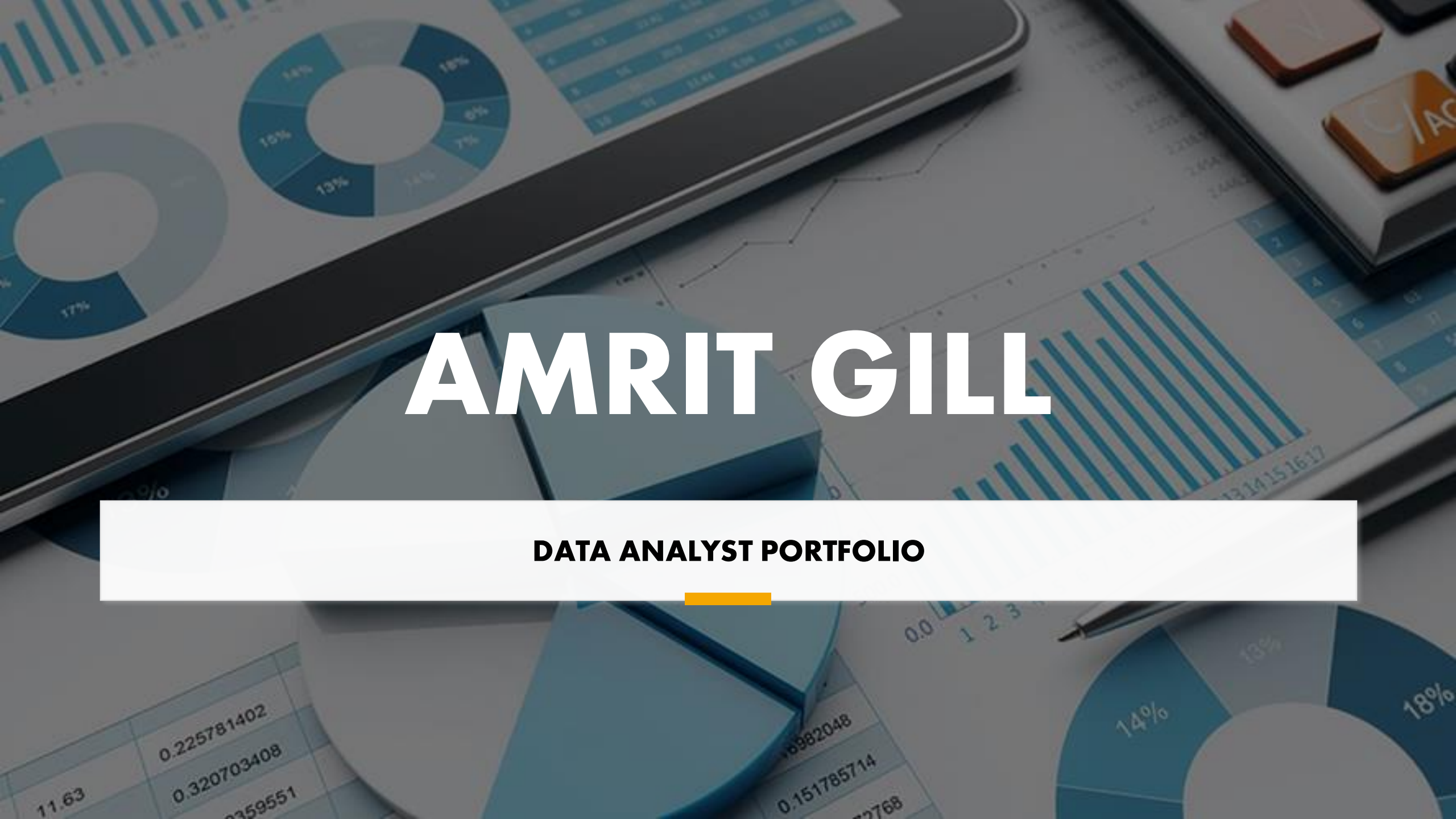


AMRIT GILL

DATA ANALYST PORTFOLIO

ABOUT ME

Hi, My name is Amrit Gill. With a solid foundation in Business Administration and a deep passion for mathematics and logical analysis, transitioning into a data analyst role felt like a natural progression in my career.

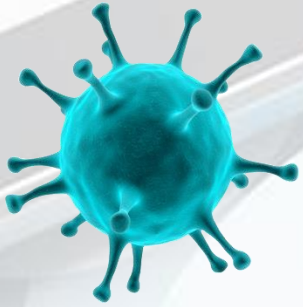
My background in Computer Engineering and experience in marketing and skills in problem-solving, decision-making and data-driven results. I thrive on breaking down complex challenges into clear, actionable insights. I'm eager to bring my analytical expertise to drive success in your organization. The following slides are showcase of my few projects from my recent training in Data Analytics.

PROJECTS

GAME CO.

GLOBAL VIDEO GAME SALES ANALYSIS

TOOLS USED:



INFLUENZA

MEDICAL STAFF ALLOCATION TO PREPARE FOR INFLUENZA SEASON

TOOLS USED:

ROCKBUSTER

ANALYSIS OF CUSTOMER RENTAL HABITS TO SUPPORT AN ONLINE STREAMING SERVICES.

TOOLS USED:

PostgreSQL

INSTACART

ANALYSIS OF CUSTOMER DEMOGRAPHICS AND PURCHASING PATTERNS.

TOOLS USED:

python

PIG E. BANK

CUSTOMER RETENTION ANALYSIS FOR GLOBAL BANK.

TOOLS USED:

INVENTORY FORECAST

ANALYZING AND FORECASTING RETAIL STORE INVENTORY DEMAND.

TOOLS USED:

python

GAME CO.

OVERVIEW:

GAME CO. IS A VIDEO GAME COMPANY WHICH WANTS TO USE DATA ANALYSIS TO UNDERSTAND THE MARKET DEVELOPMENT TO GUIDE STRATIGIES AND GENERATE GROWTH IN THE MOST SUSTAINABLE MARKETS.

DATA USED:

DATA SET COVERS HISTORICAL SALES OF VIDEO GAMES (FOR GAMES THAT SOLD MORE THAN 10,000 COPIES) SPANNING DIFFERENT PLATFORMS, GENRES AND PUBLISHING STUDIOS, DRAWN FROM WEBSITE VGCHARTZ.

OBJECTIVES:

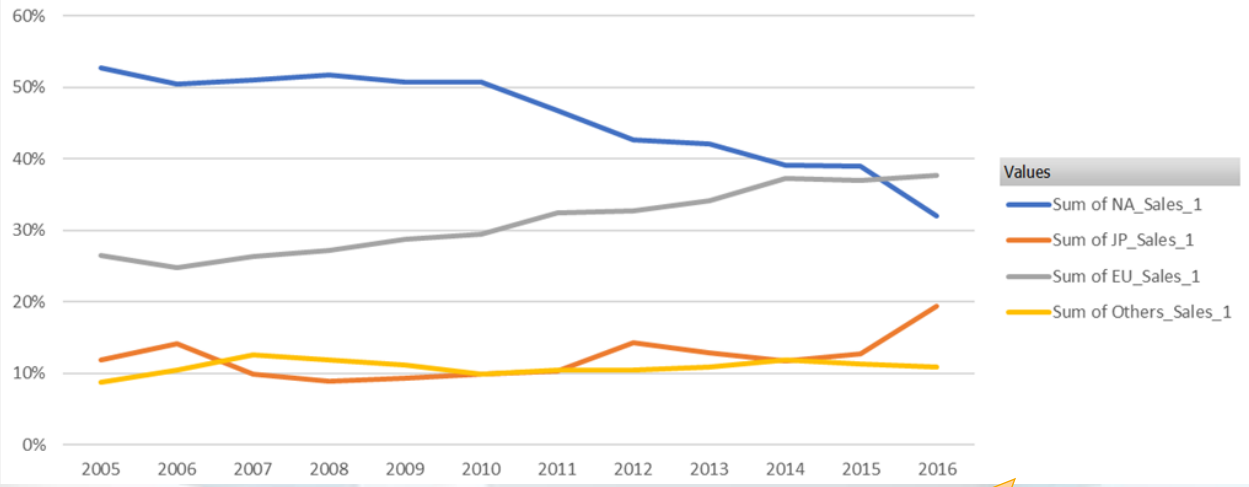
- ANALYSE GLOBAL SALES TRENDS OVER PAST YEARS ACROSS DIFFERENT GAME GENRES, TIME PERIODS AND GEOGRAPHICS REGIONS.
- PROVIDE ACTIONABLE INSIGTHS TO GUIDE GAMECO'S GAME DEVELOPMENT AND MARKET STRATEGY.

METHODS:

- GROUPING, SORTING AND FILTERING DATA IN EXCEL.
- DATA CLEANING.
- DESCRIPTIVE ANALYSIS.
- DATA VISUALIZATION.

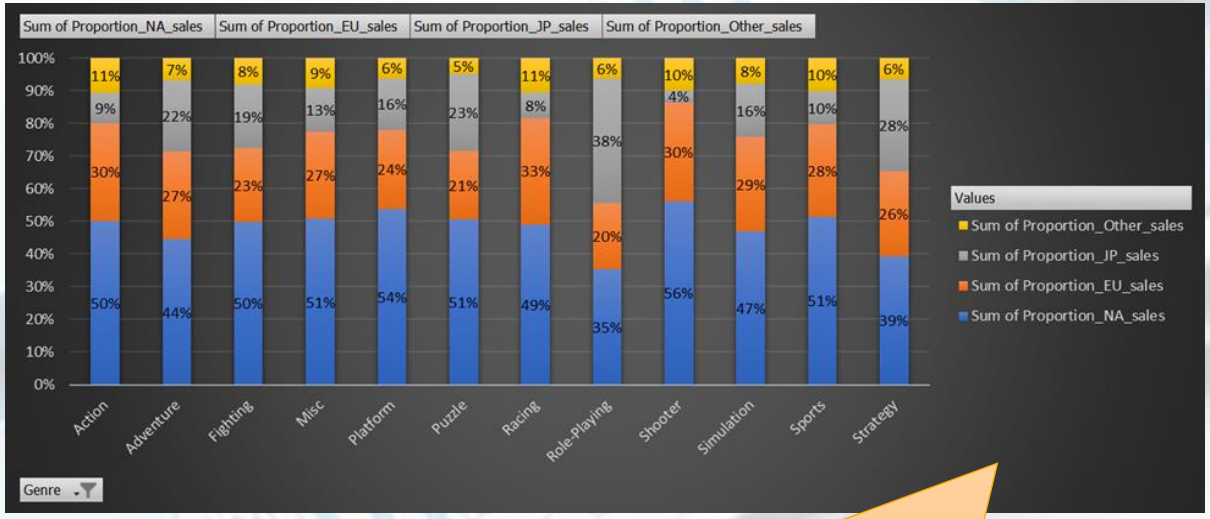
ANALYSIS AND VISUALIZATION

SALES MARKET FROM 2005-2016



THIS GRAPH SHOWS NORTH AMERICA WAS ON TOP IN 2005 & NEGATIVE CORRELATION APPEARS IN NORTH AMERICA AND JAPAN SALES

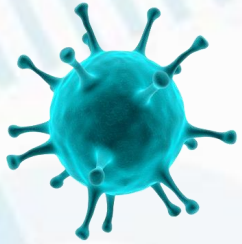
SALES BY CATEGORIES



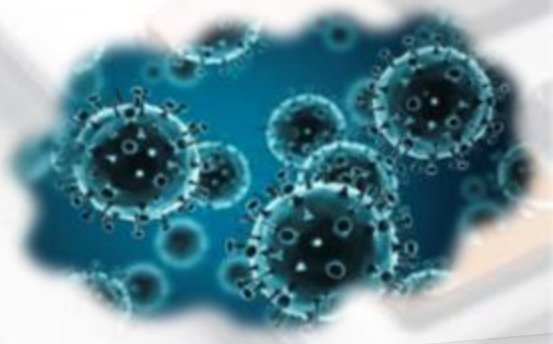
THIS BAR CHART SHOWS THAT SHOOTER, ACTION AND SPORTS TRENDING IN NORTH AMERICA AND EUROPE. WHILE JAPAN FAVORED IN ACTION AND ROLE- PLAYING GAMES.

INSIGHTS AND RECOMMENDATION

- THE SIGNIFICANT DROP CAN BE SEEN IN GLOBAL SALES AFTER 2008. NORTH AMERICA HAS REMAINED LEADER IN SALES SINCE 2015. IN 2016 SALES HAD BEEN STEADILY DECLINED.
- A POSITIVE CORRELATION CAN BE SEEN IN ALL REGIONS EXCEPT JAPAN. EUROPE SALES ARE INCREASING SINCE 2005 TILL 2016.
- PRIORITIZE THE GAMES BY GENRE IN ALL REGIONS TO INCREASE THE SALES. MAINTAIN THE UPWARD MARKET IN EUROPE TO EXCEED SALE NUMBERS.
- FOCUS ON JAPAN'S MARKET BUDGET ON ROLE-PLAYING GENRE TO GAIN POPULARITY IN 2016.
- MAINTAIN THE ALLOCATED BUDGET TO NORTH AMERICA SALES AND FOCUS ON
- OBSTACLES TO ACHIEVE THE GOALS.



INFLUENZA SEASON



OVERVIEW:

THE UNITED STATES HAS AN INFLUENZA SEASON WHERE MORE PEOPLE THAN USUAL SUFFER FROM THE FLU. HOSPITALS AND CLINICS NEED ADDITIONAL STAFF TO ADEQUATELY TREAT THE PATIENTS BY PROVIDING TEMPORARY STAFF.

OBJECTIVES:

- PROVIDE A PLAN FOR INFLUENZA SEASON FORCASTING ADDITIONAL STAFF DEMAND.
- EXAMINE TRENDS IN INFLUENZA AND USE IT TO PLAN STAFFING NEEDS ACROSS THE COUNTRY.

INTERIM REPORT

DATA:

- INFLUENZA DEATHS BY GEOGRAPHY DATASET FROM THE CDC.
- POPULATION DATA BY GEOGRAPHY, TIME, AGE, AND GENDER DATASET FROM US CENSUS BUREAU.

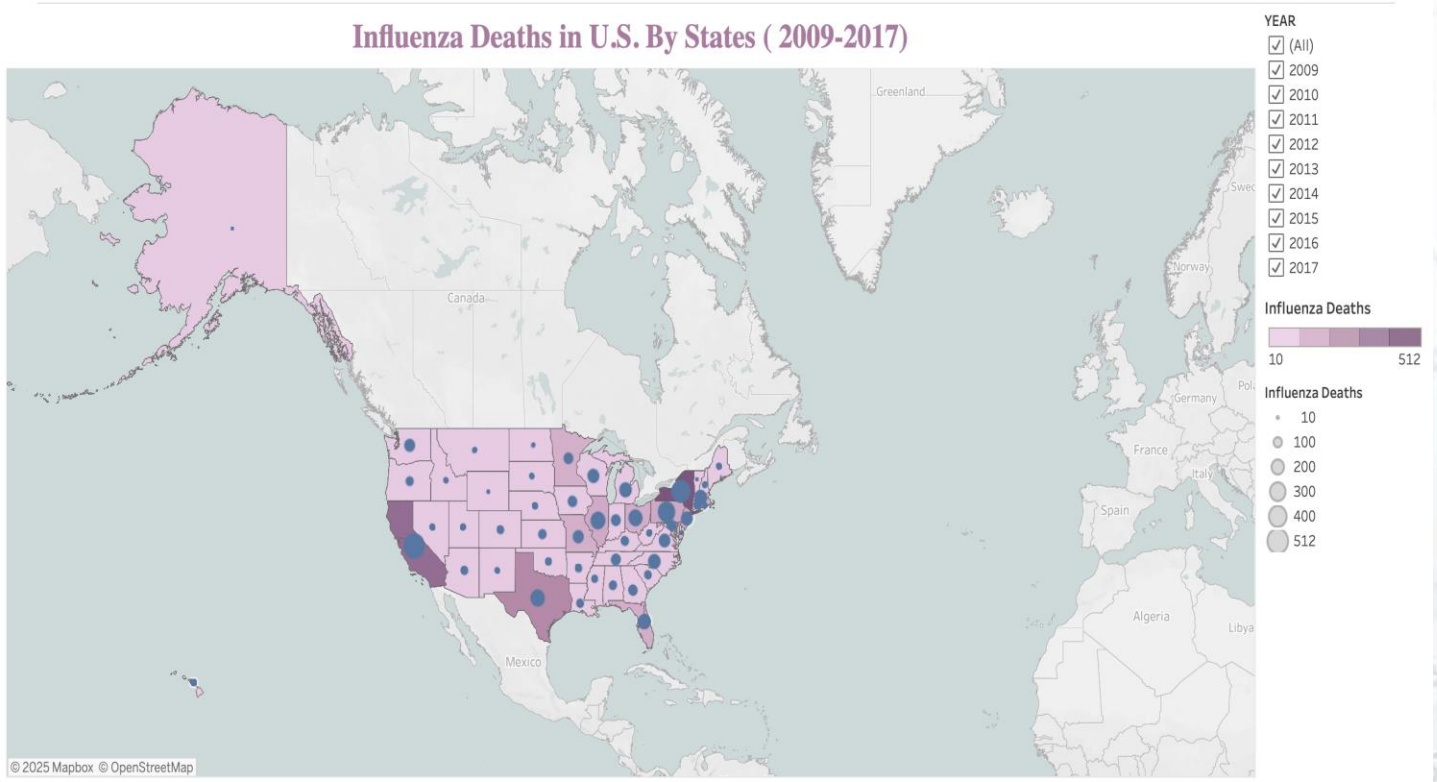
METHODS:

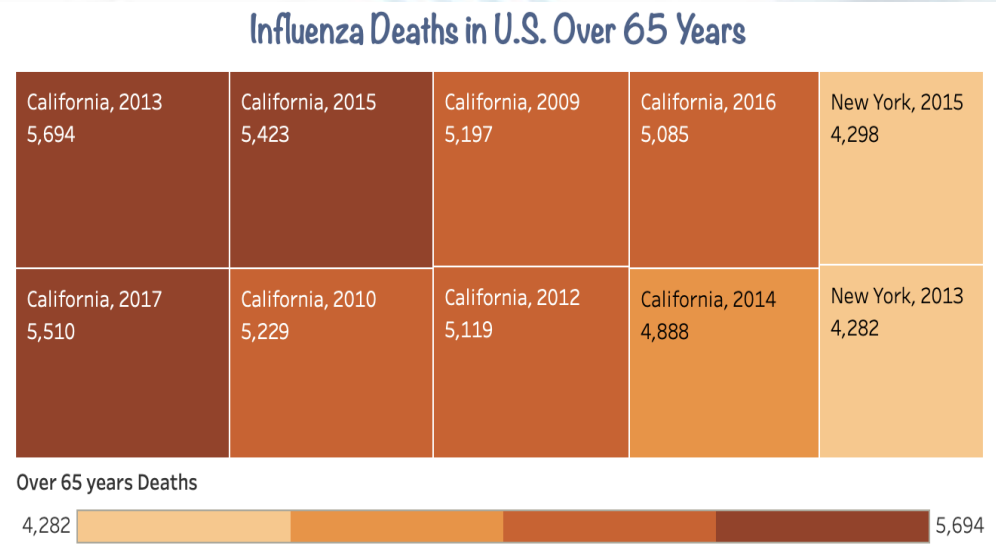
- DATA TRANSFORMATION.
- DATA INTREGATION.
- DATA PROFILING.
- STATISTICAL ANALYSIS & TESTING.
- FORCASTING & SPATIAL ANALYSIS.
- DATA VISUALISATION & STORYTELLING.

ANALYSIS AND VISUALIZATION

INFLUENZA DEATHS STORY BY TABLEAU

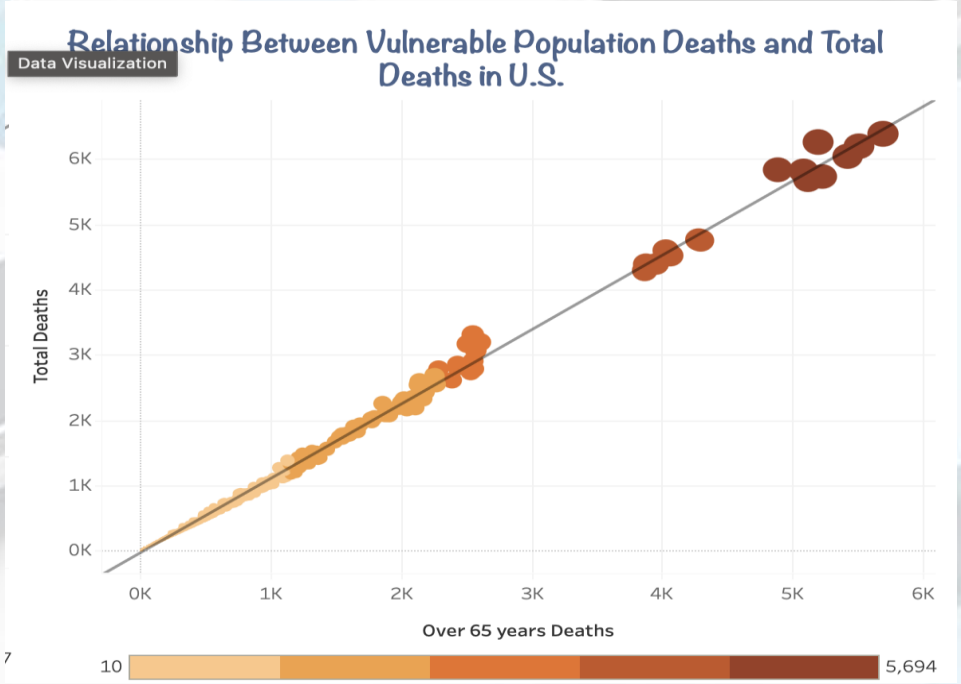
INFLUENZA DEATHS & POPULATION IN U.S. BY STATES (2009-2017)



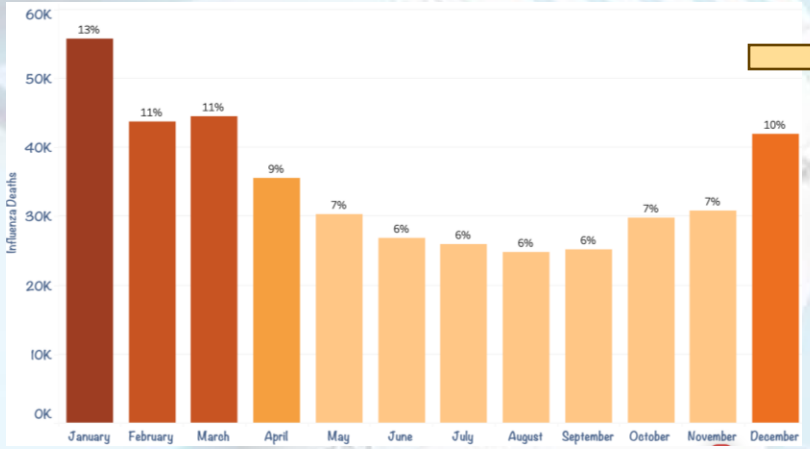


- CALIFORNIA IS THE STATE WITH HIGHEST DEATH RATE BY INFLUENZA IN U.S.
- ALASKA IS THE STATE WITH THE LEAST DEATHS.

INSIGHTS AND RECOMMENDATIONS



- **STATES SUCH AS CALIFORNIA AND NEW YORK SHOWS HIGHEST VULNERABLE DEATHS DUE TO INFLUENZA. ADVICE AN ANNUAL FLU VACCINATIONS TO HIGH-RISK GROUPS TO REDUCE THE INCIDENCE OF SEVERE ILLNESS AND DEATHS.**
- **PUBLIC HEALTH STRATEGIES FOCUS ON INCREASING VACCINATION RATES AMONG THESE POPULATION AND IMPROVING ACCESS TO HEALTHCARE SERVICES TO MITIGATE THE IMAPCT OF INFLUENZA.**



WHEN IS INFLUENZA SEASON: JANUARY, FEBURARY, MARCH AND DECEMBER.

THERE IS A POSITIVE CORRELATION BETWEEN THE NUMBER OF THE OLDER POPULATION AND INFLUENZA DEATHS, WHICH MEANS THAT IN STATES WITH A HIGHER POPULATION OF 65+ YEARS OLD, THERE WILL BE MORE DEATHS BY INFLUENZA.

ROCKBUSTER STEALTH

OVERVIEW:

ROCKBUSTER STEALTH LLC IS A GLOBAL LEADER IN MOVIE RENTAL, PLANS TO LAUNCH AN ONLINE VIDEO RENTAL SERVICE TO REMAIN COMPETITIVE AGAINST COMPANIES LIKE NETFLIX AND AMAZON PRIME. THE MANAGEMENT TEAM PLANS TO USE ITS EXISTING MOVIE LICENSES TO LAUNCH AN ONLINE VIDEO RENTAL SERVICE TO STAY COMPETITIVE.

OBJECTIVES:

- IDENTIFY TOP-PERFORMING MOVIES AND LOYAL CUSTOMERS TO MAXIMUM REVENUE.
- ANALYZE RENTAL DURATION AND GEOGRAPHICS SALES TRENDS TO OPTIMIZE MARKET STRATEGIES.

DATA:

DATASET IS RETRIEVED FROM RDBMS USING SQL, CONTAINS INFORMATION ABOUT ROCKBUSTER'S FILM INVENTORY,

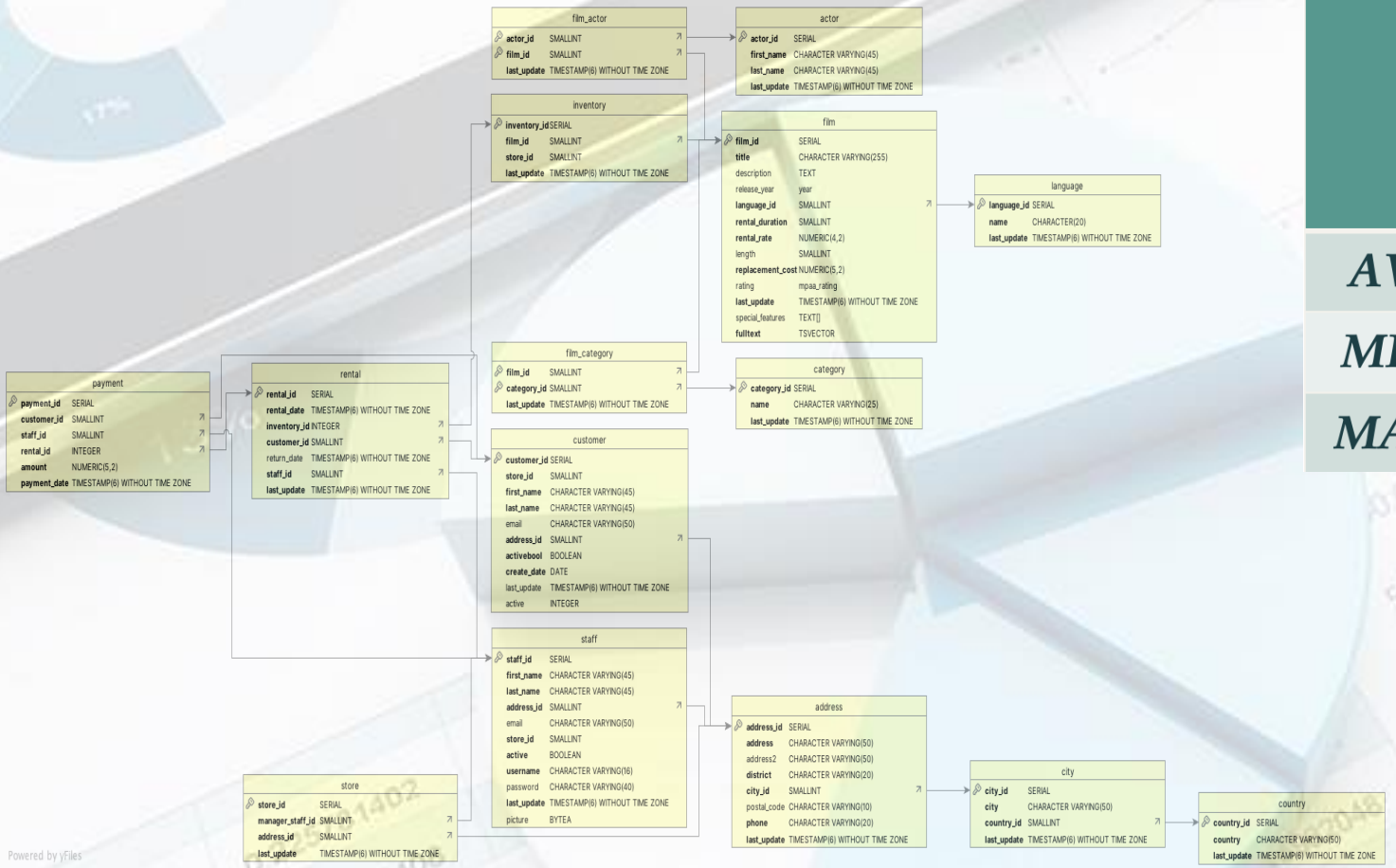
CUSTOMERS AND PAYMENTS.

METHODS:

- POSTGRESQL
- DATA CLEANING & FILTERING
- JOINING TABLES
- SUBQUERIES &CTE
- DATA VISUALIZATION

GitHub GITHUB ROCKBUSTER REPOSITORY

ERD & STATISTICS

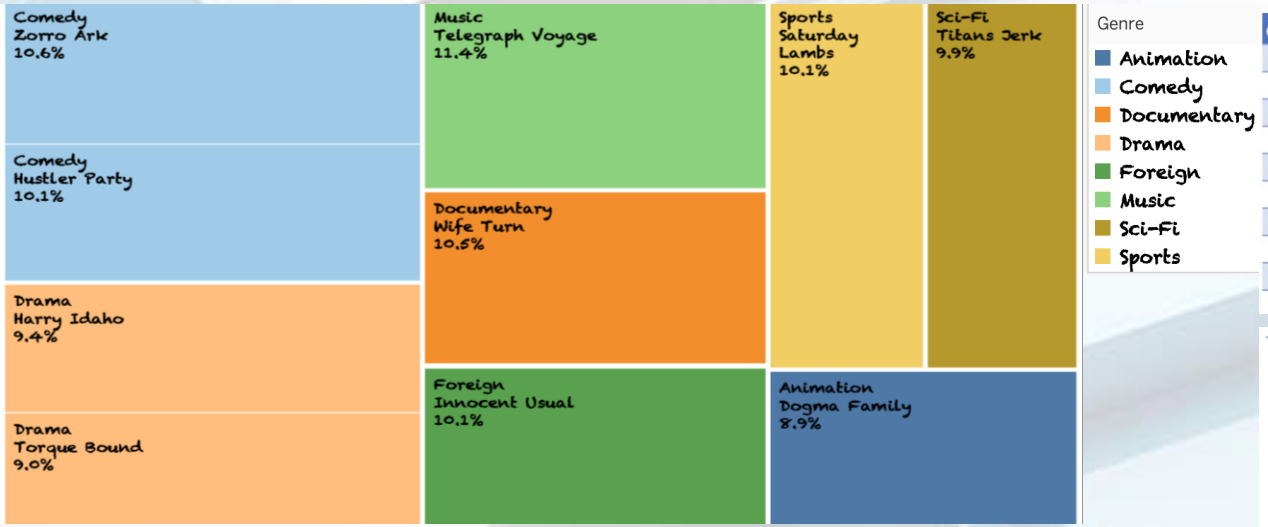


	RENTAL DURATION (IN DAYS)	RENTAL RATE ($)
AVERAGE	5 DAYS	$2.98
MINIMUM	3 DAYS	$0.99
MAXIMUM	7 DAYS	$4.99

RENTAL DATA INDICATES AN AVERAGE DURATION OF 5 DAYS, WITH PRICES RANGING FROM $0.99 TO $4.99.

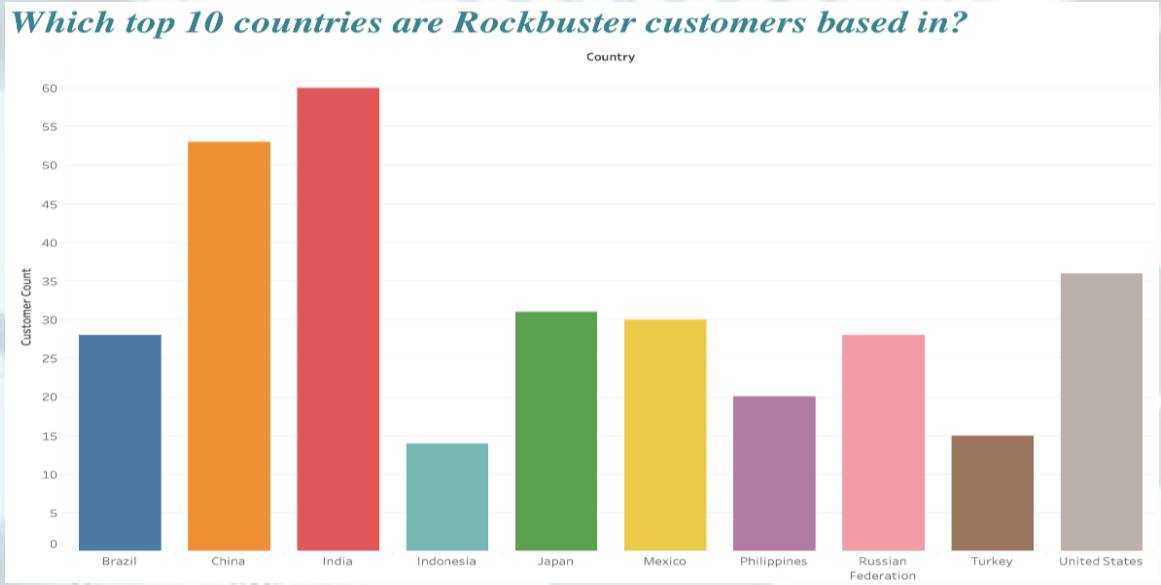
ANALYSIS AND VISUALIZATION

TOP 10 FILMS WITH MOST REVENUE BY GENRE



TOP 10 LOYAL CUSTOMERS

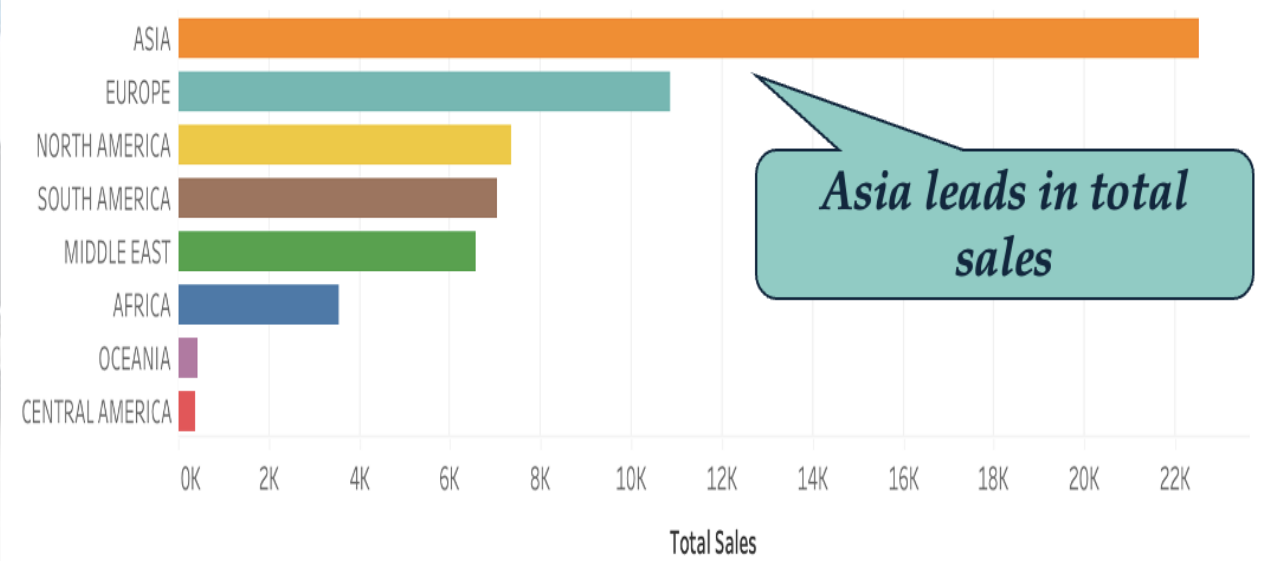
customer_id	first_name	last_name	city	country	total_amount_paid
148	Eleanor	Hunt	Saint-Denis	Runion	211.55
526	Karl	Seal	Cape Coral	United States	208.58
178	Marion	Snyder	Santa Brbara dOeste	Brazil	194.61
137	Rhonda	Kennedy	Apeldoorn	Netherlands	191.62
144	Clara	Shaw	Molodetno	Belarus	189.6
459	Tommy	Collazo	Qomsheh	Iran	183.63
181	Ana	Bradley	Memphis	United States	167.67
410	Curtis	Irby	Richmond Hill	Canada	167.62
236	Marcia	Dean	Tanza	Philippines	166.61
403	Mike	Way	Valparai	India	162.67



- THE TOP GENRES ARE SPORTS, SCI-FI, ANIMATION, DRAMA & COMEDY.
- THE HIGHEST NUMBER OF CUSTOMERS IN INDIA AND CHINA.

INSIGHTS AND RECOMMENDATIONS

TOP SALES ACROSS WORLWIDE



[PROJECT PRESENTATION]

- ROCKBUSTER SHOULD FOCUS IN MARKETING ON TOP REVENUE-GENERATING MARKETS LIKE ASIA.
- THE TOP 10 REVENUE GENERATING COUNTRIES ARE CONTRIBUTED TO OVER HALF OF THE TOTAL REVENUE. MOVIES WITH 'PG-13' OR 'NC-17' RATINGS GENERATED HIGHER REVENUE. LOWEST EARNING GENRES: MUSIC, TRAVEL, THRILLER.
- DEVELOP LOYALITY PROGRAMS AND EXCLUSIVE PROMOTIONS FOR HIGH-VALUE CUSTOMERS TO ENCOURAGE SUSTAINED ENGAGEMENT.
- CONDUCT MARKET RESEARCH TO UNDERSTAND LOW PERFORMANCE COUNTRIES AS COMPARED TO OTHERS, IDENTIFY WHETHER THIS IS DUE TO POPULATION SIZE OR ANY OTHER MARKET DYNAMICS.

INSTACART GROCERY BASKET

OVERVIEW:

INSTACART, AN ONLINE GROCERY STORE THAT OPERATES THROUGH AN APP, TO IDENTIFY KEY SALES PATTERNS AND CUSTOMER BEHAVIOR. THE ANALYSIS ADDRESSES CRITICAL QUESTIONS FROM THE SALES AND MARKETING TEAMS, INCLUDING IDENTIFYING THE BUSIEST DAYS AND HOURS FOR ORDERS, UNDERSTANDING SPENDING PATTERNS, AND EXPLORING PRODUCT POPULARITY BY DEPARTMENT.

OBJECTIVES:

- PERFORM INTIAL DATA AND EXPLORATORY ANALYSIS OF THE CUSTOMER PURCHASE HISTORY DATA TO UNCOVER PATTERN, PURCHASING BEHAVIORS INCLUDING LOYALITY, DEMOGRAPHICS AND REGIONAL DIFFERENCE, TO SUPPORT TARGETED MARKETING EFFORTS, SALE GUIDE AND PROMOTIONAL STRATEGIES.

DATA:

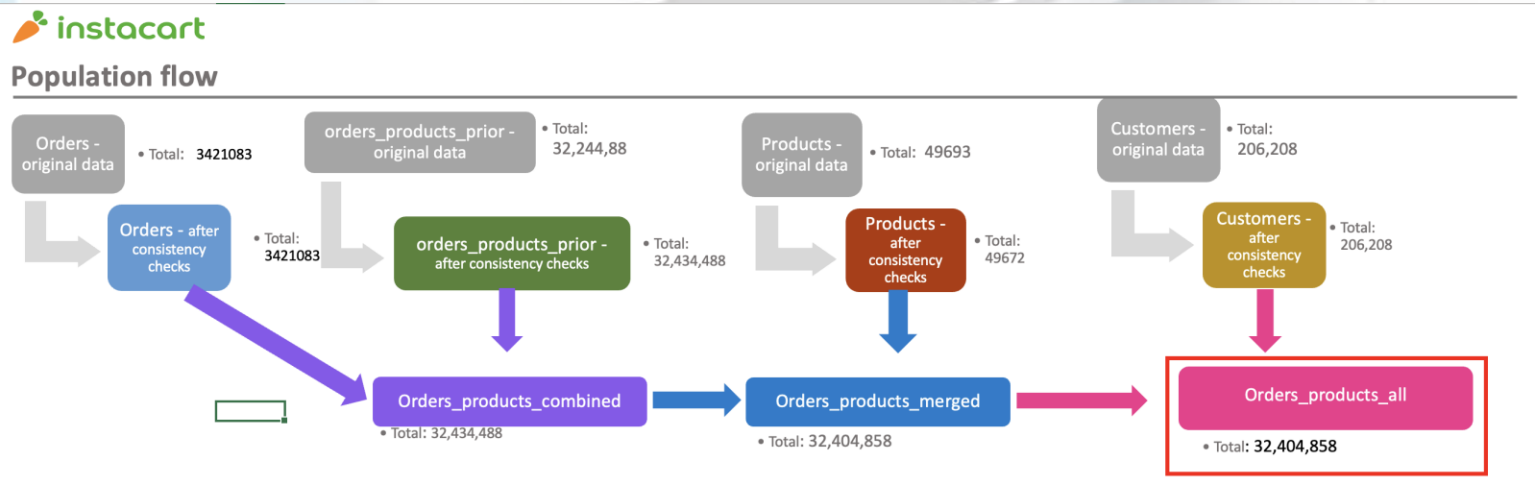
DATASET FROM INSTACART AND CUSTOMER DATASET FROM CAREERFOUNDRY: “THE INSTACART ONLINE GROCERY SHOPPING DATASET 2017” VIA KAGGLE.

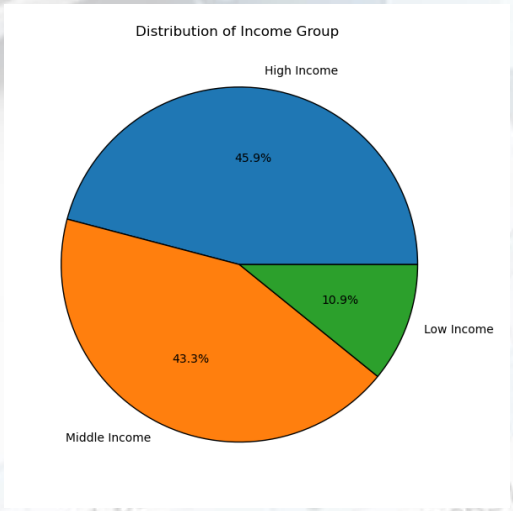
METHODS:

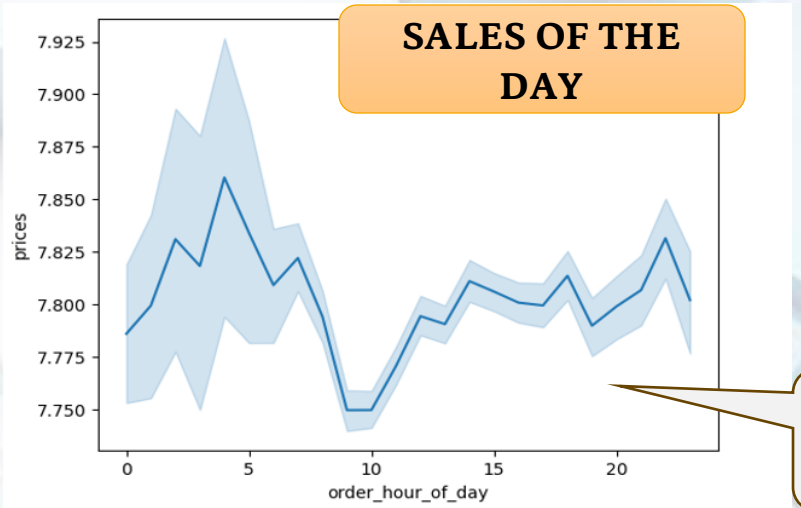
- DATA WRANGLING.
- DATA MERGING.
- GROUPING DATA.
- AGGREGATING DATA.
- EXCEL REPORTING & POPULATION FLOWS.
- DATA VISUALIZATION WITH PYTHON.

GitHub GITHUB INSTACART REPOSITORY

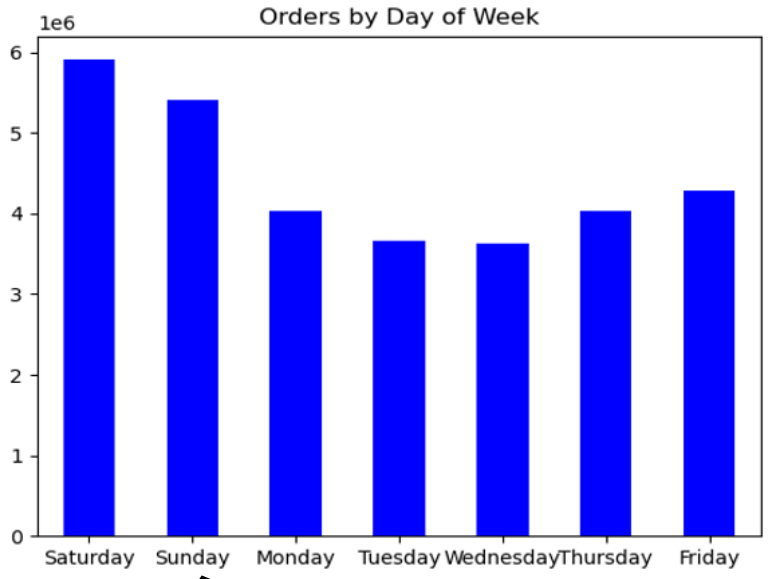
ANALYSIS AND VISUALIZATION







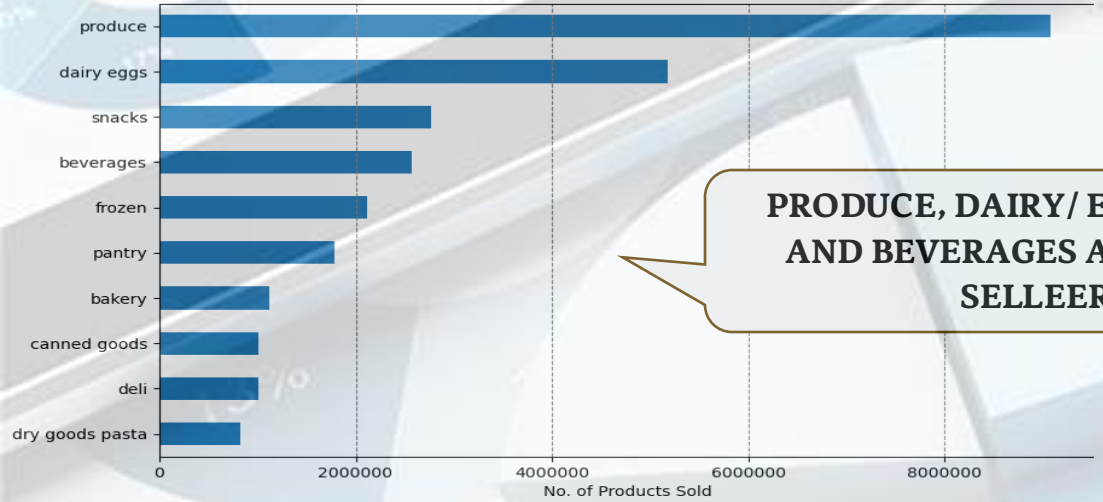
SALES ARE HIGHER IN THE EARLY MORNING, BUT LOWER ORDER VOLUMN AROUND 10 AM.



SATUREDAY AND SUNDAY ARE THE BUSIEST DAY OF THE WEEK. WHEREAS TUESDAY AND WEDNESDAY ARE THE LEAST BUSY DAY OF THE WEEK.

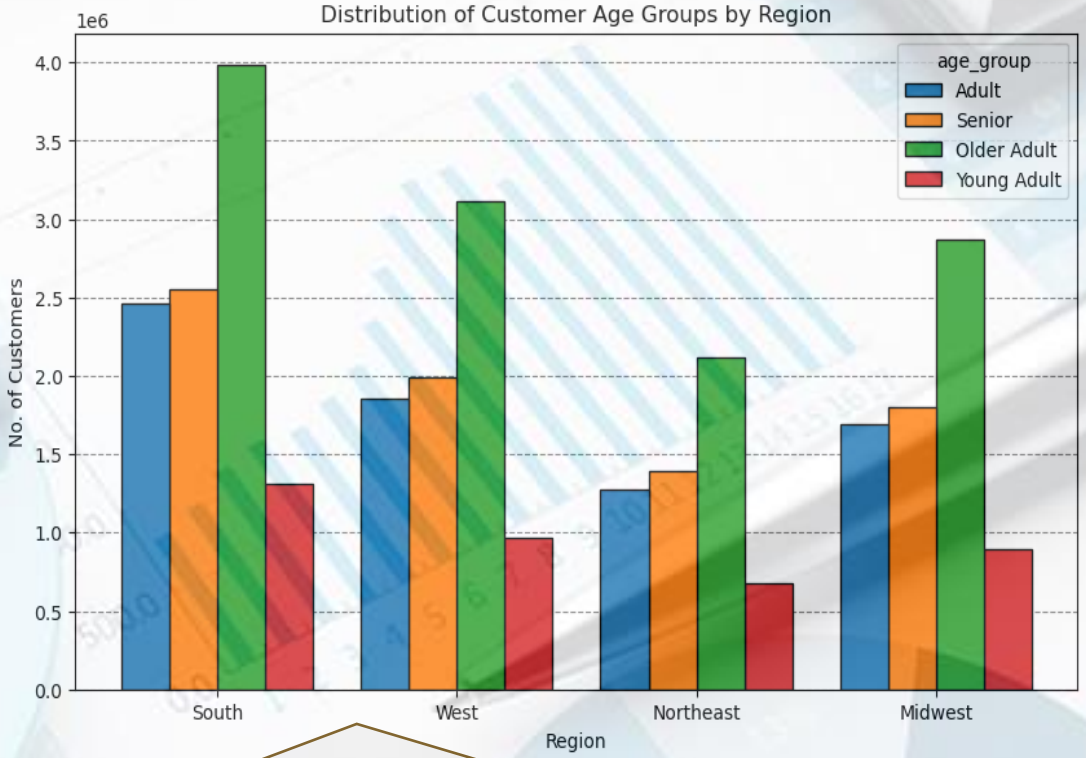
ANALYSIS AND VISUALIZATION

TOP 10 DEPARTMENT BY SALES.



PRODUCE, DAIRY/ EGGS, SNACKS AND BEVERAGES ARE BIGGEST SELLEERS.

MOST CUSTOMERS ARE CONSIDERED TO BE LOW SPENDERS.



ACROSS ALL AGE GROUPS, THE FREQUENCY OF ORDERS REMAINS CONSISTENT ACROSS U.S. REGIONS, SOUTH GENERATE HIGHEST NUMBER OF ORDERS, FOLLOWED BY WEST, MIDWEST AND NORTHEAST.

INSIGHTS AND RECOMMENDATIONS

- INSTACART SHOULD IMPLEMENT SALES AND PROMOTIONS TEND TO BE SLIGHTLY HIGHER IN THE MORNING AND ON WEEKENDS. OFFER COUPONS AND SPECIALS ON POPULAR ITEMS LIKE PRODUCE AND DAIRY CAN ENCOURAGE CUSTOMERS TO PLACE ORDERS. MEANWHILE, PROVIDING DISCOUNTS ON LESS POPULAR ITEMS, SUCH AS INTERNATIONAL PRODUCTS, ALCOHOL, AND BULK GOODS, IS A GOOD STRATEGY TO INCREASE SALES IN THOSE CATEGORIES.

- INSTACART WOULD LIKELY BENEFIT MORE FROM HAVING HIGH SPENDERS RATHER THAN FREQUENT SHOPPERS. WHILE OFFERING DISCOUNTS ON THE ENTIRE GROCERY ORDER COULD INCREASE THE AMOUNT SPENT PER ORDER, IT MIGHT ALSO REDUCE HOW OFTEN CUSTOMERS SHOP, AS THEY MAY BE LESS MOTIVATED TO ORDER FREQUENTLY IF THEY ONLY SHOP WHEN DISCOUNTS ARE AVAILABLE.

- INCREASE ADVERTISING AND PROMOTIONS IN REGIONS WHEN THERE ARE FEWER CUSTOMERS. PRIORITIZE ADVERTISING IN THE SOUTH WHILE TARGETING GROWTH OPPORTUNITY IN WEST AND MIDWEST.

EXCEL FINAL REPORT

PIG E. BANK

OVERVIEW:

PIG E. BANK, A PROMINENT GLOBAL FINANCIAL INSTITUTION, IS SEEKING ANALYTICAL SUPPORT FOR ITS ANTI-MONEY LAUNDERING COMPLIANCE DEPARTMENT.

OBJECTIVES:

- IDENTIFY THE RISK FACTOR CONTRIBUTING TO CLIENT LOSS USING DECISION TREE.
- PROVIDE RECOMMENDATION TO INCREASE CUSTOMER RETENTION.

DATA:

PROVIDED BY CAREERFOUNDRY (PIGEBANK-CLIENT-DATA)

METHODS:

- EXPLORATORY DATA ANALYSIS
- DATA MINING TECHNIQUES
- DATA CLEANING
- DATA ETHICS
- CONSISTENCY CHECKS
- DATA MODELLING
- LINEAR REGRESSION

GitHub GITHUB PIG E. BANK REPOSITORY

ANALYSIS AND VISUALIZATION

COMPARISION BY ACTIVITY

COUNT OF ACTIVE MEMBERS	Exited Customers	Active Customers
0	70.10%	43.84%
1	29.90%	56.16%

COMPARISION BY COUNTRY

COUNTRY	Exited Customers	Active Customers
France	37.75%	51.21%
Germany	36.76%	23.13%
Spain	25.49%	25.67%

COMPARISION BY GENDER

GENDER	Exited Customers	Active Customers
0	59.31%	43.38%
1	40.69%	56.62%

COMPARISION BY AGE

AGE	Exited Customers	Active Customers
18-27	1.47%	11.82%
28-37	17.16%	46.89%
38-47	43.14%	29.48%
48-57	24.51%	6.99%
58-67	13.24%	3.18%
68-77	0.49%	1.27%
78-87		0.38%

COMPARISION BY PRODUCT

NO. OF PRODUCT	Exited Customers	Active Customers
1	69.61%	29.49%
2	15.69%	66.19%
3	13.73%	1.20%
4	0.98%	3.13%

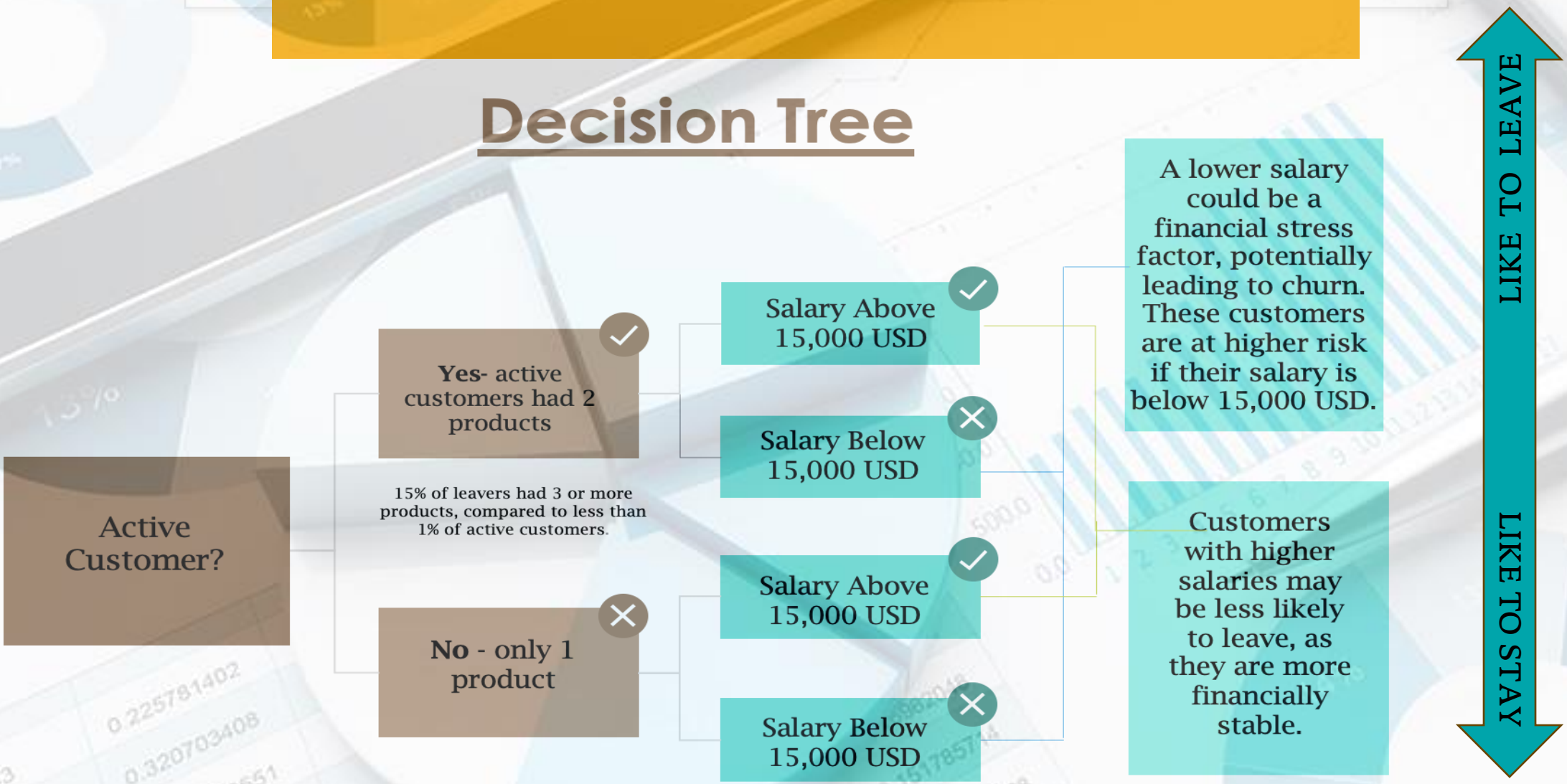
MOST OF THE CUSTOMER'S WHO LEFT THE BANK ARE BETWEEN THE AGE OF 38 AND 47 IN THE BANK. MOST OF THE CUSTOMERS WHO LEFT ARE FEMALE, ONLY HAD ONE PRODUCT AND LOCATED IN FRANCE.

RAW DATA

RISK FACTOR

ANALYSIS AND VISUALIZATION

Decision Tree



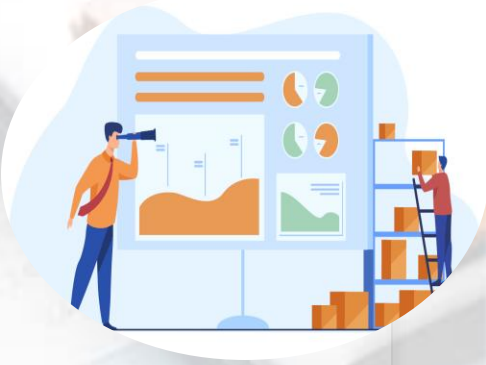
INSIGHTS AND RECOMMENDATIONS

TO IMPROVE CUSTOMER'S RETENTION, PIG. E BANK SHOULD FOCUS ON BULIDING A BASE OF ACTIVE CUSTOMER WHO USE MULTIPLE PRODUCT BY PROMOTING ENGAGEMENT ANF ACCOUNT ACTIVITY.

CONDUCT A SURVEY TO UNDERSTAND THE REASON BEHIND CUSTOMER'S INACTIVITY IN THE BANK AND PROPOSE SOME CHANGES OR NEW PRODUCT BASED ON THEIR FINANCIAL NEEDS THAT COULD INCREASE THEIR ENGAGEMENT OR TRUST.

FURTHERMORE, IMPLEMENTING CROSS-SELLING INITIATIVES CAN STRENGTHEN CUSTOMER LOYALTY BY FOSTERING GREATER ENGAGEMENT WITH THE BANK'S SUITE OF PRODUCTS AND SERVICES. THIS APPROACH NOT ONLY DEEPENS CUSTOMER RELATIONSHIPS BUT ALSO PLAYS A CRITICAL ROLE IN REDUCING ATTRITION RATES.

RETAIL STORE INVENTORY FORECASTING



OVERVIEW:

A DATASET FOR PRACTICING INVENTORY MANAGEMENT AND DEMAND FORECASTING. IT CONTAINS DAILY DATA ACROSS MULTIPLE STORES AND PRODUCTS, INCLUDING ATTRIBUTES LIKE SALES, INVENTORY LEVELS, PRICING, WEATHER, PROMOTIONS, AND HOLIDAYS.

OBJECTIVE:

1. PREDICT DAILY PRODUCT DEMAND ACROSS STORES USING HISTORICAL SALES AND INVENTORY DATA.
2. OPTIMIZE INVENTORY LEVELS BY ANALYZING SALES TRENDS AND MINIMIZING STOCKOUTS WHILE REDUCING OVERSTOCK SITUATIONS.
3. DEVELOP A PRICING STRATEGY BASED ON DEMAND, COMPETITOR PRICING, AND DISCOUNTS TO MAXIMIZE REVENUE.

DATA:

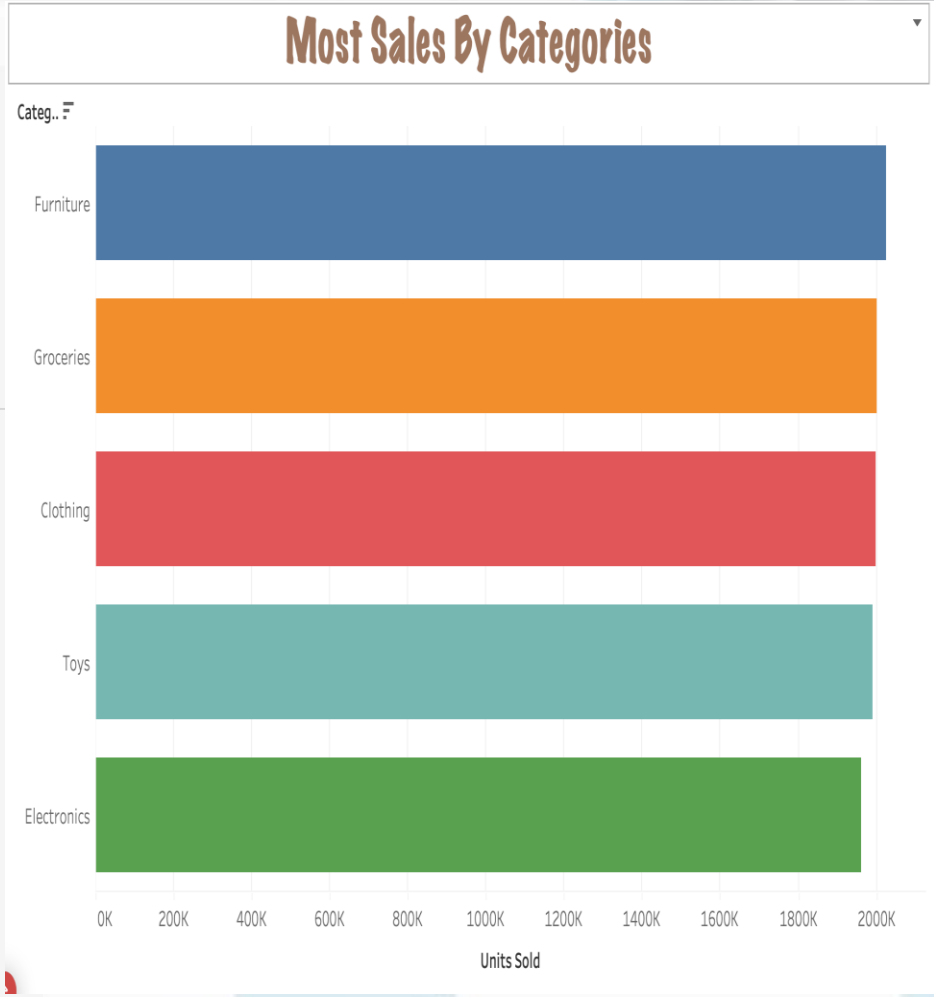
RETAIL STORE INVENTORY FORECASTING DATASET FROM KAGGLE

METHODS:

- EXPLORATORY ANALYSIS THROUGH VISUALIZATION(SCATTERPLOTS, CORRELATION HEATMAP, PAIR PLOT, CATEGORIAL PLOT).
- GEOSPATIAL ANALYSIS USING SHAPEFILE.
- REGRESSION ANALYSIS
- CLUSTER ANALYSIS
- TIME-SERIES ANALYSIS
- FINAL RESULTS.

GITHUB FORECASTING REPOSITORY

ANALYSIS AND VISUALIZATION



ANALYSIS AND VISUALIZATION

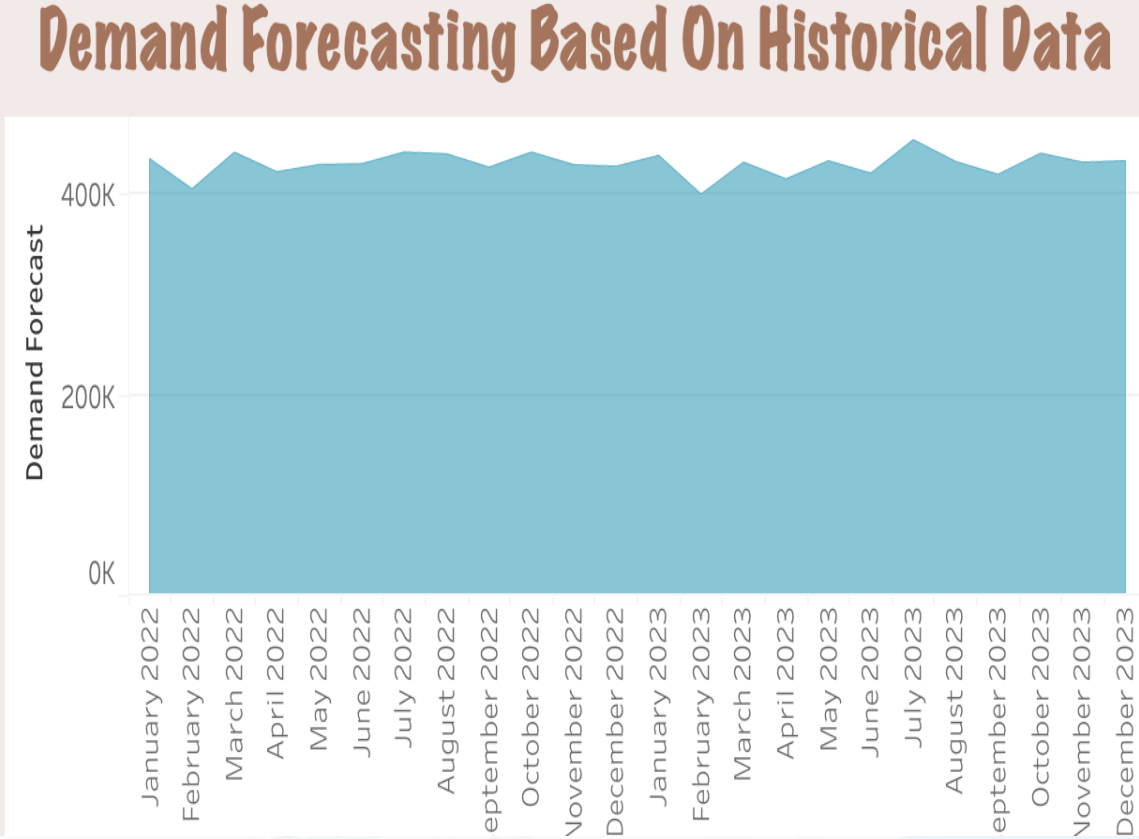
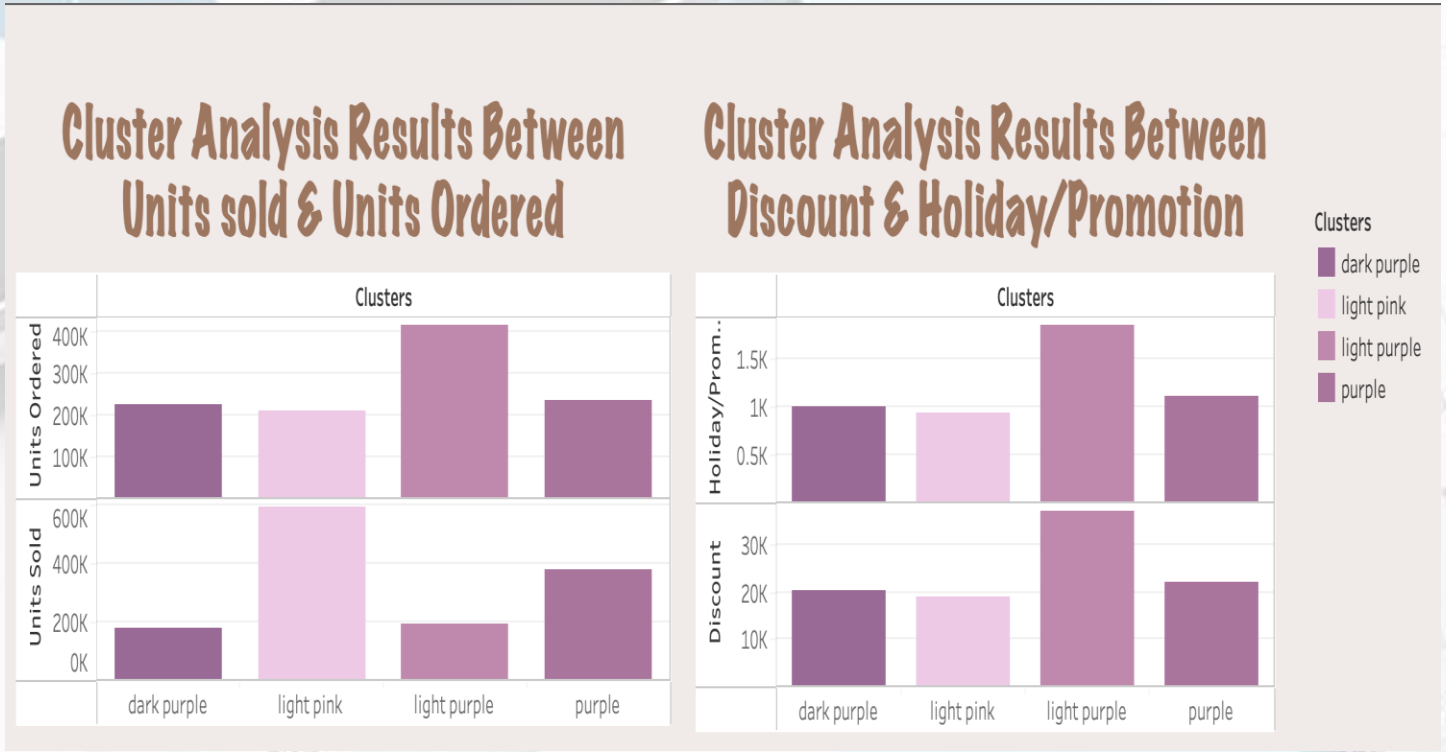


TABLEAU DASHBOARD LINK

RAW DATA

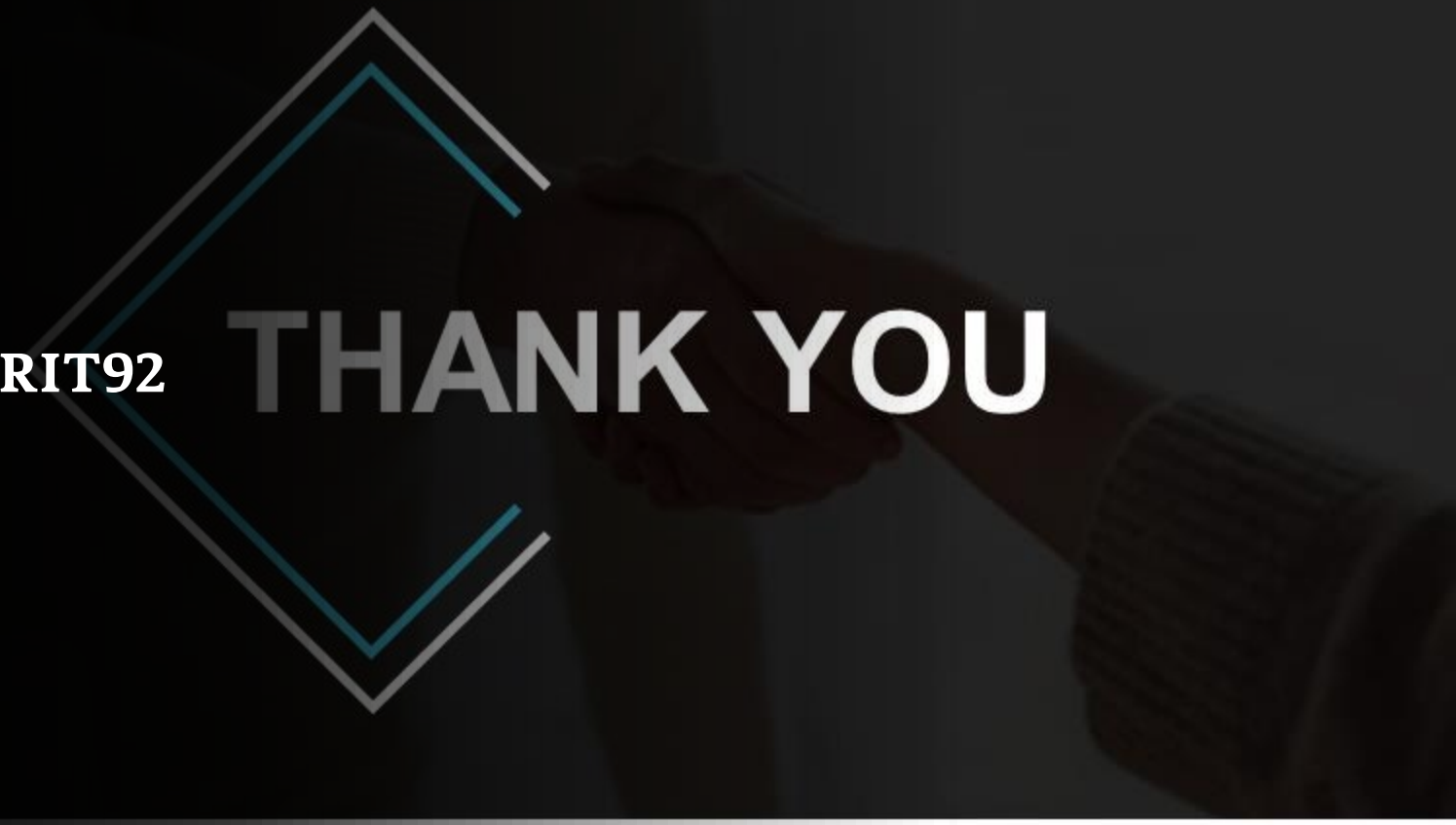
INSIGHT AND RECOMMENDATIONS



KEY FINDINGS AND SUGGESTIONS:

- ANALYSIS HAVE SHOWN THAT CORRELATION BETWEEN VARIABLES LIKE DISCOUNTS AND HOLIDAY PROMOTIONS ARE NOT INCREASING SALES IN RETAIL STORE AS MUCH AS THEY LOOKS LIKE, WHICH LEADS US TO THINK WE NEED TO TAKE EXTERNAL FACTORS THAT MIGHT INFLUENCE THE SALES IN RETAIL STORES.
- DUE TO DATA LIMITATION(WE ONLY HAVE DATA FROM JANUARY 2022 TO JANUARY 2024, WE ARE NOT ABLE TO PREDICT FUTURE SALES DURING SEASIONALITY.
- WE SHOULD SUGGEST TO PROVIDE ADDITIONAL PROMOTIONS OR COUPONS TO THOSE AREAS WHERE SALES ARE LESSER THAN OTHER AREAS.

MY CONTACT

GILLAMRIT03@OUTLOOK.COM

HTTPS://GITHUB.COM/GILLAMRIT92