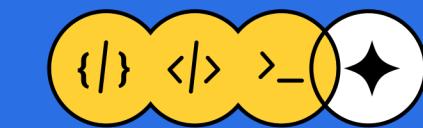
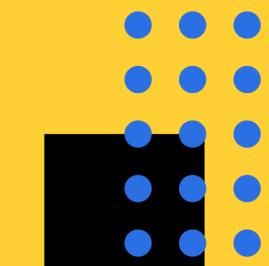
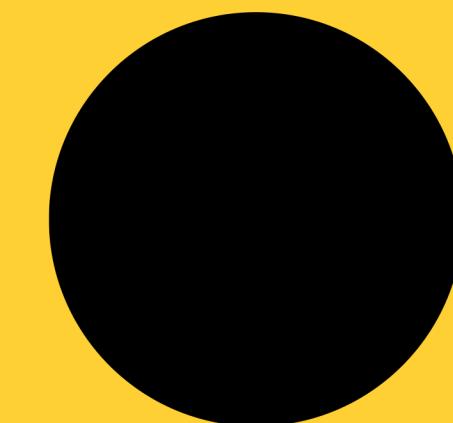


Sarah Edens

Data analytics portfolio 2024



PROJECTS



01

02

03

04

05

GameCo

Global market analysis of console industry

Influenza Season

National staff distribution project

Rockbuster

SQL based project studying video rental services and consumer behavior

Instacart

Python-based project analyzing consumer behavior & trends

Pig E. Bank

Predictive analysis of customer retention risk for finance company

01

GameCo

Video Games
Sales Analysis
& Market
Reccomendations

GameCo, an emerging video game enterprise, aims to leverage data for refining the creation of new games.

Objective: Conduct a descriptive analysis on a video game dataset to enhance comprehension of potential market reception for GameCo's upcoming releases.

GOALS



Conduct an in-depth examination of video game data from 1980 to 2016 to advise GameCo on crafting their 2017 strategy for creating fresh games and expanding the company.



Skills



- Data cleaning/grouping
- Pivot tables
- Descriptive Analysis
- Visualizations in Excel/Powerpoint



DATA & LIMITATIONS

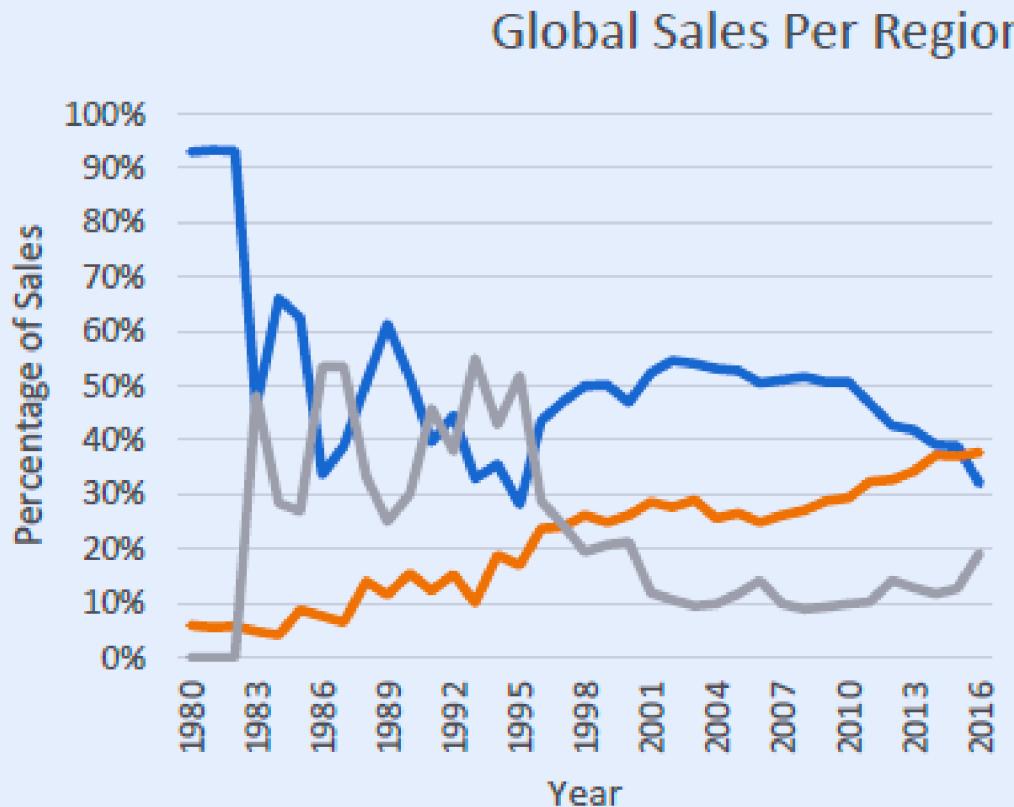
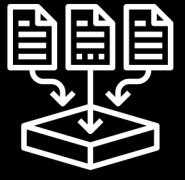
- The dataset utilized for analysis comprises historical video game sales data sourced from VGChartz, encompassing diverse platforms, genres, and publishers from 1980 to 2016.
- It monitored total units shipped rather than actual sales, thus lacking precise financial figures. Additionally, data collection was incomplete for the year 2016.

Analysis



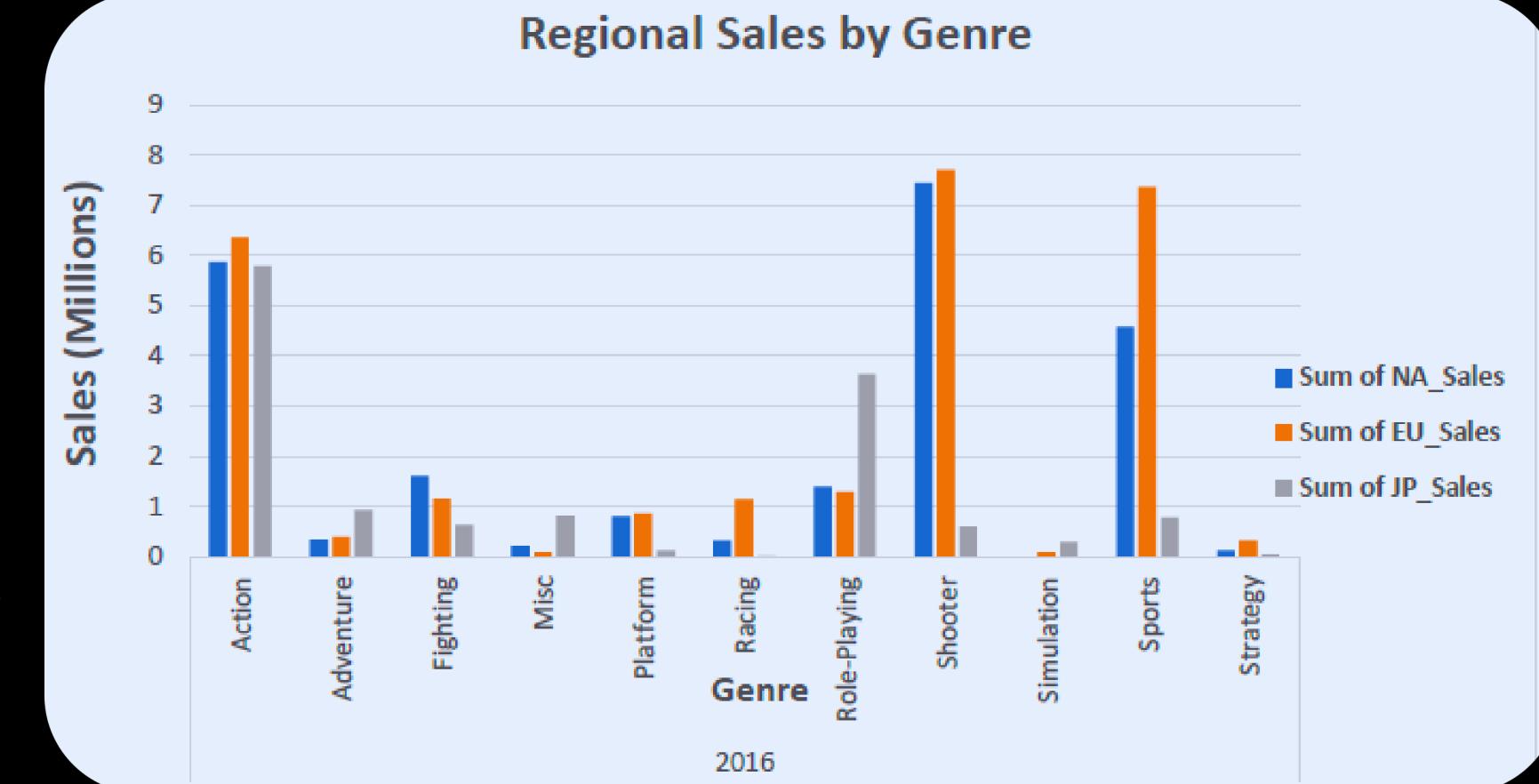
Data source:

<https://www.vgchartz.com/>



- Analyzed regional sales trends across various game genres to identify performance disparities.
- Recommend reallocating capital to optimize regional sales by focusing on genres with higher potential for growth, enhancing future sales prospects.

- Conducted regional data grouping North America Sales(NA), Europe Sales(EU), and Japan Sales(JP) analyzing years spanning from 1980-2016.
- North American sales init exhibited substantial dominance.
- Towards 2016, European sales demonstrated consistent growth, potentially surpassing North America and Japan.



Recommendations

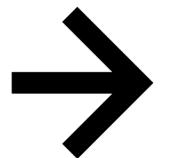
GameCo's Marketing Budget for 2017

MARKETING STRATEGY



The thorough analysis of GameCo's data yielded valuable insights into various facets such as global sales performance, regional market segmentation, and the nuanced preferences surrounding genres and platforms.

MARKETING ADAPTATION



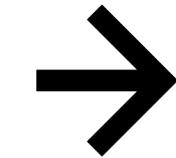
By strategically reallocating the marketing budget to address the evolving market landscape, we can effectively capitalize on opportunities presented by emerging markets while mitigating the downward trend observed in the North American market.

FOCUS ON STRENGTHS



Following years of declining sales, the markets have shown a gradual adjustment at a modest pace. Investigate the existing customer base and delve into areas where console games maintain superiority over mobile games.

PROMOTE GROWTH



Support the development of emerging markets, prioritize strong genres, and pursue opportunities for growth in additional countries.

02

Influenza Season

National staff
distribution
project



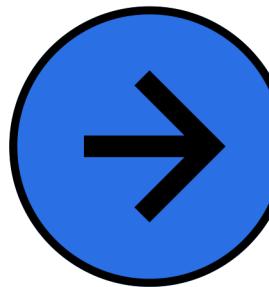
Assisting a medical staffing agency in preparing for influenza season by analyzing trends to proactively plan staffing needs nationwide.

OBJECTIVE:

FORECAST INFLUENZA TRENDS TO OPTIMIZE STAFFING FOR A MEDICAL AGENCY DURING FLU SEASON, WITH A FOCUS ON MORTALITY RATES AND AGE CORRELATION, ESPECIALLY AMONG PATIENTS AGED 65 AND OLDER.

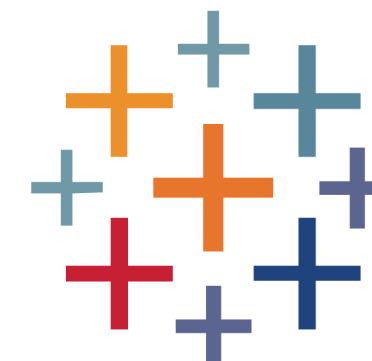
Goals

Analyze historical trends in influenza outbreaks across the United States to proactively forecast national staffing needs and devise a systematic strategy for a medical staffing agency to efficiently deploy temporary healthcare professionals during the upcoming flu season.



SKILLS

- Data cleaning
- Data integration
- Hypothesis testing
- Tableau visualizations



DATA & LIMITATIONS

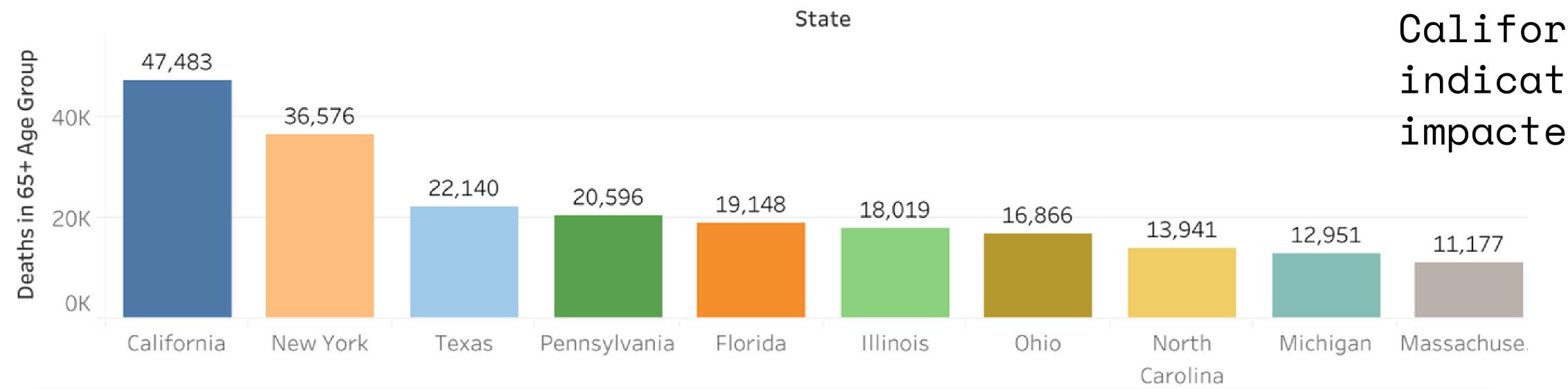
- Influenza fatality statistics (2009-2017) segmented by:
 - Location
 - Timeframe
 - Age
 - Gender
- 82% of entries suppressed for confidentiality, sourced from CDC
- Population demographics by geography obtained from the US Census Bureau.
- Death records may only reflect one cause of death, potentially resulting in an underestimation of influenza-related fatalities.
- Data spans from 2009 to 2017 and is restricted to participating clinics and hospitals.



Analysis

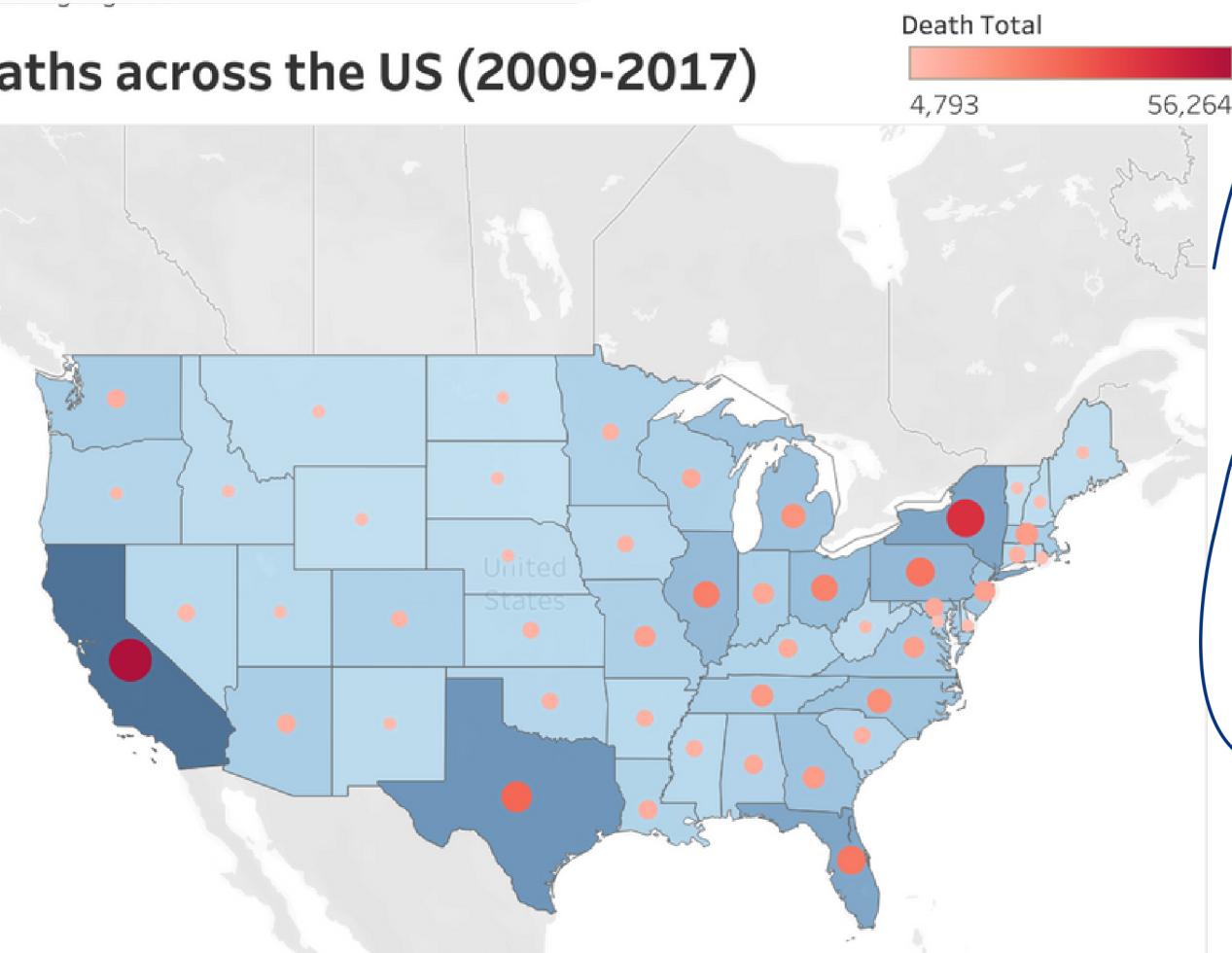


Top 10 US States Flu Mortality 65+ (2009-2017)

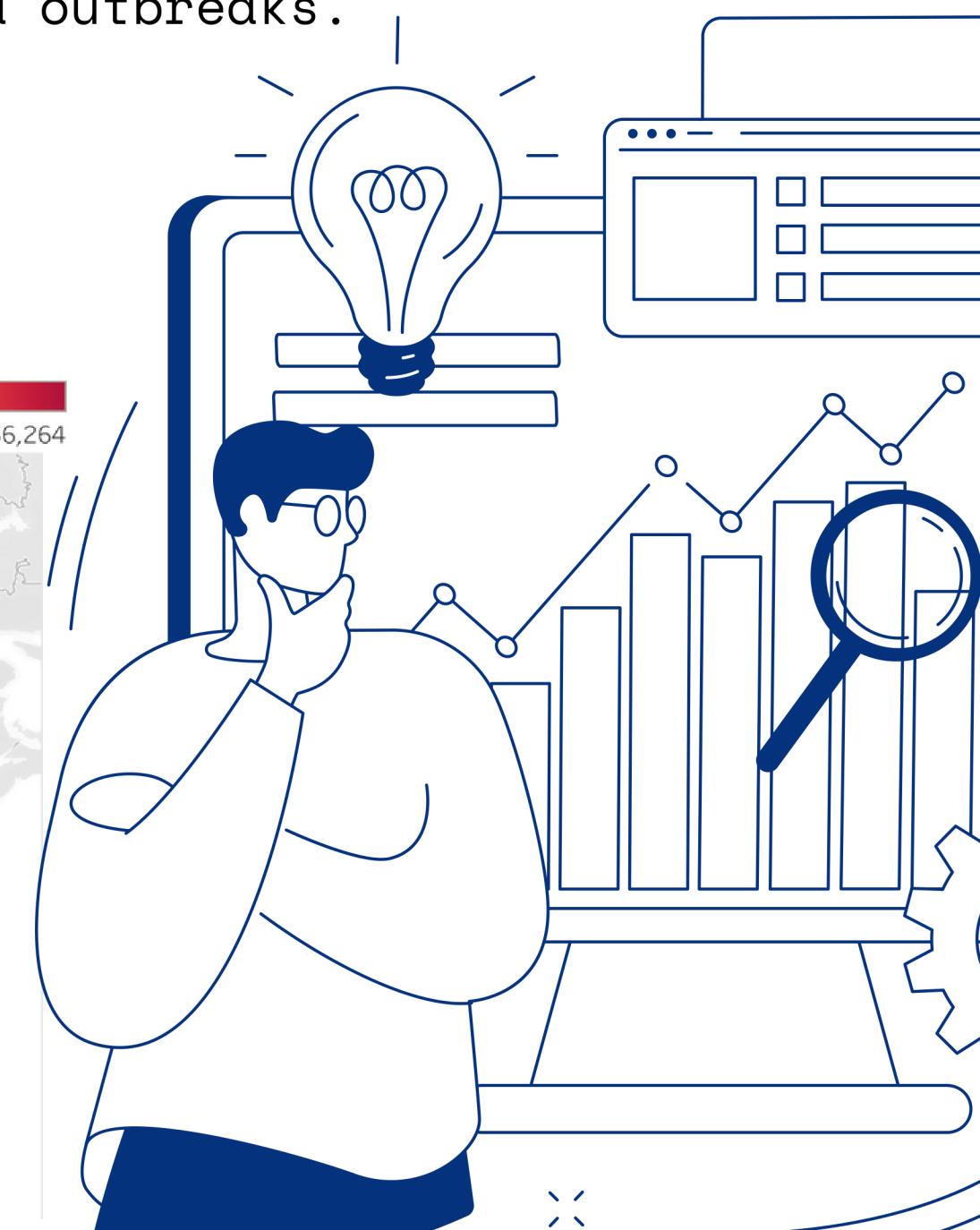


- The bar graph to the left illustrates the top 10 states with the highest influenza rates with California and New York leading the list, indicating the regions most significantly impacted by influenza outbreaks.

Deaths across the US (2009-2017)



- The heatmap to the right identifies hotspots and zones experiencing elevated influenza-related fatalities, underlining the states demanding heightened attention regarding staff allocation.

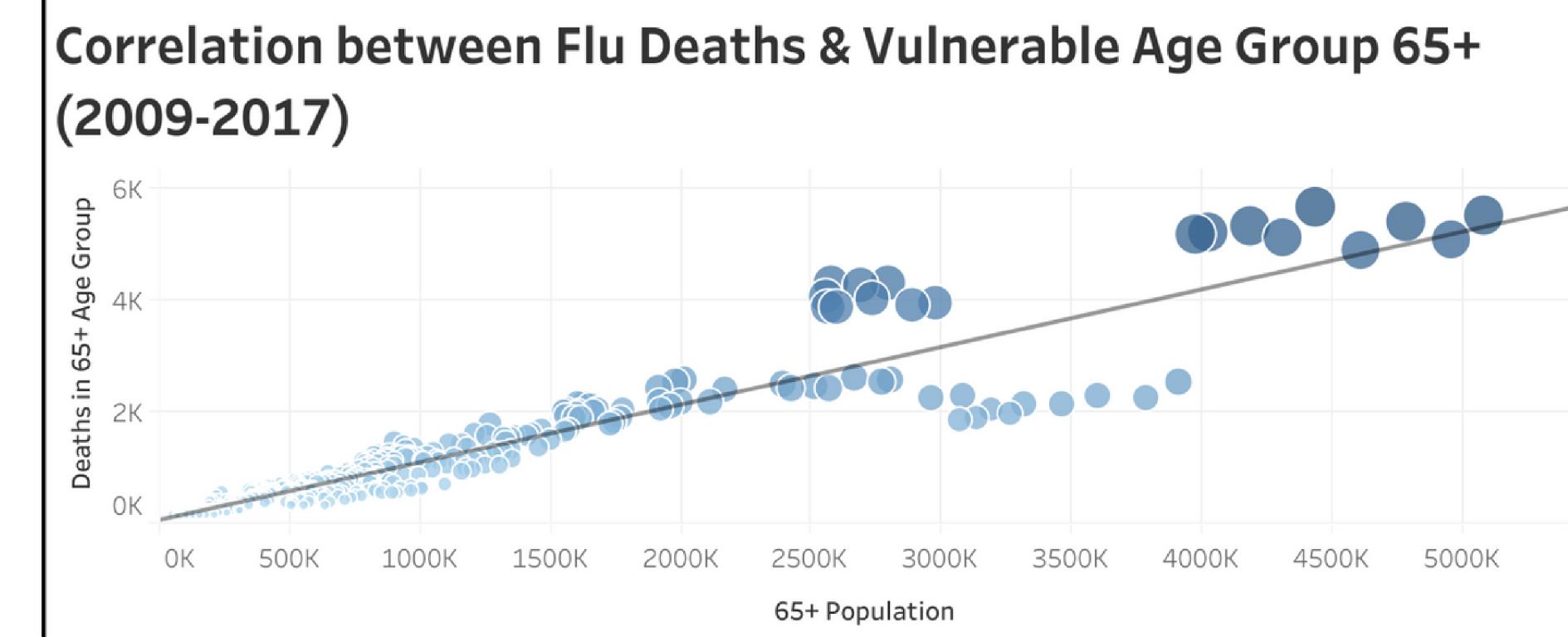
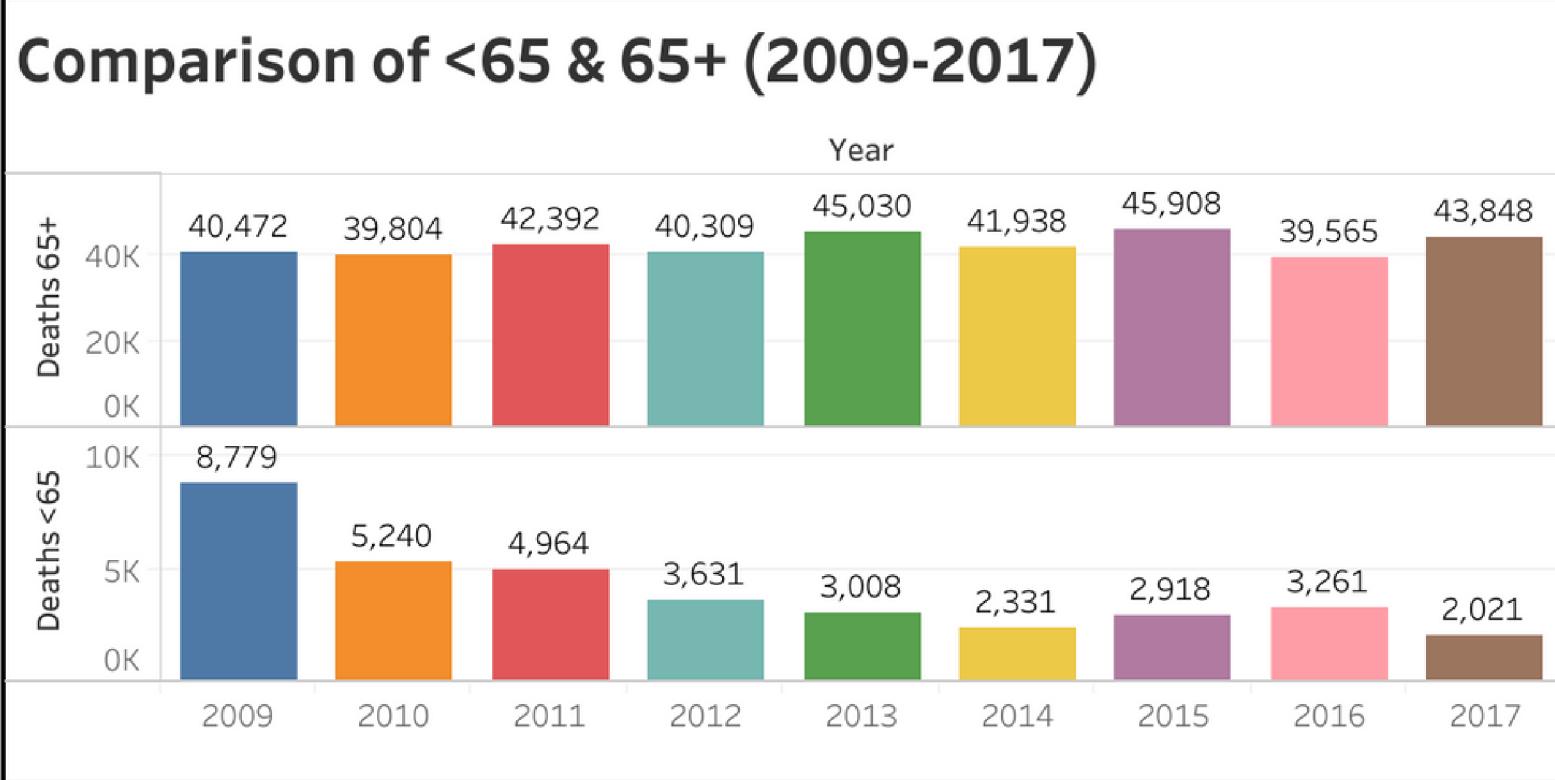


Analysis

(cont.)



DATA SOURCE:
CDC INFLUENZA DEATHS & US CENSUS BUREAU



Recommendations

PREPARING FOR INFLUENZA SEASON IN THE UNITED STATES



01.

Elevated Care for 65+:

Allocate extra healthcare staff to facilitate serving the 65+ age group flu season due to their heightened vulnerability to influenza complications.

02.

Prioritized support:

Ensure staffing resources are directed to areas with significant elderly populations, enhancing response capacity and personalized care for this at-risk demographic.

03.

Strategic state allocation:

To prepare for the influenza season, allocate more medical staff to each state based on the previous year's death count and mortality rate.

03

Rockbuster Stealth



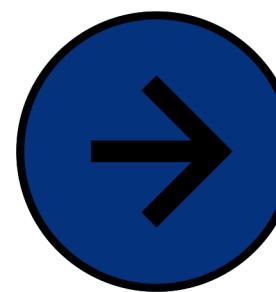
SQL data analysis for the Rockbuster management team to provide strategic insights for their new online video service, enhancing their operational efficiency.

Objective:

Conduct SQL data analysis for the Rockbuster management team to deliver strategic insights for their newly launched online video service, with the aim of optimizing operational efficiency.

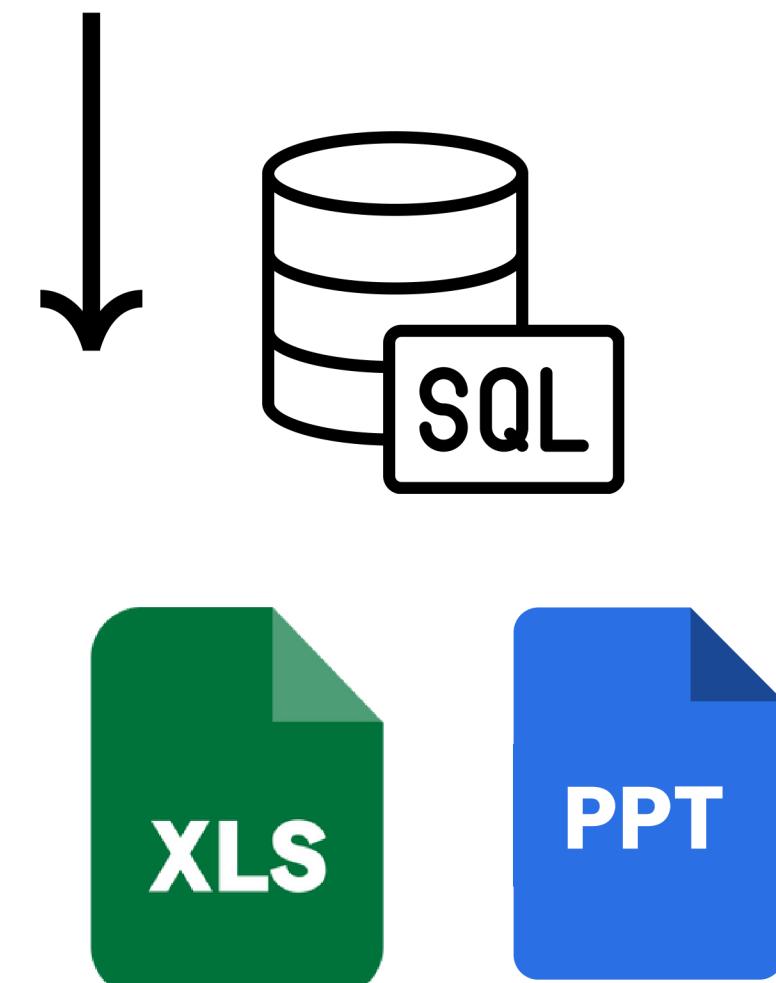
GOALS

Offer data-driven insights and solutions on pertinent aspects to guide the company's launch strategy for online video services in 2020, aiming to improve its current business position.



Skills

- Data cleaning
- Database manipulation
- SQL joins
- CTE's and subqueries
- Visualizations



DATA & LIMITATIONS

- Develop a relational database tailored for a movie rental company, integrating comprehensive data on customers, film inventory, rentals, stores, payments, etc., and compile a data dictionary outlining 15 tables. The dataset encompasses internal records of stores, customers, payments, inventory, films, and associated details.



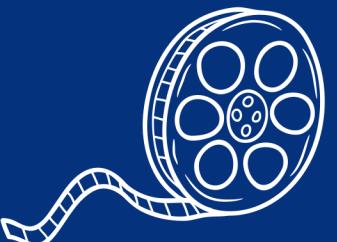
Analysis



599 Customers - 584 Active Customers



20 Genres



1000 movies



Revenue of \$61,312

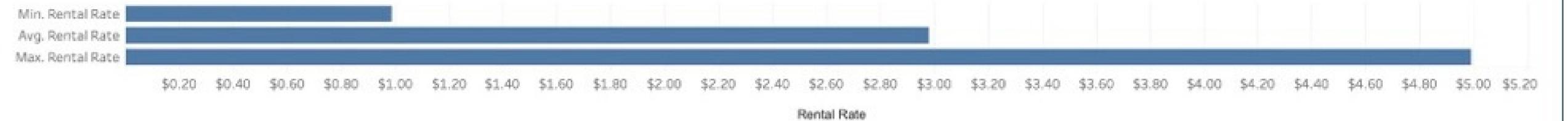


108 countries

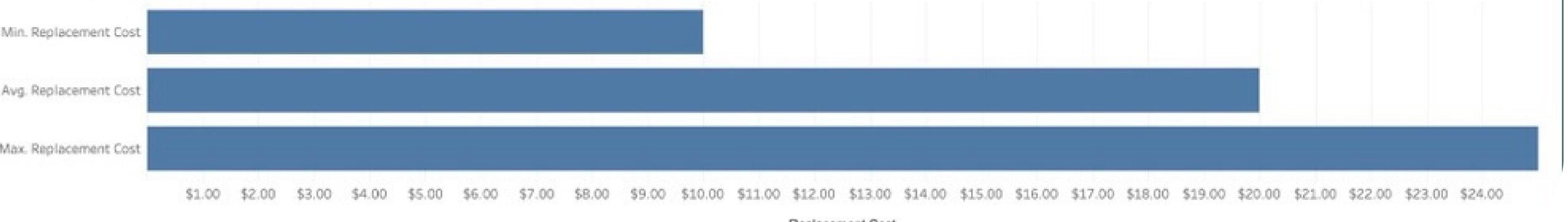
Rental Duration



Rental Rate

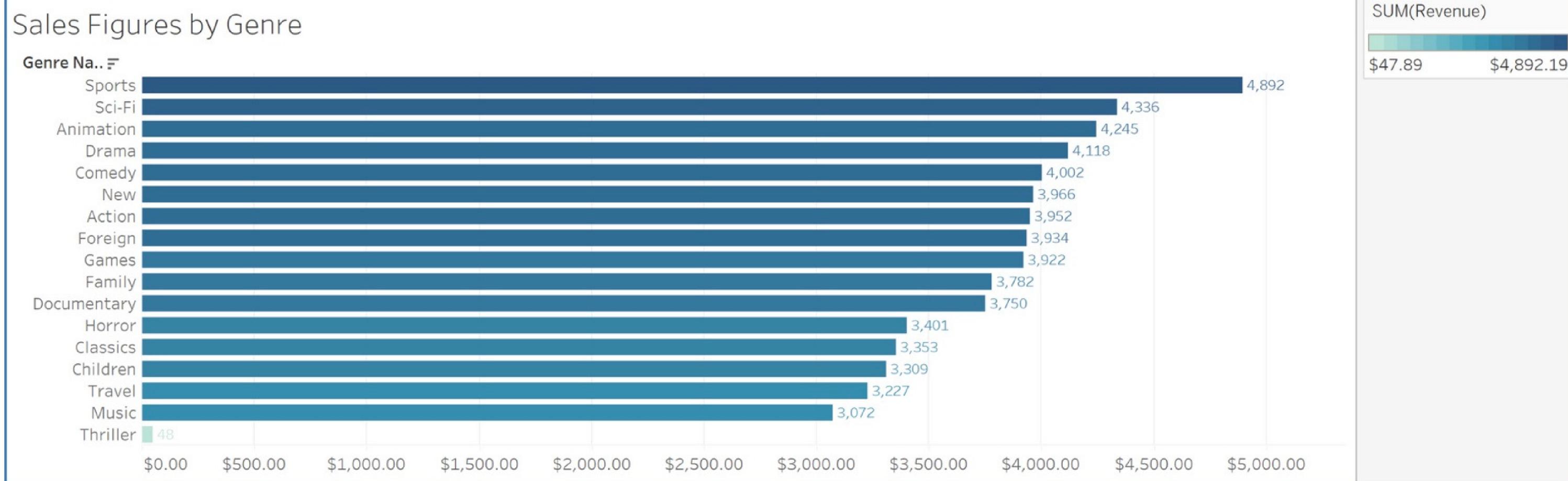


Rental Replacement Costs



Analysis

(cont.)



Recommendations

ROCKBUSTER STEALTH MARKETING STRATEGY



01.

Focus on high quality content:
Prioritize top-performing movies and
discontinue underperforming ones to
streamline the content library,
ensuring a consistent focus on quality
titles that resonate with diverse
customer bases.

02.

Strategic promotion and engagement:
Offer targeted discounts on highest-grossing movies year-round,
especially in key markets like India, China, and the U.S., to
attract new audiences and foster loyalty. Additionally, introduce
personalized discounts upon online
account sign-up to enhance customer
retention.

03.

Tailored customer experience:
Implement a special pricing
structure for High Lifetime
Value Customers (HLVC) and
expand language offerings to
include Spanish, Chinese,
Hindi, and other languages
prevalent in top-performing
markets.

04

Instacart

Python analysis



Consulting services for online grocery retailers, specializing in enhancing marketing strategies through advanced behavioral analysis and precise customer segmentation techniques.

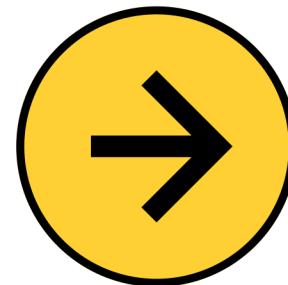


Objective:

Provide consulting services to online grocery retailers, leveraging behavioral analysis and customer segmentation to develop highly targeted marketing strategies for improved efficiency and effectiveness.

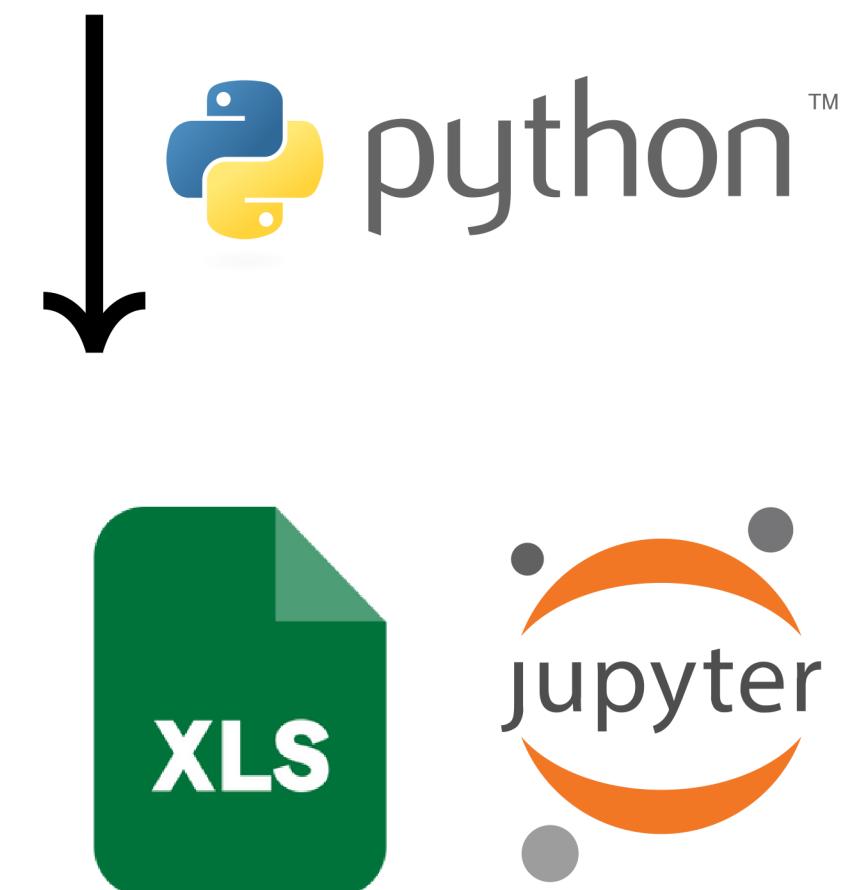
Goals

Conduct an exploratory analysis of customer behavior and sales patterns for an online grocery retailer, extracting insights to improve marketing, segmentation, and customer profiling strategies..



SKILLS

- Data wrangling/merging
- Excel reporting
- Grouping/Aggregating data



...

DATA & LIMITATIONS

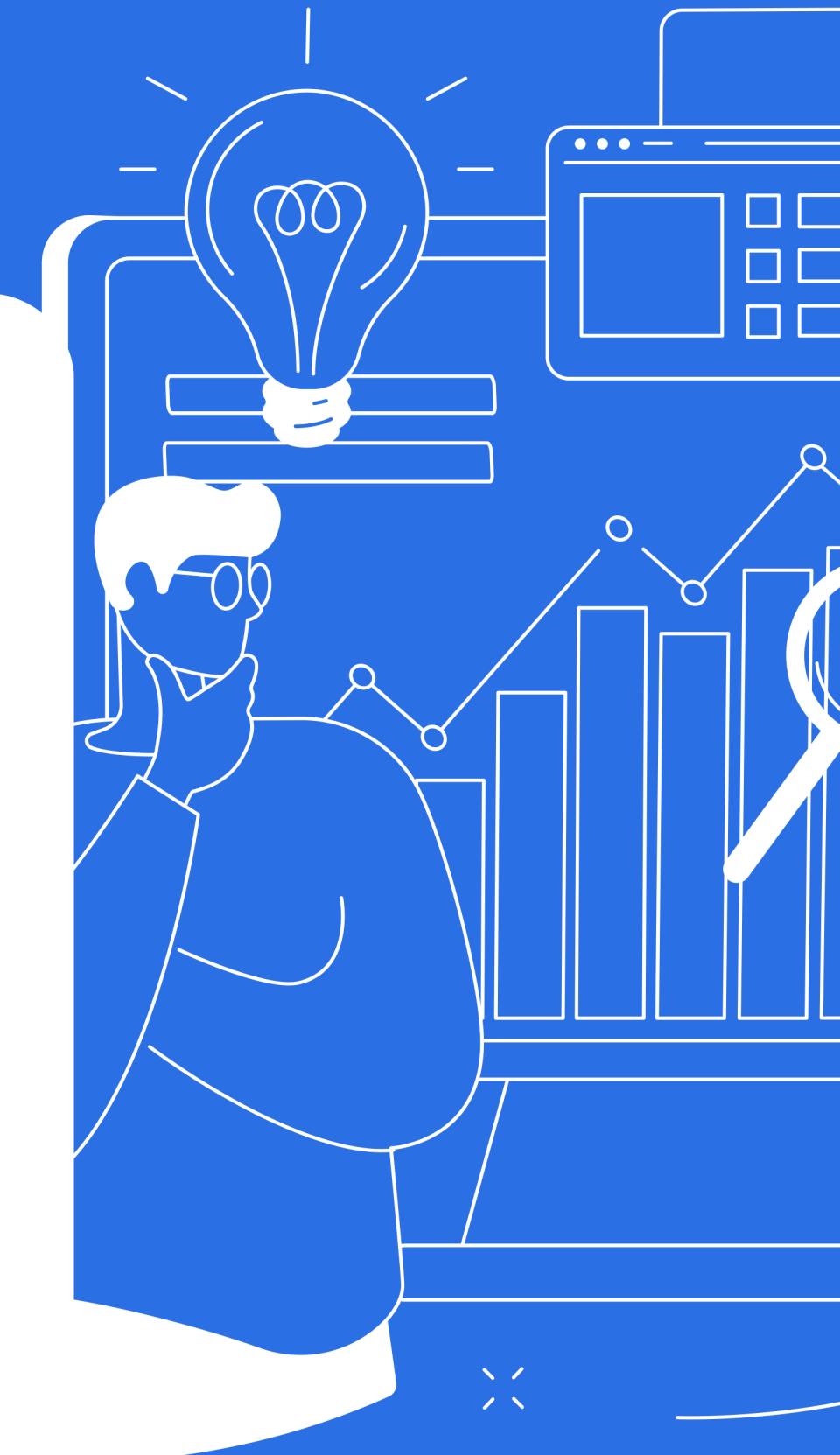
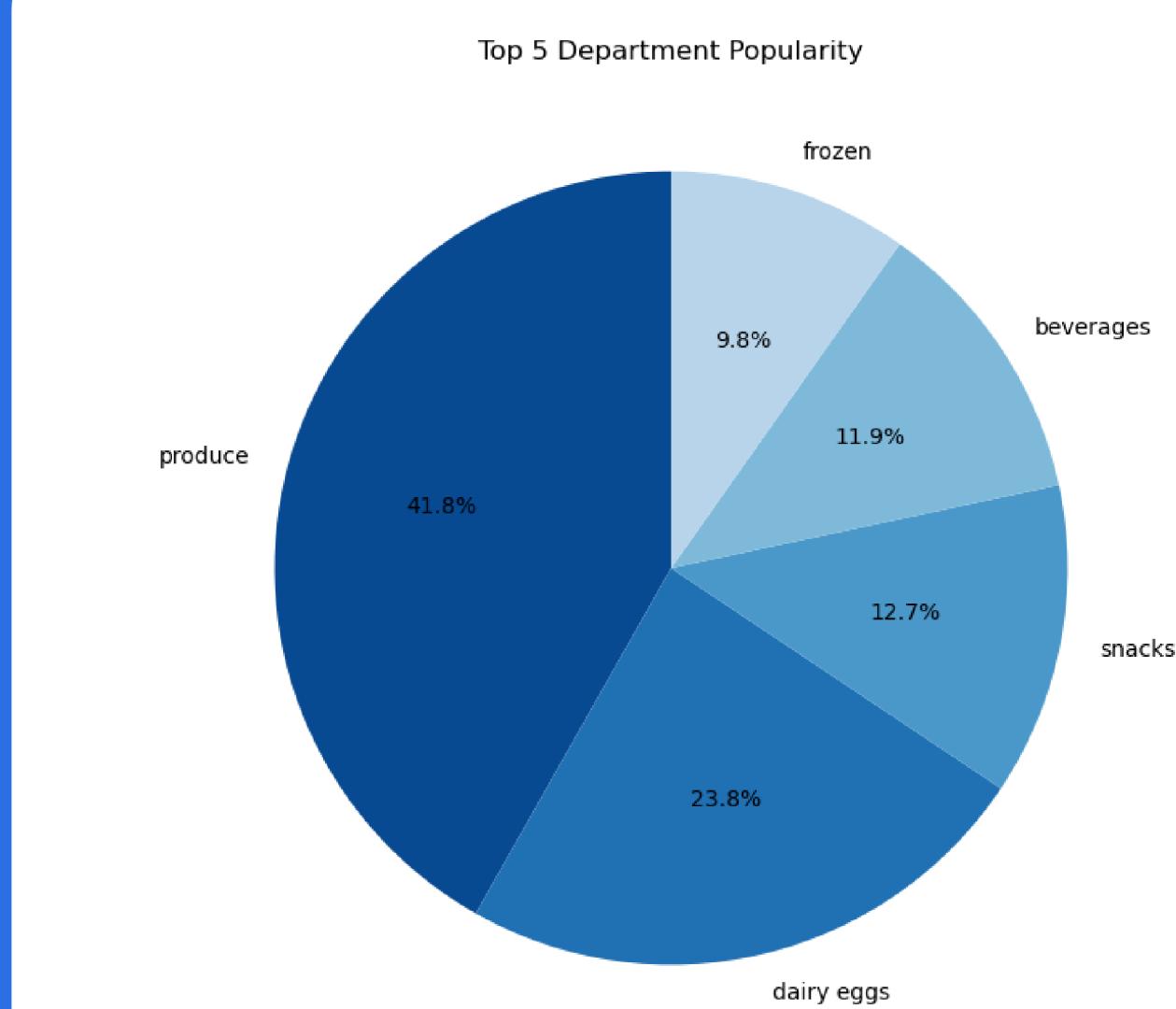
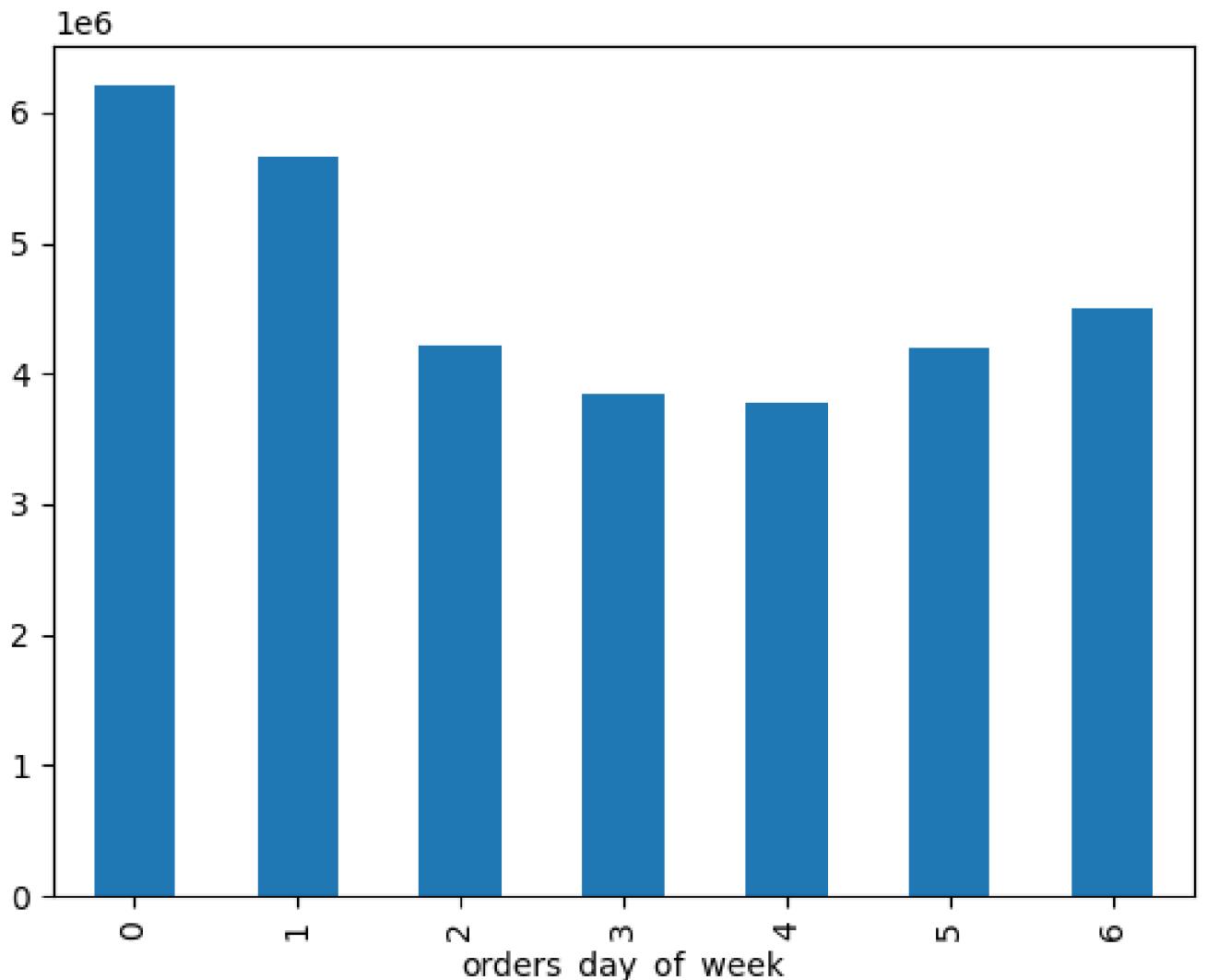
- The dataset contains only records from the year 2017. Skills required include data wrangling and merging, grouping and aggregating data, population flow analysis, and Excel reporting. Customer demographic information is limited to income, age, family size, and marital status. The datasets consist of information on orders, products, and departments obtained from real Instacart data from 2017 on Kaggle, along with fictional customer data provided by Career Foundry.

...

Analysis



Data Source: provided by Career Foundry

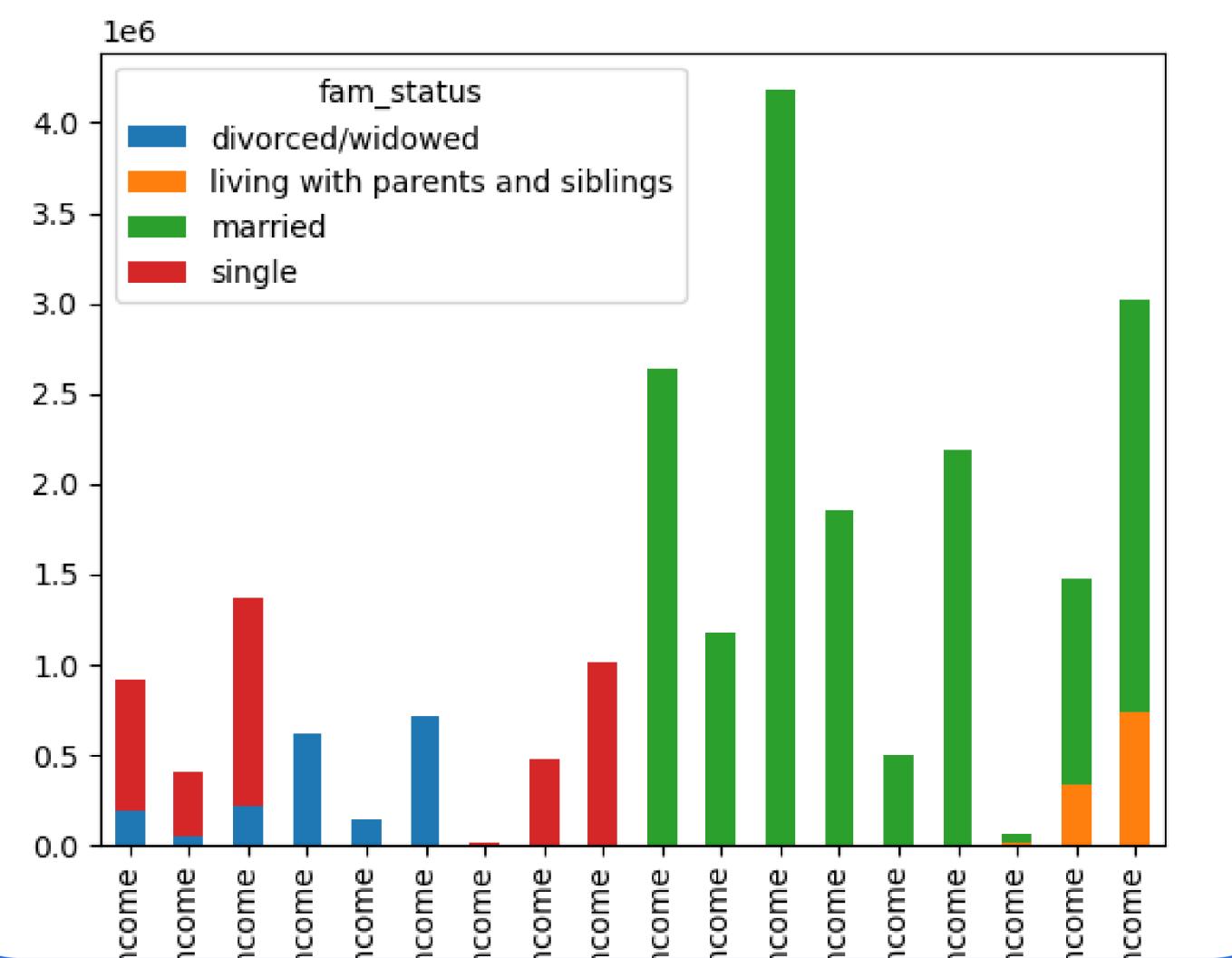


Analysis



Data Source: provided by Career Foundry

Region	High_spender	Low_spender
Midwest	155975	7441350
Northeast	108225	5614511
South	209691	10582194
West	160354	8132559



Recommendations

INSTACART GROCERY BASKET ANALYSIS



01.

The items most commonly requested include fresh ingredients such as fruits, vegetables, dairy, eggs, and meat or seafood, underscoring the importance of consistently stocking a varied and fresh inventory.

02.

While Instacart boasts its largest customer base in the South region, ordering patterns remain consistent across all regions. To achieve a deeper understanding of regional dynamics and consumer behavior, it's imperative to gather population data for more comprehensive insights.

03.

Introduce a customized pricing strategy for High Lifetime Value Customers and target marketing efforts specifically towards married couples, who typically exhibit higher spending behaviors.

05

Pig E. Bank

Big Data
Analysis

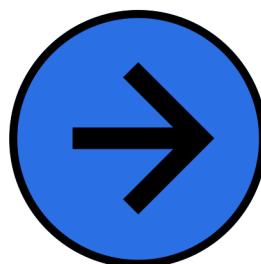
Provide analytical support to the anti-money-laundering compliance division, evaluating client and transaction risks, and improving metric reporting. Develop and enhance models to optimize compliance efficiency.

Objective:
Implementing customer retention strategies by analyzing behavior, enhancing satisfaction levels, and effectively reducing attrition rates

GOALS

Assist the company's anti-money laundering compliance department with data-driven projects to evaluate client and transaction risks, while also generating metric reports.

Additionally, contribute to refining models to enhance the efficiency of the compliance program.



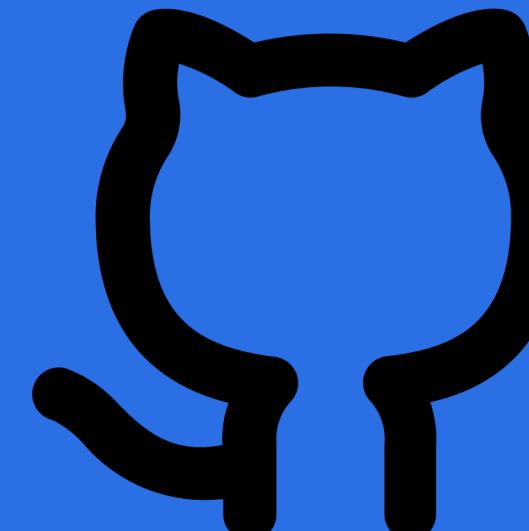
Skills

- Intro to Big Data
- Data ethics
- Datat mining
- Predictive analysis
- Forecasting



DATA & LIMITATIONS

- Data lacks timestamps, making it impossible to determine if there was a trend of customer loss within a defined period or if it's a consistent occurrence.



CONTRIBUTING FACTORS OF CLIENTS LEAVING THE BANK:

Age



Number of
products



Gender



Balance



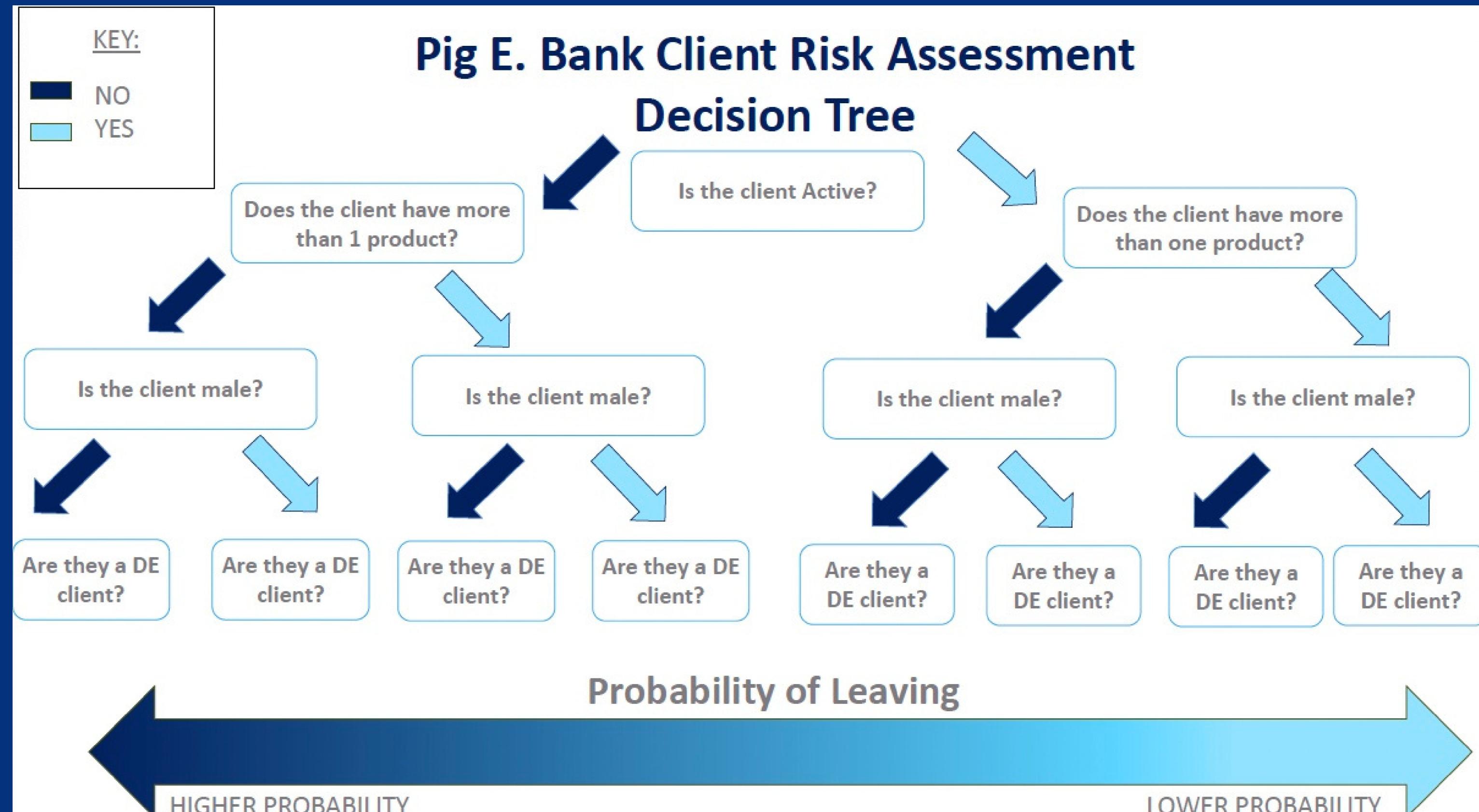
Tenure



Credit
Card

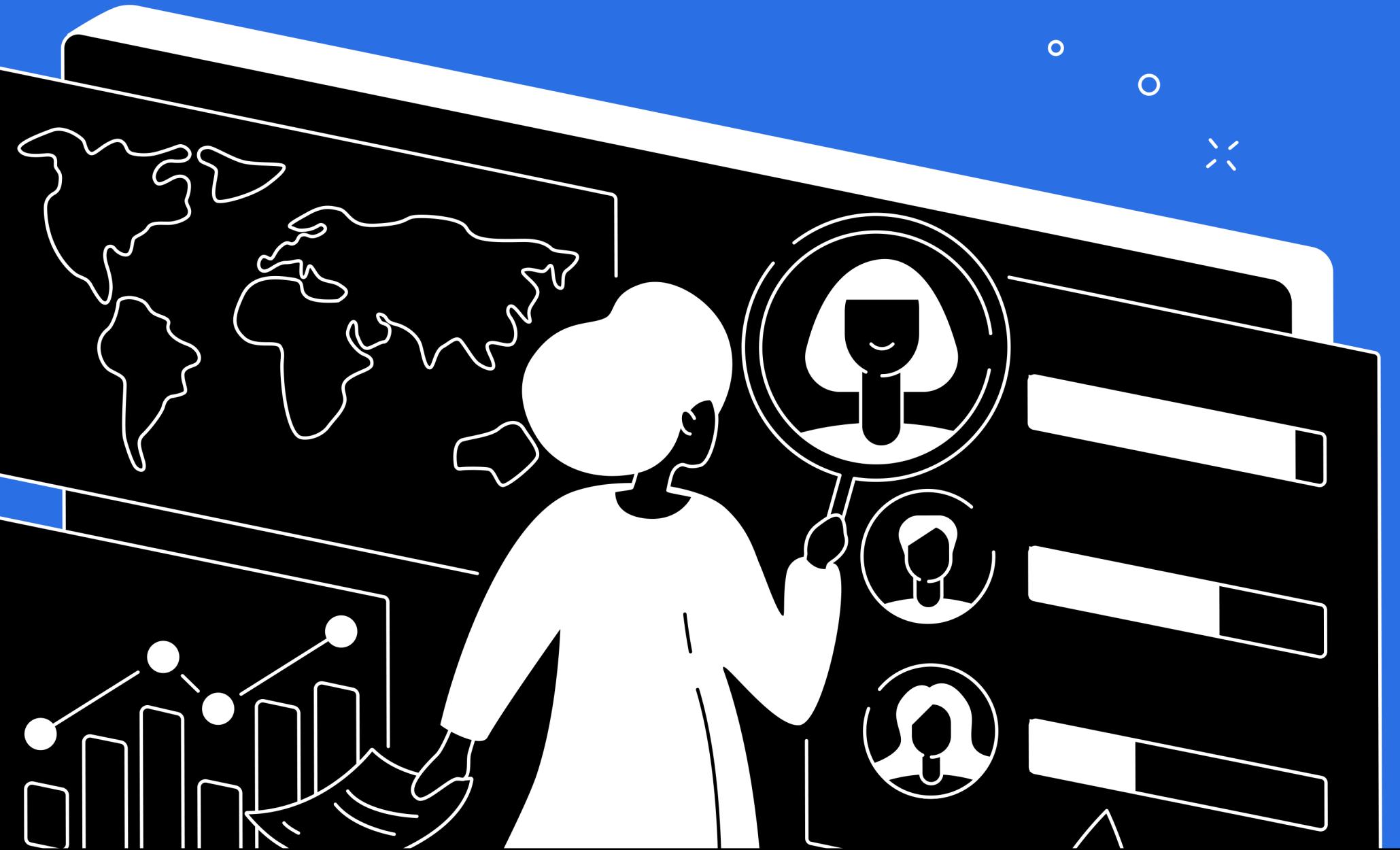


PREDICTIVE ANALYSIS



Recommendations

PIG E. BANK



01.

In order to gain deeper insights into our customers' banking requirements, it's crucial to conduct regular surveys and interviews. This ongoing feedback loop enables us to track customer satisfaction levels and foresee their evolving needs.

02.

Expanding our range of products can enhance customer loyalty, as clients utilizing multiple services are less inclined to switch providers. Consequently, broadening our product portfolio and actively marketing them could aid in minimizing customer turnover.

03.

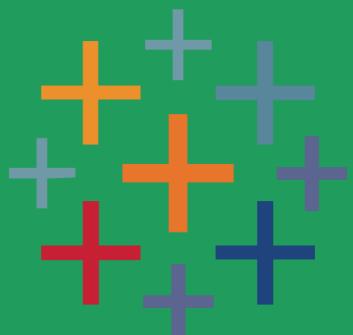
Customers who engage less are more prone to leaving. Hence, it's imperative to incentivize increased activity, such as offering enticing rewards for active participation, to enhance engagement and diminish the probability of customer turnover.

06

Housing Market Predictions

Advanced Regression Techniques

Final Project

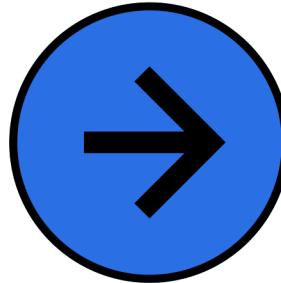


Analyzing and predicting house prices using machine learning techniques.

Objective: This project aims to utilize various features of houses to predict their selling prices accurately. The analysis covers data cleaning, exploratory data analysis (EDA), and the application of machine learning models for regression.

Goals

- Develop a model to predict house prices based on available features.
- Identify the most influential features affecting house prices.
- Provide actionable insights for real estate stakeholders.



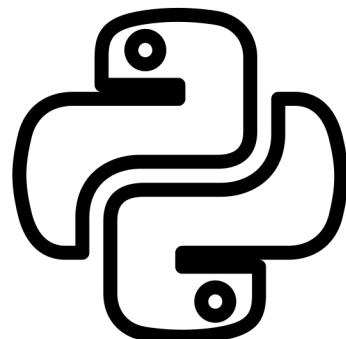
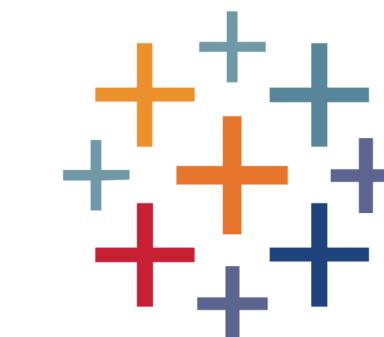
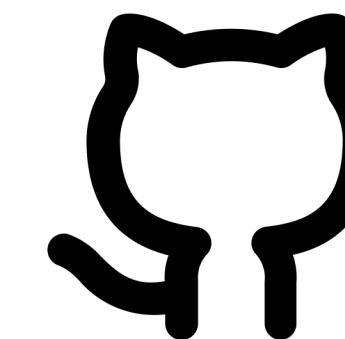
SKILLS

- Data Cleaning: Handling missing values, dealing with outliers, and encoding categorical variables.
- Exploratory Data Analysis (EDA): Visualizing data distributions, correlations, and patterns.
- Machine Learning: Implementing regression models such as Linear Regression, Random Forest, and Gradient Boosting.
- Data Visualization: Using tools like Matplotlib, Seaborn, and Tableau to create informative visualizations.



DATA & LIMITATIONS

- Missing Values: Some columns (e.g., LotFrontage, Alley) have significant missing data that needed imputation or exclusion.
- Outliers: Certain data points exhibited extreme values which could skew the analysis.
- Categorical Variables: Required appropriate encoding to be used in machine learning models (e.g., One-Hot Encoding).



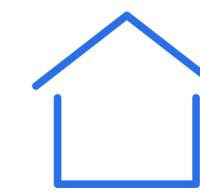
Analysis



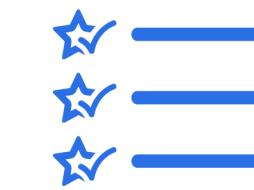
1460 HOUSES



Avg Sale Price:
\$180,921

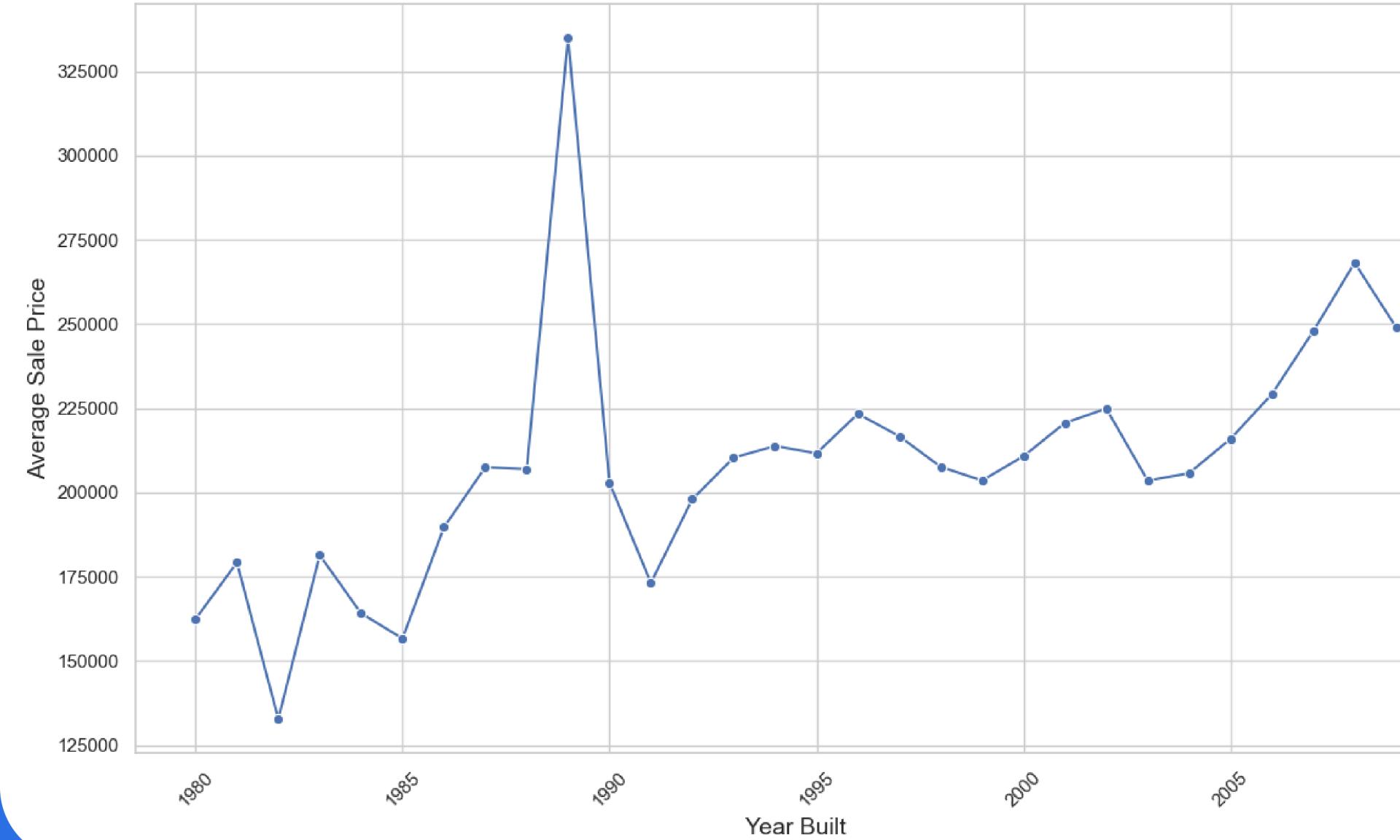


8 unique styles - 25
unique neighborhoods

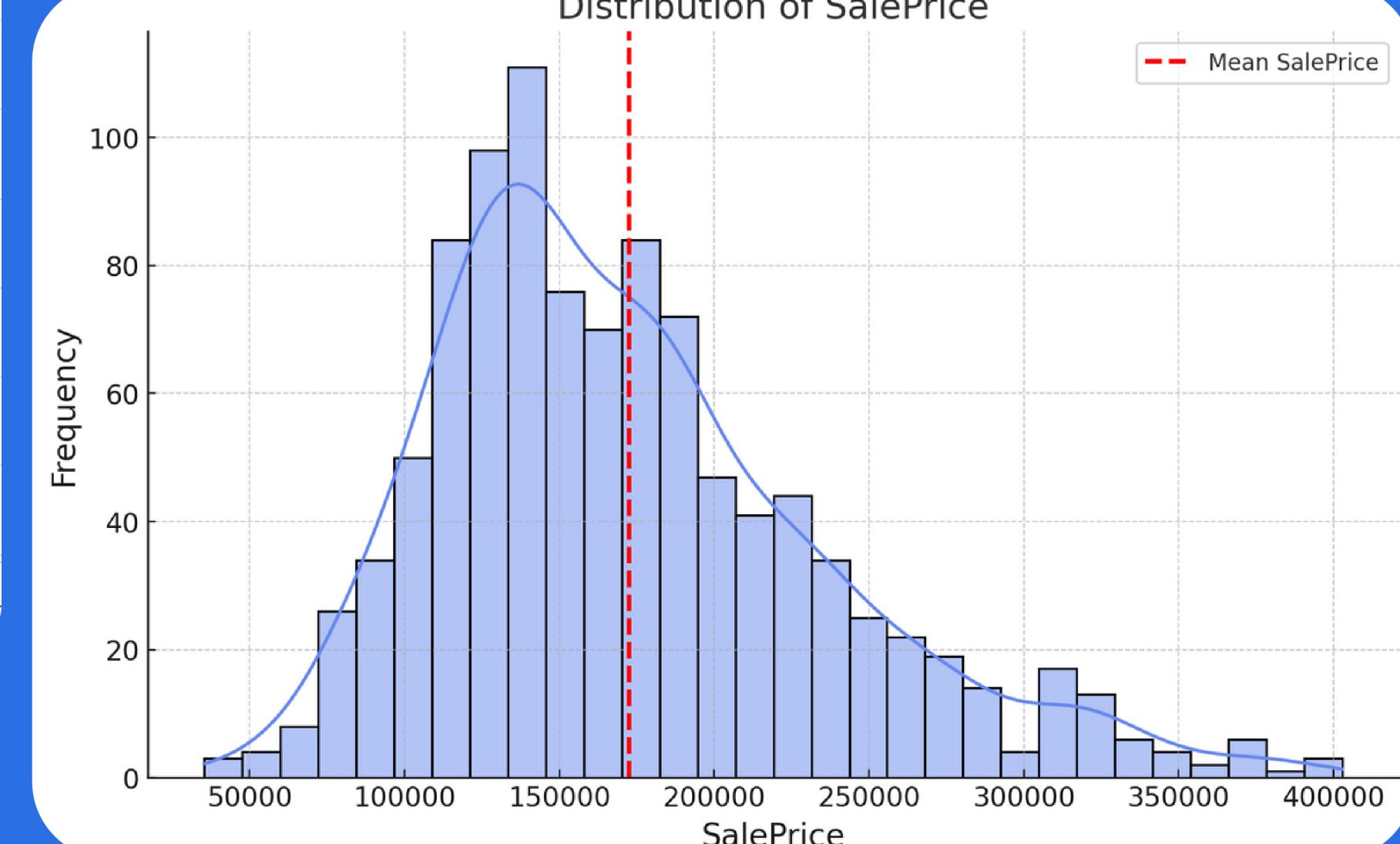


81 FEATURES

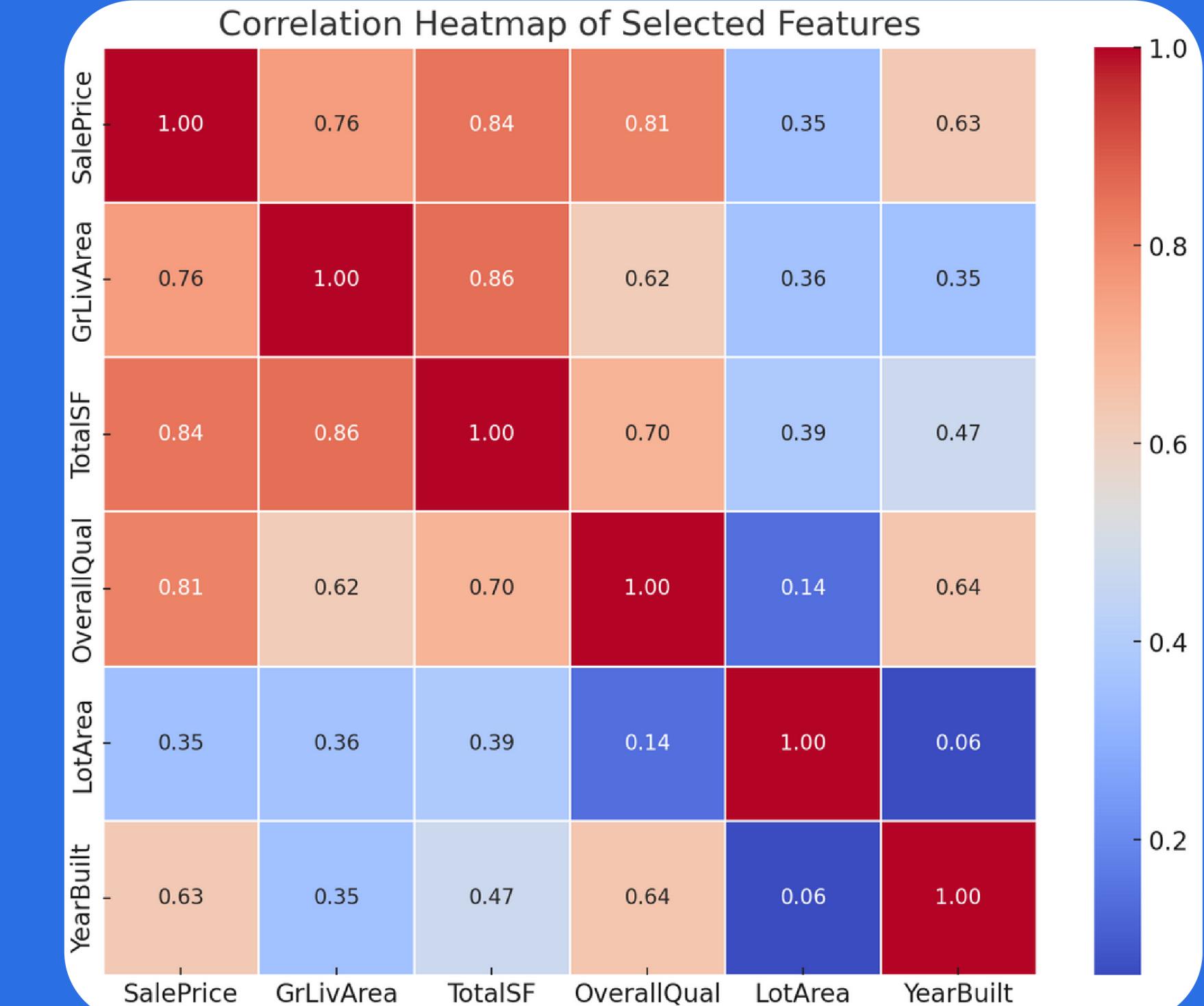
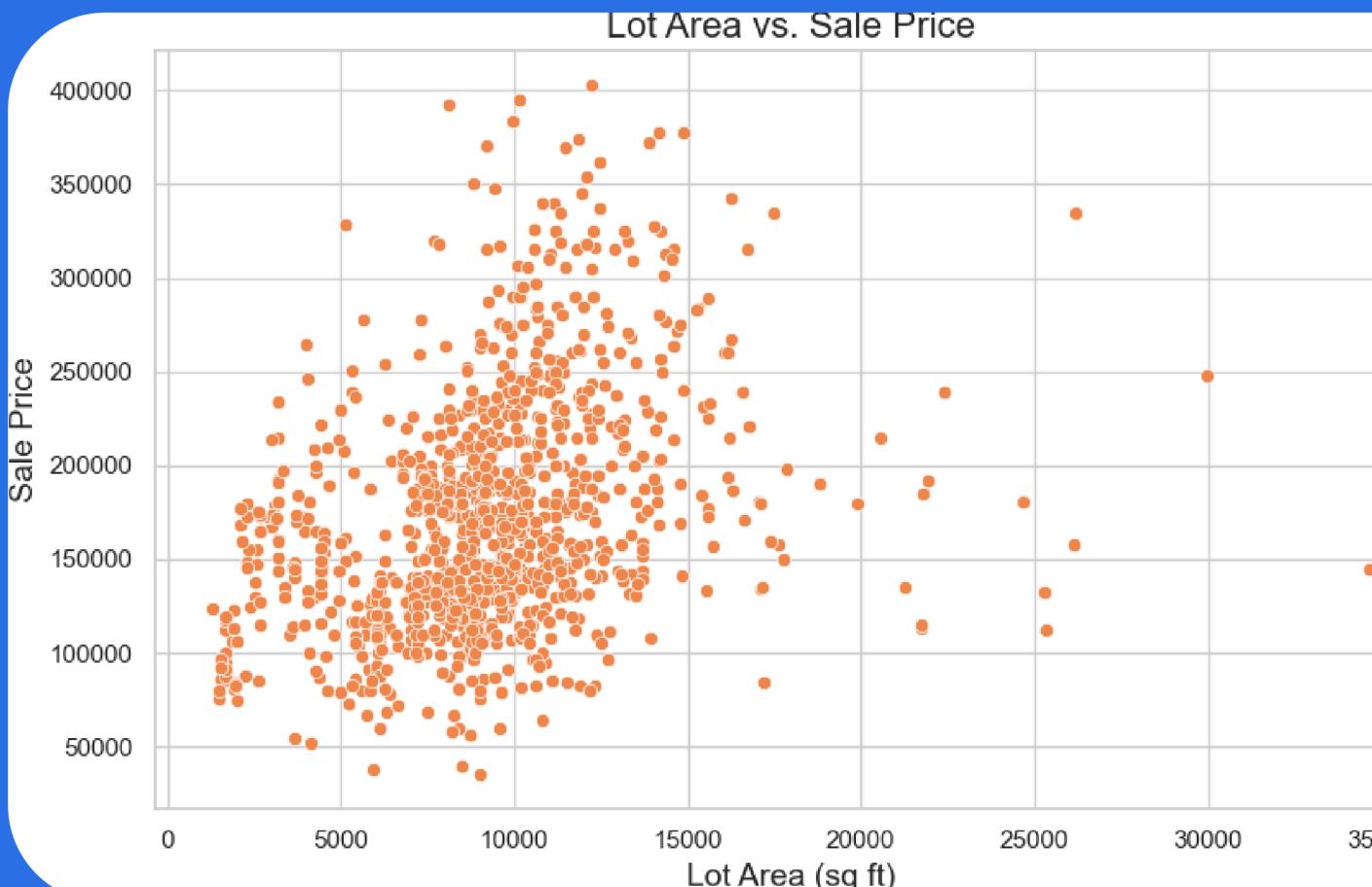
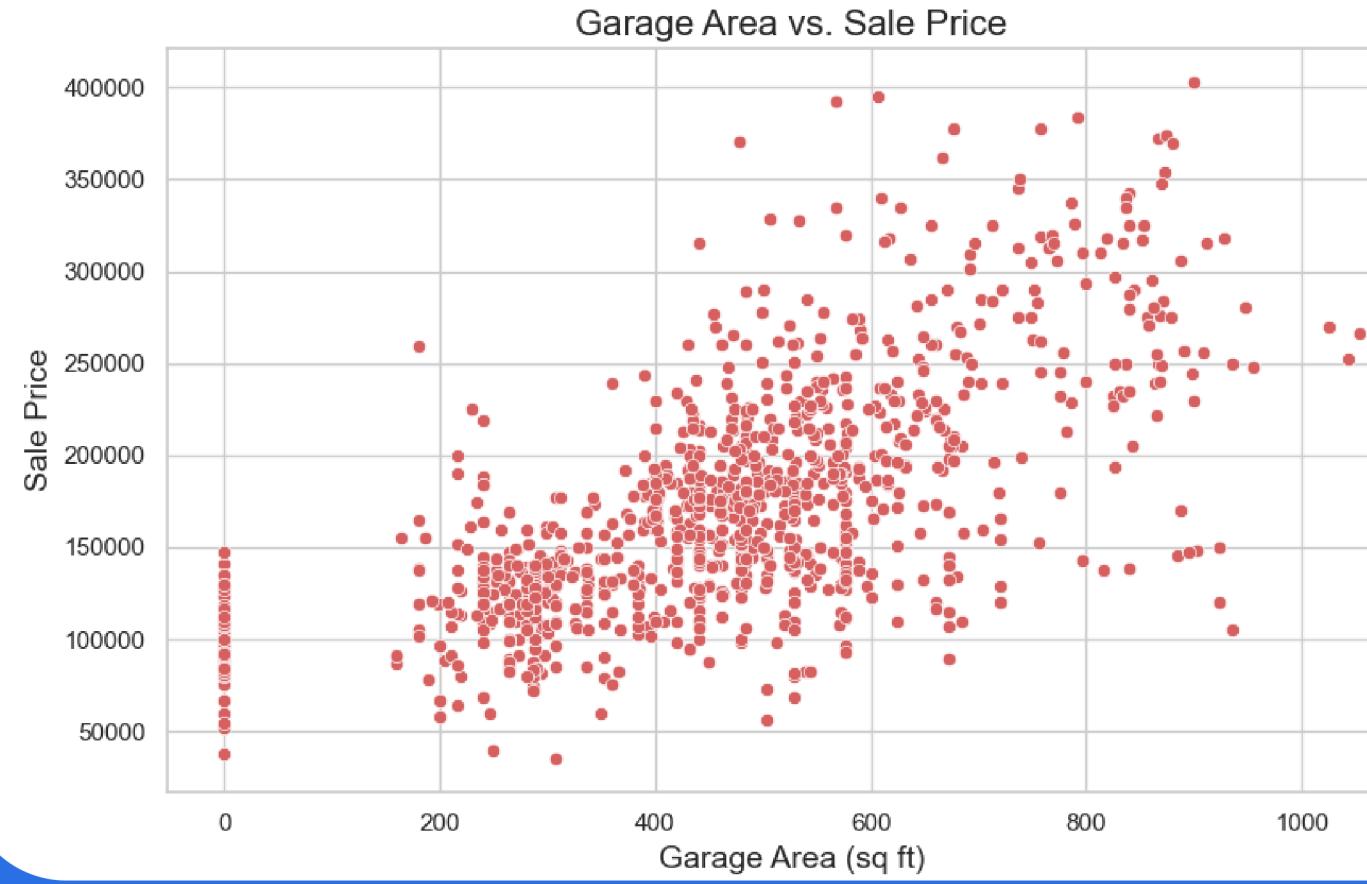
Average Sale Price by Year Built (1980-2010)



Distribution of SalePrice



Analysis



Recommendations

HOUSING REGRESSION MODEL

01.

MODEL PERFORMANCE:

Experiment with advanced machine learning techniques and hyperparameter tuning to enhance model performance. Regularly update the model with new data to maintain its accuracy and relevance, ensuring it adapts to market changes.

02.

PRIORITIZE KEY PREDICTORS:

Concentrate efforts on features with the highest impact on house prices, such as Overall Quality, Living Area, and Basement Area. These predictors should be the focus of detailed analysis and feature engineering to achieve more accurate and reliable price predictions.

03.

INCORPORATE NEIGHBORHOOD EFFECTS:

Take neighborhood and other location-based features into account when setting house prices or making real estate investments. Since these factors significantly influence house prices, incorporating them into predictive models and pricing strategies can provide more localized and accurate insights.



Let's connect!

SARAH EDENS
DATA ANALYST

[LinkedIn profile](#)

[GitHub profile](#)

[Final Project](#)

