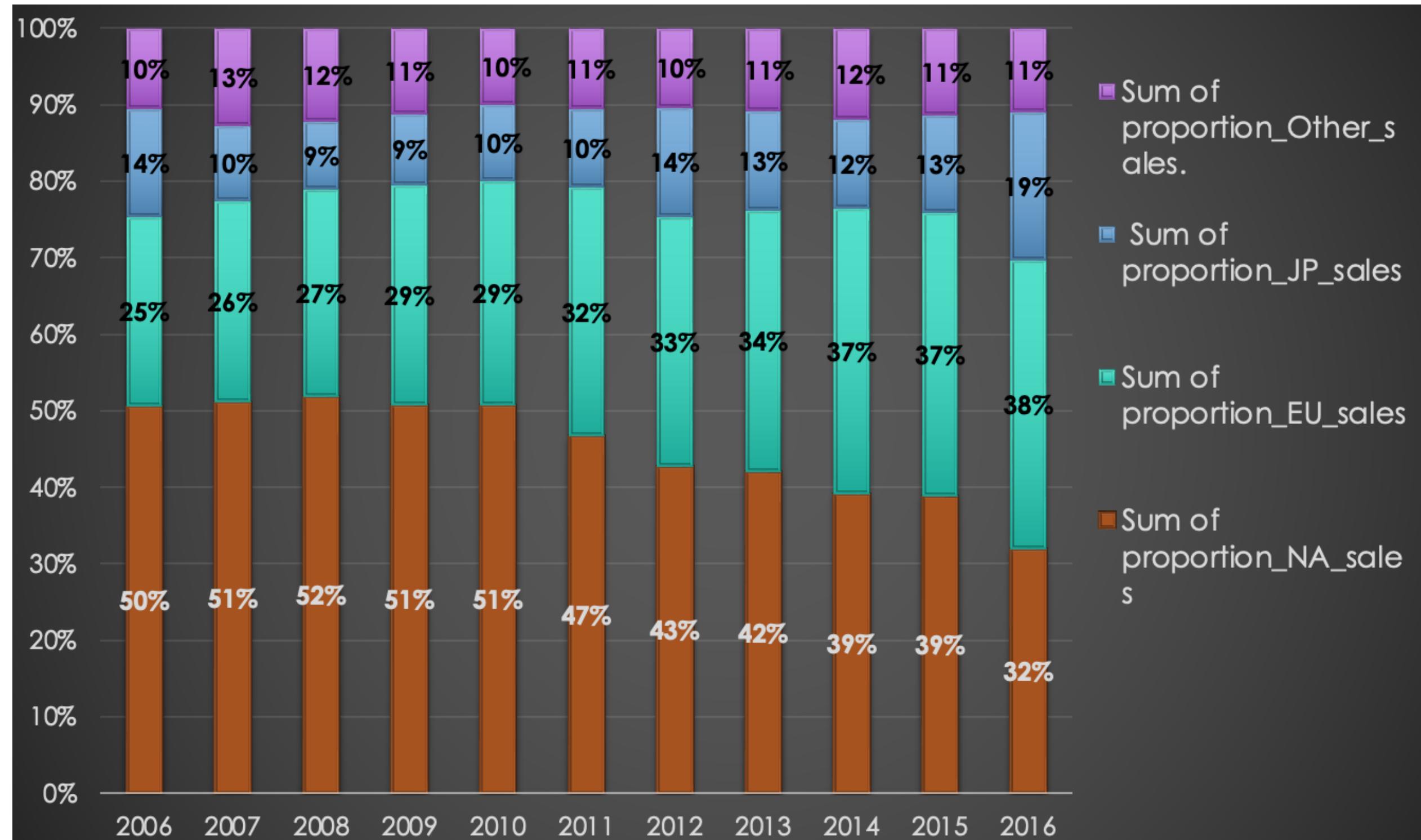
In 1983-1996 significant fluctuations in North American and Japanese sales are noticed.



Popular platforms over the last decade.

In 2016 PS4 considered as a leading platform.

VISUALIZATIONS



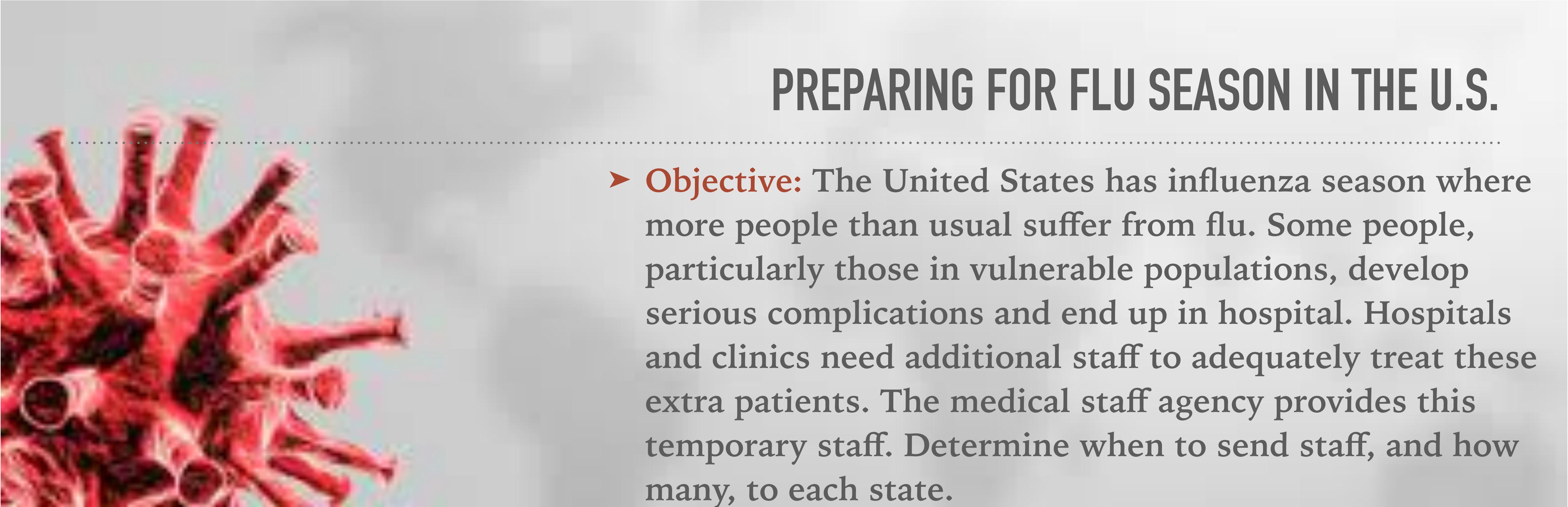
Sales in each region in proportion to global sales over the last decade.

The dynamic of NA sales decline during last 10 years whereas Japan and EU sales incline.

CONCLUSION

- NA sales were considered the highest and expected to remain in the same position. Though we can observe the decrease from 2010 and by 2015, they have almost the same percentage as EU sales.
- Based on the data, allocating marketing resources to the EU region is suggested. Since this region has the largest population, focusing on a larger population and knowing consumers' preferences may significantly increase sales.
- The dynamic of Japanese sales showed a slight increase in 2016. The Japanese market can also be included in the new marketing strategy.
- Considering popular genres, platforms, and games in these regions is suggested to increase marketing resources following the consumers' demand.



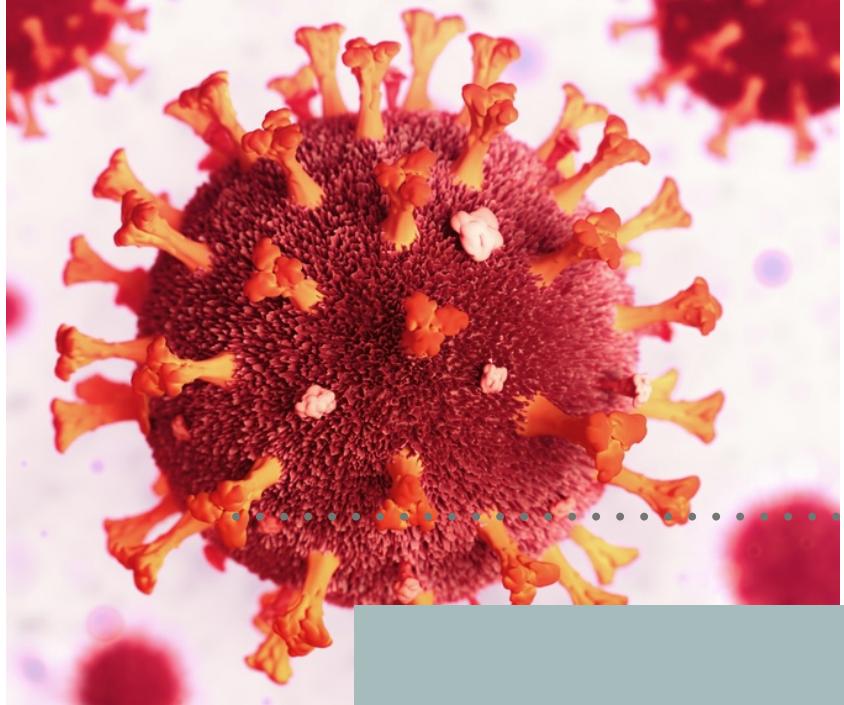


PREPARING FOR FLU SEASON IN THE U.S.

► **Objective:** The United States has influenza season where more people than usual suffer from flu. Some people, particularly those in vulnerable populations, develop serious complications and end up in hospital. Hospitals and clinics need additional staff to adequately treat these extra patients. The medical staff agency provides this temporary staff. Determine when to send staff, and how many, to each state.

► **Data:**

- Influenza deaths by geography, time, age, and gender. It is owned by the Centers for Disease Control and Prevention ([CDC](#)) through their National Center for Health Statistics. This data counts monthly influenza death rates divided by state, months, and age groups from 2009 to 2017.
- Population data by geography. This data set is an external data source. It is owned by the US Census Bureau. This data shows yearly population sizes of counties of Unites States divided by gender and age groups (under 5 years, 5-9 years, 10-15 years etc.) from 2009 to 2017.



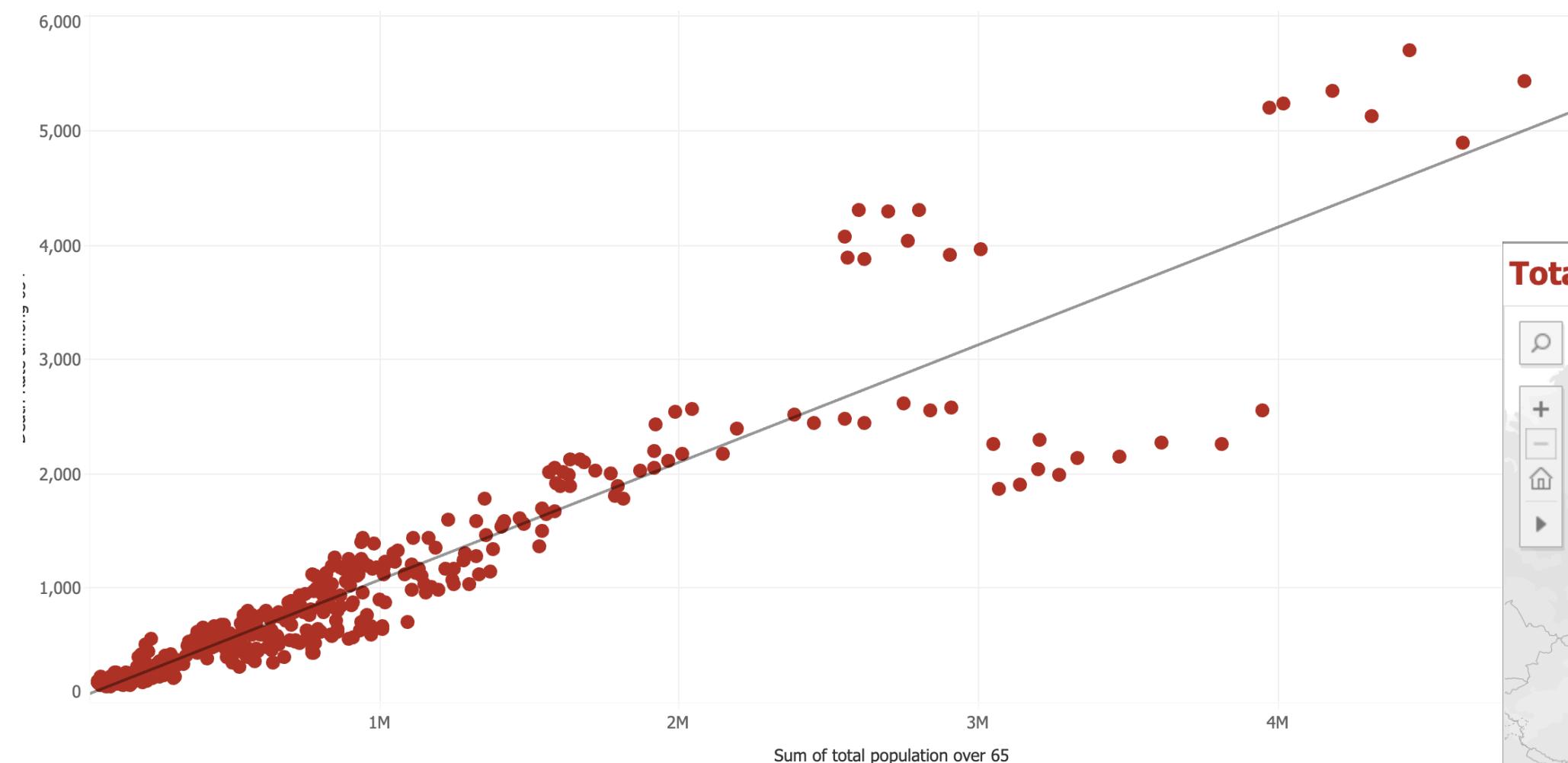
PROJECT OVERVIEW

KEY QUESTIONS:

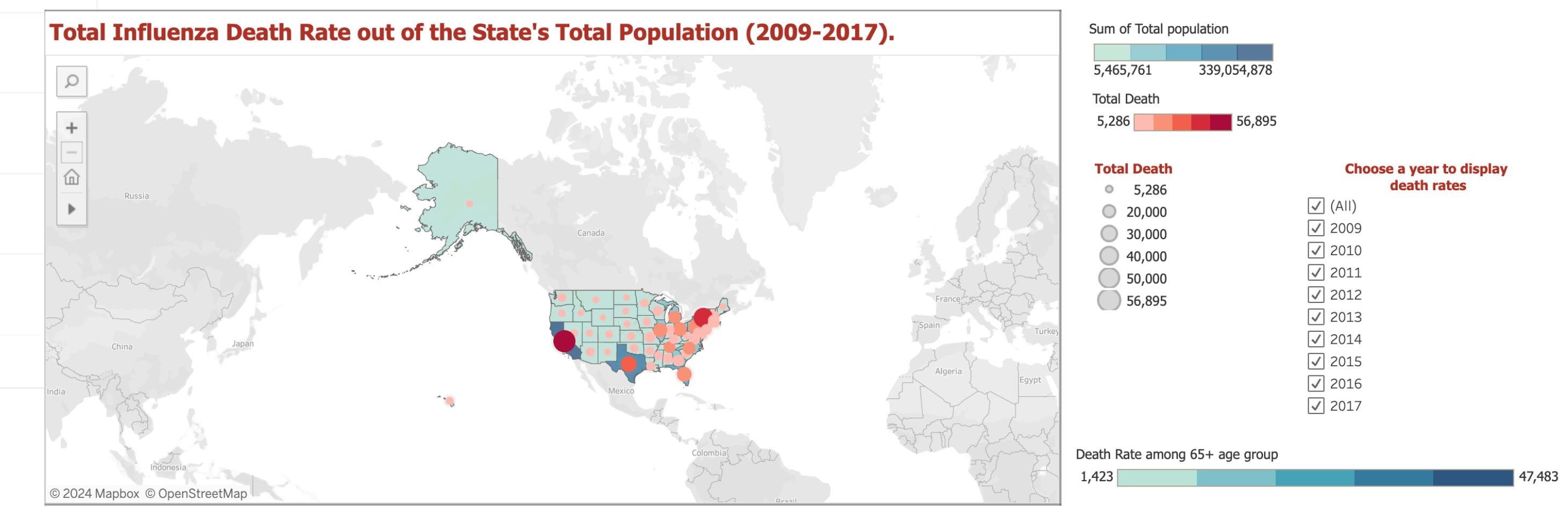
SKILLSET

- How many additional staff members need to be send to each state?
- What states have the largest vulnerable population?
- Does influenza occur seasonally, or is it present throughout the entire year? If it is seasonal, does it start and end at the same time each year in every state? Are there differences in timing between states?
- Excel
- Translating business requirements
- Data cleaning, integration, transformation
- Statistical hypothesis testing
- Visual analysis
- Forecasting Storytelling in Tableau

VISUALIZATIONS



Correlation between 65+ age group and death rates



Comparing Influenza Death Rates among 65+ years and under 64 age groups between the States.

California 47,483 9,412	New York 36,576 7,399	Pennsylvania 20,596 5,098	Illinois 18,019 5,366	North Carolina 13,941 5,006	Tennessee 11,097 4,949	New Jersey 9,651 4,175	Wisconsin 7,636	Kentucky 6,829				
		Florida 19,148 6,233	Ohio 16,866 5,508	Michigan 12,964	Georgia 10,268	Missouri 9,614 4,143	Maryland 7,571			Iowa		
	Texas 22,140 7,510		Massachusetts 11,179		Virginia 10,067	Indiana 7,822	Alabama 7,171		Louisiana			
								Kansas				

Total Influenza death rate out of the States total population. Influenza death rates comparison between the states.

CONCLUSION

[Link to the dashboard](#)

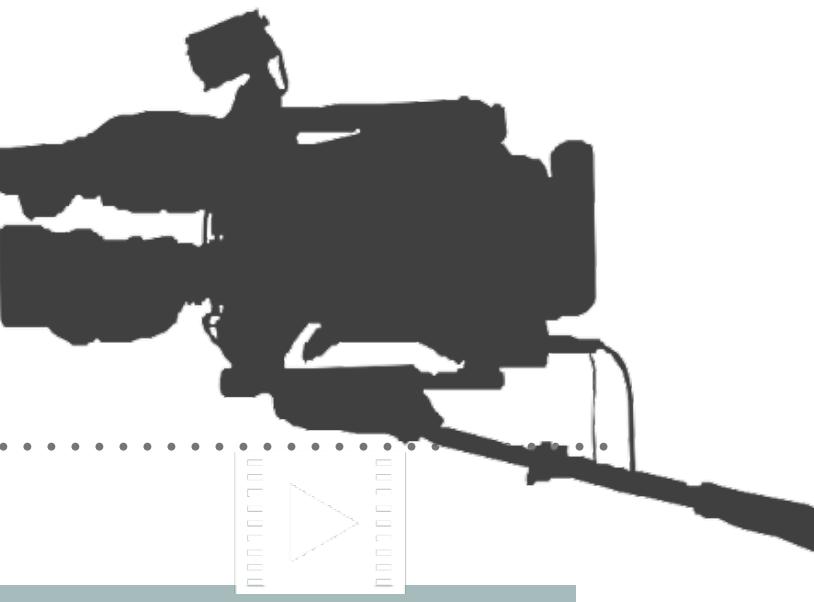
- The highest influenza death rates are among the elderly 65+ age group.
- The influenza season in most states starts in December and lasts till the beginning of April.
- The peak months are December, January and March.
- States with the highest populations have the highest mortality rates. Therefore, they require additional medical staff.
- The top 10 states with the highest population, therefore high mortality rates, need additional medical staff primarily.
- The allocation should be done from late November until the end of March.
- Further analysis should include the identification and role of effective preventive measures such as vaccination among medical staff and population, awareness-raising campaigns among vulnerable population groups, analysis of hospital quality indicators, medical staff size, etc.



ROCKBUSTER STEALTH

- **Objective:** Rockbuster Stealth LLC is a movie rental company that used to have stores around the world. Facing stiff competition from streaming services such as Netflix and Amazon Prime, the Rockbuster Stealth management team is planning to use its existing movie licenses to launch an online video rental service in order to stay competitive.
- **Data:** Rockbuster Data set contains information about Rockbuster's film inventory, customers, and payments, among other things.

PROJECT OVERVIEW



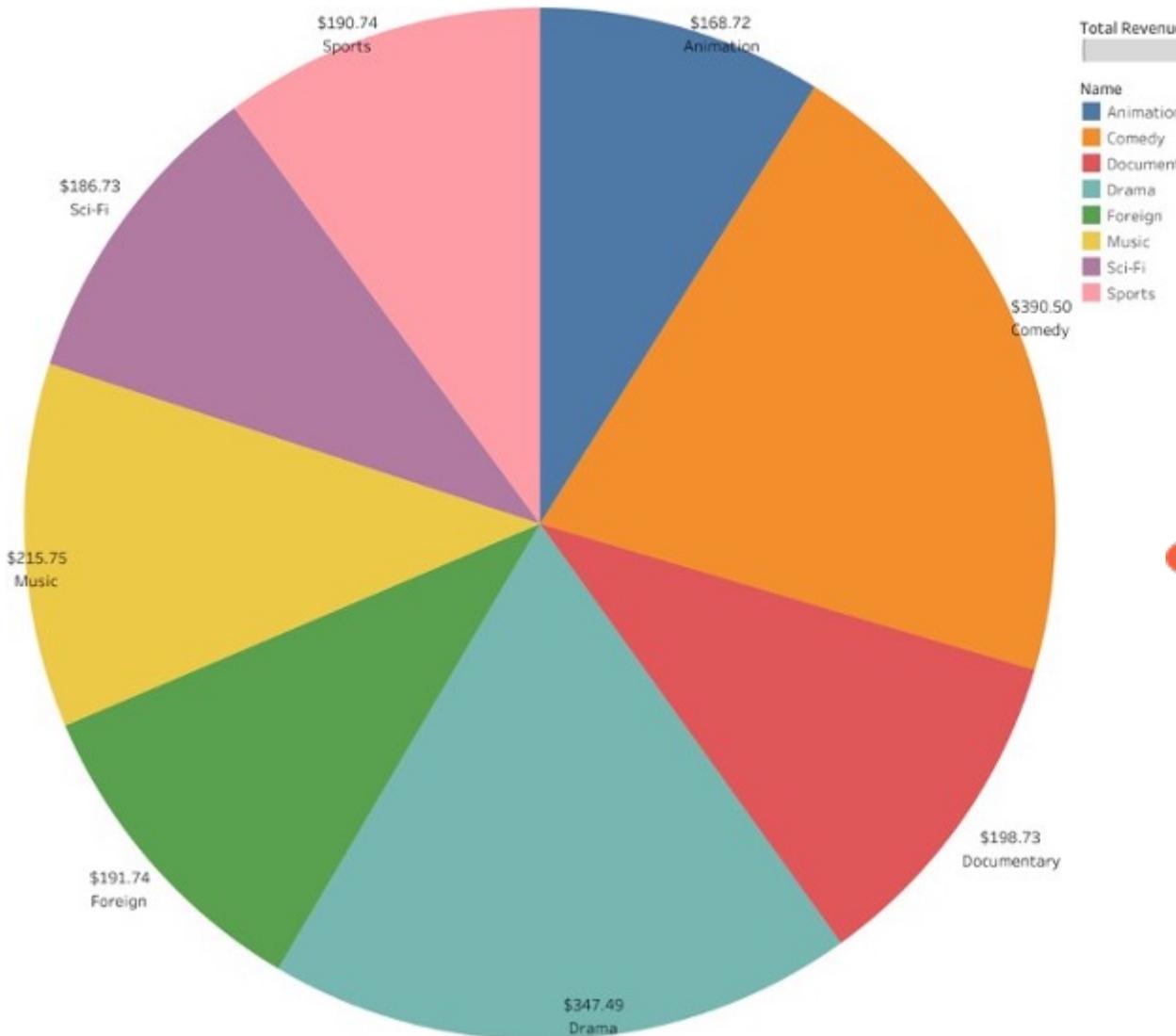
KEY QUESTIONS:

- 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?

SKILLSET

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

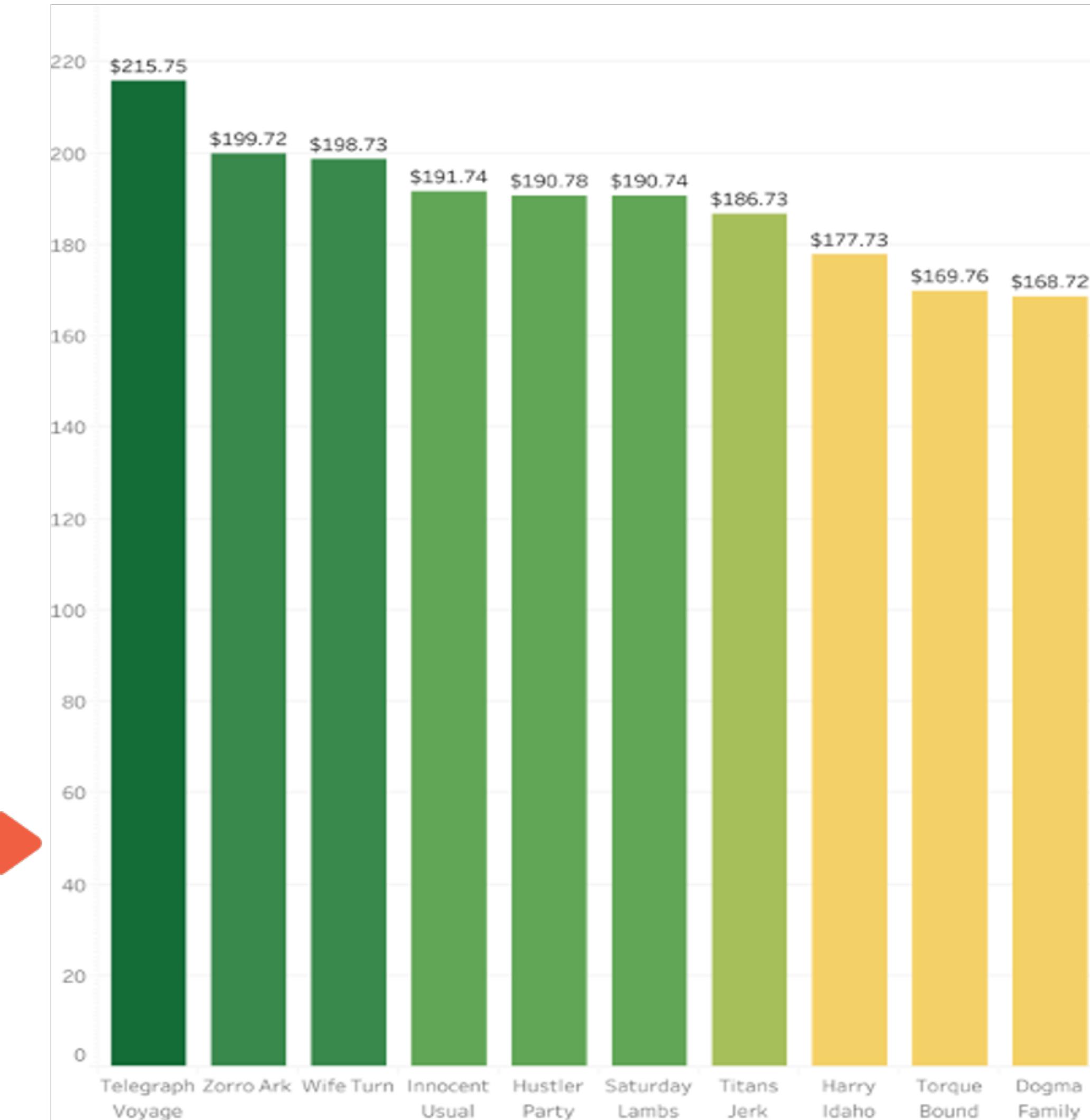
VISUALIZATIONS



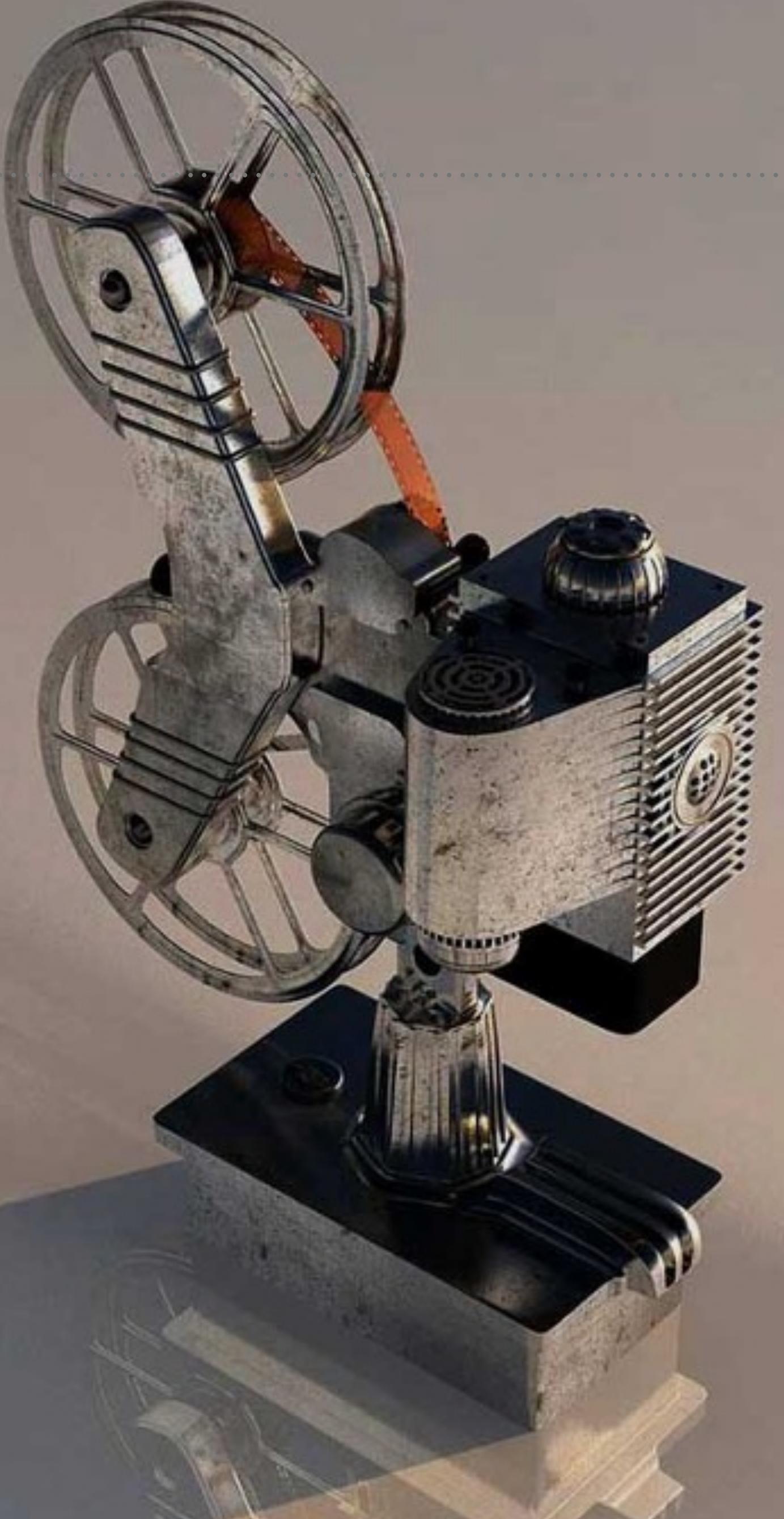
Genre	Total Revenue
Music	\$215.75
Comedy	\$390.50
Documentary	\$198.73
Foreign	\$191.74
Sports	\$190.74
Sci-Fi	\$186.73
Drama	\$347.49
Animation	\$168.72

Top 10 Popular movies and their genres by total revenue

Title	Genre	Total Revenue
Telegraph Voyage	Music	\$215.75
Zorro Ark	Comedy	\$199.72
Wife Turn	Documentary	\$198.73
Innocent Usual	Foreign	\$191.74
Hustler Party	Comedy	\$190.78
Saturday Lambs	Sports	\$190.74
Titans Jerk	Sci-Fi	\$186.73
Harry Idaho	Drama	\$177.73
Torque Bound	Drama	\$169.76
Dogma Family	Animation	\$168.72



CONCLUSION

- 
- Increase the number of movies that are most popular among customers: PG-13 movies and Sports, Sci-Fi, Animation, Drama, and Comedy genres.
 - Increase movie numbers with the most popular starring actors.
 - Understand peak rental times and customer preferences.
 - Based on new releases and popularity rate, implement pricing strategies.
 - Regularly update the movie catalog with new releases based on popularity.
 - Collect and analyze customer feedback to identify areas of improvement and preferences.



INSTACART

- **Objective:** 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.
- **Data:**
 - The Instacart Online Grocery Shopping Dataset 2017
 - Customer data set (created and included by Career Foundry for the purpose of this project)



PROJECT OVERVIEW

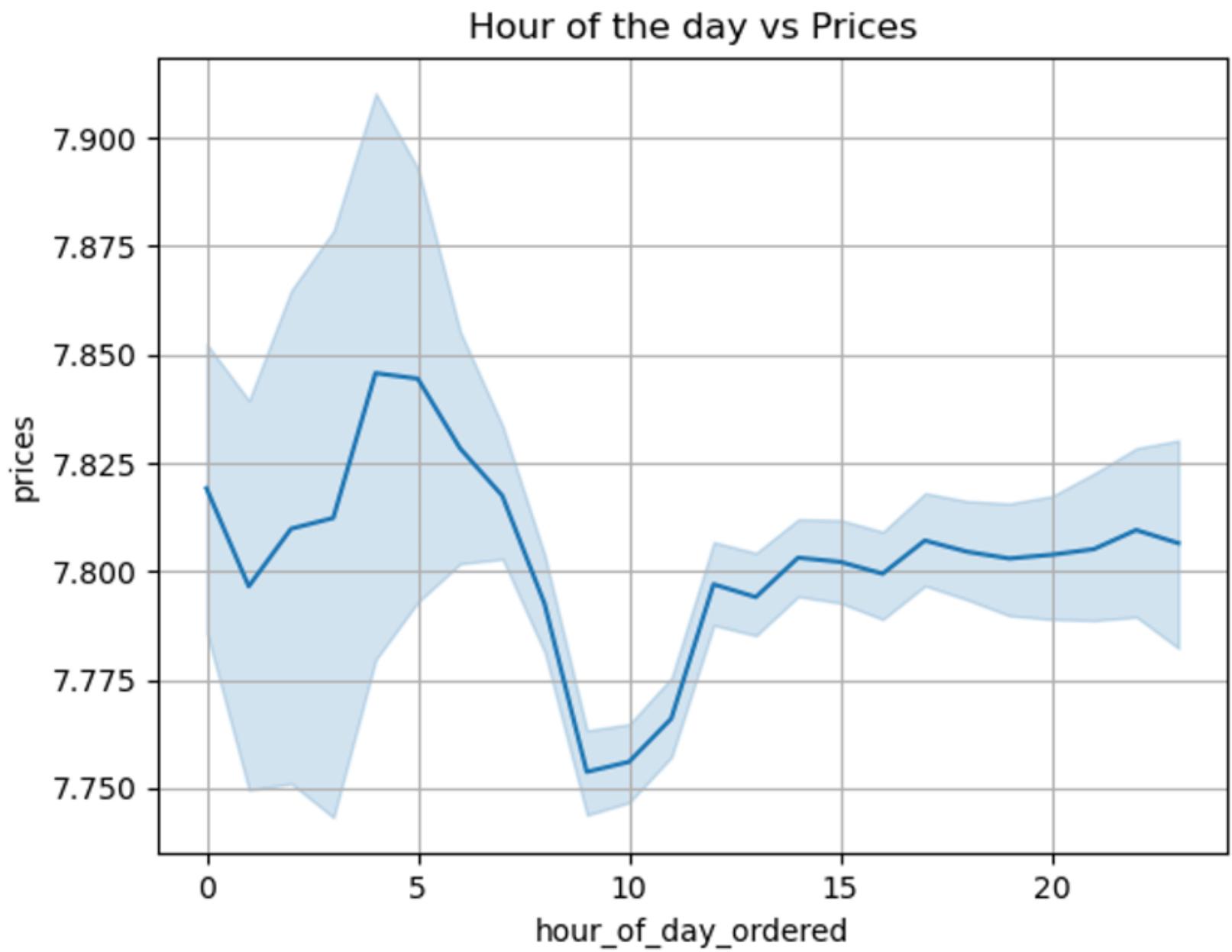
KEY QUESTIONS:

- What are the busiest days of the week and hours of the day are?
- Are there specific times of the day when people tend to spend the most money?
- Are there certain types of products that are more popular than others
- Which departments have the highest frequency of product orders?

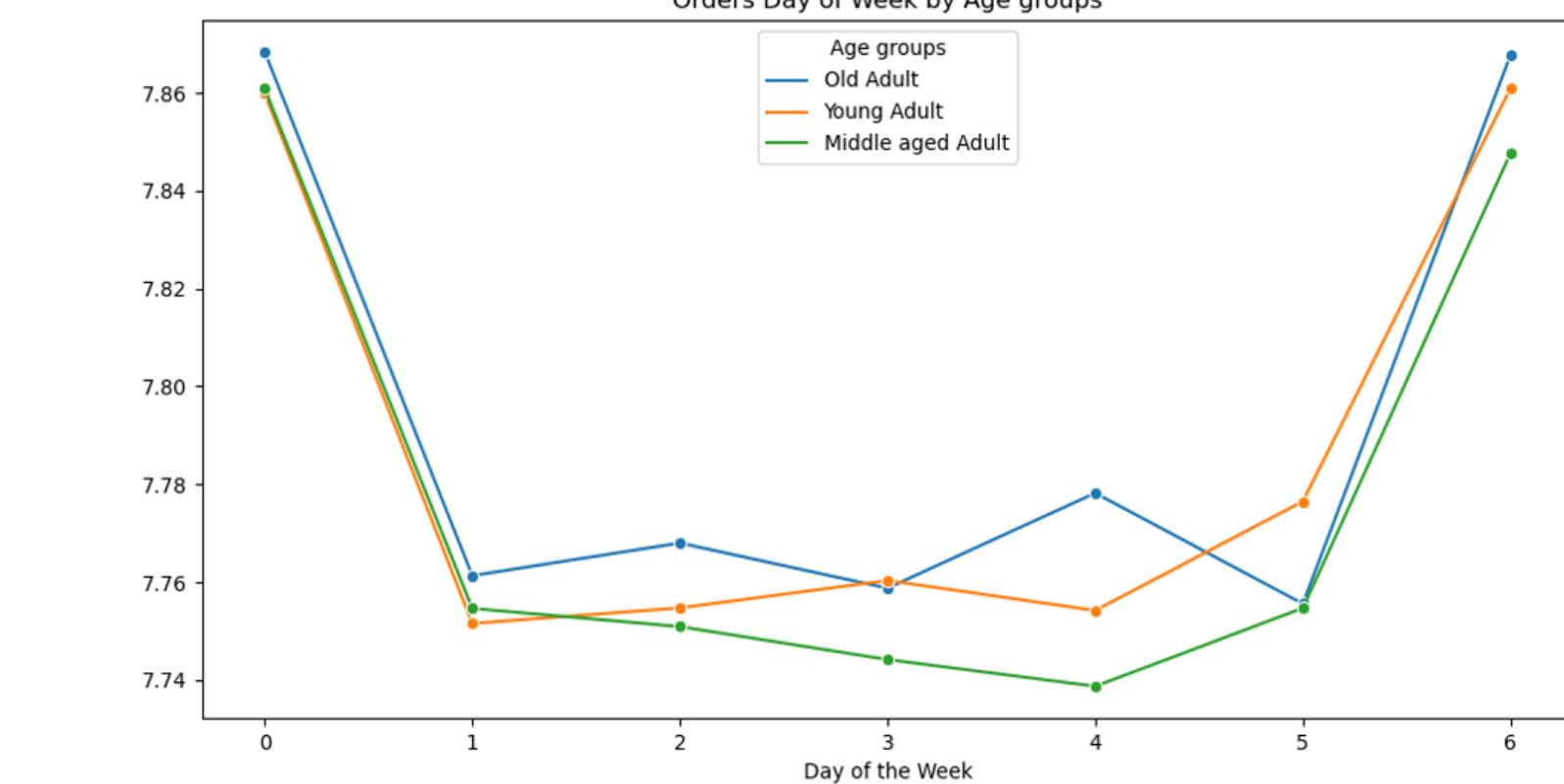
SKILLSET

- Python
- Data wrangling
- Data merging
- Deriving variables
- Grouping data
- Aggregating data
- Statistical hypothesis testing

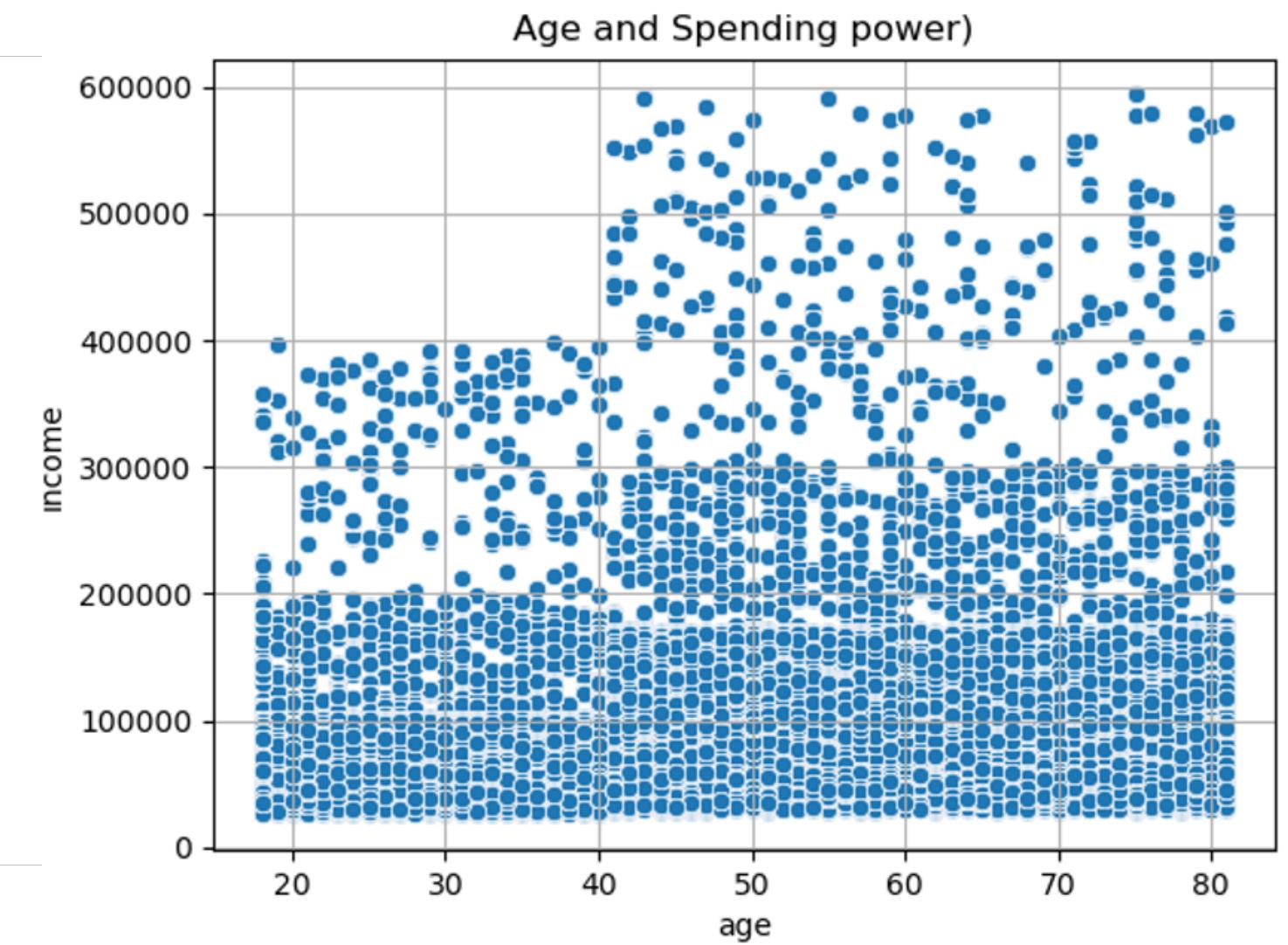
VISUALIZATIONS



Hour of the day with most orders.
Expenses are highest at 4am



Orders week of the day by Age groups. The most orders and expenditures are made on weekends (0. Sunday. 6. Saturday)



Age and spending power. Individuals under the age of 40 generally exhibit lower spending power in comparison to those aged 40 and above.



CONCLUSION

The primary customer demographic for Instacart online grocery shop comprises married middle-aged adults with dependents in the South region, who predominantly place orders during weekends. Their favorite departments include produce, dairy/egg, and snacks. To effectively target this segment, the marketing strategy should focus on personalized promotions tailored to their preferences. Additionally, optimizing their website for a user-friendly experience is crucial.

To boost weekday orders, it is recommended implementing the following strategies:

- Offer exclusive discounts, promotions, or bundled deals specifically for weekdays to incentivize customers.
- Introduce limited-time flash sales or special offers during weekdays to create a sense of urgency.
- Launch a loyalty program rewarding customers with points and discounts for consistent weekday orders.

These measures aim to not only attract weekday orders but also enhance customer loyalty through targeted promotions and a seamless online shopping experience.

LINKEDIN JOB POSTS 2023

- **Objective:** The objective of analyzing the LinkedIn job posts dataset is to identify the most in-demand job positions, understand the key factors influencing job application rates, and determine which industries are predominantly posting job opportunities. Through this analysis, we aim to gain insights into the current job market dynamics, including the types of roles sought after by candidates and the factors that drive their decision-making process when applying for jobs.
- **Data:** LinkedIn Job Postings – 2023 data set acquired from [Kaggle](#).



PROJECT OVERVIEW

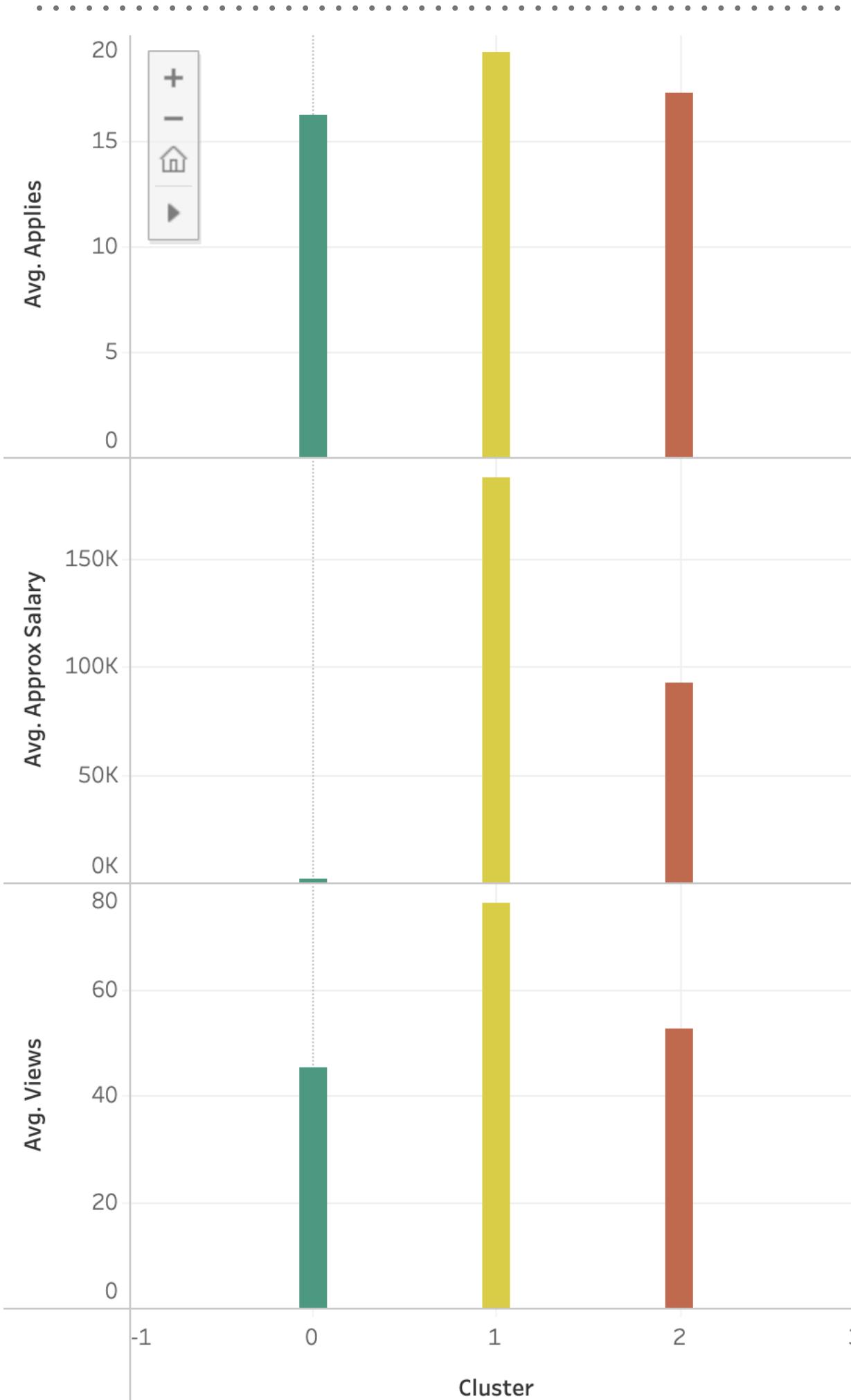
KEY QUESTIONS:

- What job titles have the most applies?
- What are top industries by job postings?
- How are job listings distributed across different experience levels?
- Hypothesis to test:
Job posts with high salaries receive more applications

SKILLSET

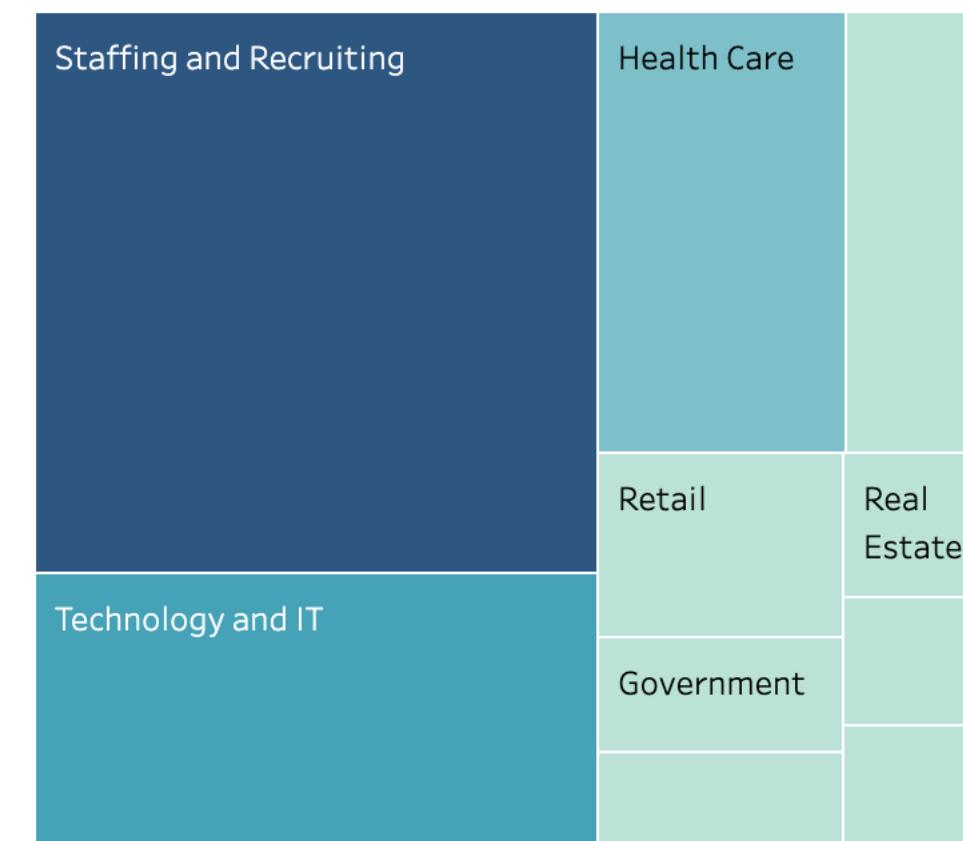
- Supervised Machine Learning:
Regression
- Unsupervised Machine Learning:
Clustering
- Time-series Analysis
- Spatial Analysis in Python

VISUALIZATIONS

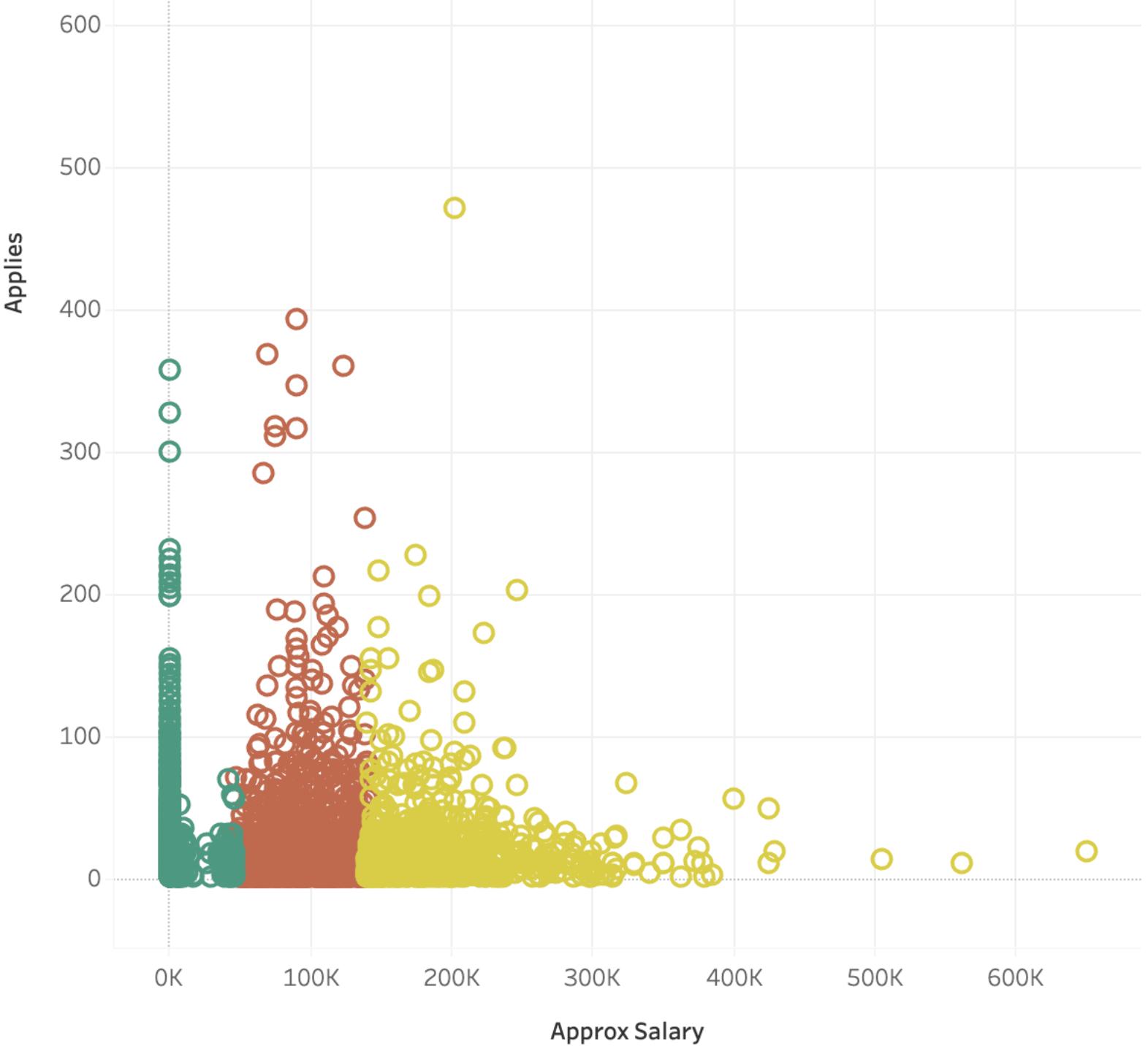
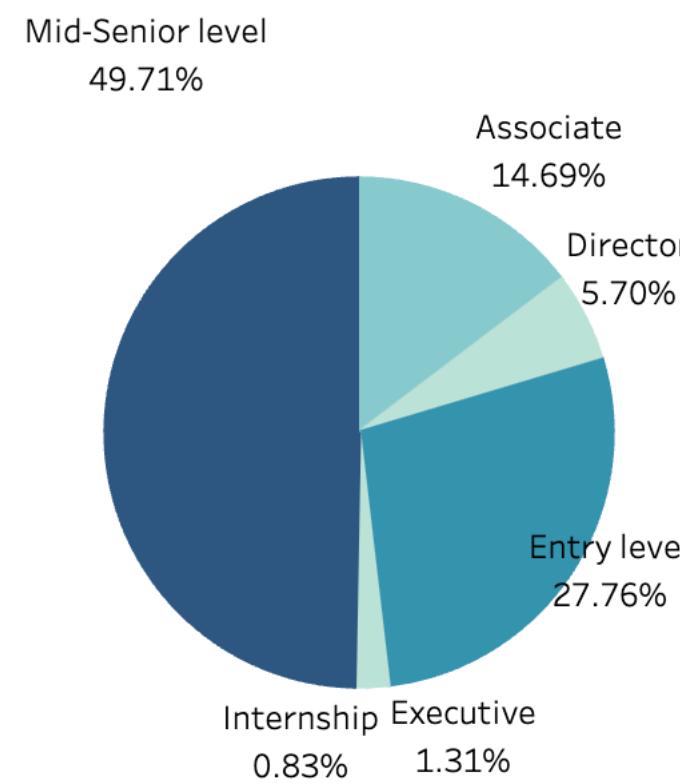


The yellow cluster appears to have a higher level of engagement, both in terms of views and applications, along with a relatively high salary. The green cluster has lower engagement, fewer views, and a notably lower average salary. The red cluster falls in between, with moderate engagement, higher views, and a mid-range average salary.

Top 10 Industries in demand (number of applies)



Listing per experience level



CONCLUSION

[Link to the dashboard](#)

Results:

- 10 Top industries in demand: Staffing and Recruiting, Technology and IT, Healthcare, Financial Services, Retail, Government, Human Resources, Real Estate, Non-Profit and Hospitality.
- Almost half (49.71%) of the job posts are mid-senior level
- Friday is the day of the week with most job listing and Monday is the least.
- Job posts with highest views are mostly applied.
- The analysis suggests that salary has minimal impact on the decision to apply to job posts. Further investigation is warranted to comprehensively examine various factors influencing application decisions, including location, benefits, experience level, and other relevant considerations.

Limitations of the case study:

- Insufficient data points, which may have limited the statistical significance of the results.
- The analysis was constrained by the availability of variables, potentially restricting the depth of the investigation.



THANK YOU

Saltanat Tulegabylova