12/01/2026

www.elialanz.com

ELIA LANZUISE

DATA ANALYTICS PORTFOLIO

ABOUT ME

I AM A DATA ANALYST WITH A STRONG FOUNDATION IN SQL, PYTHON, AND DATA VISUALIZATION, FOCUSED ON TRANSFORMING RAW DATA INTO CLEAR, ACTIONABLE INSIGHTS THAT SUPPORT BUSINESS DECISION MAKING.

MY BACKGROUND IN MANAGING HIGH PERFORMANCE KITCHEN OPERATIONS AND CONSULTING E-COMMERCE BUSINESSES HAS GIVEN ME A DEEP UNDERSTANDING OF PERFORMANCE METRICS, FINANCIAL ANALYSIS, AND CUSTOMER BEHAVIOUR.

I COMBINE ANALYTICAL THINKING WITH CREATIVITY AND PROBLEM SOLVING TO TURN COMPLEX DATA INTO MEANINGFUL INSIGHTS, USING DATA STORYTELLING AND DASHBOARDS TO SUPPORT SMARTER, DATA DRIVEN DECISIONS.

I HAVE HANDS ON EXPERIENCE WORKING WITH SQL, PYTHON, TABLEAU, GITHUB AND EXCEL ACROSS MULTIPLE END TO END PROJECTS, FROM DATA CLEANING AND ANALYSIS TO VISUALIZATION AND DASHBOARD DEVELOPMENT.

I FOCUS ON UNCOVERING PATTERNS, IMPROVING EFFICIENCY, AND HELPING BUSINESSES MAKE INFORMED, DATA DRIVEN DECISIONS.

TABLE OF CONTENTS

PROJECT 01 [ROCKBUSTER STEALTH LLC] →

CUSTOMER & REVENUE SQL ANALYSIS

IDENTIFIED KEY REVENUE DRIVERS AND HIGH VALUE MARKETS TO SUPPORT STRATEGIC DECISIONS [POSTGRE, POWERPOINT,TABLEAU]

PROJECT 02 [GAMECO LLC] →

GLOBAL SALES & MARKET ANALYSIS DASHBOARD

ANALYZED GLOBAL SALES TO IDENTIFY TOP PERFORMING REGIONS AND PRODUCT TRENDS [EXCEL, POWERPOINT, TABLEAU]

PROJECT 03 [INSTACART-MAPLEBEAR INC.] →

CUSTOMER SEGMENTATION & PURCHASE BEHAVIOR ANALYSIS

SEGMENTED CUSTOMERS AND IDENTIFIED PURCHASING PATTERNS TO SUPPORT TARGETED STRATEGIES [PYTHON, JUPYTER, PANDAS]

PROJECT 04 [CITIBIKE NEW YORK CITY] →

DEPLOYED URBAN MOBILITY ANALYTICS DASHBOARD

BUILT AND DEPLOYED AN INTERACTIVE DASHBOARD TO IDENTIFY HIGH DEMAND ROUTES [STREAMLIT, GITHUB, PYTHON, JUPYTER, MATPLOTLIB]

PROJECT 05 [KING COUNTY USA] →

REAL ESTATE MARKET ANALYSIS AND DRIVERS

IDENTIFIED KEY FACTORS INFLUENCING HOUSING PRICES AND MARKET TRENDS [PYTHON, JUPYTER, PANDAS, MATPLOTLIB]

PROJECT 06 [CDC INFLUENZA USA] →

SEASONAL TREND ANALYSIS & RESOURCE PLANNING

ANALYZED SEASONAL PATTERNS TO SUPPORT RESOURCE PLANNING AND DEMAND FORECASTING [EXCEL, TABLEAU]

PROJECT 01

ROCKBUSTER STEALTH LLC - CUSTOMER & REVENUE SQL ANALYSIS

TOOLS USED: [POSTGRESQL, TABLEAU]

DATASET:

ROCKBUSTER INTERNAL DATABASE (15 TABLES TOTAL)

Summary

This project explores Rockbuster's customer, film, and revenue data to help the company transition from its traditional store model to an online streaming platform. Using SQL, I analyzed rental trends, customer lifetime value, and genre performance across countries to identify the most profitable markets and audience preferences. The main goal was to provide data driven recommendations that support Rockbuster's digital launch strategy and maintain competitiveness in an evolving industry landscape.

PROJECT 01

ROCKBUSTER STEALTH LLC: CUSTOMER & REVENUE SQL ANALYSIS

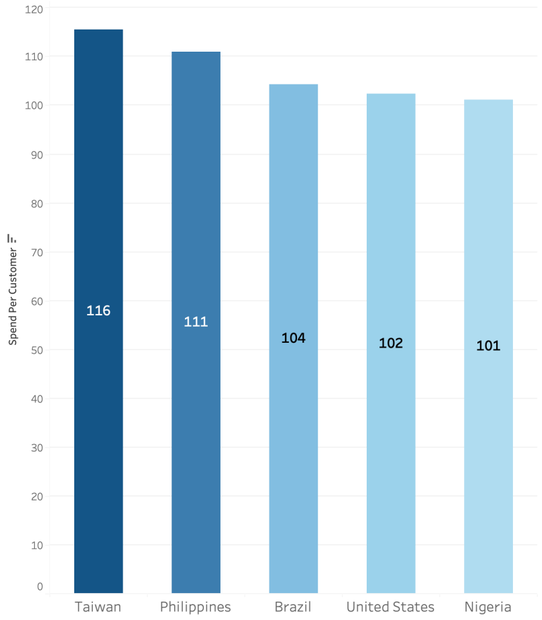
[PROCESS - ANALYSIS]

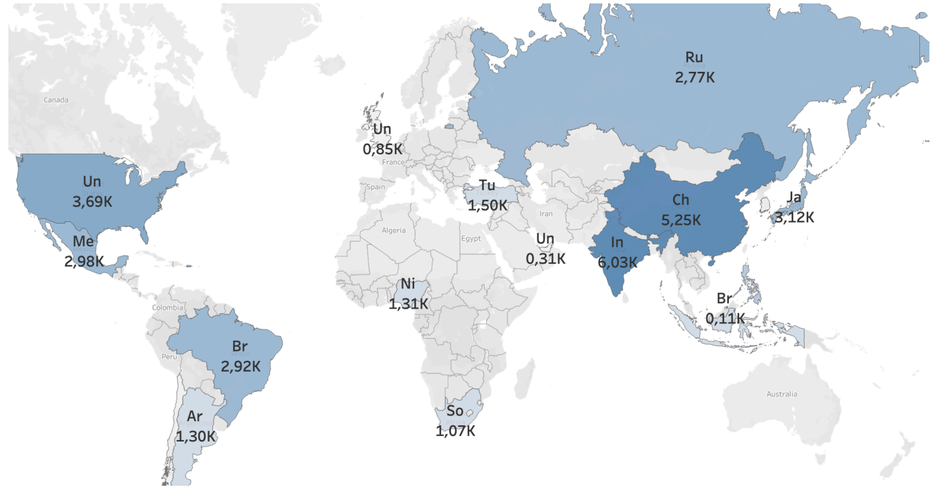
KEY STEPS TAKEN

- Loaded the Rockbuster dataset into PostgreSQL and verified data integrity.
- Joined multiple tables (customers, payments, films, rentals) to create unified datasets.
- Calculated key performance metrics such as total revenue, average spend, and rental duration.
- Segmented customers and countries by revenue contribution and spending behavior.
- Compiled results into visuals highlighting market performance and genre trends.

ROCKBUSTER STEALTH: CUSTOMER & REVENUE SQL ANALYSIS

[RESULTS - INSIGHTS]



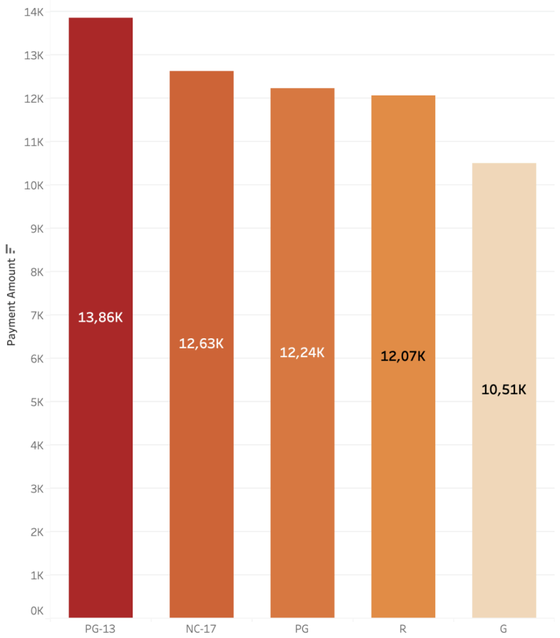


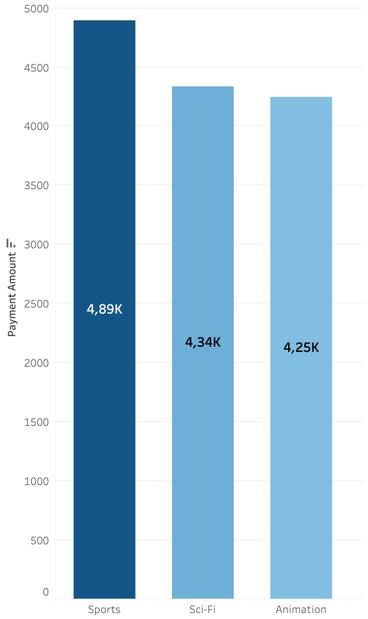
KEY FINDINGS

- India and China were the top revenue contributors, followed by the United States and major European markets.
- Taiwan and the Philippines had the highest average customer spend, indicating highly engaged niche audiences.
- Expanding regional marketing and loyalty programs in high-spending countries could boost overall profitability.

ROCKBUSTER STEALTH: CUSTOMER & REVENUE SQL ANALYSIS

[RESULTS - INSIGHTS]





KEY FINDINGS

- Sports and Sci-Fi films generated the most rental revenue, outperforming Drama and Comedy.
- PG-13 movies were the highest-earning rating, showing strong appeal among broad family audiences.
- Tailoring promotions by genre and rating preferences across top markets could maximize online sales performance.

ROCKBUSTER STEALTH: CUSTOMER & REVENUE SQL ANALYSIS

[REFLECTION - LEARNINGS]

This project deepened my understanding of SQL as a tool for uncovering business intelligence. I learned how to translate raw database outputs into strategic insights that can guide executive decisions. It also reinforced the value of linking customer and sales data to actionable market strategies.

If I were to extend this analysis, I'd integrate marketing spend, inventory data, and streaming engagement metrics to create a fuller picture of Rockbuster's global opportunities.

LINKS

https://github.com/elialanz/RockBuster_Project

https://github.com/elialanz/RockBuster_Project/blob/main/Rockbuster_Stealth_Project.pdf

PROJECT 02

GAMECO LLC - GLOBAL SALES & MARKET ANALYSIS DASHBOARD

TOOLS USED: [EXCEL, POWERPOINT, TABLEAU]

DATASET: HISTORICAL VIDEO GAME SALES DATASET (1980–2016)
SOURCED FROM VGCHARTZ

Summary

This project analyzes historical video game sales data to help GameCo understand how regional demand and genre performance have evolved over time. The goal was to identify patterns that could guide smarter marketing allocation and product launch strategies across global markets.

PROJECT 02

GAMECO GLOBAL SALES & MARKET ANALYSIS

[PROCESS - ANALYSIS]

KEY STEPS TAKEN

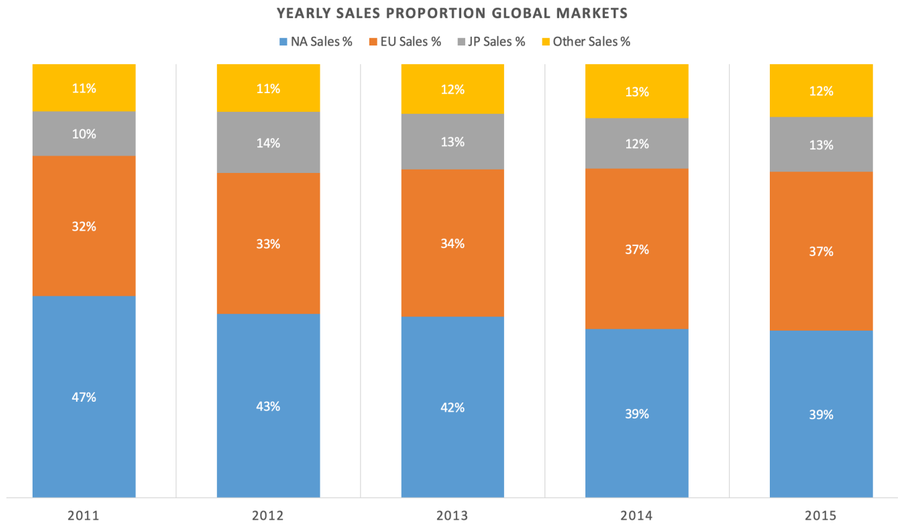
- Reviewed dataset for completeness and accuracy; handled missing or inconsistent entries.
- Applied descriptive statistics to summarize key metrics such as total and regional sales.
- Conducted trend analysis to observe changes in genre popularity over time.
- Performed regional comparisons to determine which markets showed the strongest growth.
- Built interactive visualizations in Tableau to highlight global and regional insights.

GAMECO GLOBAL SALES & MARKET ANALYSIS

[RESULTS - INSIGHTS]

Genres Sales Insights Year 2016

Genres	Global Revenue	Total Titles Release	Average Revenue Per Title
2016	70,89	343	0,21
Shooter	18,22	32	0,57
Sports	14,60	38	0,38
Fighting	3,86	14	0,28
Platform	2,07	10	0,21
Role-Playing	6,76	40	0,17
Action	19,91	119	0,17
Racing	1,64	20	0,08
Misc	1,17	18	0,07
Adventure	1,77	33	0,05
Strategy	0,50	10	0,05
Simulation	0,39	9	0,04
Grand Total	70,89	343	0,21



KEY FINDINGS

- Action and Sports games consistently led global sales, while Role-Playing titles showed regional strength in Japan.
- North America remained the largest market, but Europe showed steady growth, signaling shifting demand.
- Shooter Games are the highest revenue genre, with an average revenue of around half a million dollars per title.

GAMECO GLOBAL SALES & MARKET ANALYSIS

[REFLECTION - LEARNINGS]

This project strengthened my ability to turn large, unstructured datasets into practical insights and taught me the value of visual storytelling in data analytics.

If I were to extend this analysis, I'd incorporate additional variables such as advertising spend or release timing to better understand how external factors drive sales performance.

LINKS

https://github.com/elialanz/GameCo_Market_Analysis

https://github.com/elialanz/GameCo_Market_Analysis/blob/main/GameCo_Final_Presentation.pptx

PROJECT 03

INSTACART-MAPLEBEAR INC.

CUSTOMER SEGMENTATION & PURCHASE BEHAVIOR ANALYSIS

TOOLS USED: [PYTHON, JUPYTER, PANDAS, MATPLOTLIB, NUMPY, SEABORN]

DATASET:

- INSTACART ONLINE GROCERY SHOPPING DATASET (2017) FROM KAGGLE
- CUSTOM DATASET TO INTEGRATE FROM CAREERFOUNDRY

Summary

This project investigates purchasing patterns, customer segmentation, and shopping behaviors across millions of Instacart transactions. Using Python, I performed an exploratory data analysis to uncover trends in order timing, spending habits, price sensitivity, and customer loyalty. The aim was to provide the marketing and sales teams with actionable insights for targeted advertising and optimized product promotion strategies based on customer type, region, and time of purchase.

PROJECT 03

INSTACART-MAPLEBEAR INC.

CUSTOMER SEGMENTATION & PURCHASE BEHAVIOR ANALYSIS

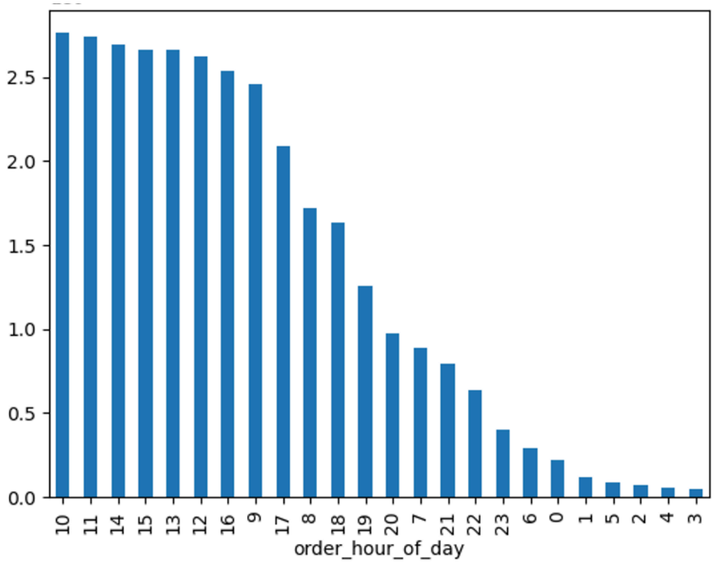
[PROCESS - ANALYSIS]

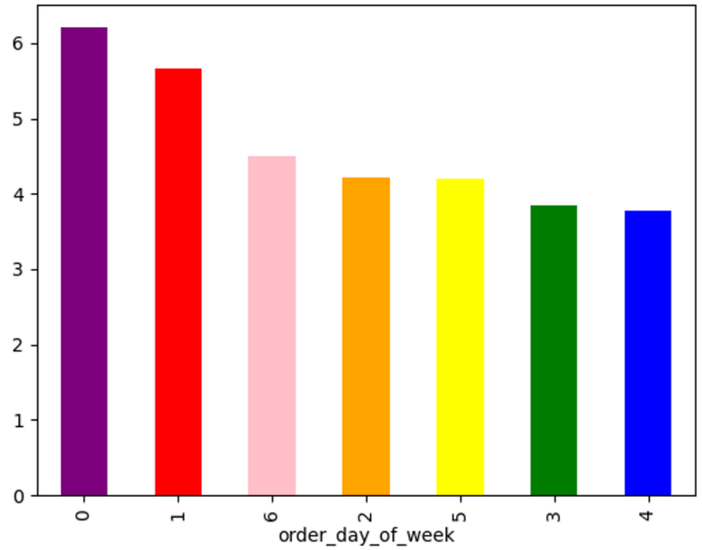
KEY STEPS TAKEN

- Imported, cleaned, and merged multiple Instacart datasets using pandas.
- Conducted consistency checks for duplicates, missing values, and mixed data types.
- Created new variables (price range, loyalty flag, region, customer profile).
- Aggregated and analyzed behavioral patterns by hour, day, and region.
- Produced visualizations highlighting key sales trends and customer segmentation insights.

INSTACART-MAPLEBEAR INC.

[RESULTS - INSIGHTS]



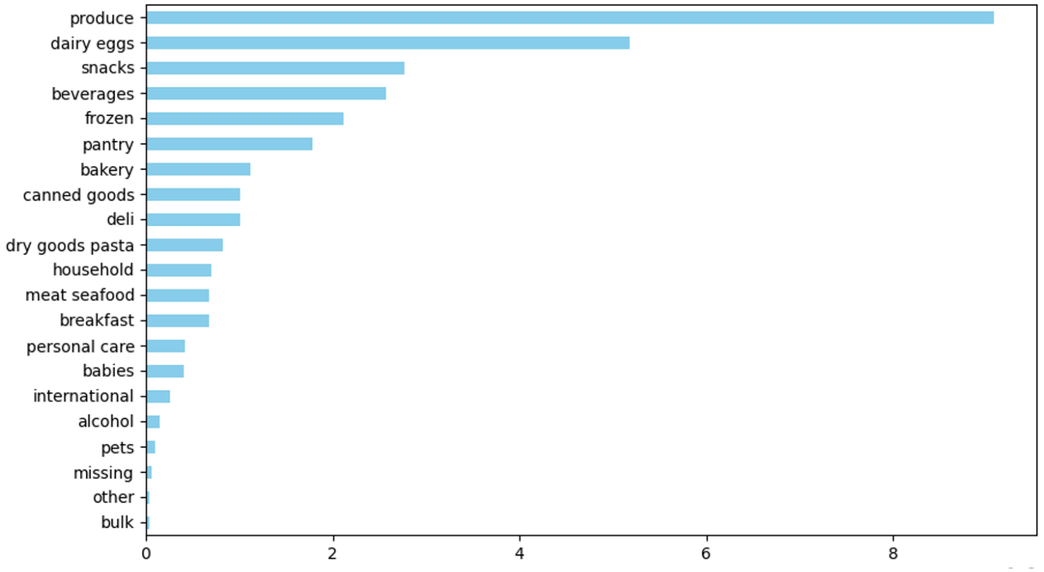


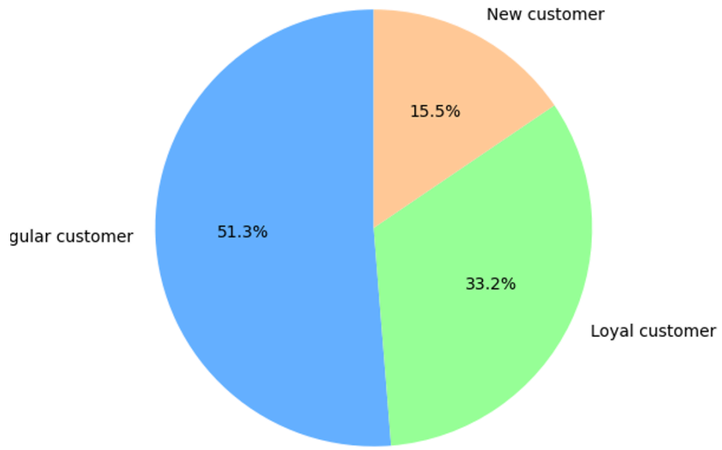
KEY FINDINGS

- Sunday and Monday are the busiest days, while late morning to early afternoon hours see the highest number of orders.
- Spending patterns closely mirror order frequency, with the peak spending between 10 AM and 2 PM.
- Advertising and promotions are best scheduled during off-peak hours (early morning or late evening) for better visibility and ROI.

INSTACART-MAPLEBEAR INC.

[RESULTS - INSIGHTS]





KEY FINDINGS

- Produce and dairy & eggs are the most frequently ordered categories, showing consistent customer reliance on fresh and staple goods.
- Regular customers make up the largest group (51%), but loyal customers order twice as frequently.
- Mid range products dominate the catalog, suggesting opportunities for premium or budget product line expansion.

INSTACART-MAPLEBEAR INC.

[REFLECTION - LEARNINGS]

This project helped me strengthen my ability to use Python for data cleaning, merging, and customer segmentation analysis. I learned to structure exploratory analysis workflows systematically, from wrangling data to visualizing trends that drive business impact.

If extended, I'd integrate external marketing and demographic data to build predictive models for personalized promotions and churn prevention.

LINKS

https://github.com/elialanz/Instacart_Basket_Analysis

https://github.com/elialanz/Instacart_Basket_Analysis/blob/main/instacart_final_report.xlsx

PROJECT 04

NEW YORK CITIBIKE USAGE ANALYSIS

TOOLS USED: [PYTHON, JUPYTER, PANDAS, MATPLOTLIB, NUMPY, KEPLER.GL, STREAMLIT]

DATASET:

- NYC CITIBIKE TRIP DATA (API SOURCED)
- WEATHER DATA (NOAA INTEGRATION)

Summary

Analyzed NYC Citi Bike trip data to identify usage patterns, station demand, and geographic distribution. Performed data wrangling, aggregation, and geospatial visualization to uncover high traffic routes and operational gaps. The goal was to support data driven decisions for bike redistribution, station planning, and service optimization.

PROJECT 04

NEW YORK CITIBIKE USAGE ANALYSIS

[PROCESS - ANALYSIS]

KEY STEPS TAKEN

- Retrieved Citi Bike data via API and structured datasets for analysis.
- Cleaned and transformed data (missing values, duplicates, data types).
- Created new variables (trip counts, time features, station pairs).
- Aggregated trips using groupby (start/end stations).
- Integrated geospatial coordinates for mapping and visualization.

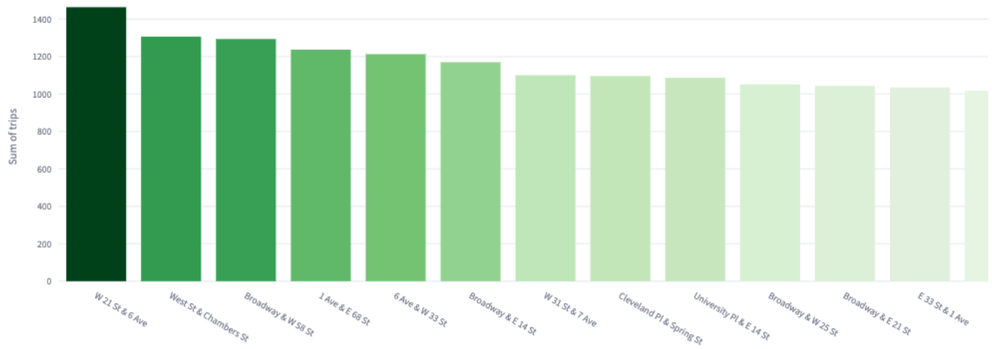
NEW YORK CITIBIKE USAGE ANALYSIS

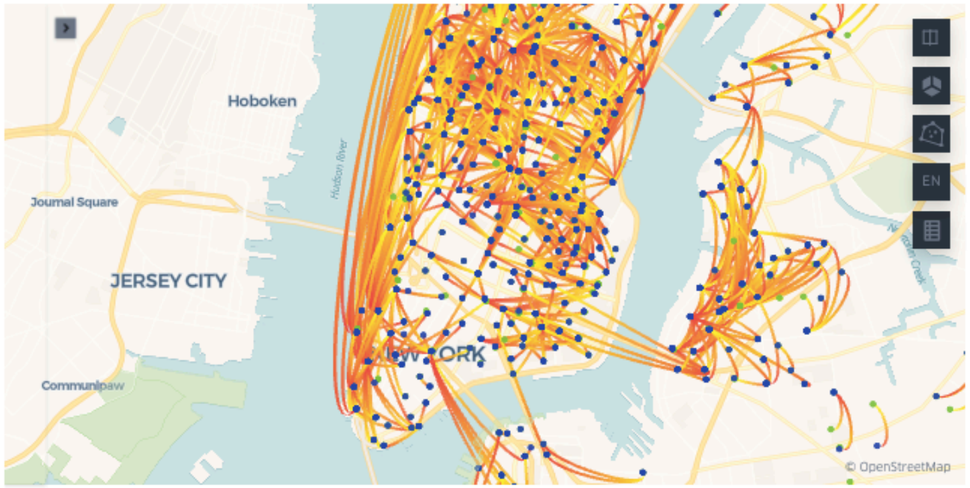
[RESULTS - INSIGHTS]

Total Bike Rides

339.03K

Top 20 most popular bike stations in New York



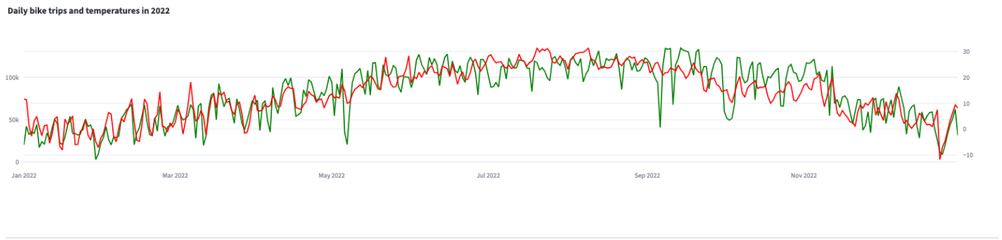


KEY FINDINGS

- High trip density observed in central Manhattan and commuter-heavy zones.
- Peak usage occurs during morning and evening commute hours.
- Popular routes connect business districts and transit hubs.
- Weekday usage exceeds weekends, indicating commuter driven demand.

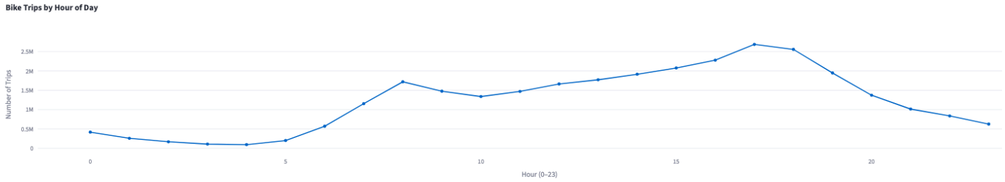
NEW YORK CITIBIKE USAGE ANALYSIS

[RESULTS - INSIGHTS]



📊 Daily Bike Trips vs Temperature (2022)

NY CitiBikes Strategy Dashboard



⏱️ Bike Trips by Hour of Day

KEY FINDINGS

- Certain stations show consistent imbalance (high departures vs arrivals).
- Tourist areas display irregular but high-volume trip patterns.
- Weather conditions impact trip frequency and duration.
- Filtering trips highlights critical high demand routes needing optimization.

NEW YORK CITIBIKE USAGE ANALYSIS

[REFLECTION - LEARNINGS]

This project strengthened my ability to work with real world API data, perform data aggregation, and build geospatial visualizations using Python. I improved my understanding of mapping tools like kepler.gl and how to translate raw data into actionable insights.

If extended, I would develop predictive models for demand forecasting and optimize bike redistribution using real-time and historical patterns.

LINKS

https://elia-nycitibike.streamlit.app

https://citibike.elialanz.com

https://github.com/elialanz/New_York_City_Bike

PROJECT 05

KING COUNTY USA - HOUSE MARKET ANALYSIS

TOOLS USED: [PYTHON, JUPYTER, PANDAS, MATPLOTLIB, TABLEAU]

DATASET:

- HOUSE SALES IN KING COUNTY, USA (2016)
- MORTGAGE INTEREST RATES FROM FED RESERVE API
- GEOJSON ZIPCODES DATASET FROM KING COUNTY GIS CENTER

Summary

This project analyzes housing prices, market segmentation, and affordability trends across King County, USA. Using Python for data cleaning, feature engineering, and exploratory analysis, I examined how factors like location, square footage, condition, and interest rates shape home values. Tableau dashboards were created to highlight pricing patterns, neighborhood differences, and key drivers of market variation.

The goal is to provide clear, data driven insights for planners, buyers, and investors to better understand demand, affordability, and long term housing trends in the region.

PROJECT 05

KING COUNTY USA - HOUSE MARKET ANALYSIS

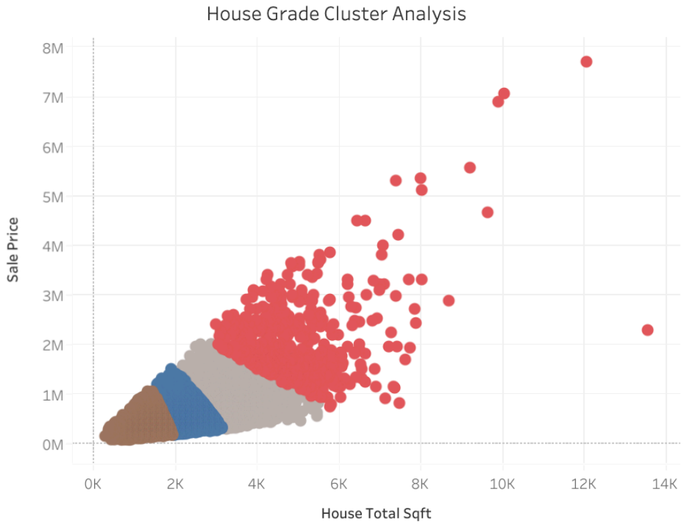
[PROCESS - ANALYSIS]

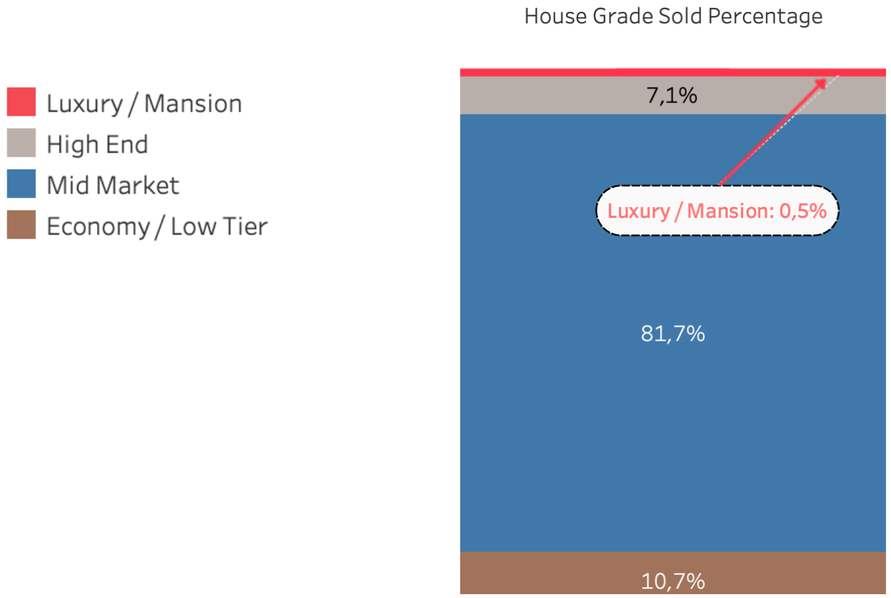
KEY STEPS TAKEN

- Cleaned, structured, and prepared King County housing, mortgage and GIS datasets for analysis.
- Performed consistency checks for missing values, outliers and unusual entries across datasets.
- Engineered new fields such as price per sqft, house grade groups and city level price indicators.
- Analysed patterns in sales volume, price trends, location differences and market segment groups.
- Built Tableau dashboards showing price drivers, regional patterns and affordability insights.

KING COUNTY USA - HOUSE MARKET ANALYSIS

[RESULTS - INSIGHTS]





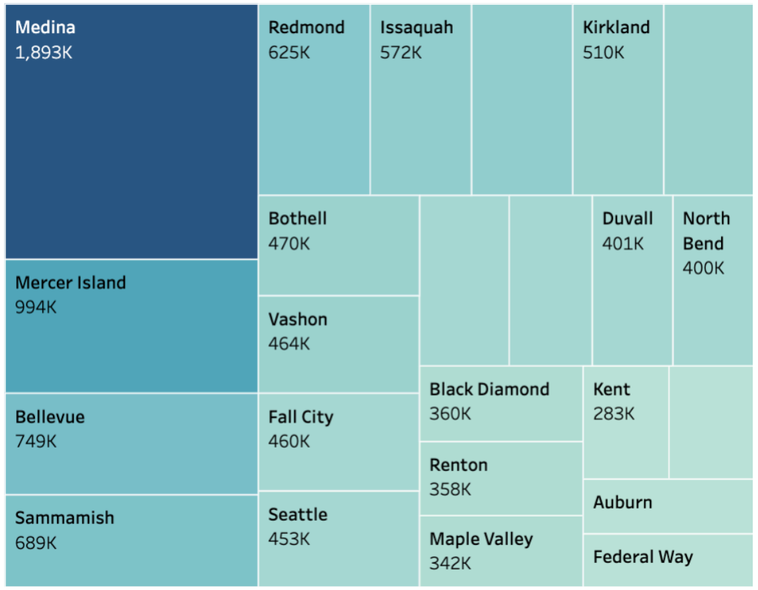
KEY FINDINGS

- Luxury and high end homes cluster at larger sizes and higher prices, forming a clear premium segment.
- Mid market properties dominate sales, while luxury homes make up less than 1% of total sales.
- House grade strongly influences price, with better rated homes consistently selling closer to the top range.

KING COUNTY USA - HOUSE MARKET ANALYSIS

[RESULTS - INSIGHTS]

MEDIAN SOLD HOUSE PRICE IN KING COUNTY CITIES



Top 5 Cheapest Cities

	Price Per Sqft	Avg. Square Feet
Federal Way	147	1,968
Auburn	149	1,955
Kent	154	1,941
Maple Valley	171	2,148
Enumclaw	172	1,832

Top 5 Cheapest Waterfront House Cities

Federal Way	243
Vashon	344
Issaquah	426
Seattle	470
Kenmore	475

KEY FINDINGS

- Premium cities like Medina, Mercer Island, and Bellevue command the highest dollars per sqft due to size and demand.
- Cities like Federal Way, Auburn, and Kent offer the lowest price per sqft, making them the most affordable options.
- Waterfront homes vary widely, with Federal Way being far cheaper than Issaquah, Seattle, or Kenmore.
- Better home condition consistently raises price per sqft, showing a clear link between maintenance and value.

KING COUNTY USA - HOUSE MARKET ANALYSIS

[REFLECTION - LEARNINGS]

This project strengthened my ability to analyse real estate datasets, combining Python and Tableau to explore price drivers, location effects, and market segmentation. I learned how to connect structural features with economic trends and how to turn raw housing data into clear, decision ready insights; despite a fairly good analysis in been done, is also important to notice that the nuances of house prices and sales are very broad and very unique, i can realise that working with real estate datasets is challenging and domain knowledge is a key factor to succeed. If extended, I'd incorporate longer time ranges, rental data, and additional economic indicators to build predictive models for price forecasting and neighbourhood trend analysis.

LINKS

GitHub Project - King County USA House Market Analysis

King County USA House Market Analysis - Tableau Visualization

PROJECT 06

MAPPING THE HUMAN COST OF INFLUENZA IN US

TOOLS USED: [EXCEL, TABLEAU]

DATASETS:

- INFLUENZA DEATHS BY US STATES (CDC)
- US POPULATION DEMOGRAPHICS (US CENSUS BUREAU)

Summary

This project explores influenza mortality data across the United States to uncover which age groups, states, and populations are most at risk. Using Tableau, I visualized over 700,000 flu related deaths recorded between 2009 and 2017, identifying high risk regions, seasonal trends, and demographic patterns. The goal was to help public health agencies optimize staffing and resource allocation during flu seasons to minimize preventable deaths and improve preparedness across hospitals and senior care facilities.

PROJECT 06

MAPPING THE HUMAN COST OF INFLUENZA IN US

[PROCESS - ANALYSIS]

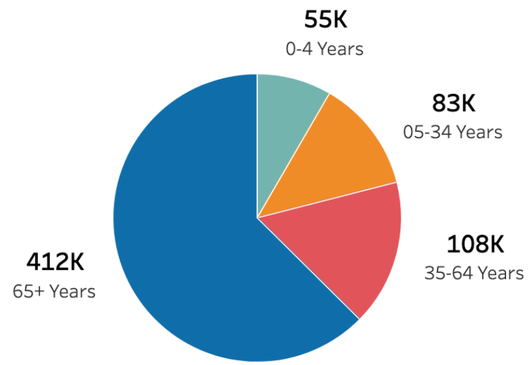
KEY STEPS TAKEN

- Imported and reviewed CDC influenza mortality data for completeness and consistency.
- Cleaned and organized the dataset using Excel (handling missing values and formatting).
- Designed multiple Tableau dashboards to visualize deaths by age, time, and geography.
- Used trend and forecast models to identify seasonal peaks and predict upcoming flu cycles.
- Documented insights and recommendations for targeted healthcare staffing strategies.

MAPPING THE HUMAN COST OF INFLUENZA IN US

[RESULTS - INSIGHTS]

Flu Deaths Totals by Age Group (2009-2017)



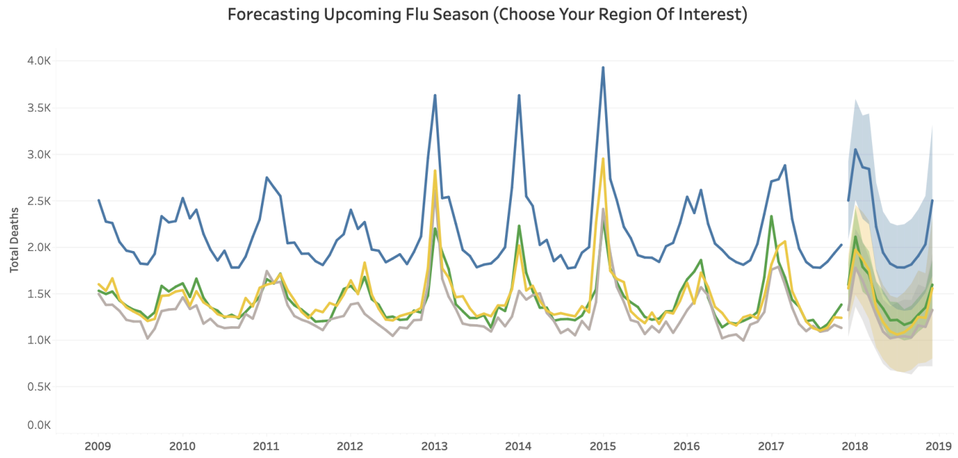
Flu Deaths Totals by Months (2009-2017)

KEY FINDINGS

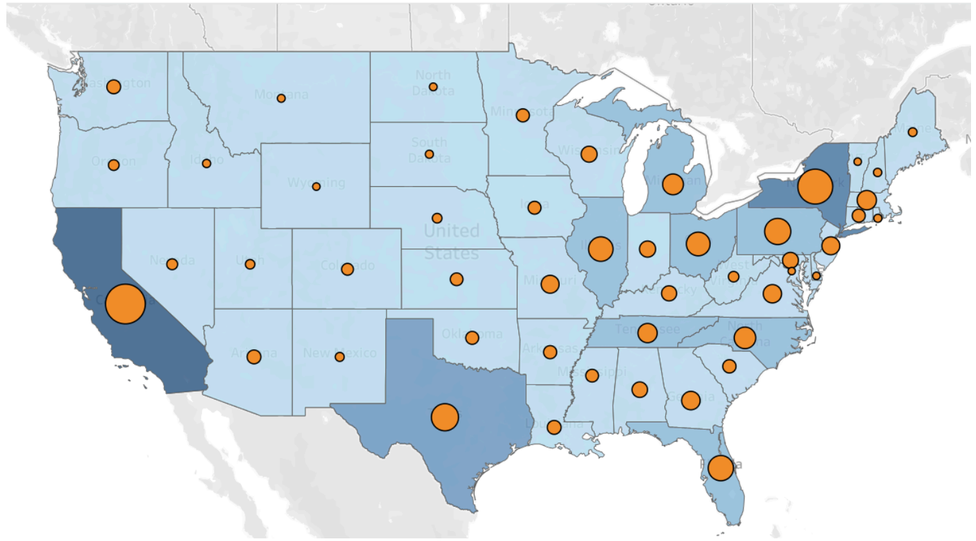
- Adults aged 65+ years accounted for the vast majority of flu related deaths (412K cases).
- Flu deaths consistently peaked between November and January, with a smaller wave in March, showing a recurring secondary surge.
- While deaths remained stable overall, years 2013, 2015, and 2017 recorded the highest mortality peaks across all age groups.

MAPPING THE HUMAN COST OF INFLUENZA IN US

[RESULTS - INSIGHTS]



Total Influenza Deaths Recorded Across U.S. States (2009-2017)



KEY FINDINGS

- States like California, Texas, Florida, and New York reported the highest flu deaths overall, both in total and among senior populations.
- The West and Northeast showed strong correlations between population density and senior death rates.
- Forecast projections indicated a continued seasonal recurrence every winter, confirming the need for resource planning based on cyclical flu patterns.

MAPPING THE HUMAN COST OF INFLUENZA IN US

[REFLECTION - LEARNINGS]

This project taught me how to translate large public health datasets into actionable insights that can directly support real world decision making. I improved my Tableau storytelling by linking data driven visuals to human impact, strengthening both the analytical and narrative clarity of my presentation.

If I were to expand this project, I would integrate hospital capacity and vaccination coverage data to better model risk factors and predict future outbreaks with more precision.

LINKS

https://github.com/elialanz/Influenza_Seasons_US

https://public.tableau.com/app/profile/elia.lanzuise/viz/MappingInfluenzaDeathsInUSA/InfluenzaDeathsInUS