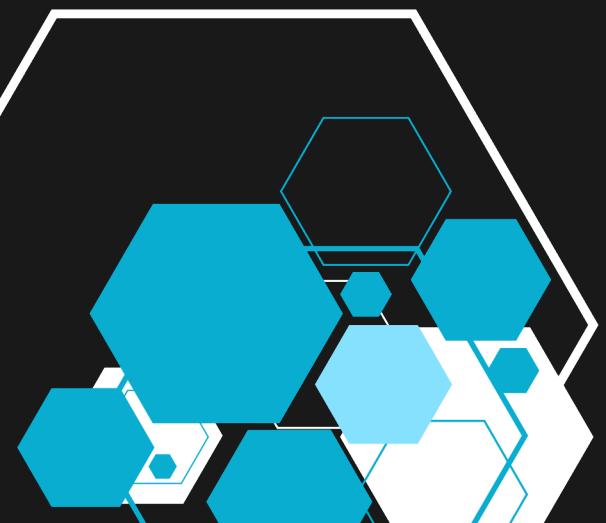


PRIYANKA MITTAL PORTFOLIO



ABOUT ME

I am Priyanka, a proficient data analyst, passionately committed to delivering transformative solutions through data-driven insights. My specialty lies in identifying correlations within datasets, transforming complex challenges into viable opportunities, and prioritizing the best interests of all stakeholders involved.

My prior engagement in the automotive navigation sector equipped me with substantial experience in enhancing product features by comprehending trends and forecasting growth. My inquisitive nature ensures a relentless pursuit of understanding and bridging knowledge gaps.

I am presently exploring opportunities in the data analysis field, aspiring to employ my expertise in scrutinizing extensive data sets. The goal is to deliver valuable insights on product feature trends that could steer business decisions towards a trajectory of success.

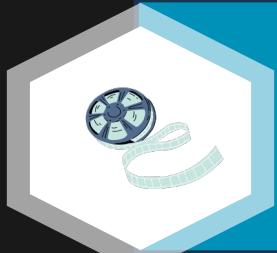


PROJECTS



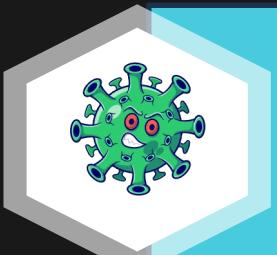
1

Instacart Basket



2

Rockbuster Stealth



3

Preparing for Influenza Season



4

Pig E. Bank



5

GameCo



INSTACART BASKET

Uncover sales patterns to suggest customer segmentation strategies and advertisement scheduling

Instacart

Company

Instacart is an online grocery store that operates through an app; allowing customers to place grocery orders and have them delivered to their homes.

Context

Instacart has a variety of customers in their database with unique purchasing behaviors. They are considering targeted marketing and need a strategy to ensure customers are advertised appropriate products.

Problem Statement

We will perform initial and exploratory data analysis of company data in order to derive insights and create suggestions for better customer segmentation.

Instacart

Perform exploratory analysis to uncover information about sales patterns and suggest strategies for segmentation and advertisement

GOAL

Determine **busiest days** of week and **hours of day** to assist in ad scheduling
Segment products using simpler **price range groupings**
Determine popular products and departments
Determine **ordering behaviors** of different customer demographics and profiles

DATA

The data used for this project included information about Instacart's customers, departments, products, and orders.

The database was derived from The Instacart Online Grocery Shopping Dataset 2017, which can be accessed [here](#)

A data dictionary can be found [here](#)

SKILLS > PYTHON

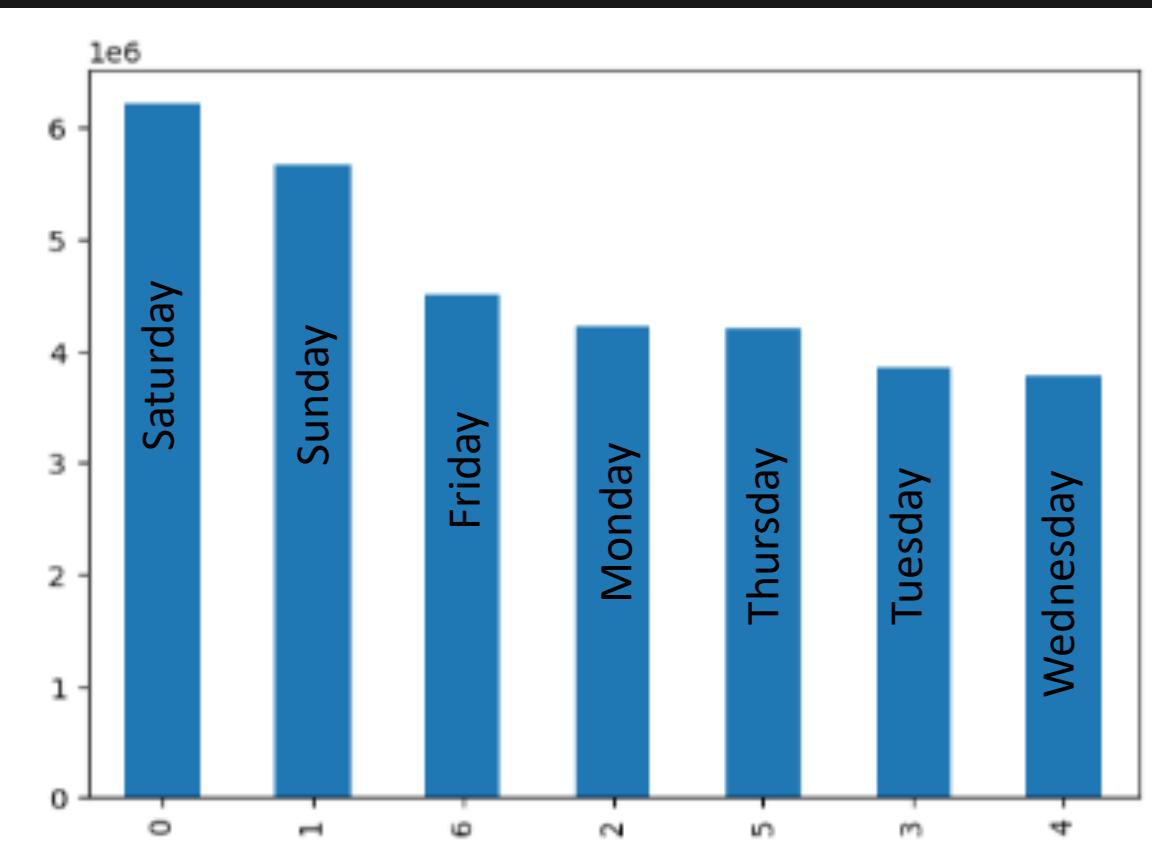
Importing libraries, including **Pandas**, **Numpy**, **Matplotlib**, and **Seaborn**
Conducting descriptive exploratory tasks
Data wrangling and subsetting
Conducting data consistency checks
Combining and exporting data
Deriving new variables using conditional logic
Grouping and aggregating data
Data visualization in Python

[Link to Project Brief](#)

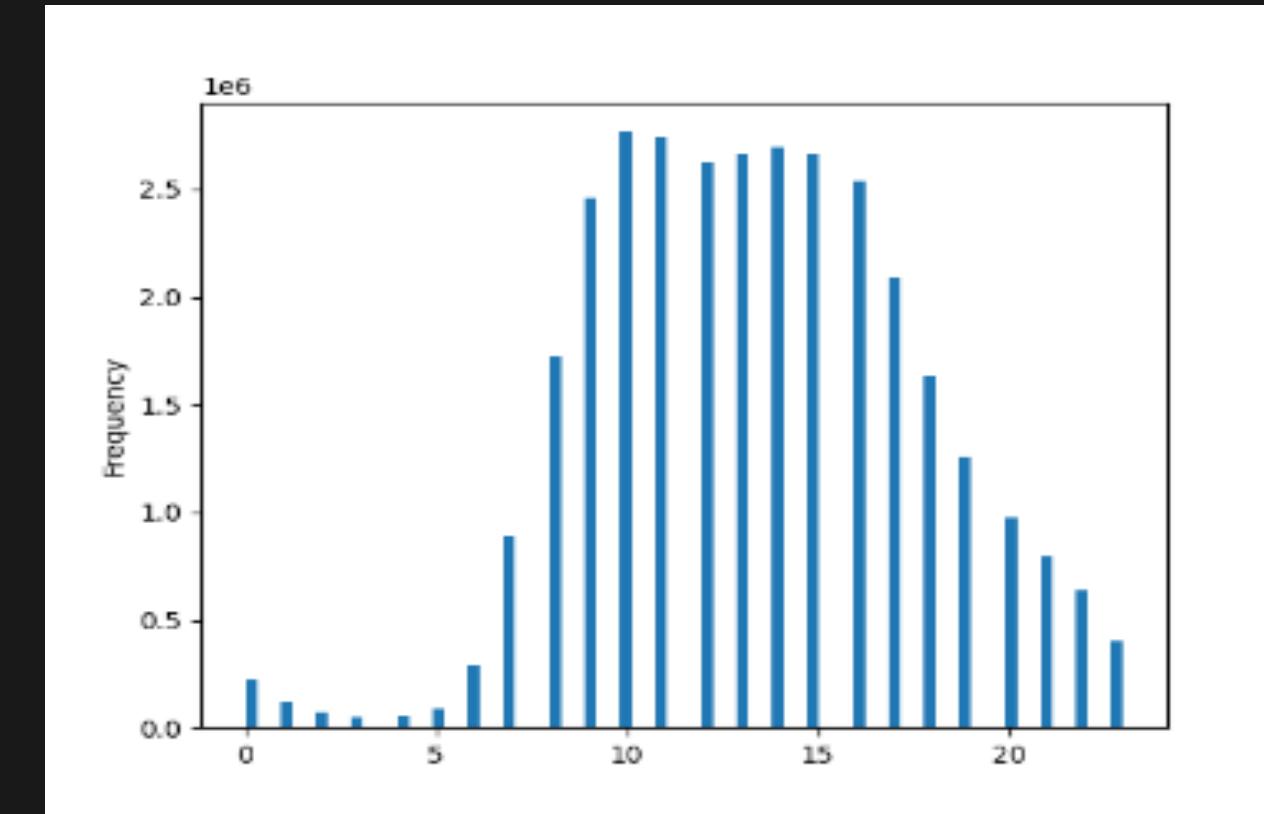
Analysis: Busiest Times

The largest number of orders occur on the weekends, and Instacart receives the most orders between 10:00 AM to 3:00 PM

Number of Orders Placed on Each Day



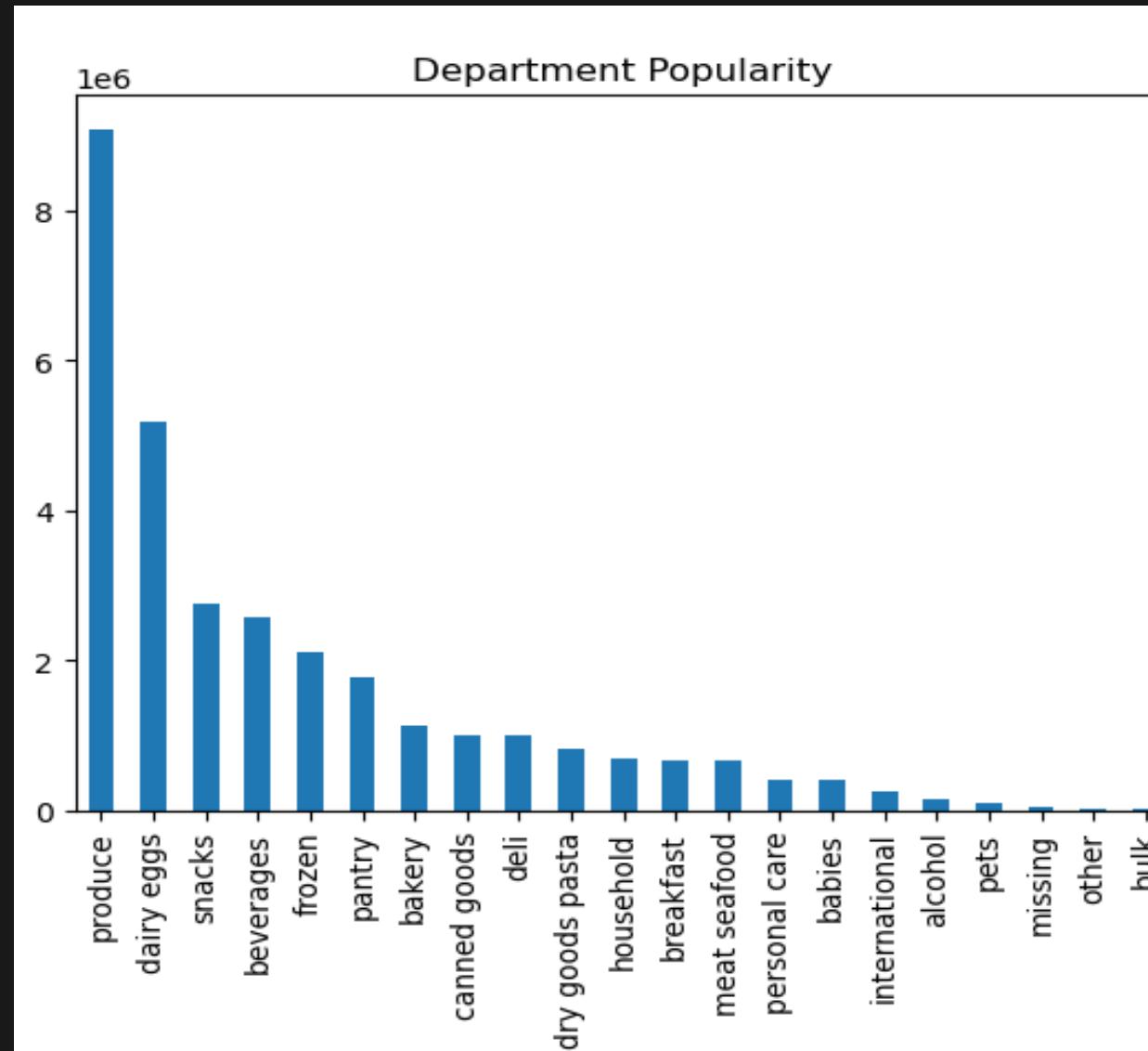
Number of Orders Placed Each Hour of the Day



There is an **Advertising opportunity** which can be scheduled in mid-week before **9 am or after 6 pm**, in order to increase sales during this time.

Analysis: Department Popularity

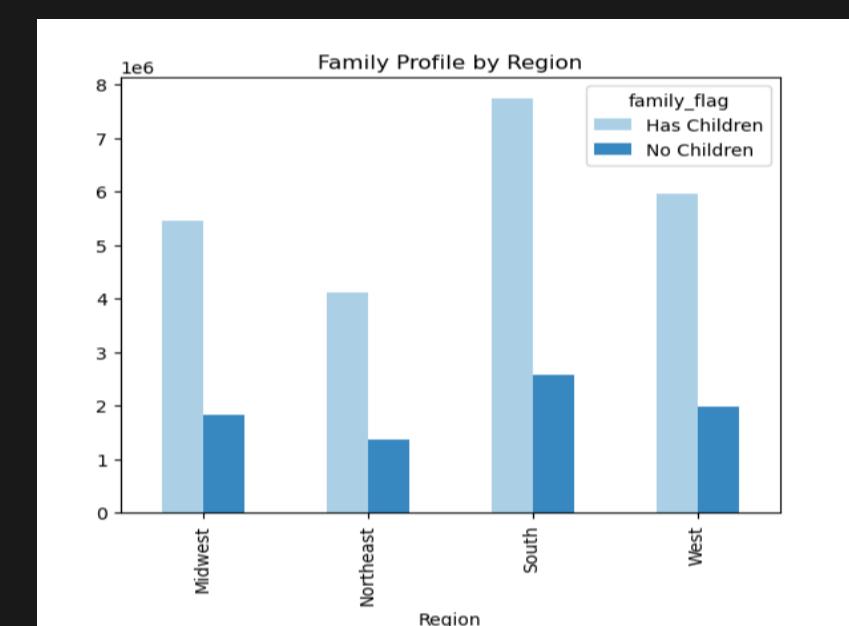
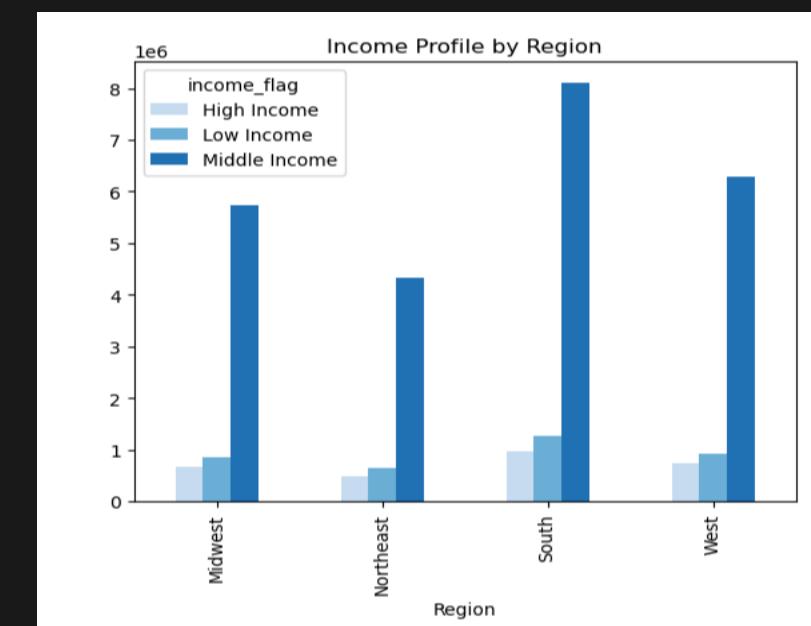
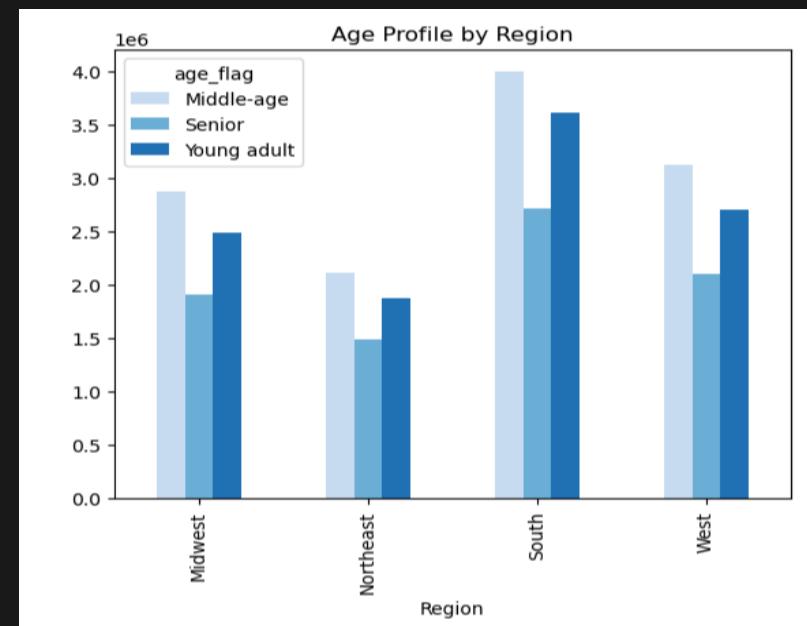
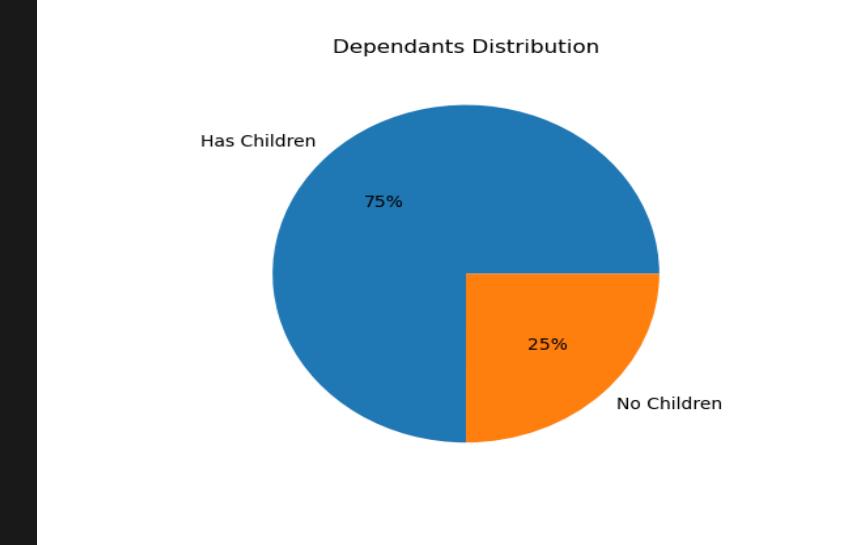
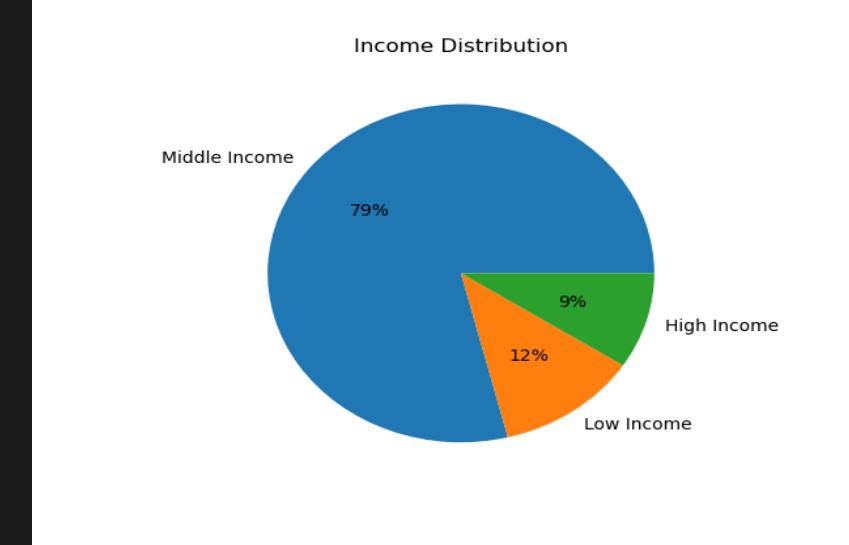
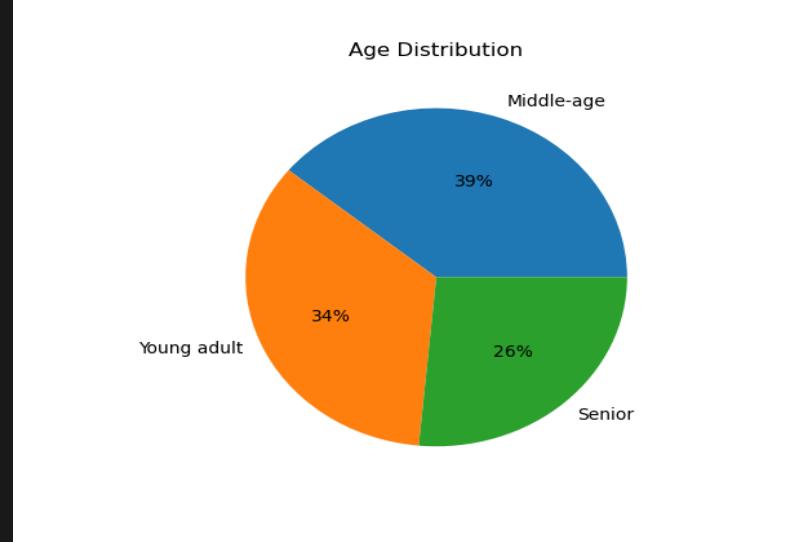
Produce, dairy/eggs, and snacks have the highest number of orders. However, alcohol, products from the babies department, and bulk products have the fewest number of days between orders, and users spend the most on meat/seafood, dairy/eggs, and bulk items.



department	Days Since Last Order			Prices			
	min	max	mean	sum	min	max	mean
alcohol	0	30	9.332472	1179302.3	1	15	8.151446
babies	0	30	9.335526	3133096.4	1	15	7.634399
bakery	0	30	10.225862	8852224.6	1	15	7.868462
beverages	0	30	10.206104	19759202	1	15	7.682624
breakfast	0	30	10.389484	5470900.4	1	14.9	8.068044
bulk	0	30	9.451048	279212.8	1.4	14.1	8.346919
canned goods	0	30	10.672517	7637778.5	1	15	7.54666
dairy eggs	0	30	10.155419	43363062	1	15	8.352176
deli	0	30	10.431782	7813188.1	1	15	7.783347
dry goods pasta	0	30	10.770432	6042181.4	1	15	7.34937
frozen	0	30	10.740142	16431751	1	15	7.738627
household	0	30	10.968798	5164912	1	15	7.379943
international	0	30	10.316078	1965702.9	1	15	7.678797
meat seafood	0	30	10.505405	10999037	8	25	16.300062
missing	0	30	9.659415	560933.3	1	15	8.660655
other	0	30	10.056842	239569.2	1.1	15	6.961995
pantry	0	30	10.220274	14289861	1	20	8.01494
personal care	0	30	10.604785	3394733.1	1	15	7.998617
pets	0	30	10.93915	734083.7	1	15	7.88786
produce	0	30	10.026067	72455791	1	15	7.980349
snacks	0	30	10.108421	11827920	1.6	7	4.275555

Analysis: Customer Segmentation

Instacart's primary user group **is middle-aged, middle-income families**. This is consistent regionally.



Recommendation for Instacart

01

Advertisement Schedule: to boost sales during lower traffic times, Instacart should schedule advertisements for Tuesdays and Wednesdays after 4:00 PM.

02

Product Popularity: because produce and dairy/eggs are already a part of most people's preferences, it is unlikely that they need to spend a lot of advertising dollars convincing people to purchase these. It may make more sense to increase advertisements of meat and seafood, where prices are higher and orders are not as high to attract more revenue.

03

Customer Profiling: Instacart's largest customer base is middle-age, middle-income families. They should decide whether to target this group further and investigate preferred products and shopping times or choose to attempt to expand their customer base.

PROJECT DELIVERABLES

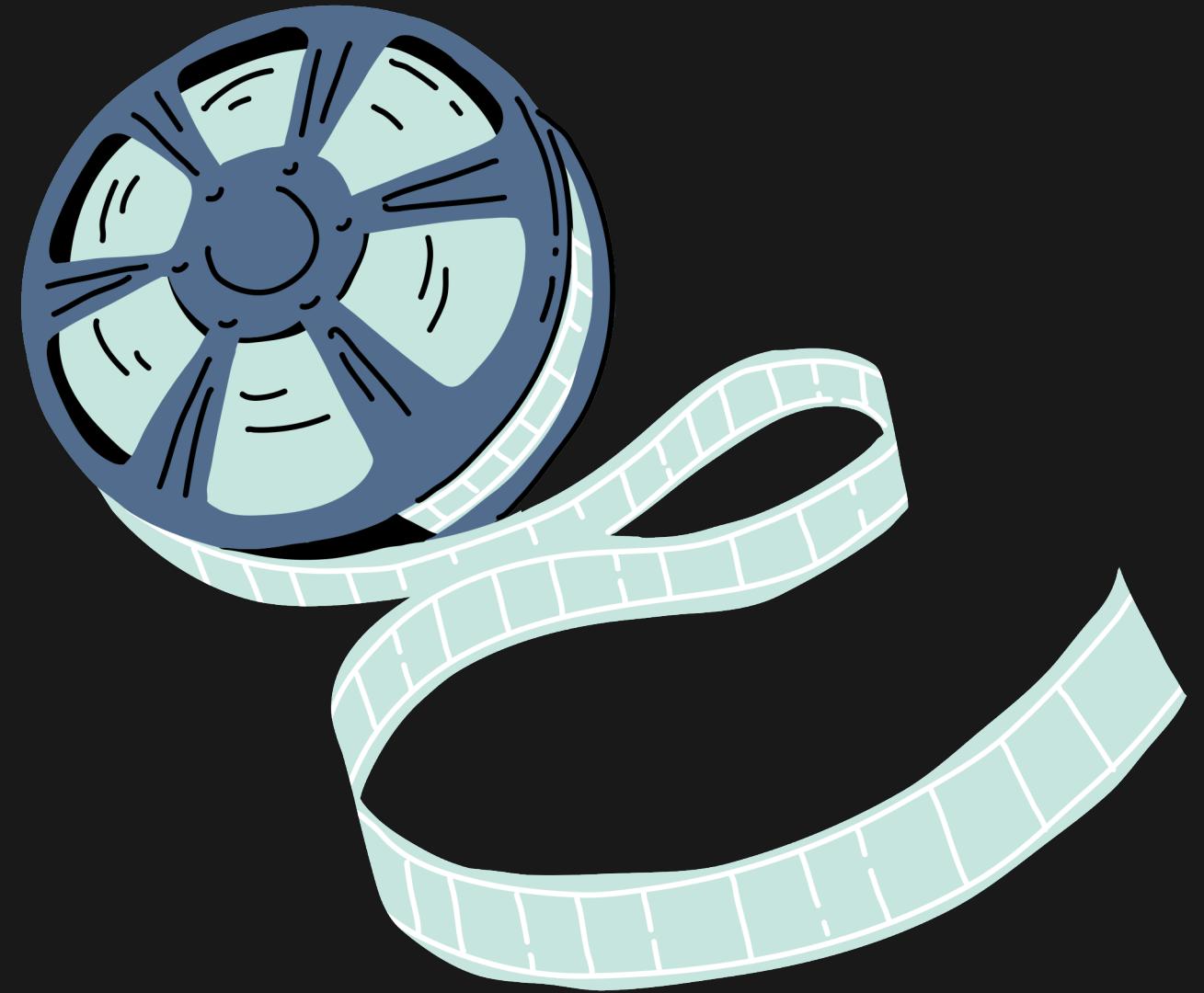
Click the icons to access the file



Project Python Scripts –
Hosted on GitHub



Final Report



ROCKBUSTER STEALTH

*Analyze database to develop competitive strategy recommendations
against streaming services*

Rockbuster Stealth

Company

Rockbuster LLC is a movie rental company that is wanting to transition to an online video rental service due to increasing competition from rival streaming services.

Context

The business intelligence department will work to create a launch strategy for the new online video rental service

Problem Statement

We will introduce data management systems and perform analysis in order to answer company questions regarding sales and customer demographics.

Rockbuster Stealth

Assist fictitious movie company Rockbuster Stealth in developing a strategy to remain competitive with online streaming services

GOAL

Analyze **descriptive statistics** of movie rentals to determine patterns and trends

Determine **customer loyalty**

Investigate regional differences in customer numbers, sales and genre preferences

Recommend strategies based on analysis

DATA

The data used for this project included information about Information provided by CareerFoundry that includes data on film inventory, customers, payments, and other information.

A complete data dictionary can be found [here](#).

SKILLS > SQL

Understand and utilize **relational databases**

Query data in SQL by **ordering, limiting, and grouping data**

Filter data using **WHERE** and **HAVING** clauses

Identify and clean dirty data

Join tables

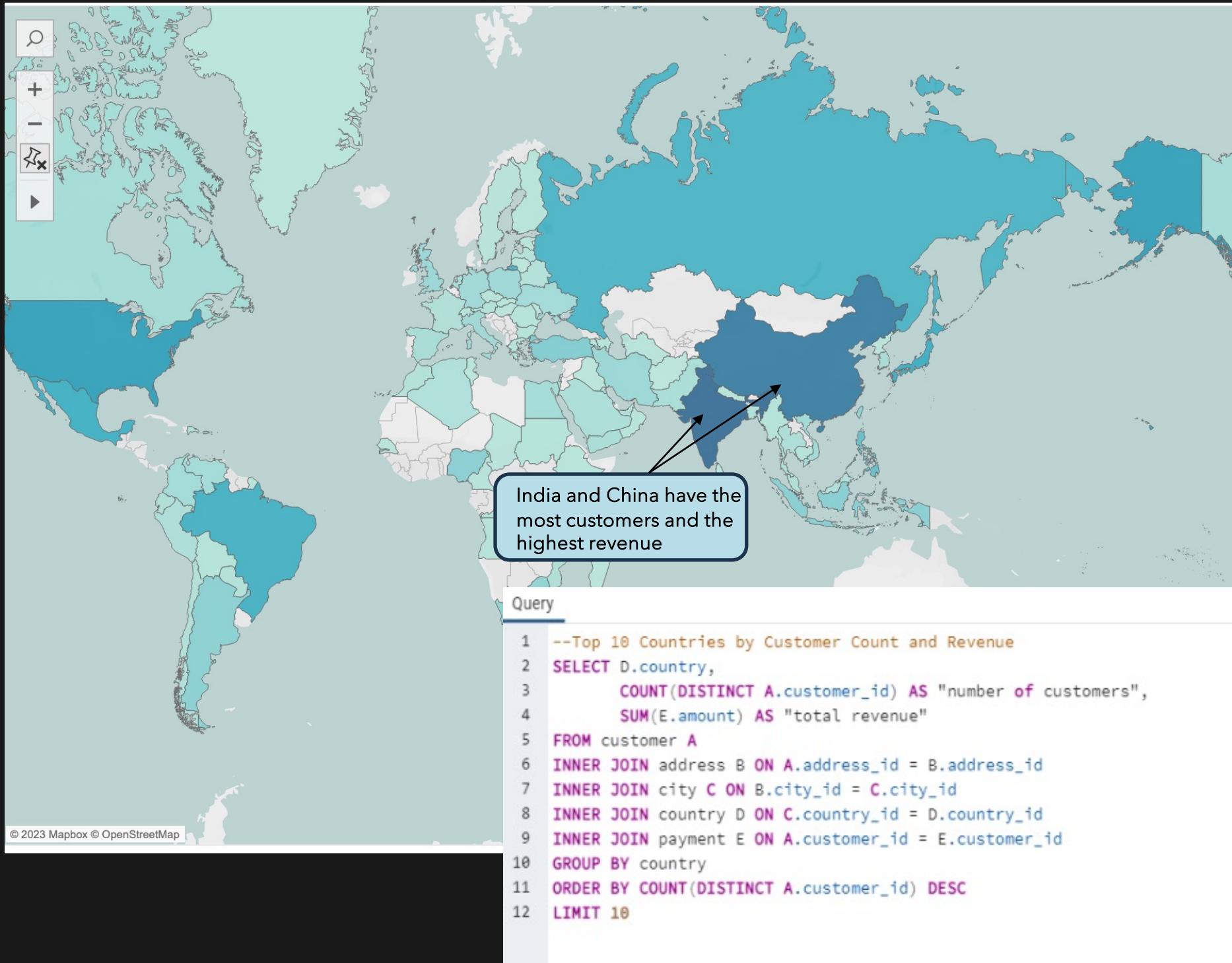
Perform **subqueries and Common Table Expressions**

Present findings

[Link to Project Brief](#)

Analysis: Geographic Distribution of Customers

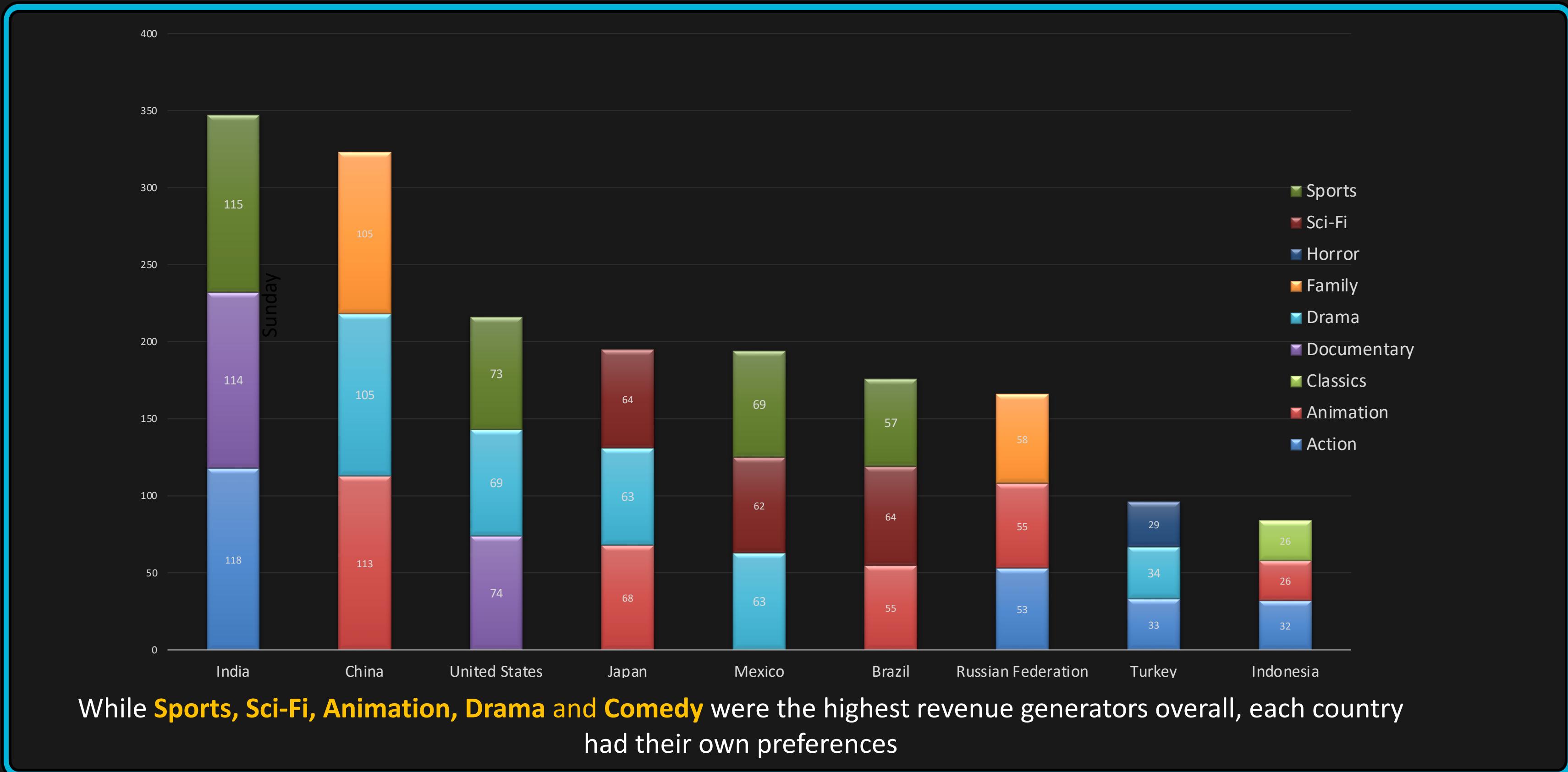
According to the data set, there are 599 customers across 108 countries, with the majority residing in Asia and the Pacific



Top Countries	Revenue
India	\$6,035
China	\$5,251
United States	\$3,685
Japan	\$3,123
Mexico	\$2,985
Brazil	\$2,919
Russia Federation	\$2,766
Philippines	\$2,220
Turkey	\$1,498
Indonesia	\$1,353

Analysis: Genre Preferences in Top 10 Countries

Each of the top 10 countries is unique in their go-to movie genres



Recommendation for Rockbuster Stealth

01

Focus marketing strategy & budget on Asia: Particularly in India, China and Japan where we already have large pools of customers and high spenders

02

Genres: While the top genres are Sports, Sci-Fi, Animation, Drama and Comedy collectively, consider marketing specific genres within each country

03

Expand our selection of movies: For example by offering a wider choice of languages and movies We currently only offer movies in English and from 2006.

PROJECT DELIVERABLES

Click the icons to access the file



Project Presentation



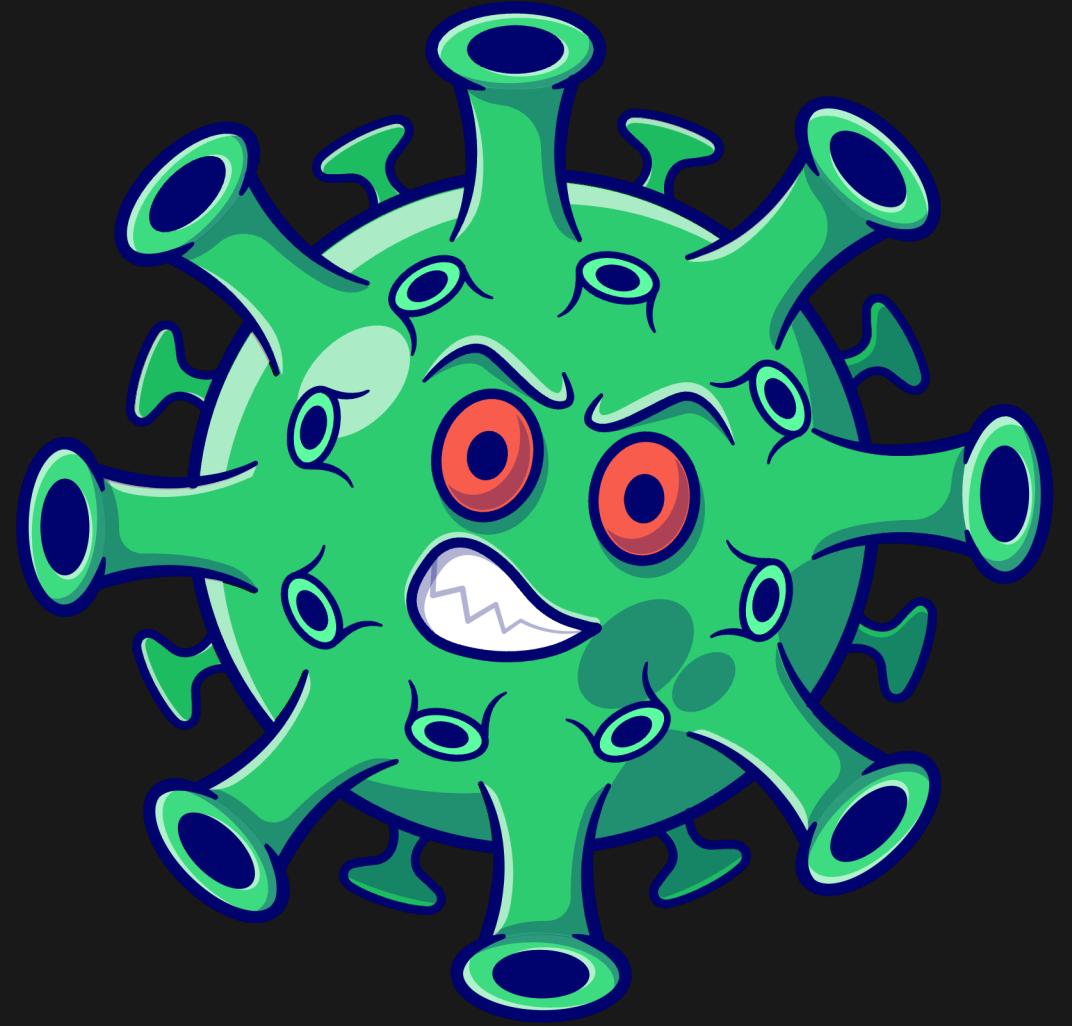
Tableau - Storyboard



Project Technical File
– Hosted on GitHub



Data Dictionary



PREPARING FOR INFLUENZA SEASON

Investigating trends to assist in staffing agency needs

Preparing for Influenza season

Company

We will be aiding a medical staffing agency that provides temporary workers to healthcare facilities.

Context

The United States has an influenza season where more people than usual contract and suffer from the flu. This results in increased complications and hospitalizations. The increase of patients will require additional medical staff.

Problem Statement

In order to properly plan for the yearly outbreak, trends in influenza will be examined and used to proactively plan for staffing needs across the country.

Preparing for Influenza season

Assist a medical staffing agency with planning the disbursement of temporary workers to clinics and hospitals during the flu season throughout the United States

GOAL

Identify who falls into the **vulnerable population** category and prioritize states with high percentages of these groups

Determine **seasonality** of influenza and variances across states

DATA

Influenza deaths by geography, time, age and gender Source: [CDC](#)

Population data by geography
Source: [US Census Bureau](#)

SKILLS > EXCEL > TABLEAU

Designing data research projects:

- Formulating **hypothesis**
- Data **profiling** and quality measures
- Data transformation & integration
- Conducting **statistical** analysis
- and **hypothesis testing**

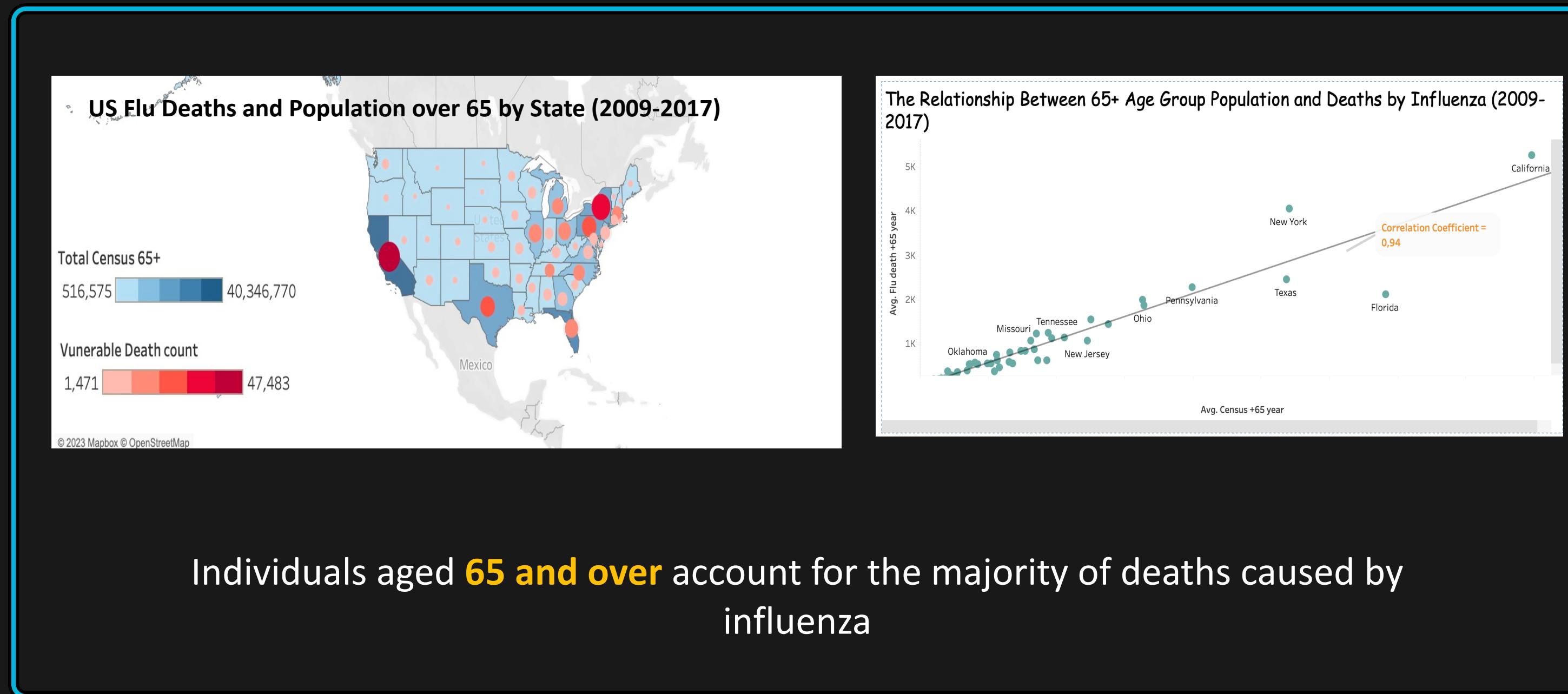
Tableau Visualizations:

- Composition & **comparison** charts
- **Temporal** visualizations
- **Statistical** visualizations
- **Spatial** analysis
- **Textual** analysis

[Link to Project Brief](#)

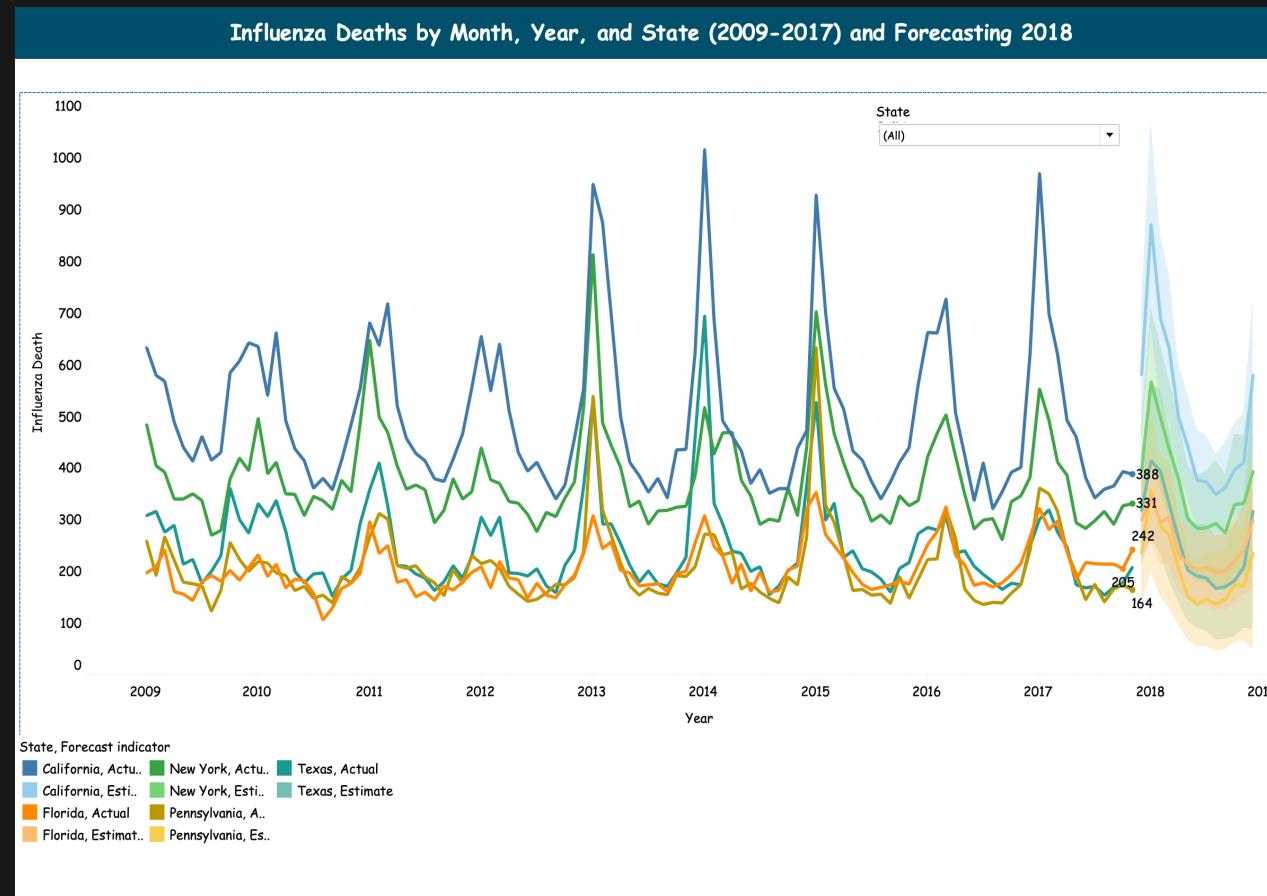
Analysis: Vulnerable Population Assessment

Determined that those age 65 and over are at higher risk for severe complications from the flu, and many of these individuals reside in California, New York, Texas, Pennsylvania, and Florida

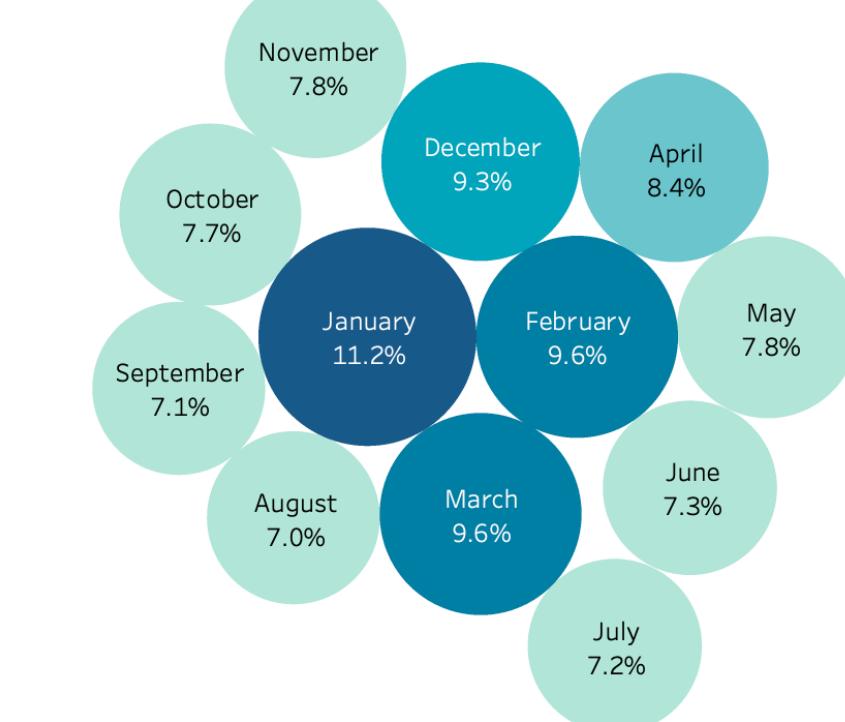


Analysis: Seasonality of Influenza

Winter months, from December to March have historically seen the highest number of flu deaths



Percentage of Total Flu Deaths for Each Month (Out of the Total Year)



There is **very little variability** between states in months with high flu deaths, which leaves no opportunity for staggered staffing plans

Recommendation for 2018 flu season

Recommendation:



Our forecast suggests a similar pattern of flu seasonality to the past nine years. The recommendation is to deploy additional medical staff to California, New York, Florida, Texas and Pennsylvania from December to March.



Next Steps:

1. Conduct further analysis on other vulnerable populations, as well as medical staff availability and hospital capacities.
2. Monitor the success of the project for future years.

PROJECT DELIVERABLES

Click the icons to access the file



Interim Report



Project Technical File



Presentation -Tableau Story board



PIG E. BANK

Dive into Data Ethics issues and Data Mining

Pig E. Bank

Company

A well-known global bank needs help with its anti-money-laundering compliance department.

Context

We have been hired to help the bank in running their compliance program more efficiently.

Problem Statement

We will help build and optimize models that assist the bank's compliance department assess client risk. We will help the bank predict client loyalty based on various factors.

Pig E. Bank

Use principals of data ethics to assist Pig E. Bank in navigating challenging issues and begin exploring the use of data mining and predictive analysis.

GOAL

Use **decision tree algorithms** to determine the probability of Pig E. Bank's customers leaving the bank
Explore **ethical issues** within Pig E. Bank's operations

DATA

Pig E. Bank's client data set, found. [here](#)

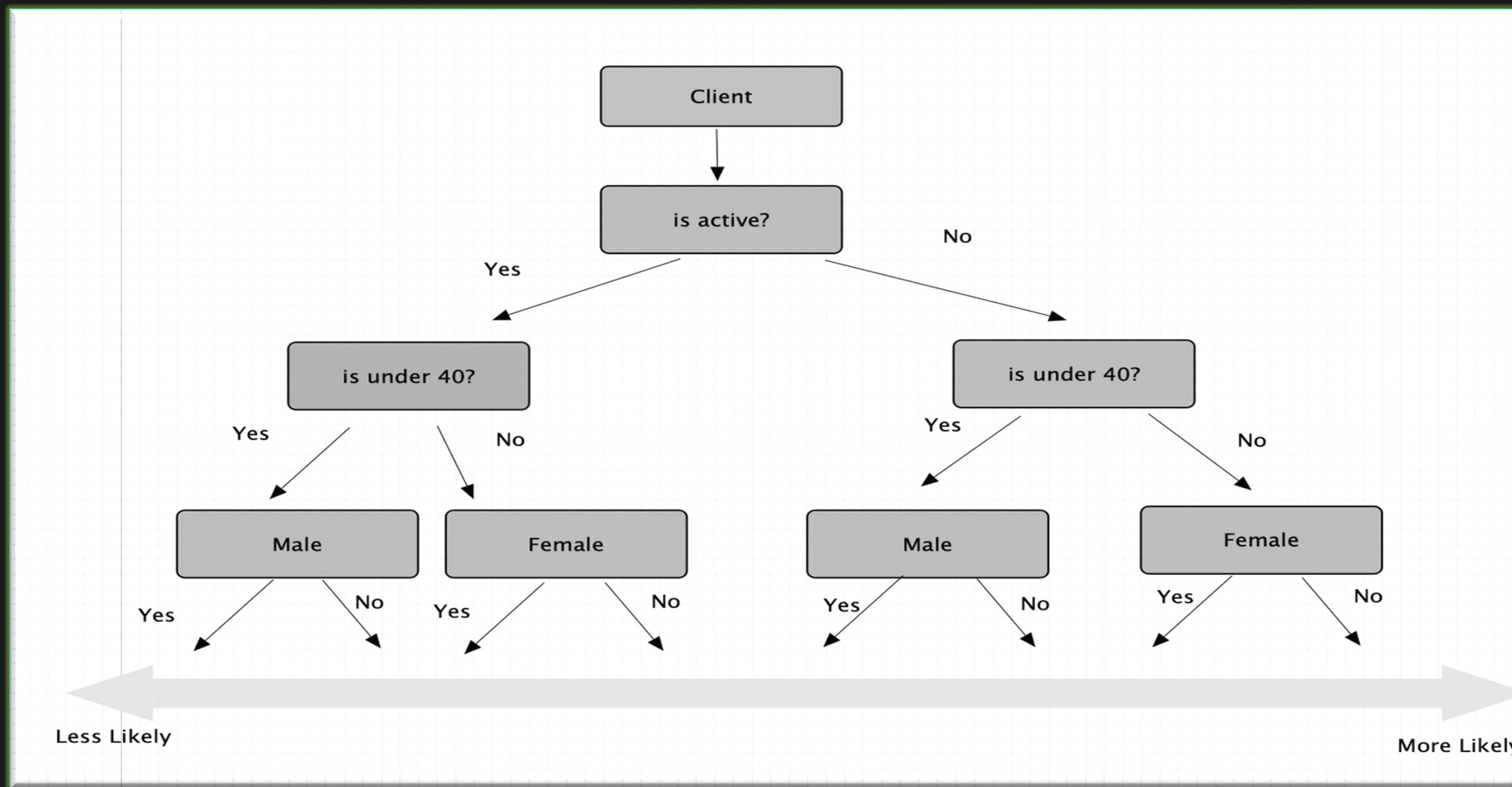
SKILLS

Explore data ethics issues such as **data bias, security & privacy**
Begin exploring **data mining** and **usage of decision trees**
Understand and utilize **CRISP- DM methodology**
Utilize **time-series analysis**

Analysis: Probability of Clients leaving

According to a descriptive analysis on Pig E. Bank's data, the likelihood of a client leaving can be estimated using the below decision tree algorithm

Likelihood of Client leaving Bank



Additional Analysis Files & Deliverables

Click the icons to access the file



Data Bias CaseStudy

Exploration into various data bias type



Descriptive Analysis

Descriptive analysis of Clients's Pig E. Bank data



Data Privacy & Security Case Study

Senario analysis for Pig E. Bank data privacy and security issues



Time Series Analysis

Stationary vs Non-stationary exploration, moving average forecasting and introduction into Forecasting model



GAME CO.

*Market trend analysis for video game
development and sales*

GameCo.

Company

GameCo is a new video game company that is interested in using data to influence the development of new games and how they will fare in the market.

Context

GameCo is interested in developing new games. In order to optimize their marketing, GameCo is interested in which markets to advertise to.

Problem Statement

GameCo wishes to analyze previous game sales in order to gain deeper insights. They are interested in exploring what variables impact a game's sales including: publisher, time, geographic region, and gaming category.

GameCo.

A fictitious video game company interested in descriptive analysis on market data to successfully develop new games

GOAL

Determine genre popularity trends
Determine largest publisher competitors
Analyze market trends to determine video game popularity over time
Uncover geographic sales differences

DATA

Data set provided through VGChartz, found [here](#)
Includes units of games sold from 1980 to 2016, represented in millions

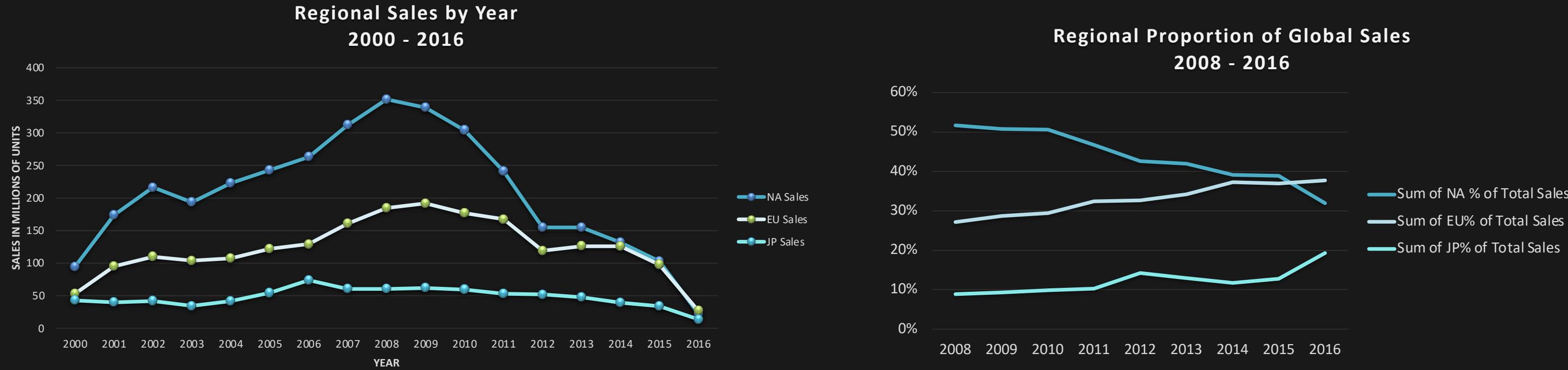
SKILLS > EXCEL

Data cleaning techniques :**Grouping and summarizing** data through pivots tables and filtering
Conducting **descriptive analysis**
Visualizing insights through **scatterplots, box and whisker plots, and bar and column graphs**

[Link to Project Brief](#)

Analysis: Sales Trends

Video game sales varied over time regionally with a sharp decline over the past year



NORTH AMERICA

-78% Total sales decline from 2015

Both total and proportional sales have been on the decline since 2009

EUROPEAN UNION

-73% Total sales decline from 2015

Proportional of global sales has been on a steady incline since 2008—surpassed North America this year

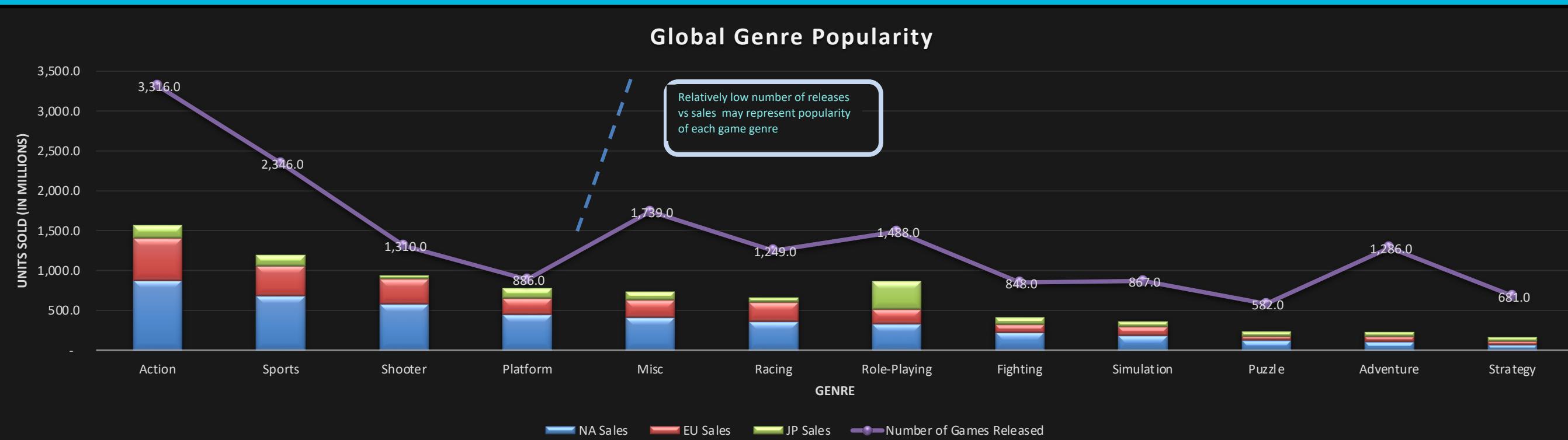
JAPAN

-59% Total sales decline from 2015

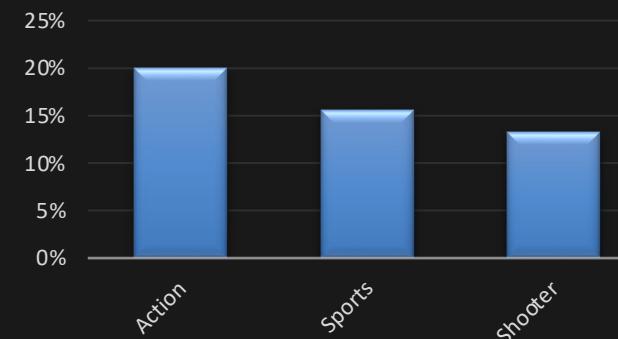
Trendin upwards in their contribution to global sales – could represent a market opportunity

Analysis: Genre Popularity

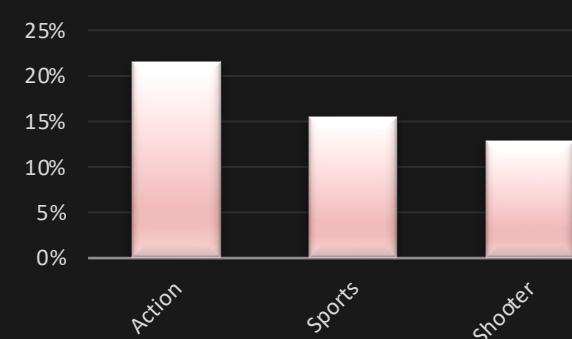
Genre preferences are unique by region, but differ most in Japan ,Action games have higher game release numbers than any other genre, which is important in considering its popularity



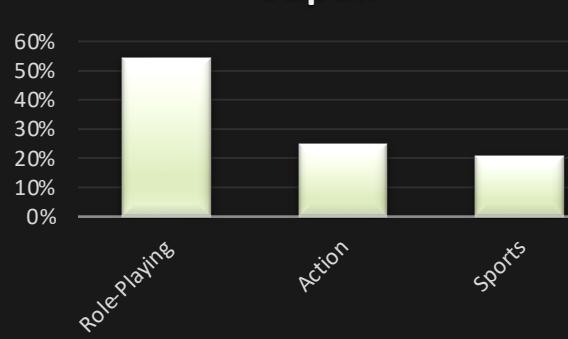
Top Genres in North America



Top Genres in the European Union



Top Genres in Japan



Recommendation for Game CO

01

Market Trends: The EU has increased their market share the most steadily – focus marketing budget on this consistency

02

Genres: Consider Shooter & Sports genres for North America and European Union development, and Role- Playing genres for Japan

03

Platforms: Play station is the most consistently popular platform across all regions and should be prioritized if planning to develop games for one platform only

PROJECT DELIVERABLES

Click the icons to access the file



Project Presentation

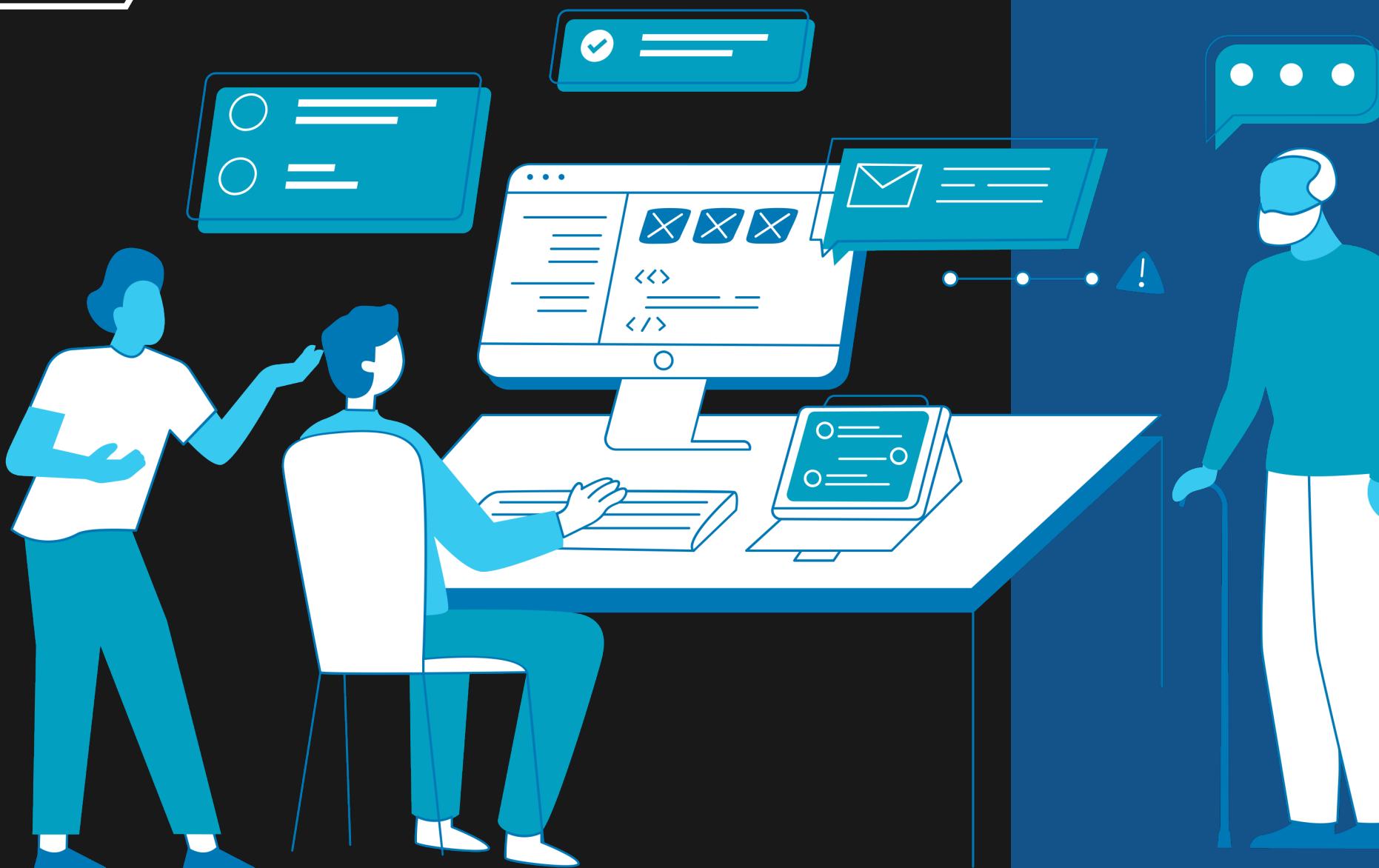


Project Technical File



Project Reflections

THANK YOU



Do you have any questions?

priyanka08101991@gmail.com

