



JESUS BARCENAS

DATA ANALYTICS PORTFOLIO

ABOUT ME

I am an Operations and Events coordinator for the Russell House Student Union, and I also serve as the AV/ Tech lead. I have always had a love for tech and found myself drawn to data analytics after a Power BI presentation on our student population. My degree in Marketing/Management and data concentration helped me when starting my journey in developing my skills in data cleaning, wrangling, and integration. I was able to gain skills in descriptive analysis, statistical hypothesis testing, and data visualization. Over the past 8 months while working full time, I've been able to use Excel, SQL, Python, and Tableau to analyze data and create meaningful visualizations/ insights on how data can help improve businesses in all fields.



TECH SKILL SET: PYTHON, TABLEAU, EXCEL, JUPYTER NOTEBOOK, POSTGRE SQL, POWERPOINT, ANACONDA, DBVISUALIZER

PORTFOLIO PROJECTS

GAMESCO

Descriptive analysis on global video game sales to inform development into new games

RockBuster Stealth LLC.

Business Analysis for a movie rental company who needs help with the launch of a new streaming service

P.E. Bank

Leverage data insights to advance the efficiency of an anti-money laundering model, using big data concepts

Influenza Season

Assisting a staffing agency in determining medical staff needs for US regions during influenza season

Instacart

Exploratory analysis on customer database to derive insights on overall customer behavior

Future project



GAMESCO

The executives at a new video game company wants to use data to inform the development of new games. Specifically, they are interested in the popularity of game genres and regional differences in sales.

GAMESCO BACKGROUND

Objective

Perform analysis on historical data to inform future budget for development of new games

Assumptions

Sales for the varying geographic regions have stayed the same over time and budget should stay the same

Skills

Filtering\cleaning data
Pivot tables
Data visualization through excel
Descriptive analysis
Presenting Results

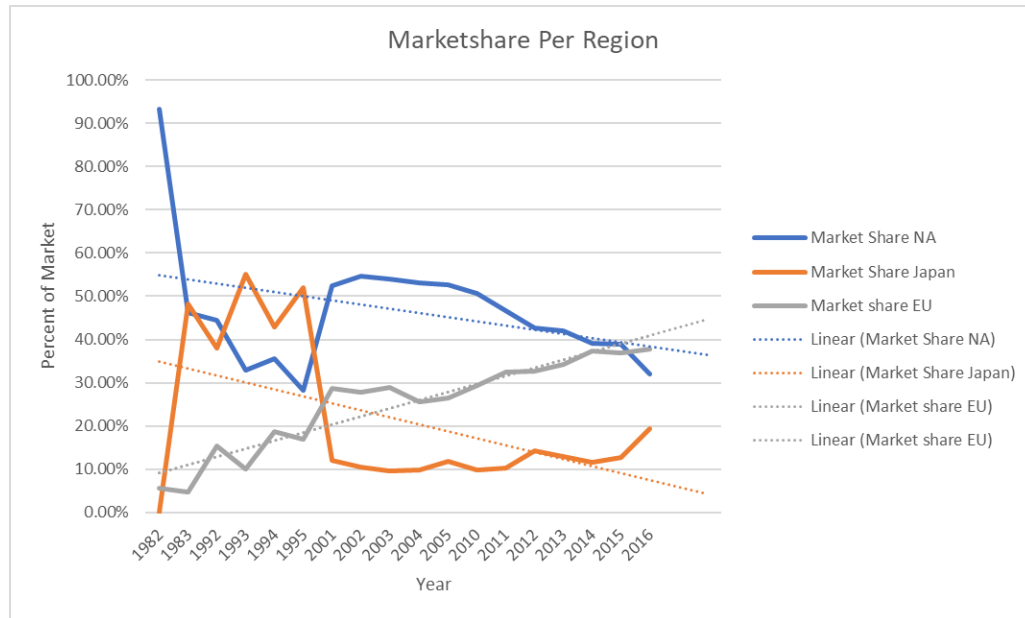
Data source: [Here](#)
Data set: [Here](#)

Tools

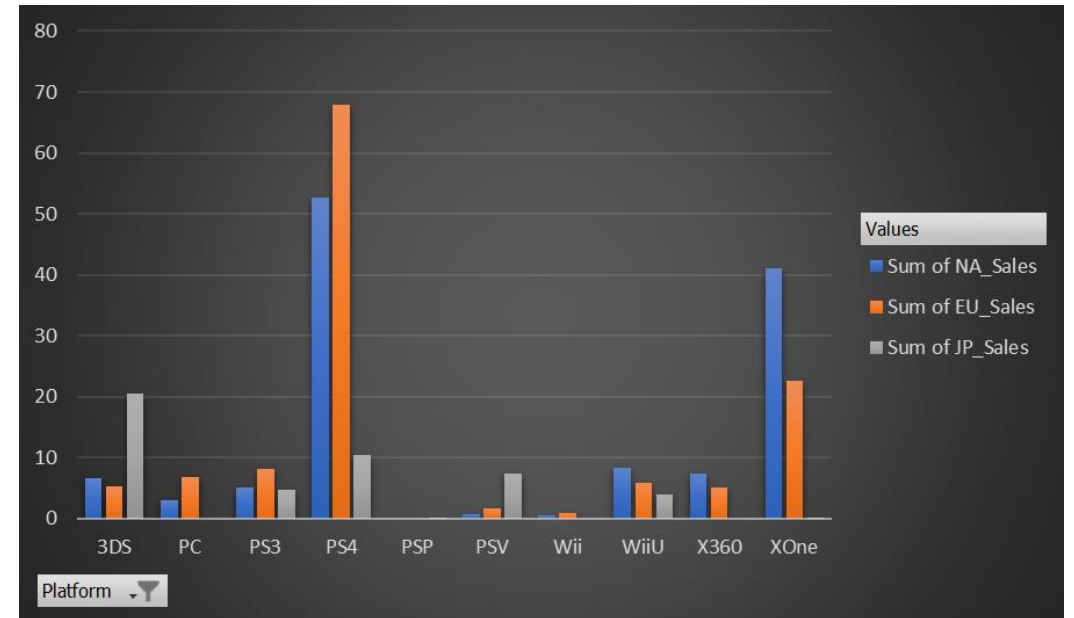


GAMESCO ANALYSIS

Market share per region



Console performance



SUMMARY



Retail Video Game Sales are decreasing

This could be due to the digital marketplace taking from retail sales.



Japan's market share is increasing at a solid rate

Despite historic downward trend, Japan is increasing its market share and has potential



PS4 and Xbox one are top performing consoles

These are the next generation consoles and should be the focus of a majority of our funds



EU has surpassed NA as top performer

EU sales hit 37% of market share while NA dropped to 31%



Different Regions have similar high performing Genres such as sports, shooters, and roleplaying

These genres are our target areas, can be successful in multiple regions.

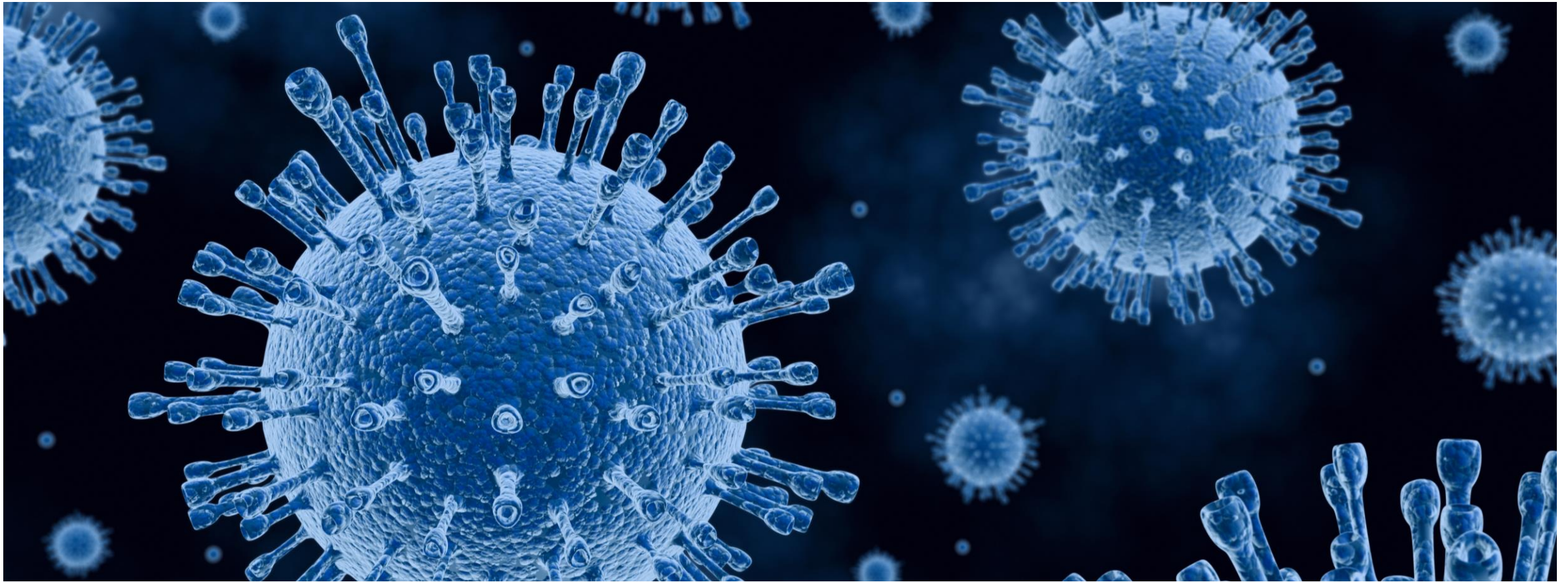


There is still much potential in this space if we target the right areas

The data we have collected will allow us to create a strategy based on past trends and give us the greatest chance at success.

RECOMMENDATION

- Based off the data collected, a 40/40/20 split would be the best way to allocate funding.
- 40% to Europe and North America since they hold the largest market share and 20% to Japan, which is growing but still the smallest market.
- Focus should be developing games on PS4 and Xbox one within the top performing genres for each region. 40/40/20 split is recommended with 20 % going to the 3ds since it is the most popular console in Japan.
- Optimize development of Action/shooter genre for NA and EU
- RPG for Japanese market



INFLUENZA SEASON: MEDICAL STAFFING AGENCY

A medical staffing agency needs to proactively plan for how many temporary workers to send to clinics and hospitals are the United states during influenza season to keep up with high demands

INFLUENZA SEASON BACKGROUND

Skills

Objective

Find and use publicly available data to determine the hot zones to send extra staff to each state.

Assumptions

Vulnerable populations suffer the most from the flu and are likely to end up in the hospital. Flu shots decrease the chance of contracting the flu

Hypothesis testing
Data sourcing
Data profiling/integrity
Data integration(vlookup)
Data transformation
Statistical Analysis
Video presentation
Visual Analysis
Tableau Dashboards and storys

Data source: [Influenza deaths by geography and age - CDC](#)
Data source: [Population data by geography - US census Bureau](#)

Tools

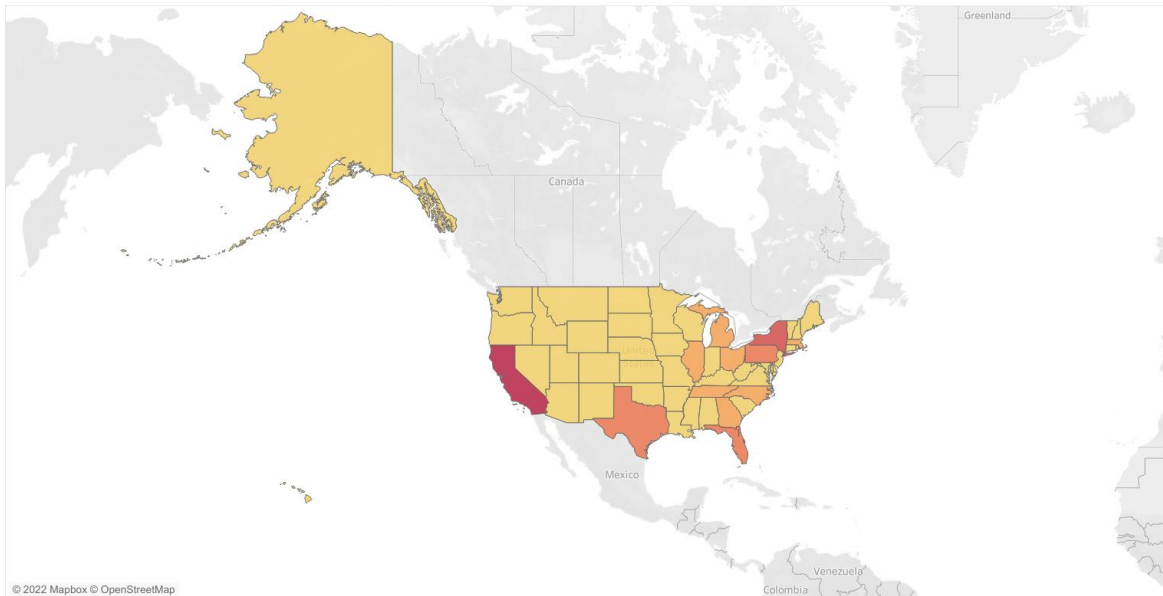


INFLUENZA SEASON ANALYSIS

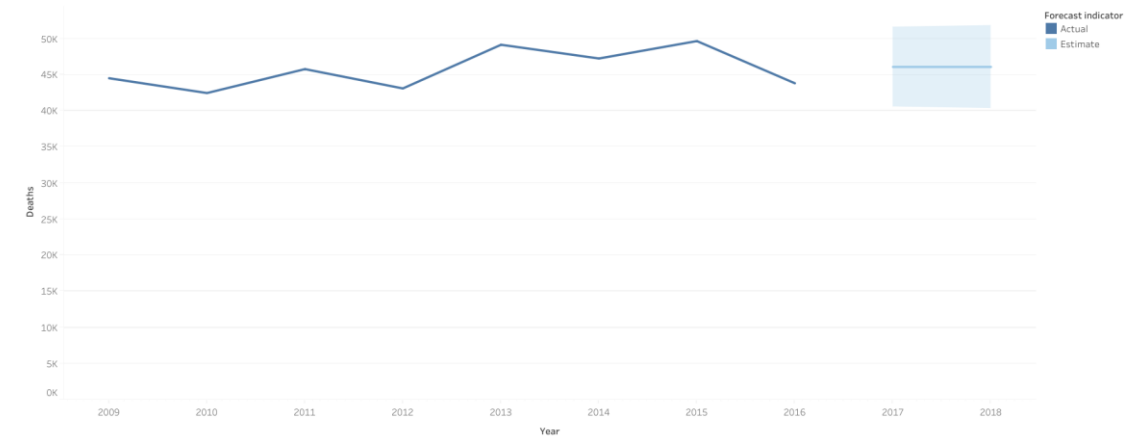
The most vulnerable age group was 55+ and the most impacted states were California, New York, Texas, and Florida

The forecast shows that deaths will increase by a couple of thousand over the next 2 years.

Flu Deaths of Population 55+ (2009 - 2017)



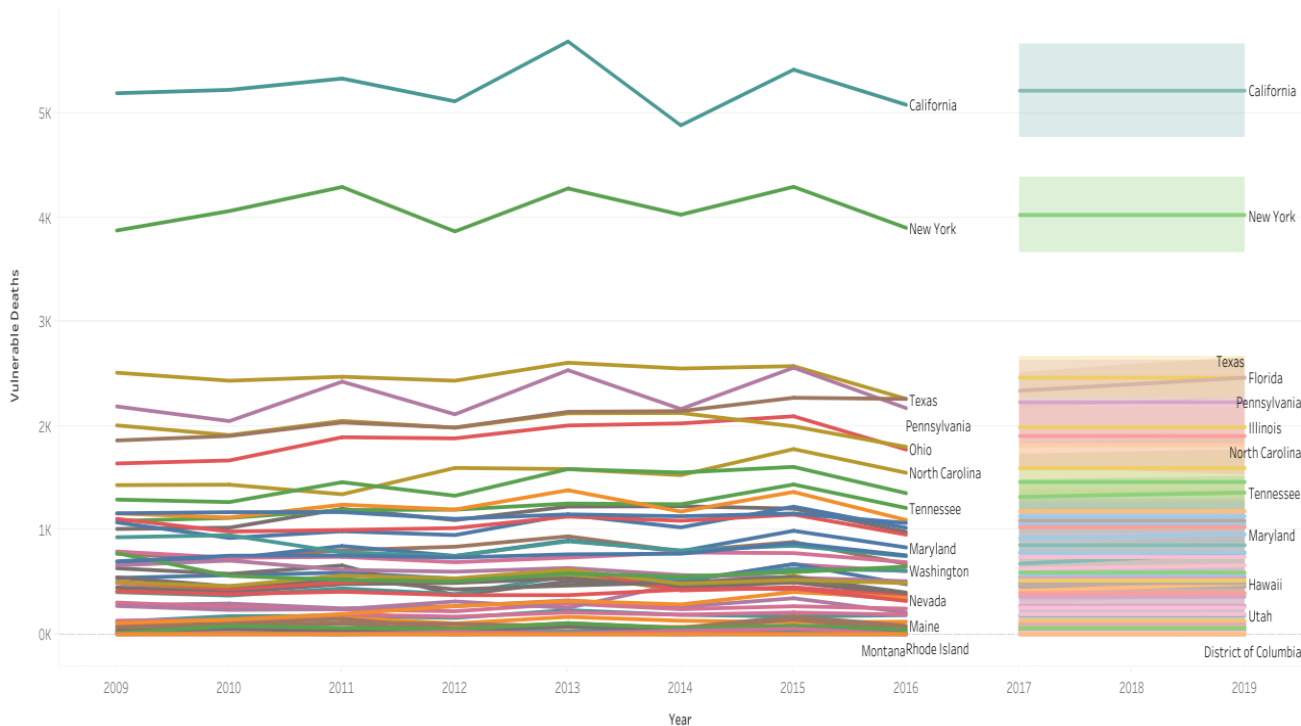
US Influenza Rate by Year Forecast.
Actual Data from 2009-2016, Forecasted Estimates from 2017-2018



RECOMMENDATIONS

Forecast of Deaths per state

Forecast of Vulnerable Flu Deaths (2017-2019)



- Staffing is needed in California, New York, Texas, Pennsylvania, and Florida the most.
- Further Analysis is can be done with for deaths per count in each state.
- Vaccinations of vulnerable population before peak influenza season is key.

- Tableau Dashboard [Here](#)



ROCKBUSTER STEALTH LLC. MOVIE RENTAL COMPANY

A movie rental company is shifting into the online streaming service and needs an analysis on their existing customer base to create and effective lunch strategy.

ROCKBUSTER STEALTH LLC. BACKGROUND

Objective

Assist business intelligence department with launch strategy by providing customer data insight

Key Questions?

Which customers have the highest customer base?

Do sales vary between regions?
What are the statistics on rentals?

Skills

Relational Database management systems

ELC interpretation and usage
SQL queries

Merging, filtering and analyzing

Data dictionary

SQL Data cleaning

Data source: [Here](#)

Tools



PostgreSQL

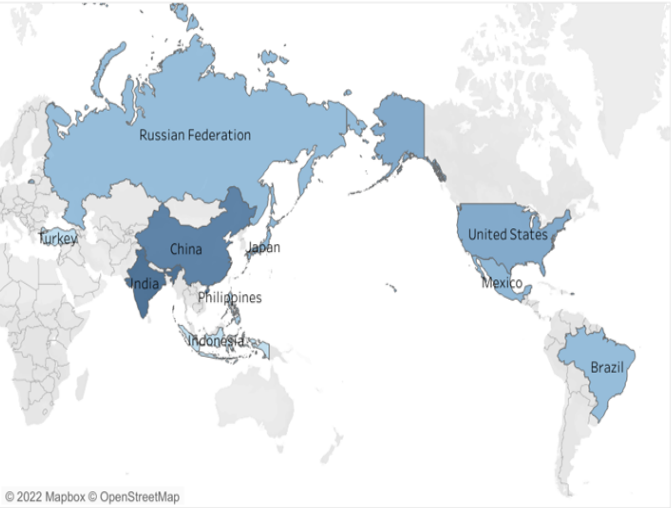


ROCKBUSTER STEALTH ANALYSIS

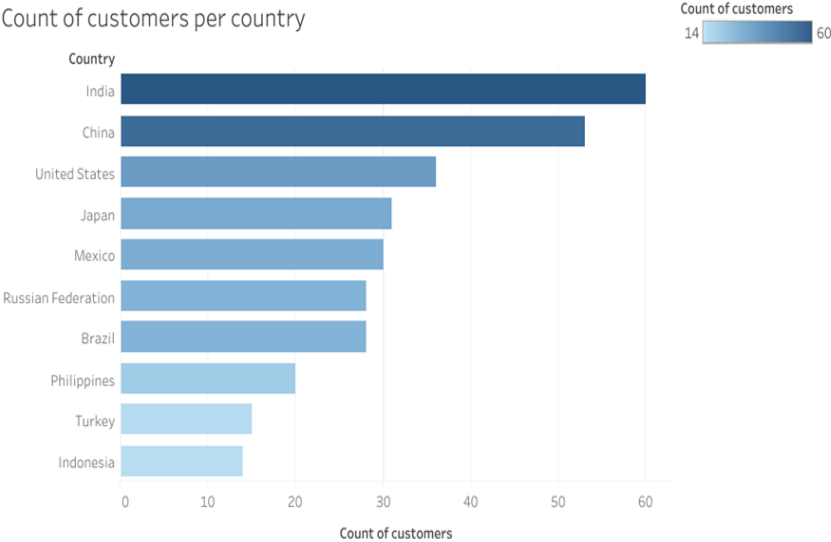
Where are Rockbuster Stealth's Customers Located?

These countries contain a large portion of RockBuster's customers. It may be beneficial to introduce/test a streaming service in these countries since there is already an established customer base.

Customer count by Country map



Count of customers per country



	Rent al Rate	Rent al Dur ation	Movi e leng th	Repl ace men t cost
Min	.99	1 days	46 min	9.99
Max	4.99	7 days	185 min	29.99
Avg	2.98	5 days	115 min	19.99

RECCOMENDATIONS

- Travel movies average the highest rental rates and duration – add more to the inventory
- The 3 biggest markets are India, China, and the U.S focus on implementing streaming service where customer base is established
- Sports, Sci-fi, and animation genres are the top 3 genres in payment amount so focus on adding more to inventory/streaming service
- Top customers – Incentive/loyalty program could prove beneficial in keeping and attracting new customers. Current customers get discount on streaming service for a limited time
- Expand inventory to newer movies.
- Streaming will reduce replacement cost

- Data Dictionary [Here](#)
- Full presentation [Here](#)



INSTACART GROCERY DELIVERY SERVICE

The online grocery delivery app is looking to uncover more information about their customer base, along with purchasing behaviors in order to optimize their targeting marketing strategy.

INSTACART BACKGROUND

Objective

Uncover information about sales patterns and suggest strategies to better segment marketing strategies and target the right customer profile with the correct ads.

Key Questions

What are the busiest days/hours?
What types of products are most popular?
Who uses our product the most?

Skills

Relational Database management systems
Python/Jupyter Notebook
Importing servers
Cleaning, filtering, integrating data
New variable derivation
Data grouping
Data Aggregating
Data Visualization

Data source: Instacart

"The Instacart Online Grocery Shopping Dataset 2017", Accessed from <https://www.instacart.com/datasets/grocery-shopping-2017> on

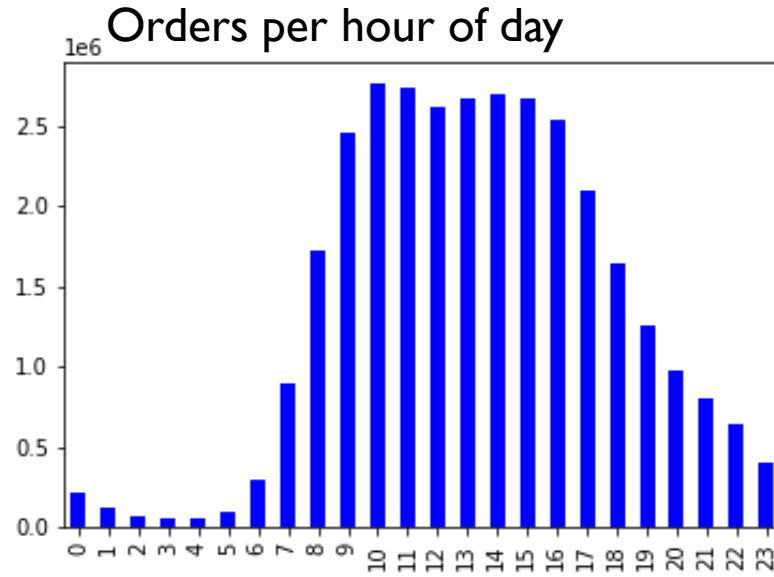
Customer data: [Here](#)

Tools

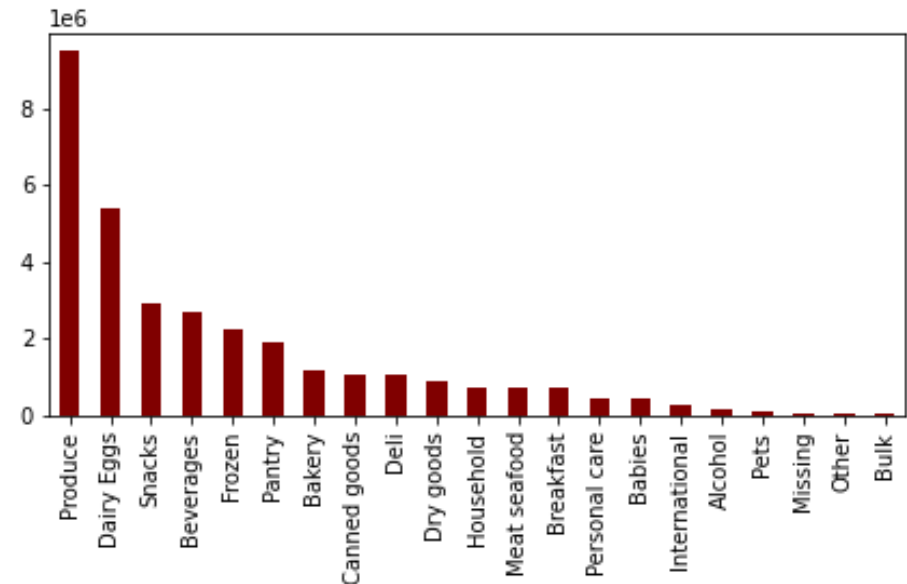


INSTACART ANALYSIS

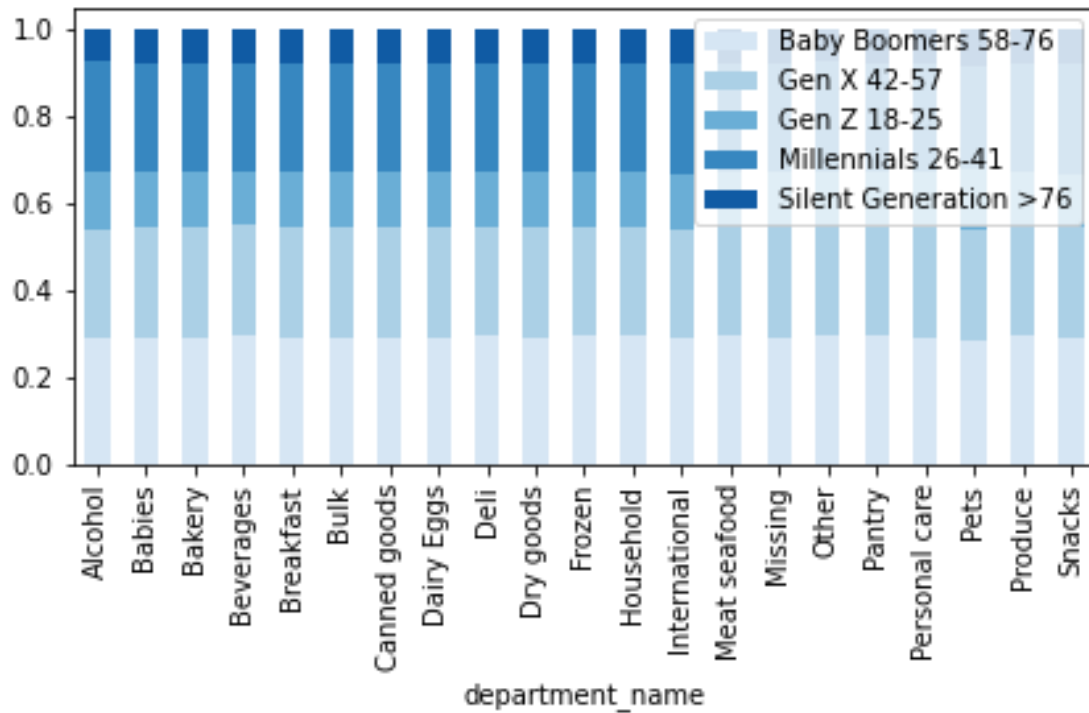
Orders during hour of the day – we can see that most orders come between 10 am and 4 pm so around lunch time.



Highest Department orders: Produce, dairy, eggs, snacks, and beverages.



CONCLUSION



- Busiest days are Friday – Sunday
- Busiest hours are – 10 am – 4pm
- Ads should be run during morning to increase usage for hungry customers.
- Most customers are low spenders less that \$10 dollars
- New customers spend more money at first.
- Gen x, millennials, and baby boomers use the app the most
- Full report [Here](#)



P.E BANK GLOBAL FINANCE SERVICE

A global finance service is building a prototype model to predict and prevent customer departure from the bank

P.E BANK BACKGROUND

Tools

Objective

Analyze probability of customer departure based on profile attributes for the customer retention project

Key Questions?

What are the leading indicators that a customer will leave the bank

Skills

Identifiable information (PII)

Decision Tree

Predictive Analysis

Time series Analysis

Data Mining

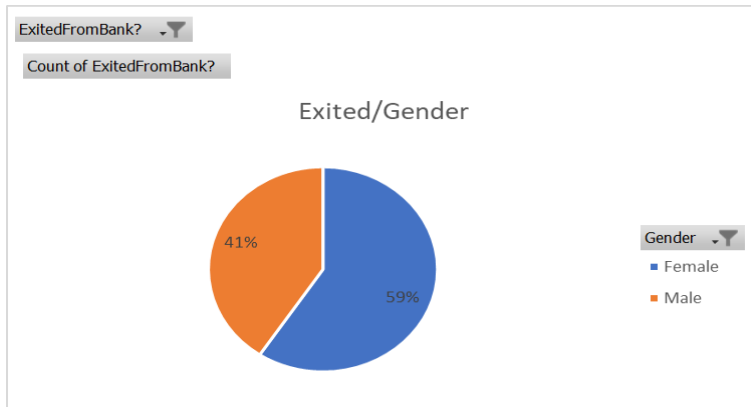
Forecasting

Data Security/Privacy

Data source: [Here](#)
(mock data set, names are fictional)



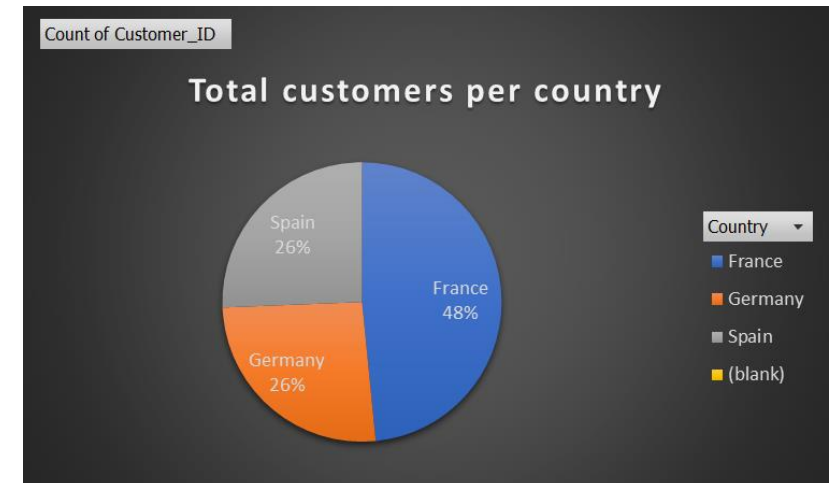
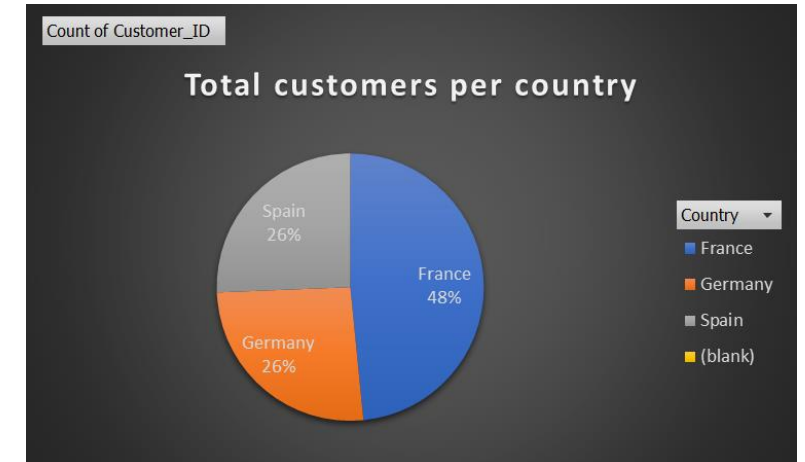
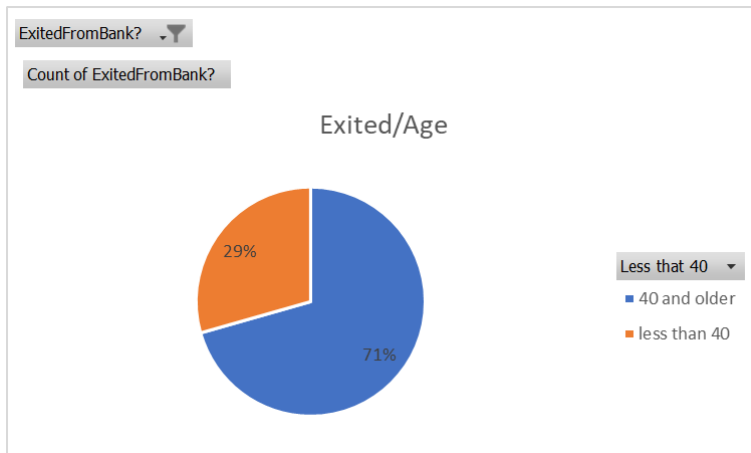
P.E BANK ANALYSIS



Factor 1: Age was the number one factor in those customers who exited the bank. Those 40+ accounted for 71% of people who left.

Factor 2: Gender had a large difference in those who left as well. 59% of those who left were women.

Other factors: Balance of account, those who had \$100,000-150k made up 48% of those who left. Finally, country had was very telling. Germany held the lowest customers by they left at the 2nd largest percentage by only 1%.



CONCLUSION AND RECOMMENDATIONS

- Leading factors of customer exit
 - - Age – 40+ years equaled a higher exit rate
 - - Gender – females had a larger exit rate
 - -Balance – certain account balances had larger exit rates
 - - Country – Germany held the 2nd highest exit rate despite being the smallest customer base.
- Focus on why Germans are leaving the bank. Perhaps operations in the country are not as streamlined as in others?
- Full report [here](#)



WORLD HAPPINESS INDEX ANALYSIS

In this case, I used the world happiness index to determine what are the leading factors contributing to overall happiness in countries.

WORLD INDEX ANALYSIS BACKGROUND

Tools

Objective

Analyze key factors contributing to higher ladder scores in highest ranked countries

Key Questions?

What are the leading indicators that a country will be happy? GDP?
Freedom of choices? Life expectancy
Ladder score is based on a scale of 1-10.

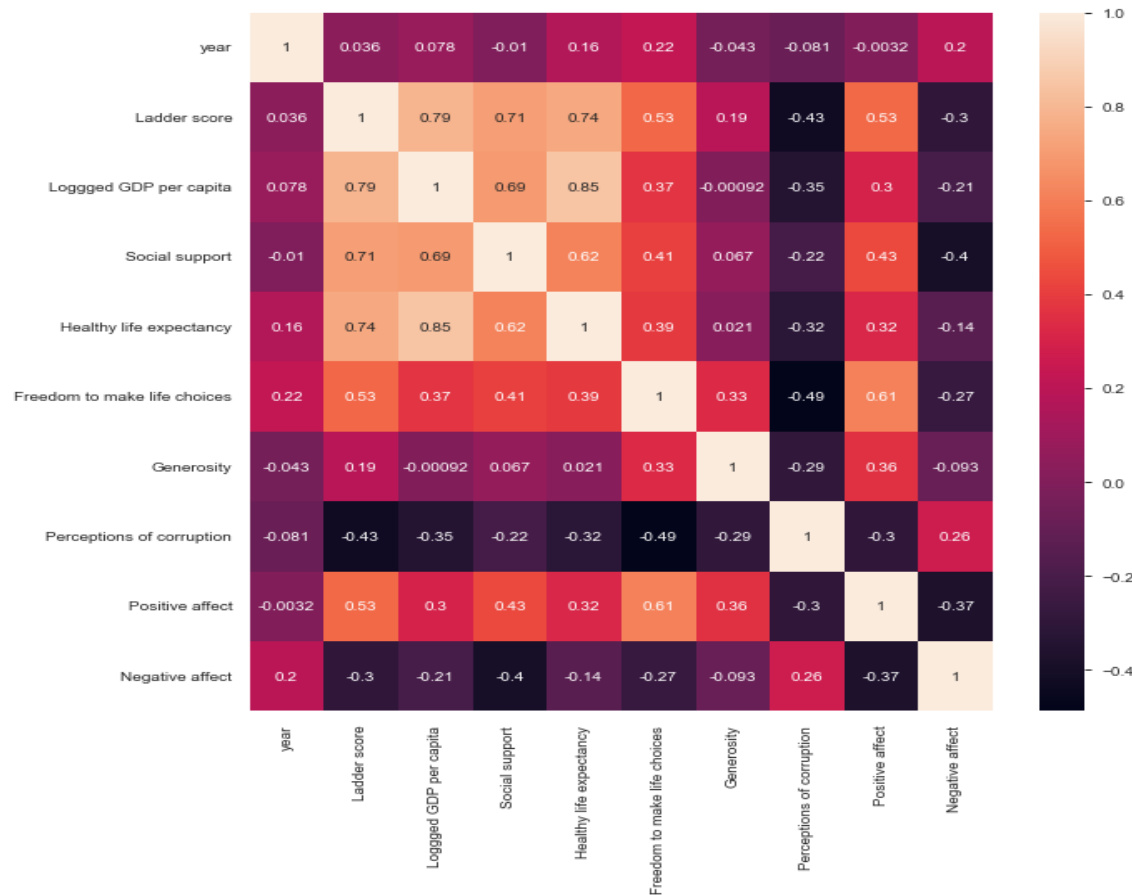
Skills

Data cleaning, Merging
Regression
Cluster analysis
Correlation matrixes
Visualization through python
Time series Analysis
Data Mining
Supervised/unsupervised machine learning: linear regression

Data source: [Here](#)
(mock data set, names are fictional)



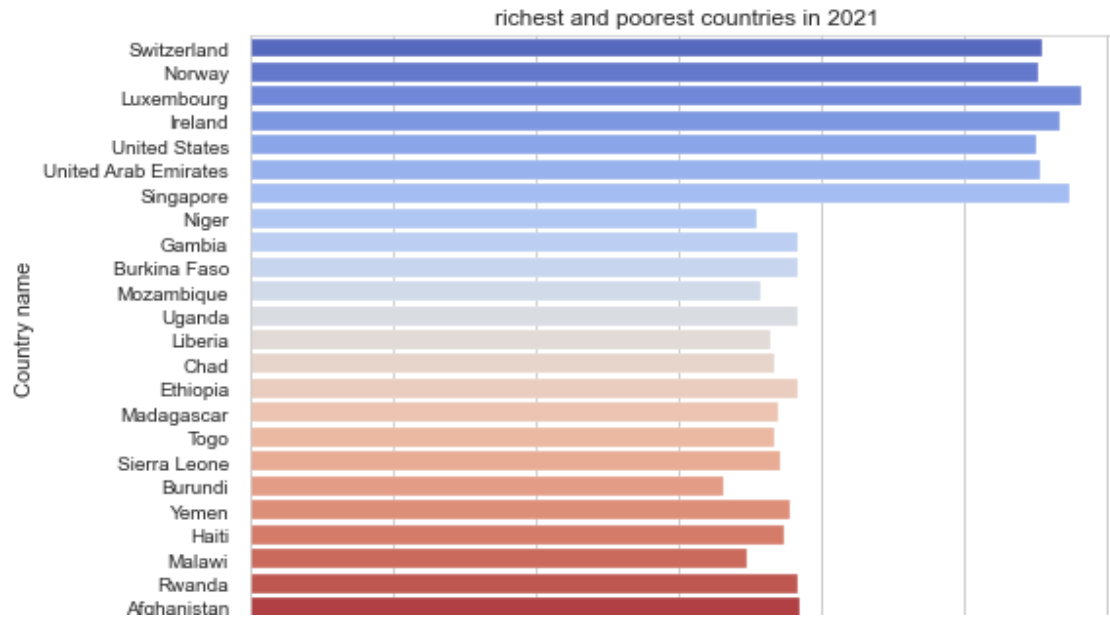
FACTOR CORRELATION



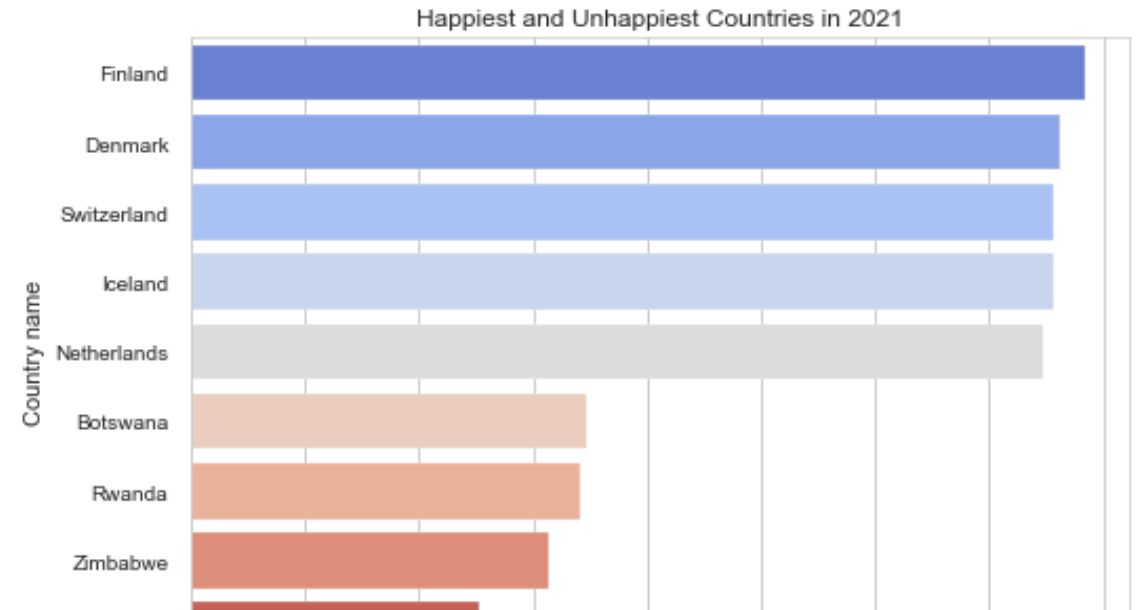
- Here we have a correlation matrix to find strong relationships between certain factors that pertain to each countries.
- GDP, Life expectancy, social support, freedom of choice all have strong correlations in respect to one another and overall Ladder score.

HAPPINESS AND GDP – HIGHEST AND LOWEST COUNTRIES

GDP has a high affect on a populations ladder score

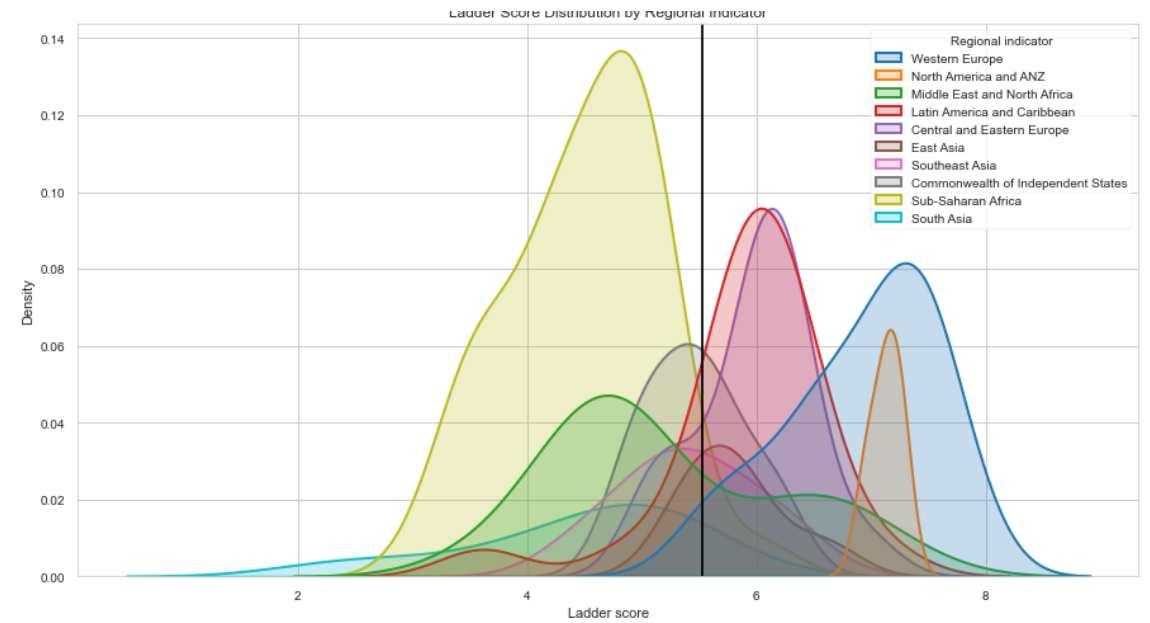
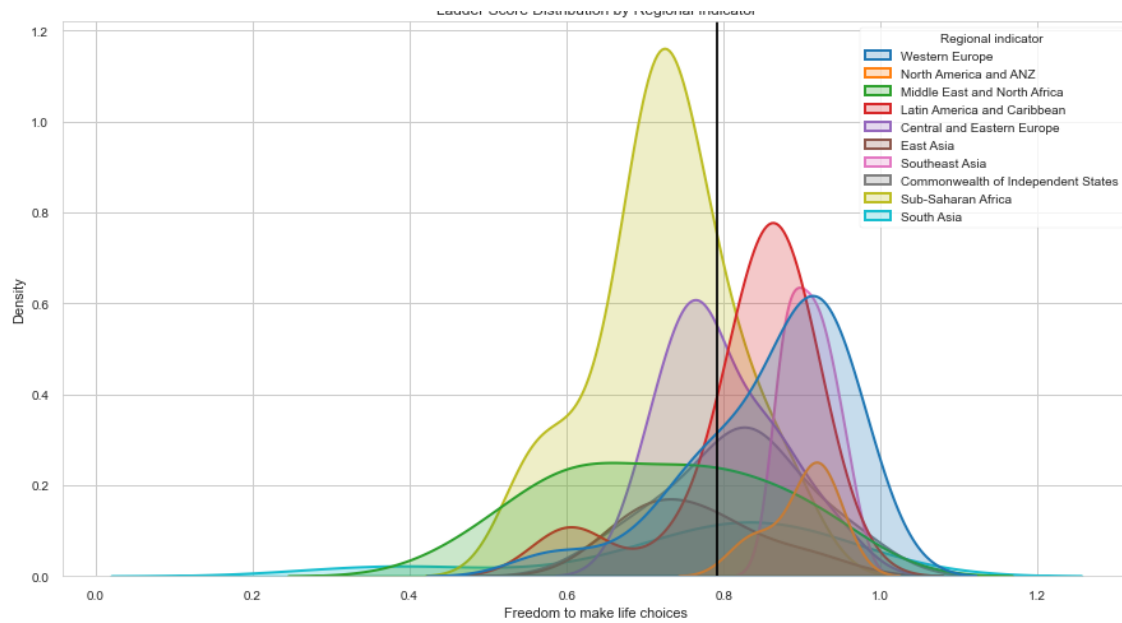


GDP is not the only factor as Finland and Denmark are not in the top 5 for GDP but rank highest in happiness.



FREEDOM AND LADDER SCORE BY REGION

As we see most westernized countries lead in freedom of choice and happiness.



CONCLUSION AND RECOMMENDATIONS

- Leading factors are GDP, Life expectancy, and Social support
- We are able to see what countries are developing through the data of the past 15 years
- Finland and Denmark rank highest in ladder score, would be beneficial to see what cultural factors are pushing that score.
- Data can lead to targeting support missions to help developing countries.



THANK YOU

CONTACT ME @

JBARCENAS2112@OUTLOOK.COM

GITHUB:

[HTTPS://GITHUB.COM/JESUSB21](https://github.com/JESUSB21)

LINKDIN: [HERE](#)