Data Analytics portfolio

Kanchan Malge

Projects

01. Game Co

Analyzing global video game sales.

02. Influenza Preparation

Planning for the upcoming flu season..

03. Rockbuster

Creating a strategy for the launch of a new online video service. .

04. Instacart

Analyzing sales data to uncover sales patterns..

05. Global Bank

Provide analytical support to its anti-money-laundering compliance department..

06. World Happiness Report(2015-2019)

Aims to assess and understand the state of global happiness and well-being.

01. GAMECO

Game Co

Tools Used:

Microsoft Excel

Microsoft Powerpoint

Objectives:

To perform a Descriptive Analysis of a video Game Data Set to assist GameCo in predicting how their new games will perform in the market
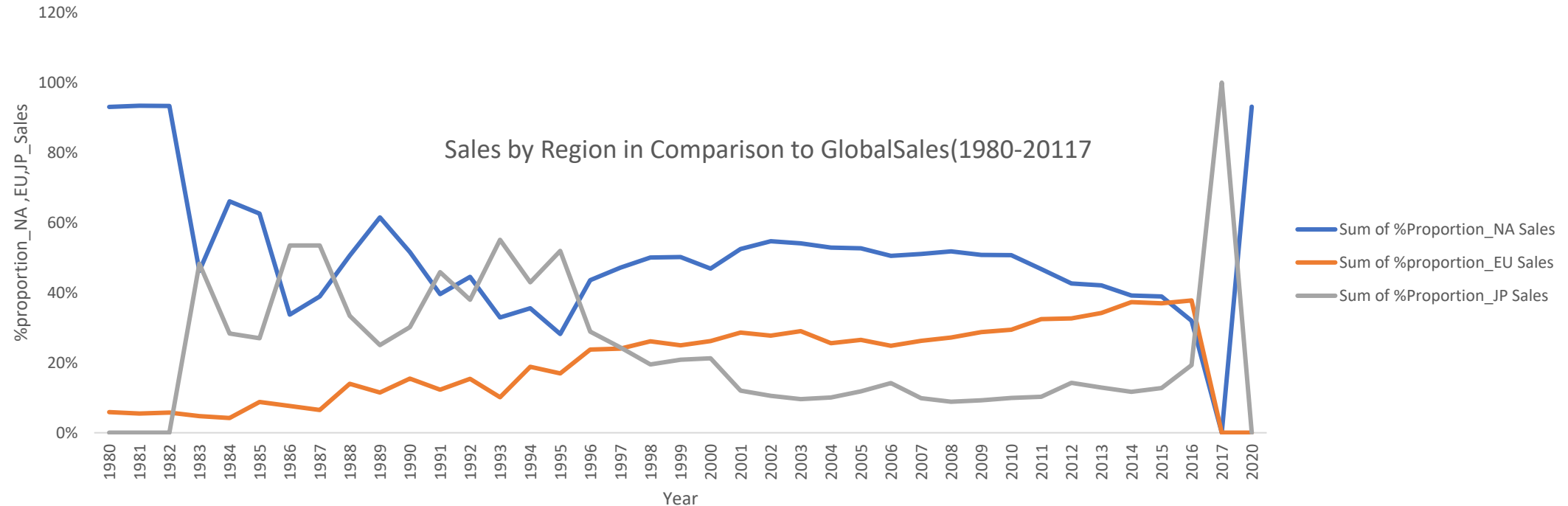
Skills Used:

Grouping Data

Summarizing Data

Descriptive Analysis

Visualizing Results

Presenting Results

GAME CO

Sales Across Geographical Regions

Sales by Region in Comparison to GlobalSales(1980-20117)

- Sum of %Proportion_NA Sales
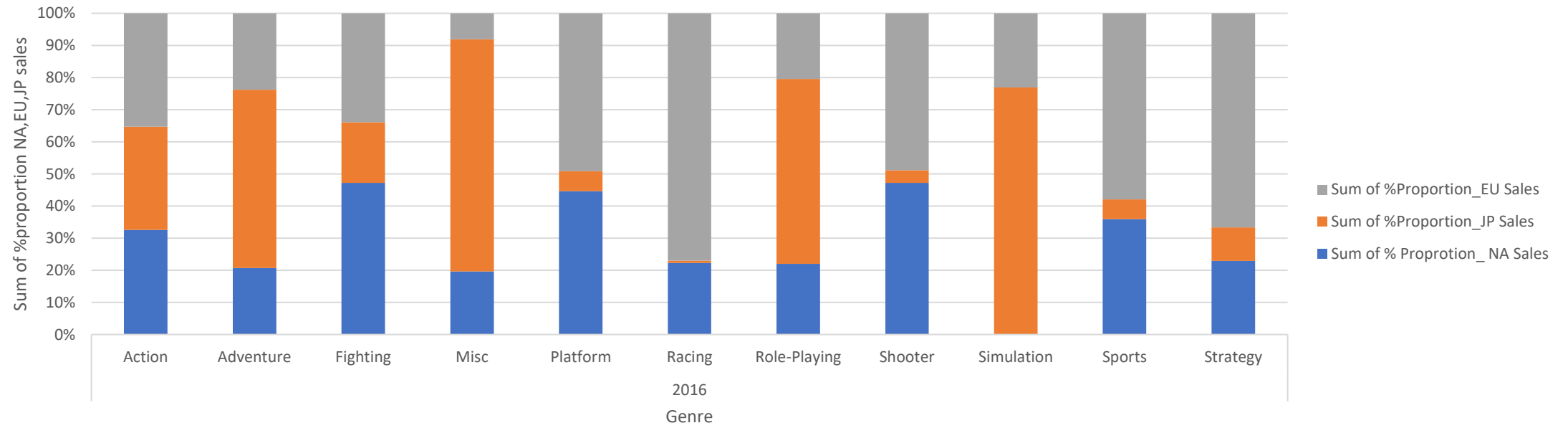- Sum of %proportion_EU Sales
- Sum of %Proportion_JP Sales

Each region does not stay constant over the years as you see in the graph. NA Sales(North Ameica) have the highest sales. JP Sales(Japanese) have the lowest sales. From the year 1995-2014, JP Sales gradually decreased.
This analysis can help identify areas for improvement, adjust strategies, and make more informed decisions for future planning.

GAME CO

Sales By Genre & Geographic Region 2016

Legend:
- Sum of %Proportion_EU Sales
- Sum of %Proportion_JP Sales
- Sum of % Proprotion_ NA Sales

Y-axis: Sum of %proportion NA,EU,JP sales (0% to 100%)

X-axis categories (2016 Genre): Action, Adventure, Fighting, Misc, Platform, Racing, Role-Playing, Shooter, Simulation, Sports, Strategy

2016
Top Selling Genre NA Sales – Shooter & Fighting
Top Selling Genre EU Sales – Racing
Top Selling Genre JP Sales - Simulation

Final Results

European Market as largest market	Genre differences	North American market decreasing in sales

- Adjust the marketing schedule to reflect this change in sales.

- Japanese market also increasing, investigation of how to better capture that market.

- Adjust marketing to reflect that the NA and EU markets sell differently than the JP market.

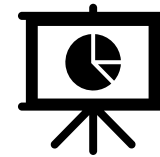- More research into Misc genre as it is top performing.

- More research as to why the NA market is losing market share.

- Advertising or marketing budget to help increase sales.

GAME CO

Project Deliverables
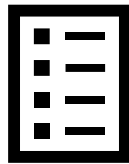
Project Brief

Project Presentation

02. INFLUENZA PREPARATION

Influenza Preparation

Tools Used:
Microsoft Excel
Microsoft Powerpoint
Tableau

Objectives:
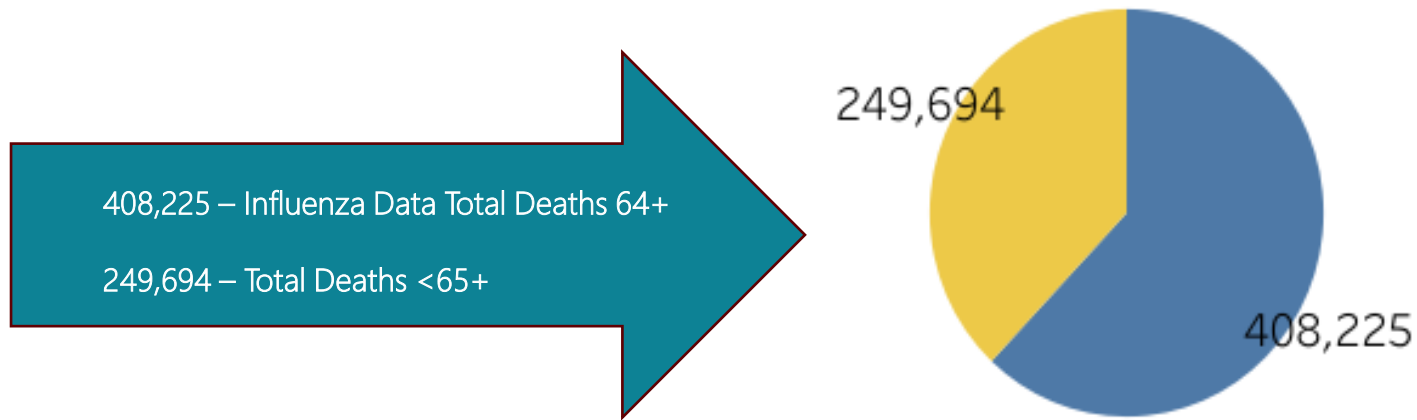Determine when to send staff, and how many, for each state

Skills Used:
Data Cleaning
Data Integration
Data Transformation
Descriptive Analysis
Forecasting
Visualization
Storytelling with Tableau

Influenza Preparation

Influenza Related Deaths

408,225 – Influenza Data Total Deaths 64+

249,694 – Total Deaths <65+

249,694

408,225

Knowing that those over 65 had a significantly higher probability of dying meant that we needed to determine which states had the most residents in this age range.

Influenza Preparation

Influenza Related Deaths

Heat Map Of Influenza Rate By State

To see if population and mortality were related, we also wanted to compare the total number of deaths by state. The majority of deaths occur in the top states(Georgia,Texas, Virginia,Kentucky, Missouri, Illinois, North California, Iowa, India.) for people over 65.

Influenza Preparation

Vulnerable Population By State(Top 10)

Vulnerable Population By State(Top 10)
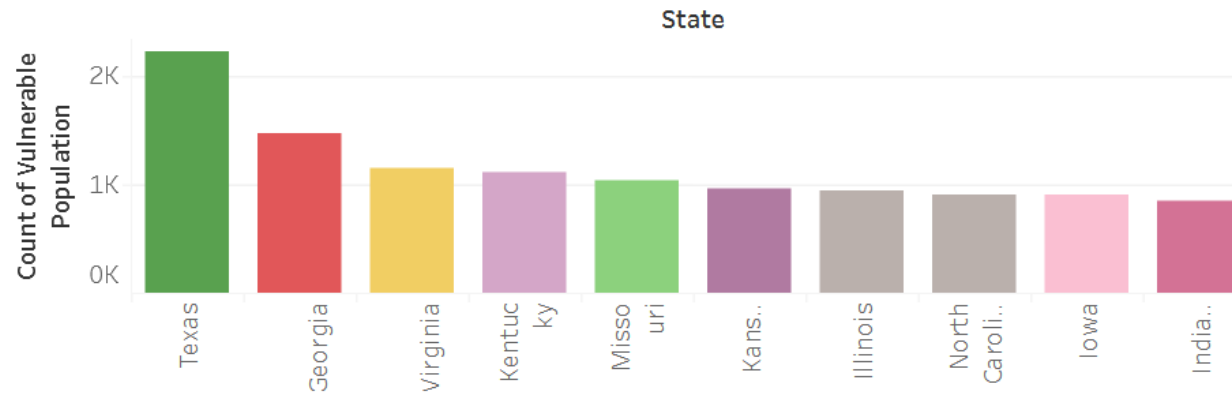
The Top 10 states with the highest counts allow stakeholders to prioritize resource allocation healthcare planning & and targeted interventions in areas with the highest concentration of vulnerable individuals.

Influenza Preparation

## Final Results

People over 65 at a
greater risk of death

Target states with a high population over 65Target
states with a high
population over 65

States with larger populations of people 65+ should
take top priority when sending out staff.

- Texas, Georgia, Virginia, Kentucky and Missouri
have the highest populations of people over 65.

- These states need to be made a priority when
sending staff.

Influenza Preparation

Project Deliverables
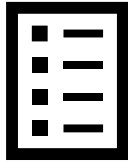
Final Report

Tableau Flu Dashboard

03.ROCKBUSTER

Rockbuster

Tools Used:

Microsoft Excel

Microsoft Powerpoint

Objectives:

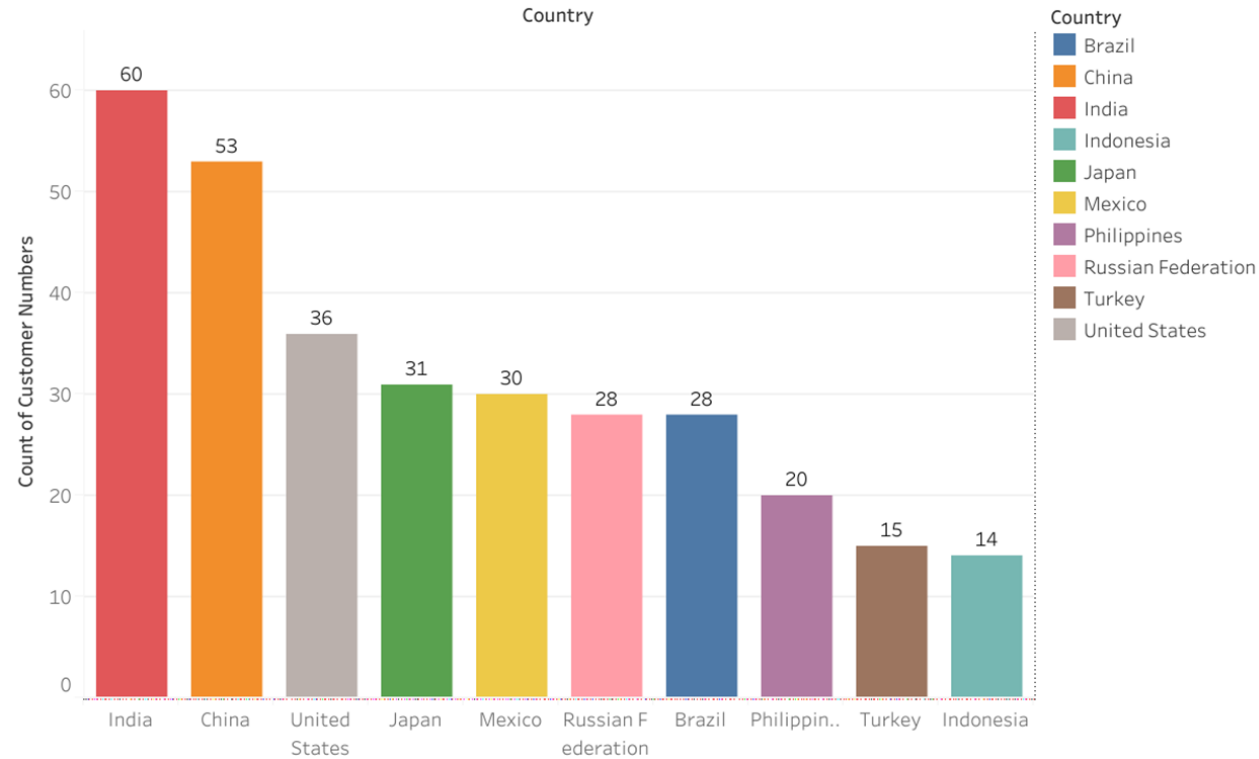To help with the launch strategy for the new online service

Skills Used:

Relational Database

Database Querying

Filtering

Cleaning & summarizing

Joining Tables

Subqueries

Common Table Expressions

Rockbuster

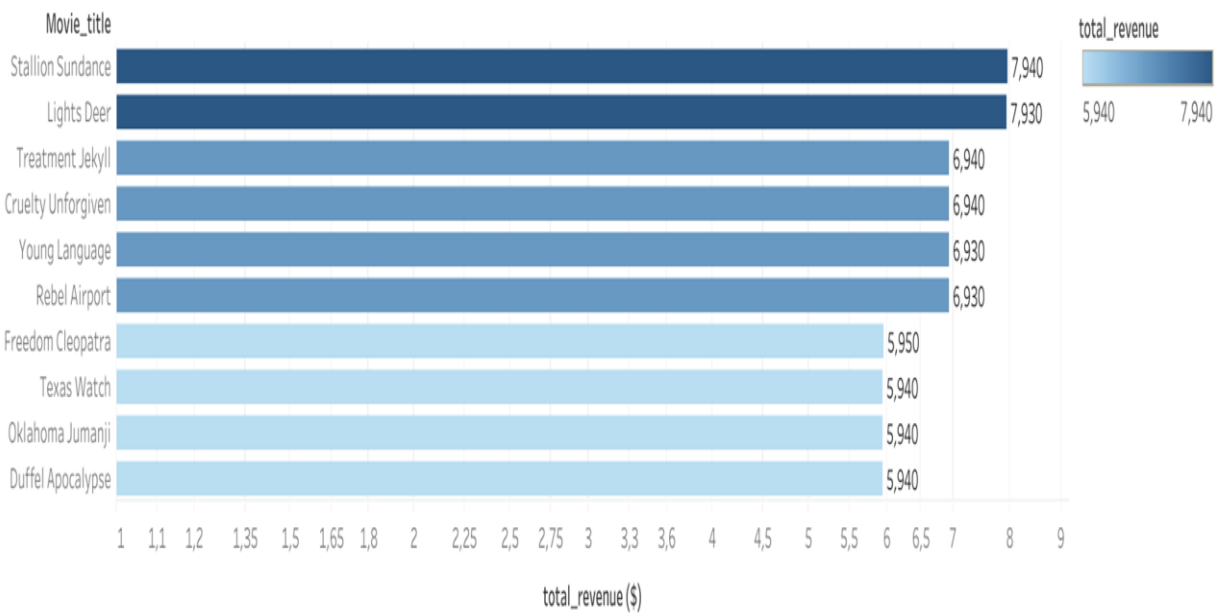## Customer By Countries

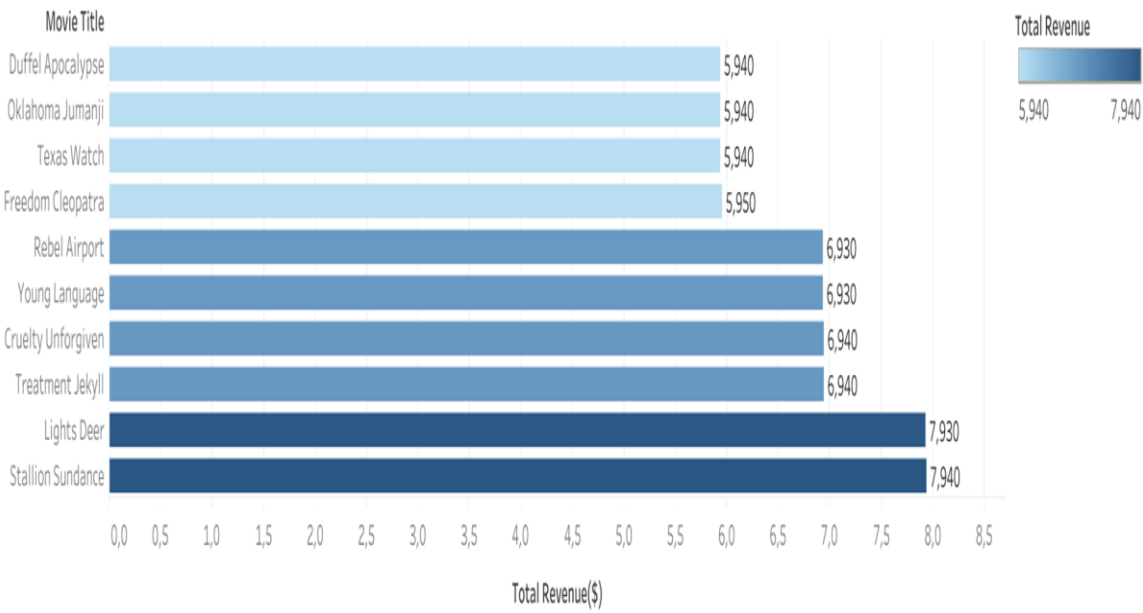Top 10 Countries for Customer Numbers

The analysis shows that most customers are located in India.

Rockbuster

Most/Least Revenue Gain

Top 10 Movie Title contributed the most Revenue Gain

Movie_title	total_revenue
Stallion Sundance	7,940
Lights Deer	7,930
Treatment Jekyll	6,940
Cruelty Unforgiven	6,940
Young Language	6,930
Rebel Airport	6,930
Freedom Cleopatra	5,950
Texas Watch	5,940
Oklahoma Jumanji	5,940
Duffel Apocalypse	5,940

total_revenue
5,940   7,940

total_revenue ($)

Top 10 Movie Title that contributed the Least Revenue Gain

Movie Title	Total Revenue
Duffel Apocalypse	5,940
Oklahoma Jumanji	5,940
Texas Watch	5,940
Freedom Cleopatra	5,950
Rebel Airport	6,930
Young Language	6,930
Cruelty Unforgiven	6,940
Treatment Jekyll	6,940
Lights Deer	7,930
Stallion Sundance	7,940

Total Revenue
5,940   7,940

Total Revenue($)

The Most Revenue Earning Title is
Stallion Sundance & Least
Revenue Earning Value is $5,940.

Rockbuster

Project Deliverables

Excel Report

Data
Dictionary

Presentation

04.INSTACART

Instacart

Tools Used:
Microsoft Excel

Microsoft PowerPoint

Python

Panda

Jupyter Notebook

Objectives:
To perform an initial data and exploratory analysis of the data from the online grocery store, Instacart, to derive insights and suggest strategies.
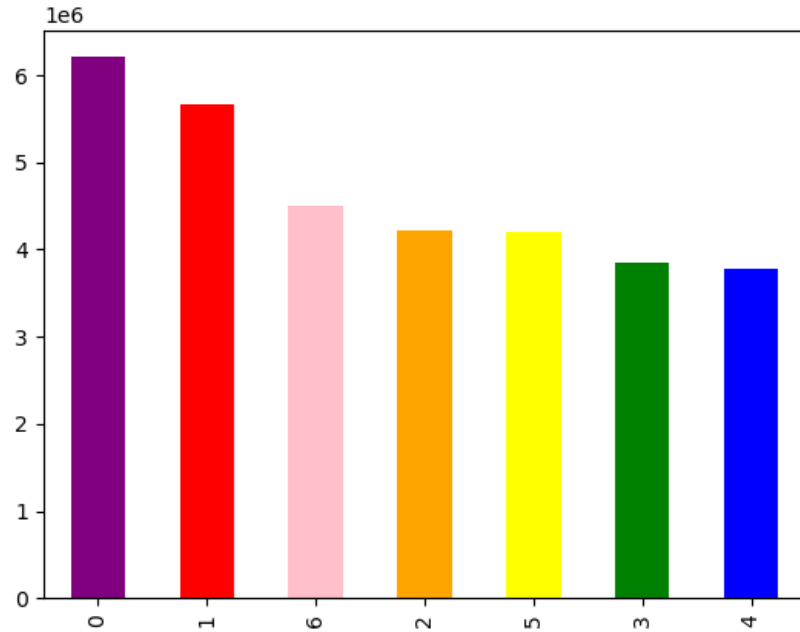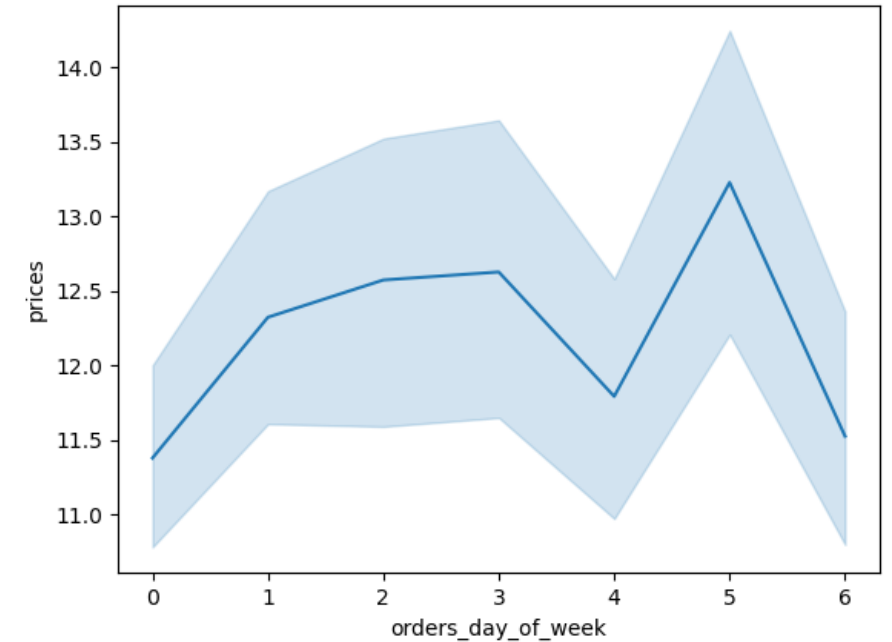
Skills Used:
Data wrangling

Data merging

Deriving variables

Grouping data

Aggregating data

Reporting in Excel

Population flows

Instacart

Frequency by Day of week & Hour of Day

Days of Week
   Legend
0 - Saturday
1 - Sunday
2 - Monday
3 - Tuesday
4 - Wednesday
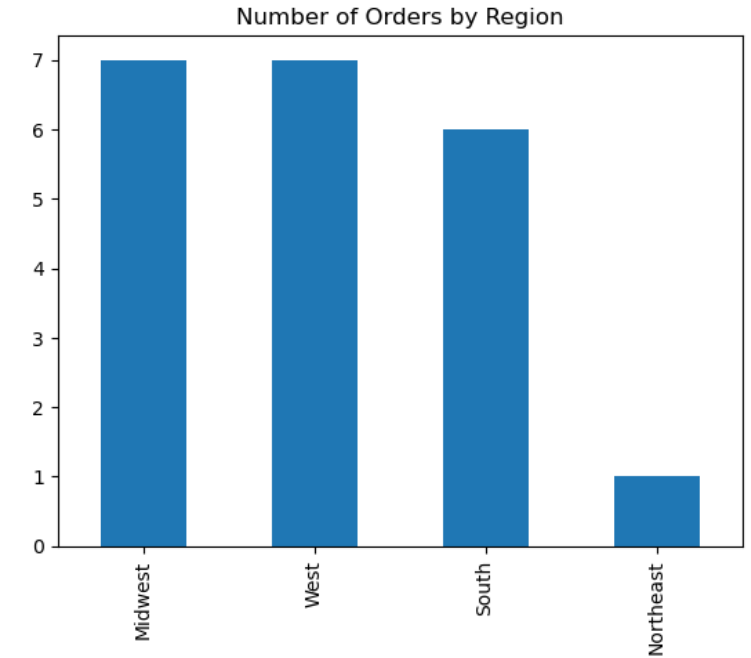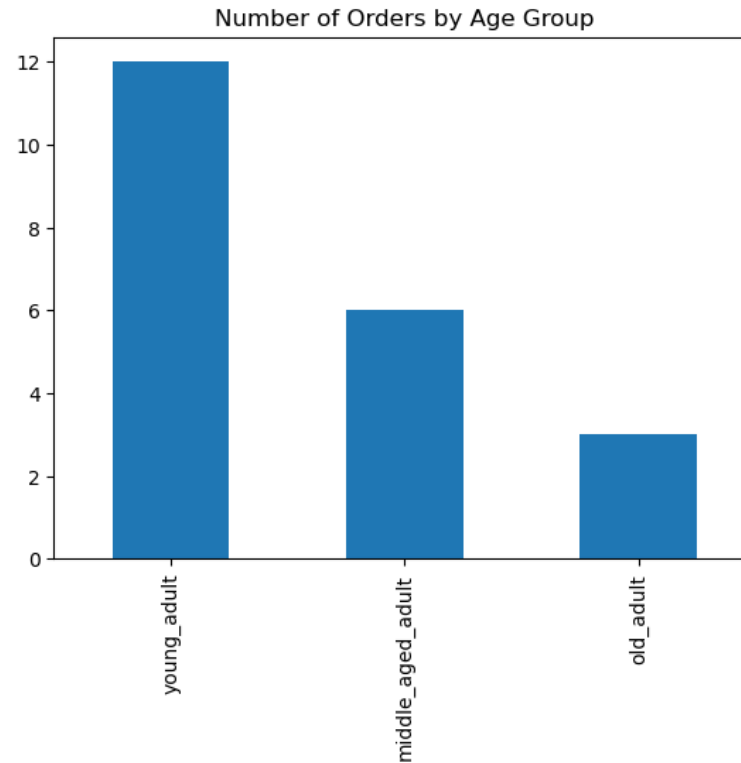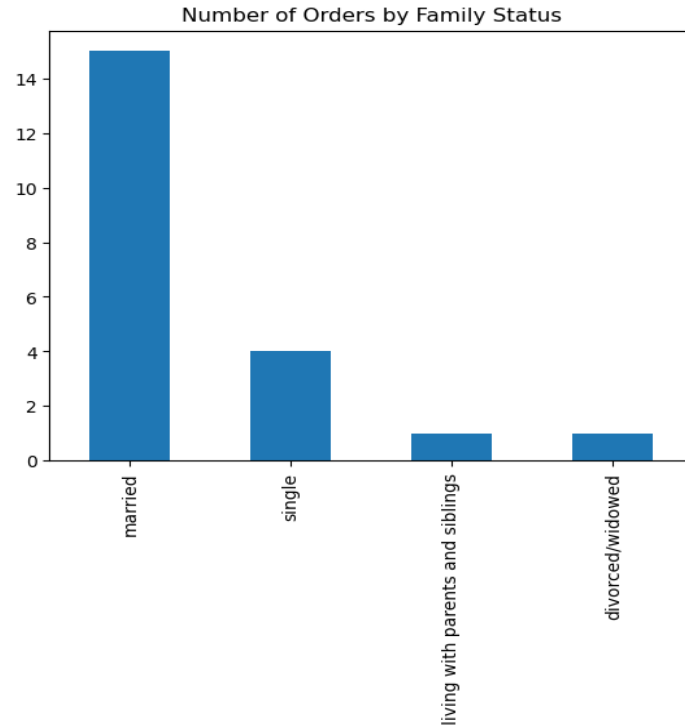5 - Thursday
6 - Friday

Saturday has the most orders, follow by Sunday and then Friday. Wednesday has the least followed by Tuesday and Monday.

Customers are spending more on Tuesday and Thursday.

Instacart

Frequency by Marital Status, Age & Region

We also wanted to consider who the primary clients were and where they were situated. We divided the nation into regions in accordance with Wikipedia, and then we divided the consumer base by age, whether they were parents or childless. This will make it easier to decide who and where to market certain products to.

Instacart

Final Results

Certain days and times have more sales than others

Young-Adult-People and Midwest are the biggest customer

- Mon-Wed have the lowest amount of orders and should be prioritized for advertising
- Ads should also be considered for the morning hours to encourage sales during the slow times.

- Talk to marketing about a new customer program for the Northeast and to childless people to improve the customer base.
- Target child products only to those with children.

Instacart

Project Deliverables

Excel Report

Python Scripts

05.GLOBAL BANK

Global Bank

Tools Used:
Microsoft Excel

Microsoft Powerpoint

GitHub

Objectives:
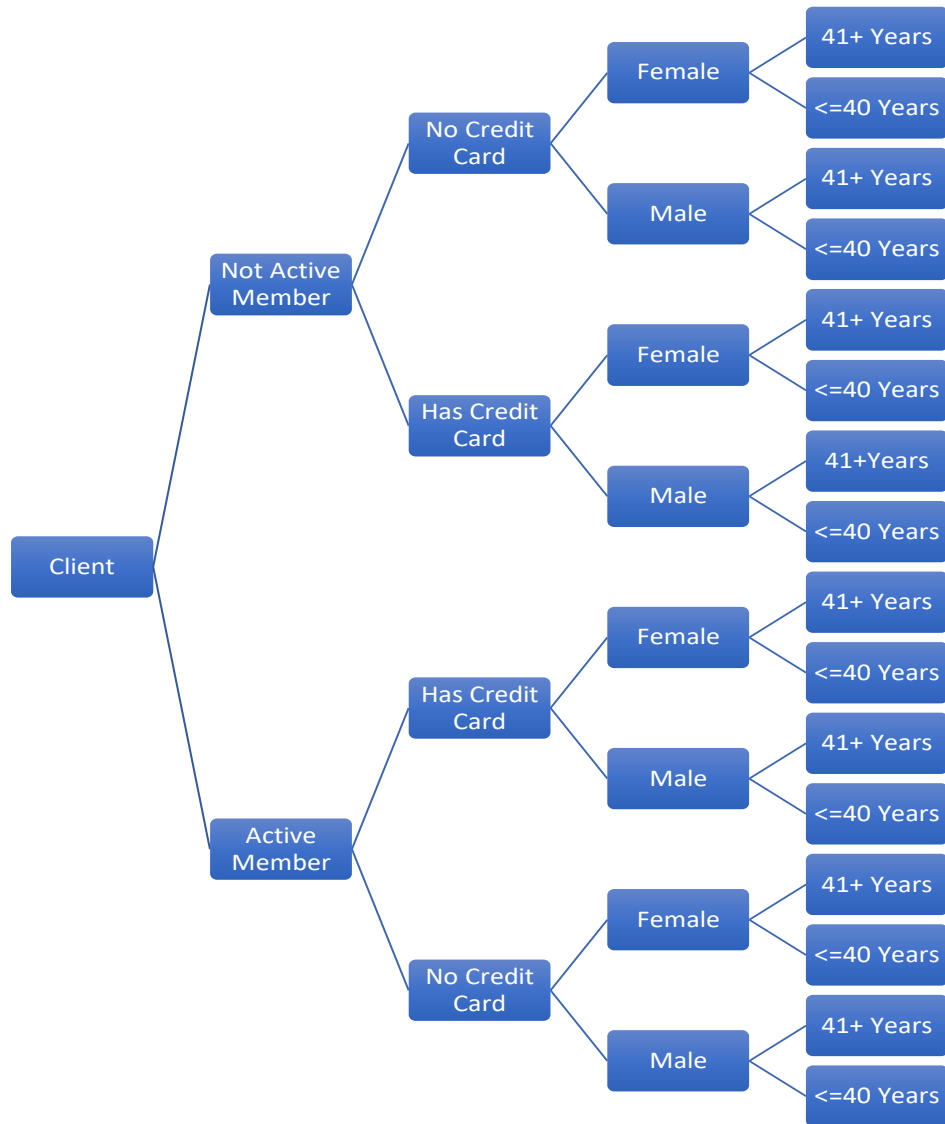To identify any indicators that would explain why customers are leaving the bank.

Skills Used:
Big data

Data Ethics

Data Mining

Predictive analysis

Time series analysis and forecasting

Global Bank

Decision Tree

```
Client
├── Not Active Member
│   ├── No Credit Card
│   │   ├── Female
│   │   │   ├── 41+ Years
│   │   │   └── <=40 Years
│   │   └── Male
│   │       ├── 41+ Years
│   │       └── <=40 Years
│   └── Has Credit Card
│       ├── Female
│       │   ├── 41+ Years
│       │   └── <=40 Years
│       └── Male
│           ├── 41+Years
│           └── <=40 Years
└── Active Member
    ├── Has Credit Card
    │   ├── Female
    │   │   ├── 41+ Years
    │   │   └── <=40 Years
    │   └── Male
    │       ├── 41+ Years
    │       └── <=40 Years
    └── No Credit Card
        ├── Female
        │   ├── 41+ Years
        │   └── <=40 Years
        └── Male
            ├── 41+ Years
            └── <=40 Years
```

High Probability

Low Probability

This decision tree shows the propensity for a client to leave Pig E. Bank.

Global Bank

Final Results

Over and under 40	Male or Female	Active or Inactive
Exiting clients are 45 years old on average. Older customers stop using the bank.	Clients tend to leave more frequently for females than for males	Inactive clients leave Pig E. Bank

Global Bank

Project Deliverables

Github

Excel Report

Final Report

06.World Happiness Report(2015-2019)

World Happiness Report (2015-2019)

Tools Used:

Microsoft Excel

Microsoft Powerpoint

Tableau

Python

Objectives:

Aims to assess and
understand the state of global
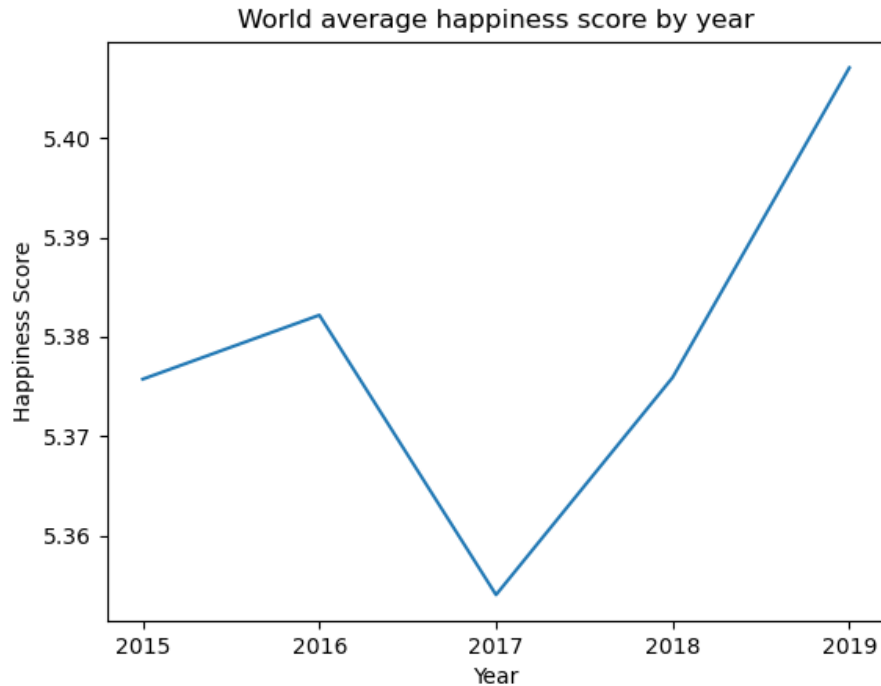happiness and well-being.

Skills Used:

Data Collection

Data Preprocessing

Data Visualization

Data Analysis

Machine Learning

Statistical Testing

Linear Regression

Time Series Analysis

Geospatial Analysis

Tableau

World Happiness Report(2015-2019)
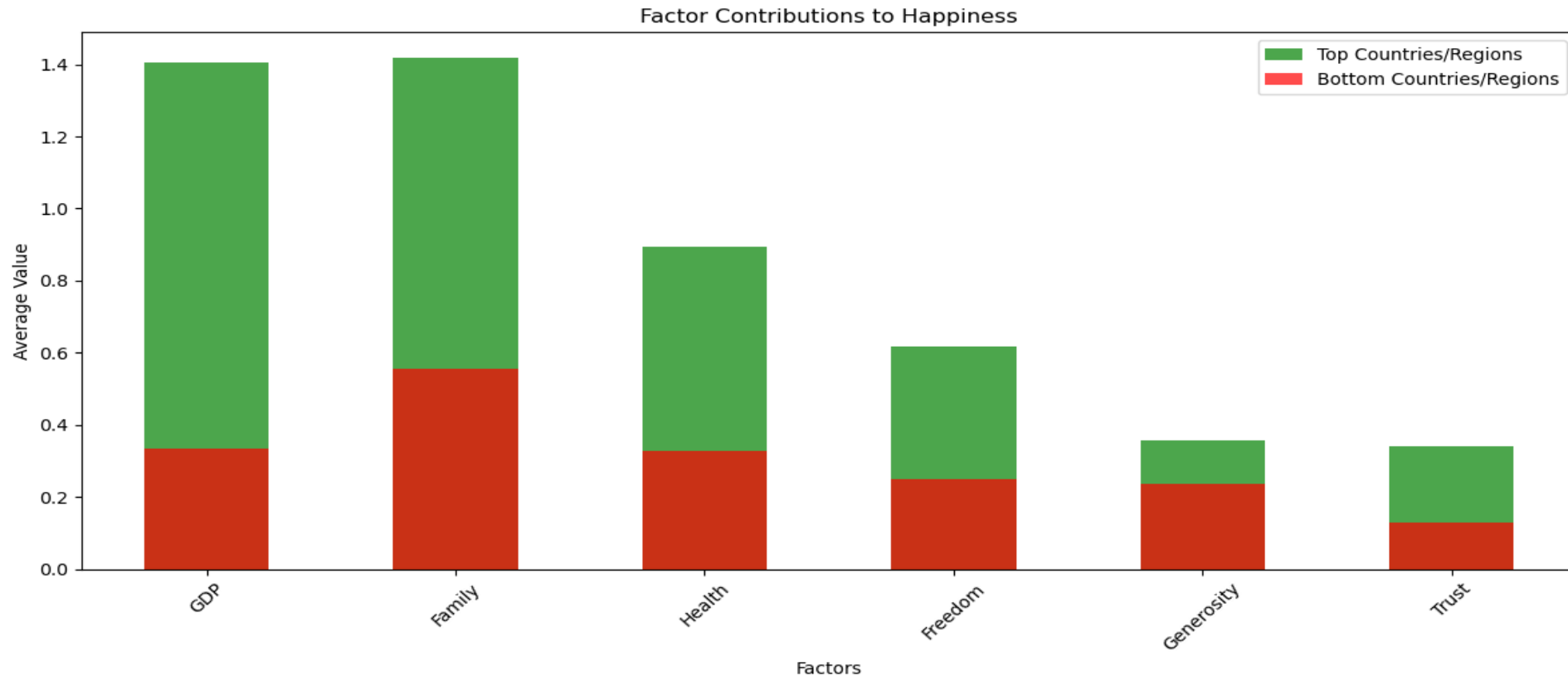
Happiness Scores Overview

By observing the plot, you can see whether there are any noticeable changes or trends in global happiness scores during 5 years.

The scatterplot shows the relationship between the actual scores (y_test) and the predicted scores (y_pred) from our linear regression model. Each point on the plot represents an observation from the test set. The red points represent the actual scores, while the green line represents the regression line, which is the model's best attempt to predict the scores based on the independent variable(s).

World Happiness Report(2015-2019)

Factor Contributions to Happiness

•Green bars represent the average factor values for the top countries/regions with high happiness scores.
•Red bars represent the average factor values for the bottom countries/regions with low happiness scores.

Final Results

What Are the 10 Happiest Countries in the World?

Country	
Australia	10.000
Canada	6.800
Denmark	2.200
Finland	3.600
Iceland	3.200
Netherlands	6.200
New Zealand	8.200
Norway	2.800
Sweden	8.600
Switzerland	3.600

What Are the 10 Unhappiest Countries in the World?

| Country	
Djibouti	126.0
Gambia	120.0
Hong Kong S.A.R..	71.0
North Macedonia	84.0
Northern Cyprus	61.0
Oman	22.0
Puerto Rico	15.0
Somaliland region	94.0
Swaziland	118.0
Taiwan Province ..	33.0

The analysis of the World Happiness Report from 2015 to 2019 provides valuable insights into global well-being trends, regional disparities, and the factors influencing happiness. These findings can serve as a basis for informed decision-making and future research in the field of happiness and well-being.

World Happiness Report(2015-2019)

Project Deliverables

Github
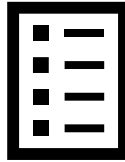
Project
Brief

 Tableau Dashboard

Thank You