Zosya Trimbacher

Data Analyst

z.trimbacher@gmail.com

LinkedIn

GitHub

Tableau

About me

Hi, I’m Zosya Trimbacher — a data-driven problem solver with a background in operations and executive support in fast-paced tech environments.

After nearly five years at Delivery Hero, where I worked closely with global leadership and helped streamline internal processes, I decided to pivot toward data analytics to deepen my ability to uncover insights and support smarter decision-making.

Since then, I’ve developed a strong foundation in tools like SQL, Tableau, Excel, and Python, and I’ve built project work focused on operational performance, reporting, and customer analysis.

My goal is to apply both my operational experience and analytical skills in data-focused roles.

Projects Overview and Tools

Influenza Season

To help U.S. healthcare providers plan staffing for flu season by identifying when and where patient volumes spike.

Data cleaning, Descriptive statistics, Pattern recognition, Strategic recommendations.

Rockbuster Stealth

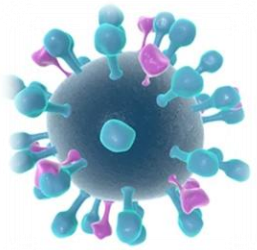
To support Rockbuster's shift from DVD rental to online streaming by identifying top markets, valuable customers, and profitable genres.

Descriptive statistics, Revenue analysis, Customer profiling, Strategic targeting.

Instacart

To help Instacart improve customer segmentation and marketing by analyzing order patterns, spending behavior, and product preferences.

Data merging, Wrangling, Profiling, Descriptive analytics.



Influenza Staffing Analysis

A U.S. based medical staffing agency used data analysis to anticipate demand during flu season. The goal was to help allocate limited resources, nurses, doctors, and physician assistants across all 50 states to minimize under and overstaffing, especially in high-need areas.

OBJECTIVES

- Identify when and where influenza cases peak across the U.S.
- Classify states by vulnerability to flu complications.
- Recommend a data-driven staffing plan to optimize resource allocation.

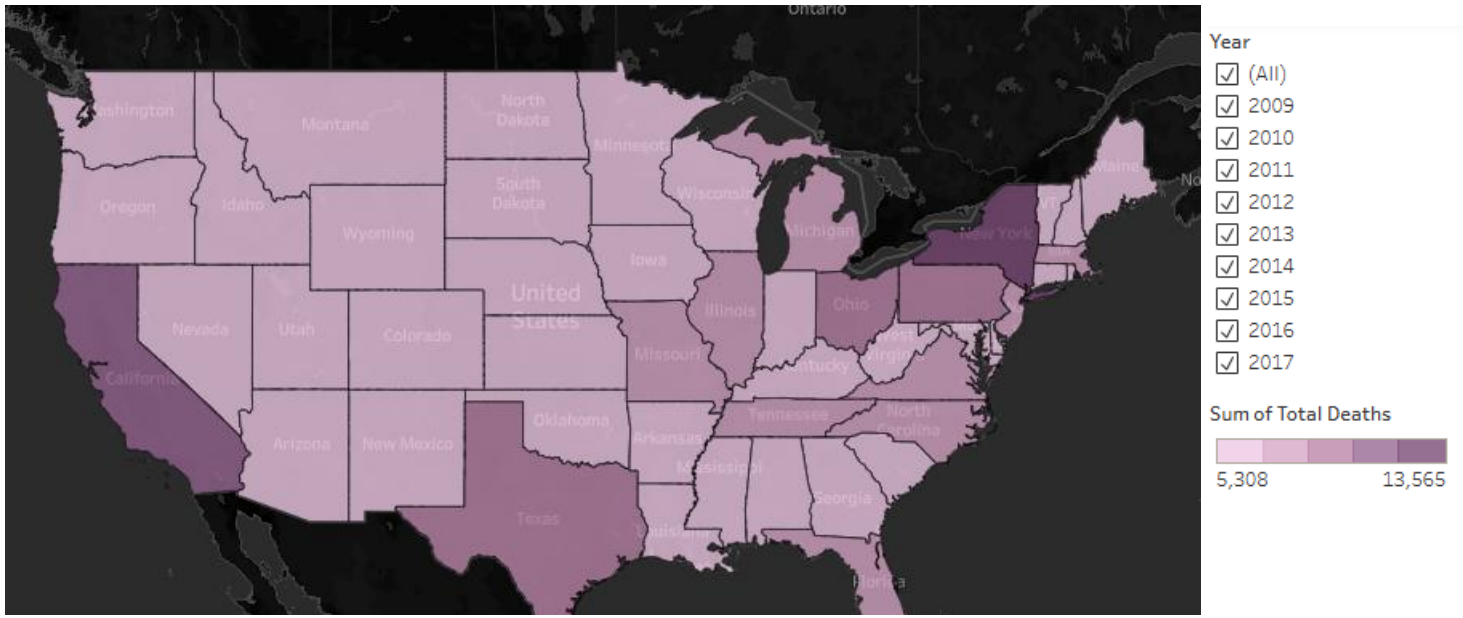
METHODS

- Cleaned and merged CDC influenza death data with U.S. Census demographics.
- Conducted exploratory data analysis on age groups and mortality patterns.
- Formulated and tested a hypothesis (65+ more likely to die from flu) using a one-tailed t-test.
- Measured correlation between elderly population and flu deaths.
- Categorized states into low, medium, and high need tiers based on vulnerability.
- Created spatial and temporal visualizations in Tableau for stakeholder use.

DATA

Sourced from CDC Influenza Deaths and U.S. Census Population Data. Link to Tableau story is here.

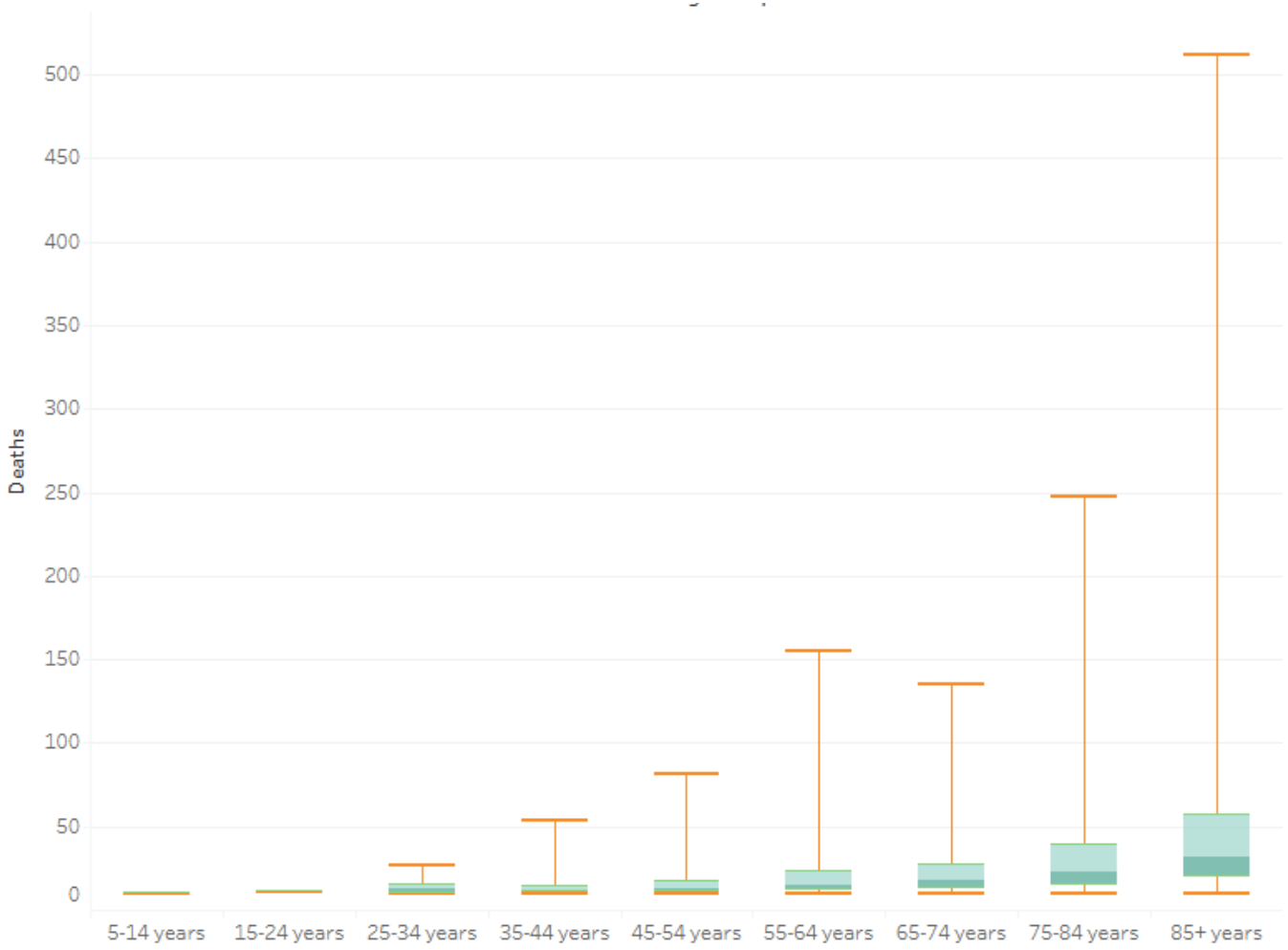
Total Influenza deaths in the United States between 2009 - 2017



This map highlights the states with the greatest healthcare demand.

- **California, New York, and Texas** show the highest flu death rates from 2009–2017.
- These states may face greater pressure on healthcare systems during flu season.
- Understanding geographic hotspots helps prioritize where staff should be sent first.
- By adjusting for population size, we can spot real demand, not just big states.

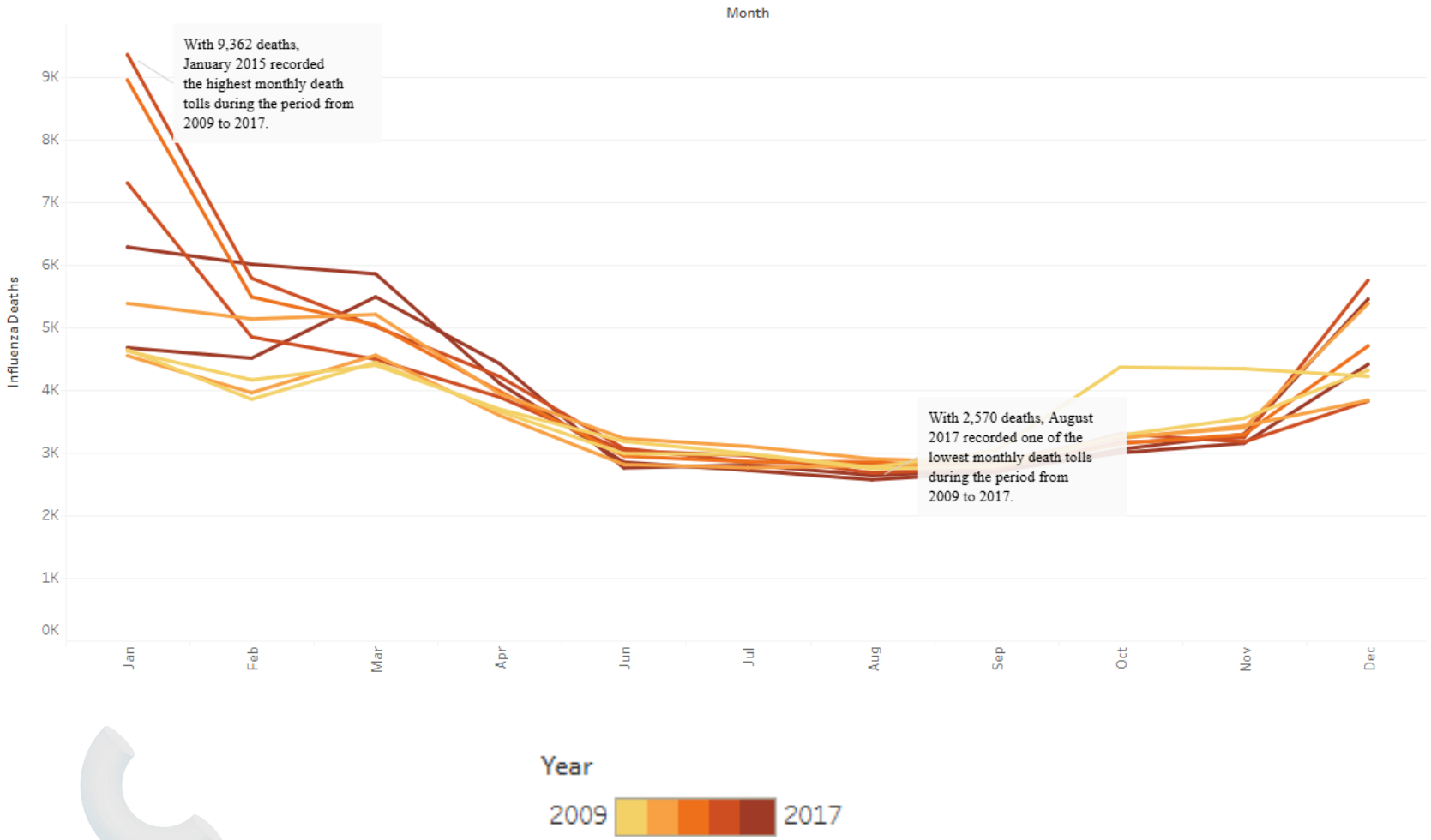
United States Death by Age Groups (2009 – 2017)



This box plot highlights the age groups most at risk.

- Deaths from influenza increase significantly in older age groups.
- The 65–74, 75–84, and 85+ brackets show the widest spread and highest death counts.
- Results vary across states, indicating regional differences in elderly vulnerability.
- Staffing plans should prioritize areas with large senior populations.

Influenza Activity by month (2009 – 2017)



This line chart reveals when flu hits hardest.

- Flu deaths peak between December and February, with January often the worst month.
- Activity starts rising in October, drops off by April, and stays low from May to September.
- These trends are key for staffing — hospitals should be fully prepared from November to March.
- Planning around this pattern ensures support arrives before the seasonal spike.

Recommendations

- Prioritize staffing in California, New York, and Texas during peak months (Dec–Feb).
- Add staff in states with large senior populations, who face the highest risk.
- Plan ahead using seasonal trends to avoid staffing gaps during flu surges.
- Stay flexible and adjust coverage if new hotspots emerge.

Influenza Season (Tableau + stats project)

Project Challenges & Reflections

Challenge

- Missing values in CDC data.
- Dataset tracked only deaths, not hospitalizations.
- Seasonal spikes varied across states.

Solution

- Combined CDC data with Census demographics.
- Ran t-test to confirm elderly risk.
- Grouped states into tiers and visualized in Tableau.

Reflection

- Clear seasonal patterns supported staffing plan.
- Limitation: deaths-only data underestimated demand.
- Next time: add hospitalization or Influenza like illness visit data.

Rockbuster Stealth Analysis

Rockbuster Stealth is transitioning from DVD rentals to a global streaming platform. This analysis supports their 2020 strategy by identifying high-performing markets, top revenue drivers, and loyal customer segments using SQL.

OBJECTIVES

- Identify top and bottom revenue-generating movies.
- Locate high-value customers and their countries.
- Analyze average rental durations and genre performance.
- Support strategic decisions on market targeting and customer retention.

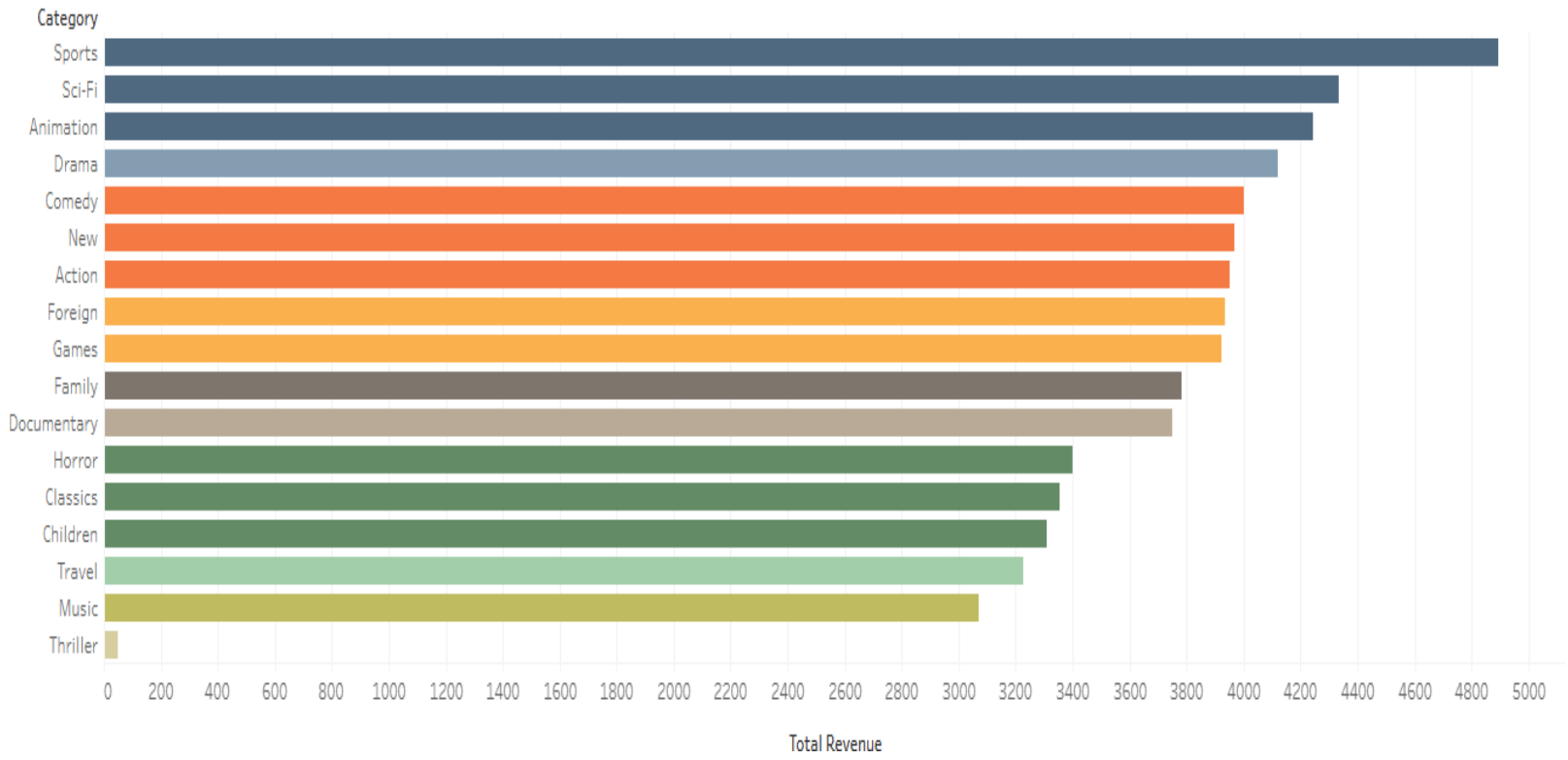
METHODS

- Used SQL queries to extract and analyze customer, payment, and rental data.
- Created CTEs and subqueries to explore patterns in revenue and customer behavior.
- Conducted aggregation, filtering, and joins to answer business questions.
- Visualized results in Excel and PowerPoint for business stakeholders.

DATA

Sourced from Rockbuster's internal PostgreSQL database. Tables included: films, categories, customers, payments, rentals, countries, cities. Link to my GitHub is here.

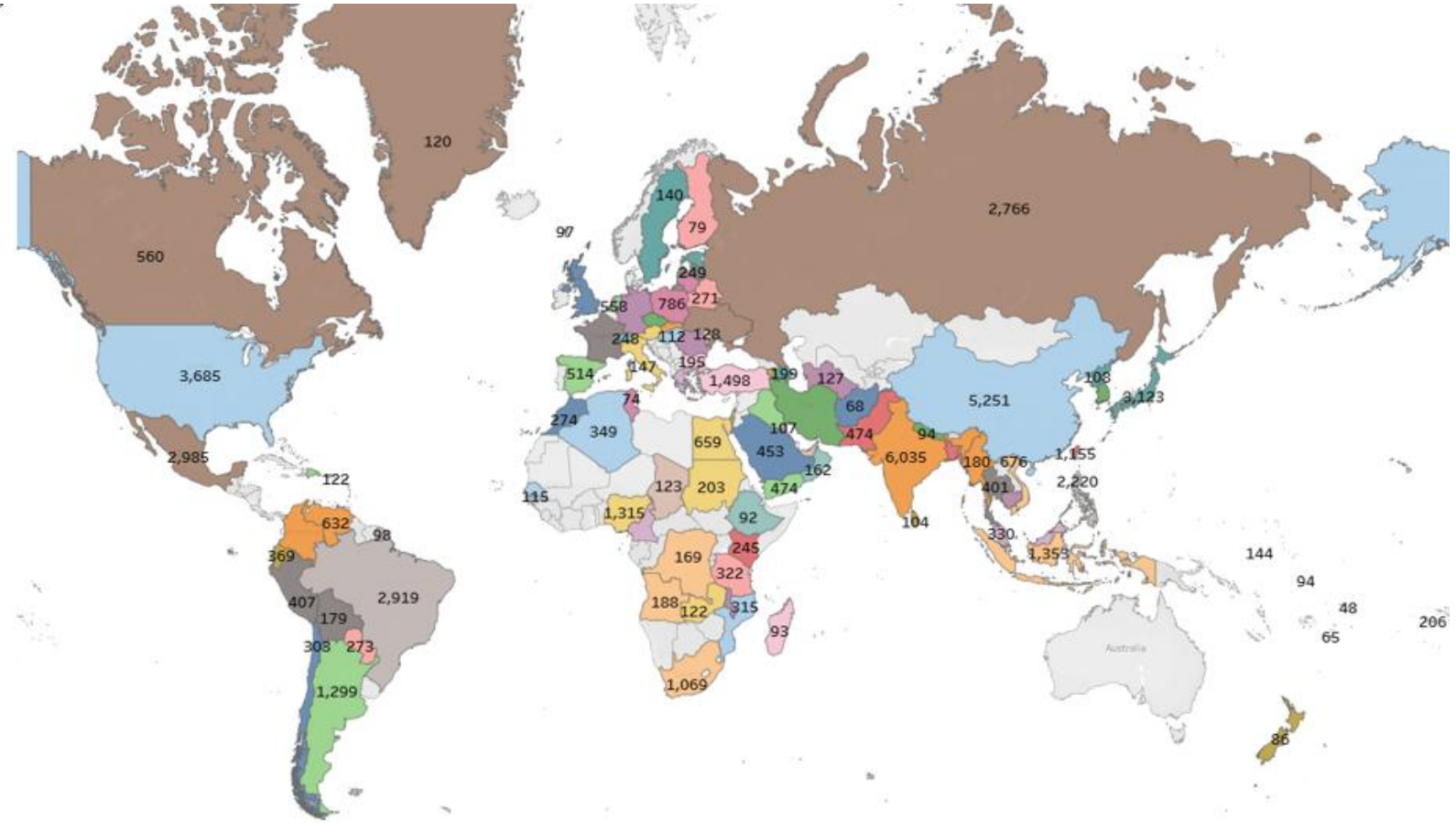
Movie categories bringing in most revenue



This chart shows which movie genres drive the most revenue.

- Sports and Sci-Fi are the top-earning categories, outperforming all others.
- Other strong contributors include Animation, Drama, and Comedy.
- This insight supports smarter catalog strategy as the company transitions to streaming.

Rockbuster’s customers location & revenue generated by each country.

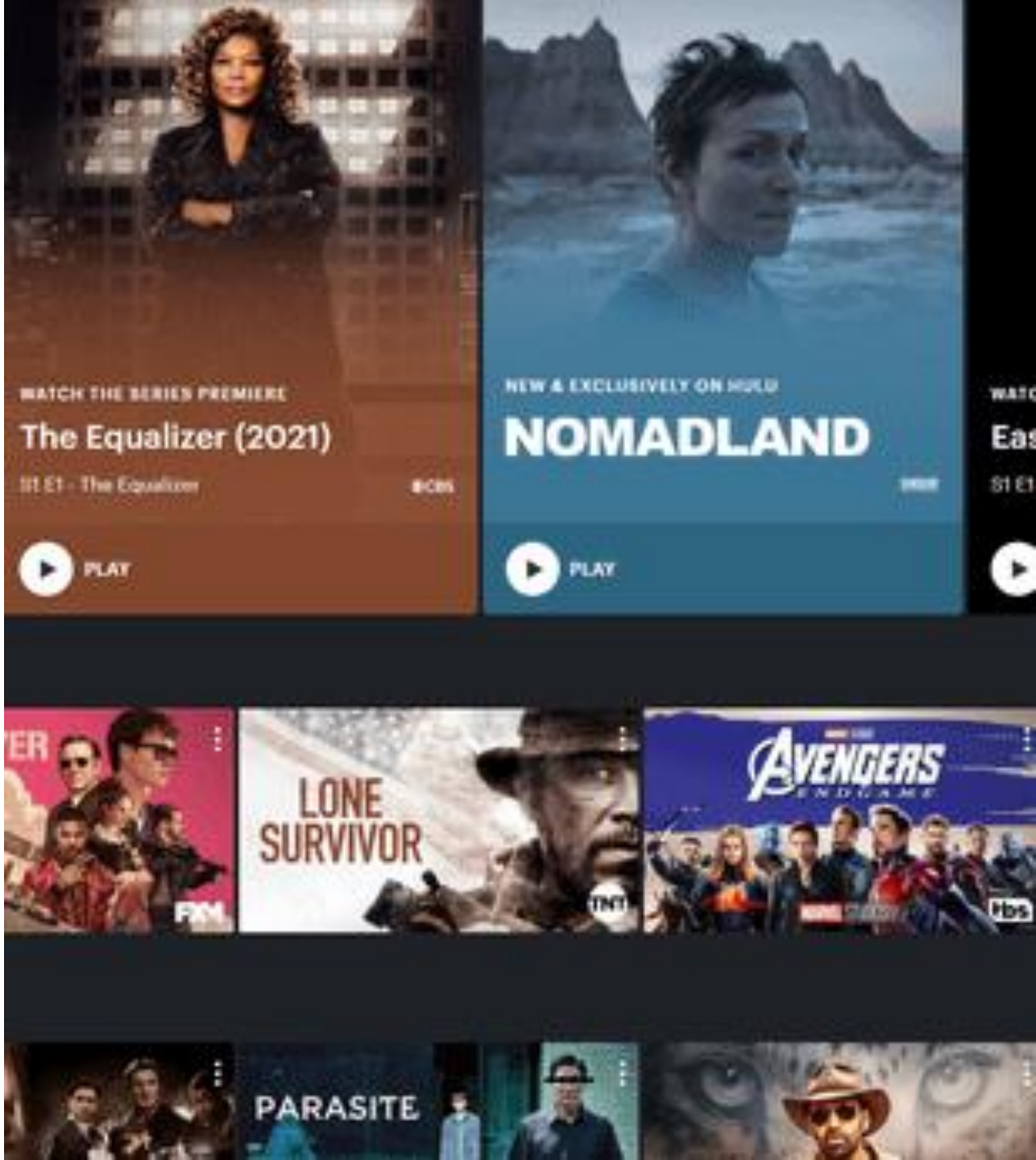


This map highlights Rockbuster’s strongest revenue regions.

- India ($6,035), China ($5,251), and the U.S. ($3,685) lead in total revenue.
- Other key markets include Japan, Mexico, and Brazil.
- These regions represent the best opportunities for an initial streaming rollout.
- Global sales insights help prioritize countries for future investment and expansion

Recommendations

- Target top markets: Launch first in India, China, and the U.S., where spending and customer value are highest.
- Promote high-performing genres like Sports, Sci-Fi, and Animation to maximize engagement.
- Reward loyalty: Introduce incentives for top-paying customers to drive retention.
- Refine catalog: Phase out low-revenue titles and focus marketing on profitable content.
- Monitor regional behavior to fine-tune pricing and tailor the streaming experience.
- Start with pilot cities, expanding based on revenue potential and customer density.



Rockbuster Stealth (SQL project)

Project Challenges & Reflections

Challenge

- Dataset spread across multiple tables.
- Joins sometimes produced overly broad results.
- Hard to extract insights cleanly.

Solution

- Used CTEs and subqueries to simplify queries.
- Applied filtering and aggregation for focus.
- Targeted high-value customers, markets, and genres.

Reflection

- Structured SQL approach worked well for clarity and reusability.
- NULLs and inconsistent rental durations slowed progress.
- Next time: build ERD upfront + optimize queries earlier.

Instacart Analysis

U.S. online grocery company used data analytics to uncover customer behavior patterns, identify peak purchasing times, and create marketing strategies tailored to different user segments.

OBJECTIVES

- Identify the busiest days and times for grocery orders to inform ad scheduling.
- Understand when customers spend the most to guide product promotions.
- Determine the most frequently purchased product types.
- Segment customers by behavior and demographics to support targeted marketing.

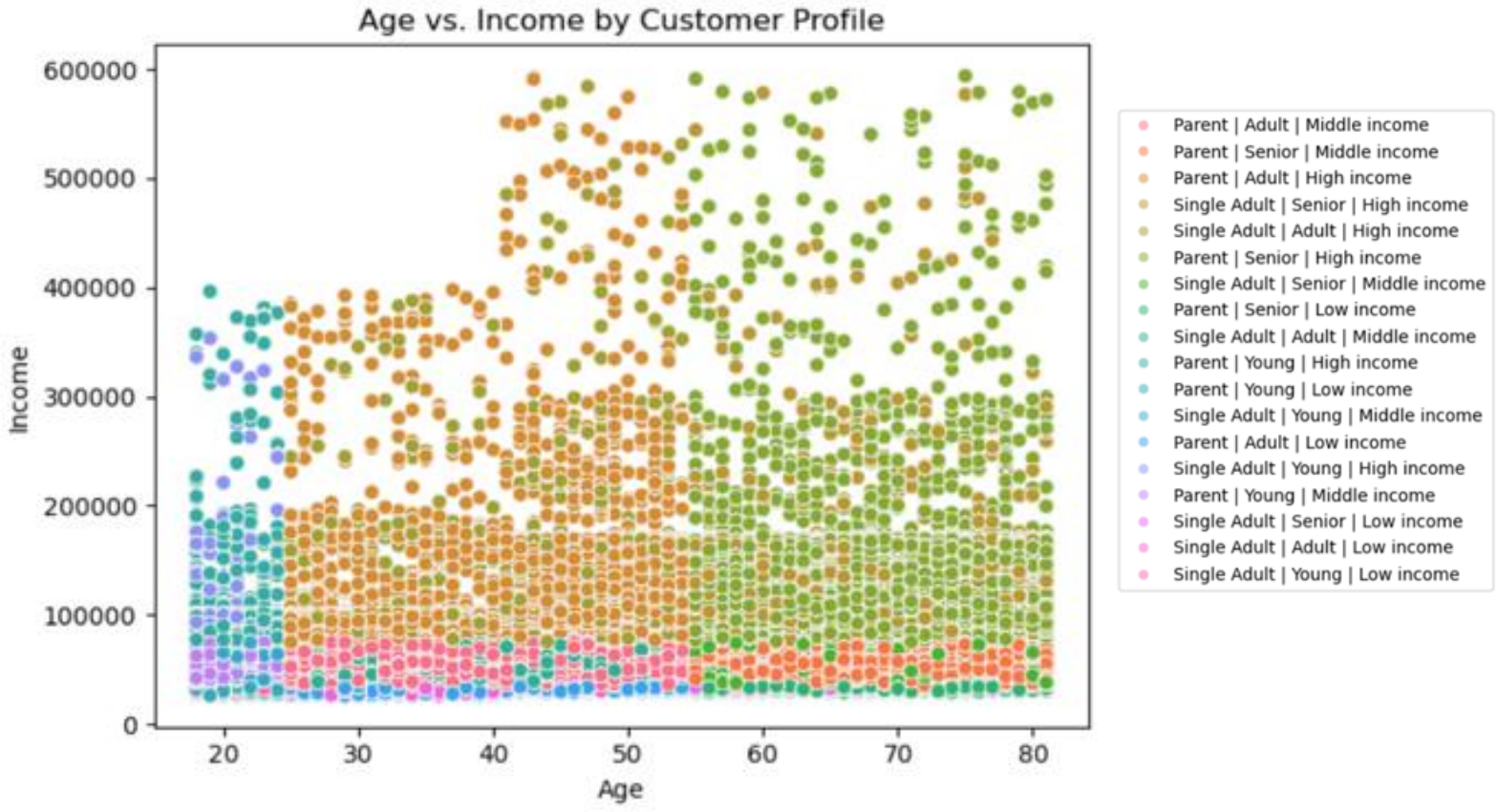
METHODS

- Imported and cleaned Instacart dataset using pandas.
- Conducted consistency checks (duplicates, missing values, mixed types).
- Merged and streamlined data for analysis.
- Created new columns using conditional logic, functions, and flags (e.g., loyalty, spend).
- Segmented customers by profile (e.g. single adult, parent, low spender).
- Aggregated key metrics by customer groups, time, and product categories.
- Built visualizations (bar, line, histogram, scatter) to highlight trends in orders, spend, and product demand.
- Analyzed regional behaviors and generated insights using Python.

DATA

Sourced from the Instacart Online Grocery Shopping Dataset 2017 (via Kaggle) . Link to my GitHub is here.

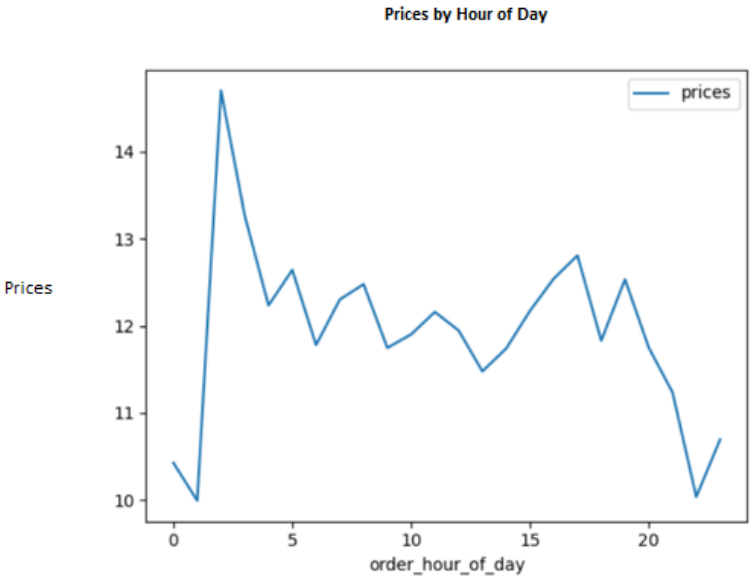
Age by Income Relationship Comparison



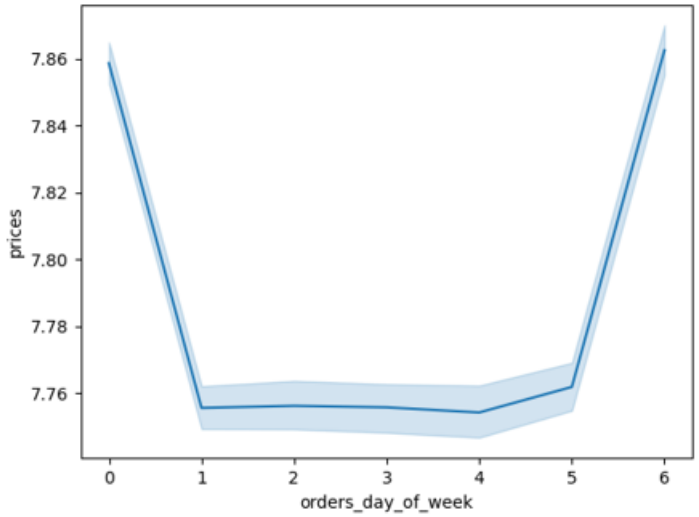
This chart highlights the relationship between age and income across different customer profiles.

- There's no strong linear correlation between age and income; earnings vary across all age groups.
- Most incomes cluster below $200,000, but higher-income groups are seen among middle-aged and senior adults.
- Younger customers tend to fall into low- and middle-income brackets.
- Profiles show broad diversity, with middle-income adults and parents making up a large portion of the customer base.

Prices by hour of day and day of the week



Notes: *To keep the line chart clear, I grouped data by hour and used average prices instead of raw values.*

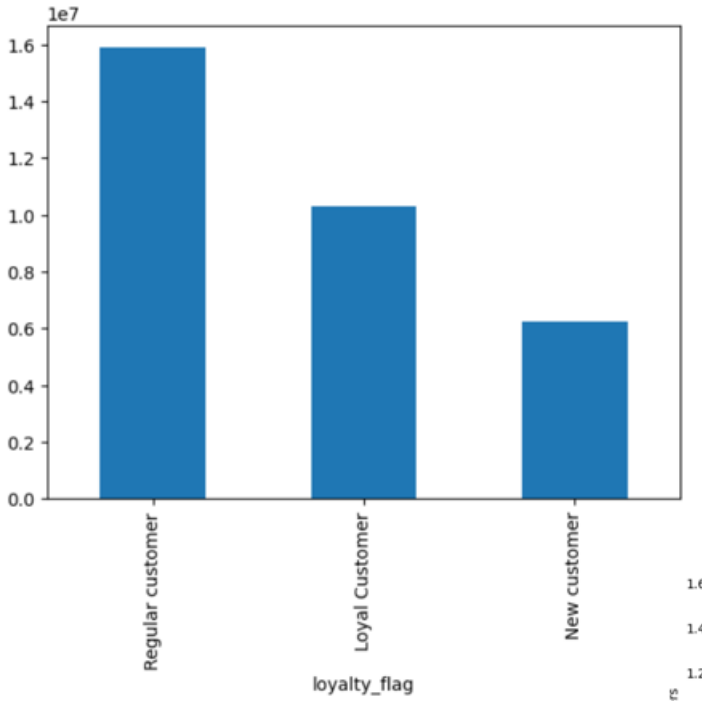


This analysis reveals patterns in when customers tend to spend the most money, based on average prices by hour and day.

- Spending peaks early in the morning, with the highest prices around 4:00–5:00 a.m., followed by a sharp drop by 9:00–10:00 a.m.
- A second upward trend occurs in the evening between 8:00–10:00 p.m.
- Weekend pricing is highest on Saturday and Friday, while Monday marks the lowest average prices.
- Tuesday through Thursday show relatively stable, mid-range pricing activity.

Customer Loyalty Insights

1st



2nd

The **first chart** highlights the distribution of Instacart users by loyalty status, showing that:

- Regular Customers make up the largest segment, indicating frequent but not highly committed users.
- Loyal Customers form a strong core group of repeat shoppers.
- New Customers, while smallest, suggest ongoing platform growth.

The **second chart** explores ordering habits based on loyalty:

- Low spenders dominate across all loyalty categories, regardless of loyalty level.
- While Regular Customers have the highest count of high spenders, the difference is small, suggesting high spenders are a minority overall.

Recommendations

- Optimize timing strategy: Run ads and promos on quieter days like Wednesday and during low-engagement hours (6–8 am, 5–10 pm). Avoid overcrowding peak times like weekends and midday hours.
- Target high-value shoppers: Early mornings (4–5 am) and late evenings (8–10 pm) see spending spikes. Use these insights to push premium products and bundle offers.
- Prioritize mid-tier pricing: Most products fall in the $5–15 range. Bundle common items and offer promotions for high-end categories like pasta sauces or imported goods.
- Promote top departments: Focus marketing on Produce, Dairy & Eggs, Snacks, Beverages, and Frozen. Use incentives in underperforming departments to stimulate trial.
- Focus on retention: Regular customers form the biggest segment. Offer loyalty upgrades and tailored onboarding for new customers to increase return rates.
- Encourage spending growth: All customer types are mostly low spenders. Implement tiered perks and spend-based incentives to nudge users up the value chain.
- Segment by family size: Segment marketing for families with 0–1 vs. 2+ dependents. Promote bulk offers to larger households and personalized bundles to small families.
- Refine demographic targeting: Middle-aged users dominate, but higher incomes are common among 40+ users. Build personalized campaigns for seniors and budget-conscious young adults.

Instacart (Python + segmentation project)

Project Challenges & Reflections

Challenge

- Very large dataset (3M+ rows).
- Duplicates, missing values, inconsistent types.
- Hard to build meaningful profiles initially.

Solution

- Cleaned data with pandas consistency checks.
- Derived loyalty + spend flags for segmentation.
- Grouped customers by profile and visualized behaviors.

Reflection

- Customer profiles worked well for marketing insights.
- Large dataset slowed operations at times.
- Next time: sample earlier + optimize functions.

Thank you