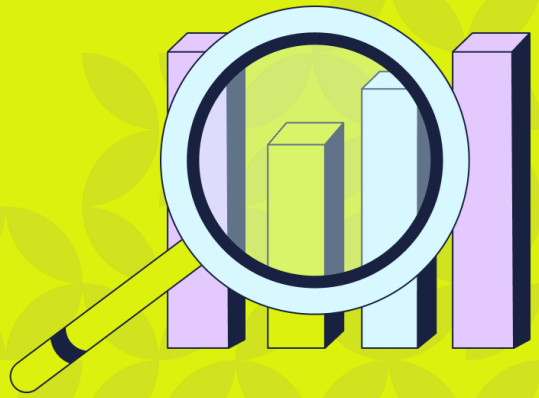
Smartwatch Health Tracker Business Expansion in the US: A State-by-State Analysis

Mc Gee A. Acebedo

A comprehensive analysis to identify the state with the best potential for business expansion.



refocus

Project description

A. Data Analysis Overview:
The analysis involved examining various datasets to gather insights for making informed business decisions. The datasets included:

- Competitor Companies' Profit and Expenditures: This dataset provided information on the profit and spending of competitor companies.
- Corruption Conviction per Capita: The dataset captured the level of corruption in each state, measured as convictions per capita.
- Health Spending: Data on average, minimum, and maximum spending on healthcare-related products and services per person.
- State Income: Information on average, minimum, and maximum income per person in different states.
- Population: The population figures for each state.

B. Addressed Business Questions:
The analysis aimed to determine the ideal state for expanding the Smartwatch Health Tracker business, considering factors such as lower corruption levels and higher income in potential markets.

refocus

Project description

C. Data Analysis Steps:

Using PostgreSQL:
1. Calculated the percentage of income and corruption convictions per capita in relation to the total amount in each state.
2. Identified states with the highest and lowest average income.
3. Identified states with the highest and lowest corruption conviction rates.

Using Excel:
1. Created a scatter plot of state income versus corruption convictions per capita to explore any relationship between the two variables.
2. Calculated the correlation coefficient between state income and corruption convictions per capita using the CORREL() function.
3. Generated a bar chart depicting the average income for each state.

Using Power BI:
1. Visualized the corruption conviction rates across states using a map of the USA. Additional statistics, such as average conviction per capita and population, were included.
2. Created a heatmap to visualize the correlation between state income and corruption convictions per capita.
3. Generated a table displaying the top 5 states with the highest and lowest income.
4. Created a map illustrating the geographical distribution of healthcare spending across different states.
5. Visualized the population density for each state using a map.
6. Generated a table highlighting the top 5 most populated states and the top 5 states in terms of health spending.

refocus

Project description

Using Python:
1. Accessed the SQL database to create a scatter plot comparing state income and corruption convictions per capita.
2. Renamed columns to enhance readability.
3. Conducted tests to determine the correlation between state income and corruption convictions per capita using the '.select_dtypes(include=[np.number]).corr()' method.
4. Created a histogram with a Kernel Density Estimate overlay to showcase the income distribution across all states.
5. Calculated the average, minimum, and maximum corruption levels for different regions of the country.
6. Created a grouped bar chart to visualize convictions per capita in various US regions, displaying the average, minimum, and maximum values.
7. Analyzed financial viability metrics of competitor companies by creating visualizations to showcase the following:
   a. Percentage of administration, marketing, and R&D spending by state.
   b. The top 10 states with the highest profit.
   c. The top 10 states with the lowest profit.
   d. Identified the highest correlation between research and development, administration, marketing spending, and profit.
   e. Created a scatter plot to explore any correlation between average income and the number of convictions per capita.

8. Utilized a multiple linear regression model to identify potential factors contributing to higher profits in a state.

These comprehensive analyses allowed for data-driven decision-making and provided insights into the most suitable state for expanding the Smartwatch Health Tracker business while considering corruption levels and income.

refocus

Postgres

By joining the state_income and corruption_convictions_per_capita tables:

- It was observed that there is a negative correlation between avg. income of US States and Conviction Per Capita.
- This leads us to conclude that it would be good to invest in states with Higher income because they tend to have lower corruption rates.
- Honestly, it was a challenge to derive this observation from SQL. Visualizing this on a chart would be a better option.

```
SELECT state_income.state_usa, state_income.average_income,
concat(round((average_income / sum(average_income)
OVER()*100),2), '%') AS "pct_of_total_avg_income",
corruption_convictions_per_capita.convictions_per_capita,
concat(round((convictions_per_capita / sum(convictions_per_capita)
OVER ()*100),2), '%')
AS "pct_to_total_convictions"
FROM state_income
INNER JOIN corruption_convictions_per_capita
ON state_income.state_usa = corruption_convictions_per_capita.state_usa
ORDER BY "pct_of_total_avg_income" DESC
```

50 rows returned

	state_usa character varying	average_income numeric	pct_of_total_avg_income text	convictions_per_capita numeric	pct_to_total_convict text
1	Maryland	89392	2.83%	1.38	1.63%
2	Massachusetts	82427	2.61%	2.27	2.68%
3	New Jersey	81740	2.59%	1.90	2.24%
4	California	80440	2.55%	1.09	1.29%
5	Connecticut	79287	2.51%	2.01	2.37%
6	New Hampshire	78676	2.49%	0.51	0.60%
7	Hawaii	78084	2.47%	0.43	0.51%
8	Washington	77338	2.45%	2.53	2.99%
9	Alaska	76440	2.42%	1.06	1.25%
10	Colorado	76240	2.41%	0.80	0.94%
11	Virginia	75417	2.39%	1.57	1.85%
12	Utah	72558	2.30%	1.13	1.33%
13	Minnesota	72027	2.28%	0.68	0.80%
14	Illinois	70387	2.23%	1.27	1.50%
15	New York	70137	2.22%	1.59	1.88%
16	Rhode Island	64962	2.06%	8.35	9.85%
17	Pennsylvania	65135	2.06%	1.25	1.48%
18	Delaware	64040	2.03%	1.08	1.27%
19	North Dakota	63715	2.02%	0.57	0.67%
20	Oregon	63835	2.02%	1.87	2.21%
21	Texas	63656	2.02%	1.82	2.15%
22	Wisconsin	63795	2.02%	1.09	1.29%
23	Iowa	62075	1.97%	0.58	0.68%
24	Arizona	62283	1.97%	1.40	1.65%
25	Michigan	61347	1.94%	1.00	1.18%
26	Wyoming	60434	1.91%	1.03	1.22%
27	Nebraska	59929	1.90%	0.57	0.67%

refocus

Postgres

Still using both the tables mentioned

- We found out that the states with the highest income is Maryland while West Virginia is in the other side of the spectrum.

Highest Income States

10 rows returned

	state_usa character varying	average_income numeric
1	Maryland	89392
2	Massachusetts	82427
3	New Jersey	81740
4	California	80440
5	Connecticut	79287
6	New Hampshire	78676
7	Hawaii	78084
8	Washington	77338
9	Alaska	76440
10	Colorado	76240

Lowest Income States

10 rows returned

	state_usa character varying	average_income numeric
1	West Virginia	46254
2	Mississippi	47131
3	New Mexico	48701
4	Arkansas	48829
5	Kentucky	50675
6	Louisiana	50686
7	Alabama	51113
8	Oklahoma	51424
9	South Carolina	52536
10	Idaho	53545

refocus

Postgres

Still using both the tables mentioned

- Looking at the corruption conviction rate we see Hawaii with the lowest while Rhode Island had the worst.

Lowest Corruption Conviction Rate

10 rows returned

	state_usa character varying	pct_of_convictions_per_capita text
1	Hawaii	0.51%
2	Vermont	0.52%
3	Maine	0.57%
4	New Hampshire	0.60%
5	Nebraska	0.67%
6	North Dakota	0.67%
7	Iowa	0.68%
8	Minnesota	0.80%
9	Colorado	0.94%
10	South Dakota	1.03%

Highest Corruption Conviction Rate

10 rows returned

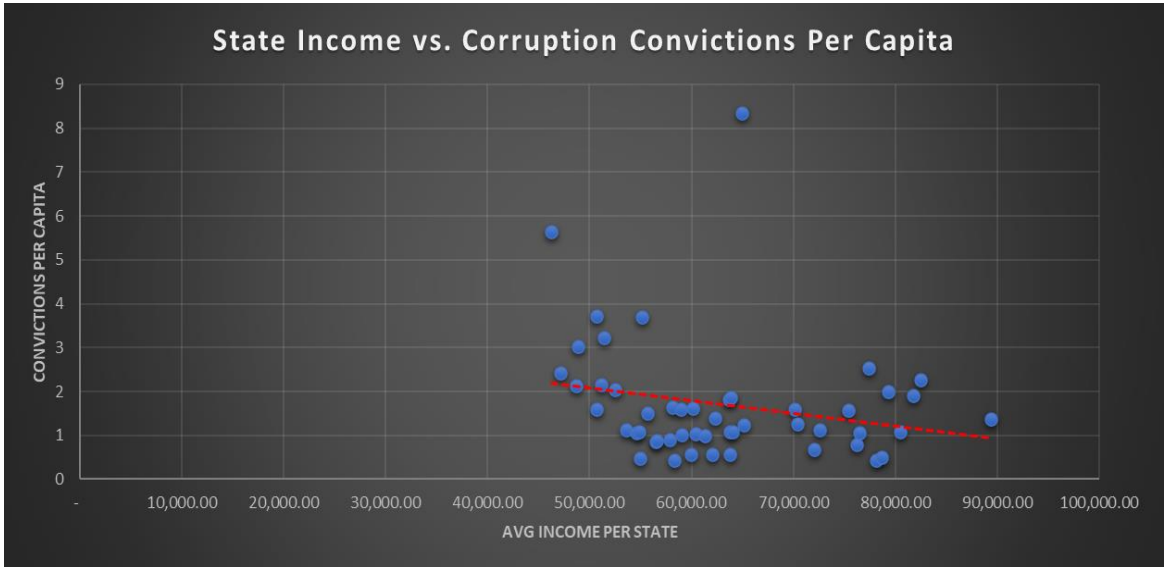
	state_usa character varying	pct_of_convictions_per_capita text
1	Rhode Island	9.85%
2	West Virginia	6.66%
3	Louisiana	4.39%
4	Tennessee	4.36%
5	Oklahoma	3.81%
6	Arkansas	3.56%
7	Washington	2.99%
8	Mississippi	2.87%
9	Massachusetts	2.68%
10	Alabama	2.54%

refocus

Excel

The analysis revealed a correlation coefficient of -0.22628231 between state income and corruption convictions per capita. This correlation suggests a tendency for lower corruption convictions per capita in states with higher income, although the correlation is weak.

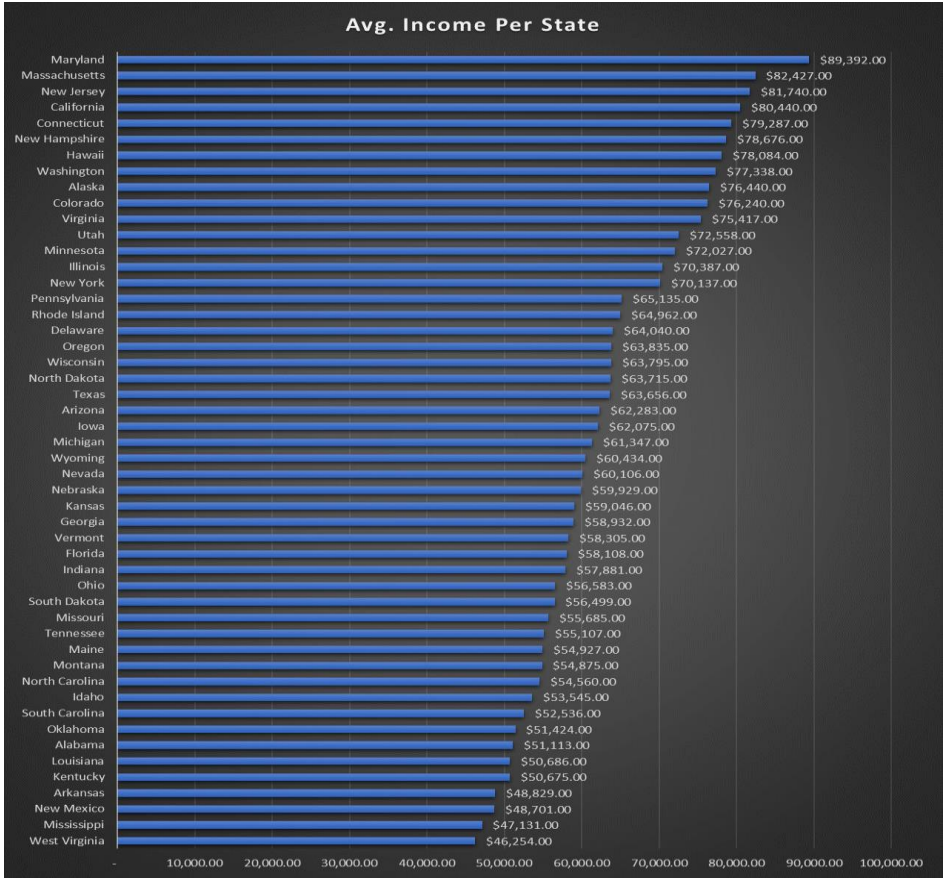
It implies that states with higher incomes may exhibit relatively lower levels of corruption. It is worth noting that correlation does not imply causation, and there may be other reasons here that contribute to this relationship.



refocus

Excel

The bar chart would show us the Avg. Income Per State in descending order.

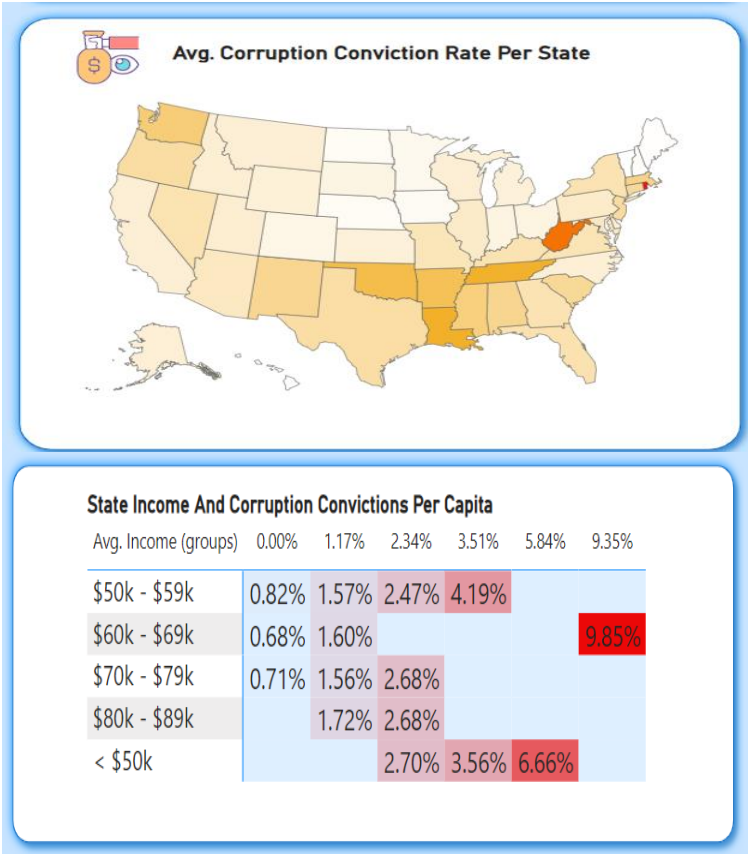


refocus

Power BI

The map shows the avg. corruption conviction rates per state. White to Red shows the level of corruption conviction rates. Rhode Island appears the highest followed by West Virginia.

The State Income and Corruption Convictions Per Capita heatmap further supports the observation that states with lower income (<$50k) have higher corruption conviction rates.



refocus

Power BI

The top 5 states have an avg. income of $413k while the bottom 5 income states have only $241k.

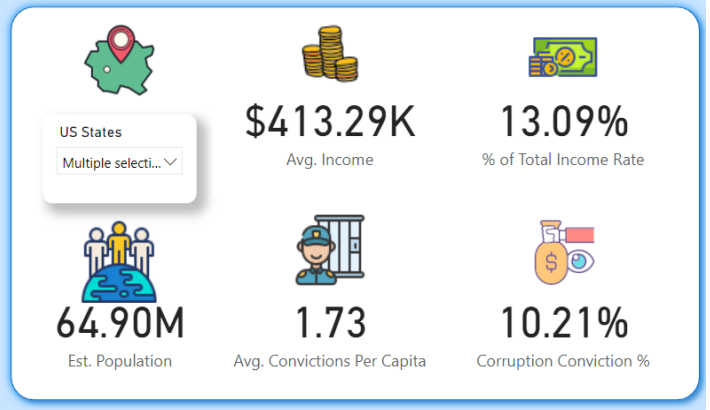
As observed, the higher income states have lower conviction rates.

Top 5 Income States

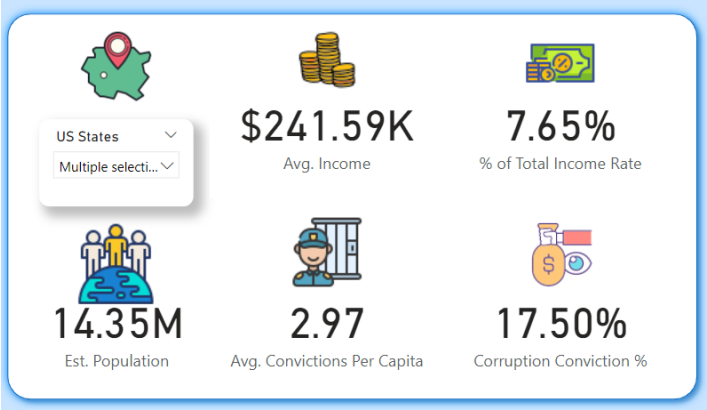
US States	Avg. Income
Maryland	$89,392
Massachusetts	$82,427
New Jersey	$81,740
California	$80,440
Connecticut	$79,287

Bottom 5 Income States

US States	Avg. Income
West Virginia	$46,254
Mississippi	$47,131
New Mexico	$48,701
Arkansas	$48,829
Kentucky	$50,675



VS



refocus

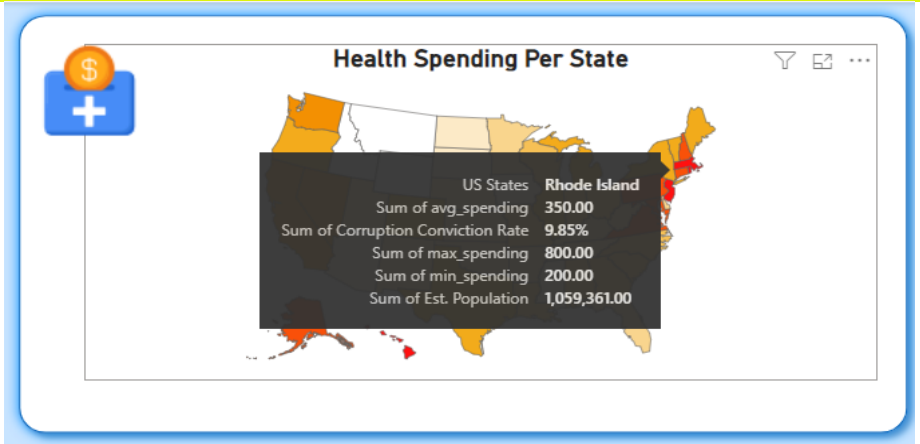
Power BI

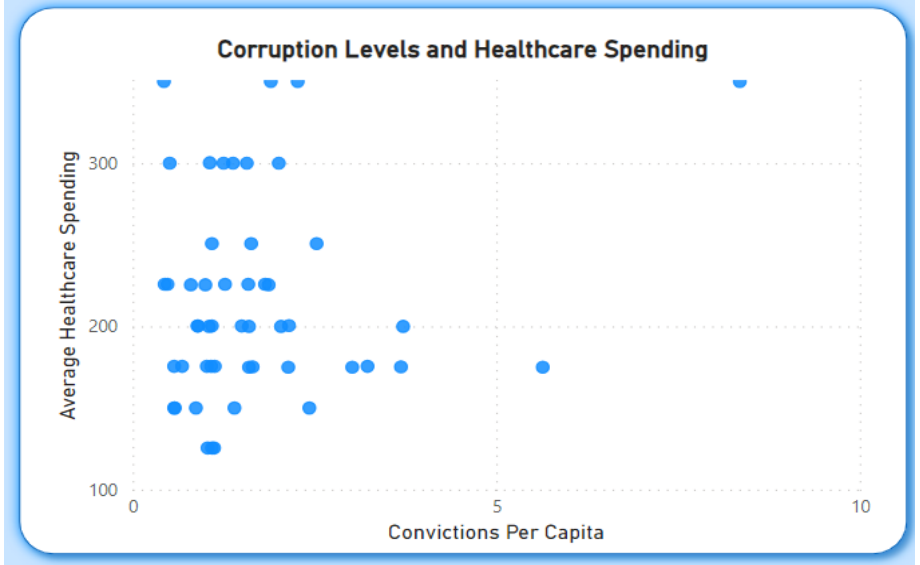
The map shows the Health Spending Per State with tooltips that would show possible correlation between different data points.

In short though, there is a slight positive correlation between the Health spending and corruption Conviction.

* A calculation would be provided in the Python section.

refocus



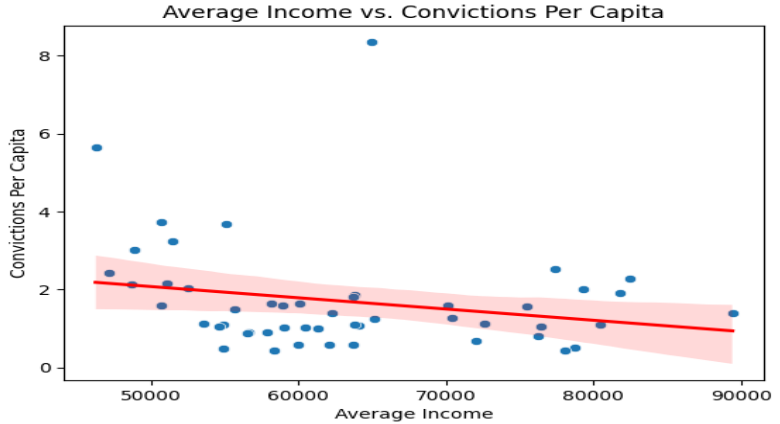


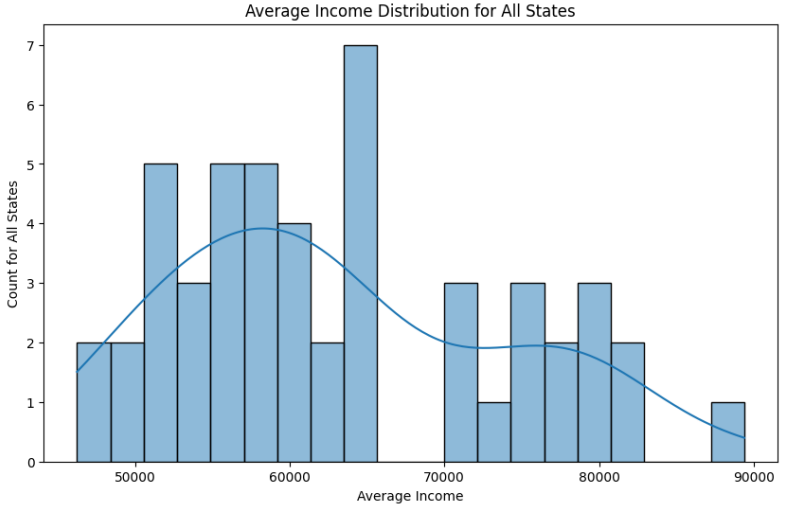
Python

The analysis revealed a correlation coefficient of -0.23 between state income and corruption convictions per capita. This correlation suggests a tendency for lower corruption convictions per capita in states with higher income, although the correlation is weak.

This histogram would show us that the average income is between for states is between $50k to $60k

Correlation Coefficient = -0.23





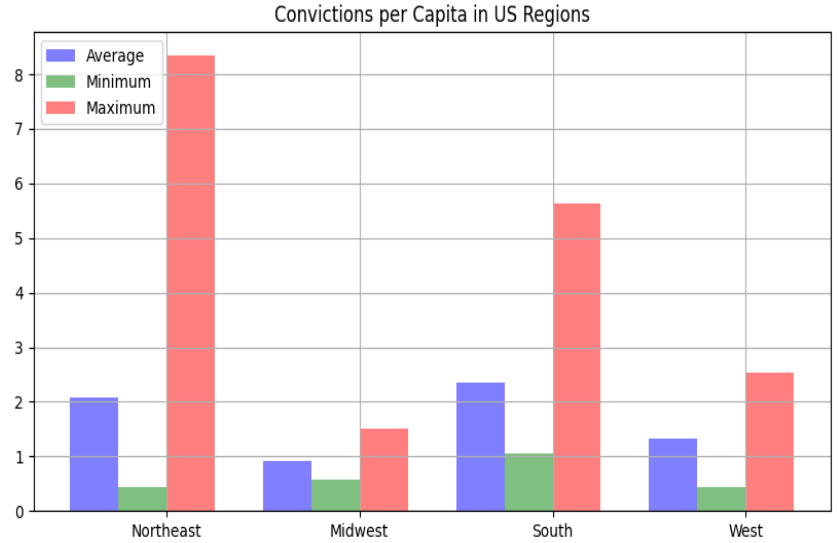
refocus

Python

This table provides a comparison of conviction rates per capita in different US regions.

It suggests that the South region has relatively higher conviction rates on average, while the Midwest region tends to have lower conviction rates. The Northeast and West regions fall somewhere in between, with varying levels of conviction rates. Further analysis and contextual information would be required to understand the factors influencing these regional differences in conviction rates.

	Regions	Average	Minimum	Maximum
0	Northeast	2.088889	0.44	8.35
1	Midwest	0.911667	0.57	1.50
2	South	2.354375	1.05	5.64
3	West	1.332308	0.43	2.53



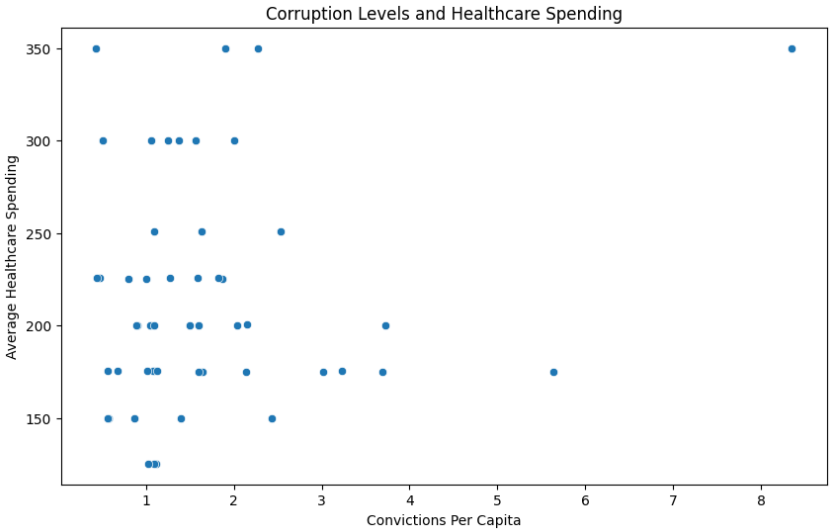
refocus

Python

```
The correlation coefficent of 0.17383131384294861 between health spending and corruption is positive, but not very strong.
```

To answer if there's a correlation between corruption levels and healthcare spending in different states, we looked at both the "health_spending" and "corruption_convictions_per_capita" datasets and plot them in a scatter plot.

We also computed for the correlation coefficient. Which tells us that 0.17 between health spending and corruption is positive, but not very strong.



refocus

Conclusion

In the evaluation of the state with the greatest potential for business expansion, we considered the following:

1. Average Income: The income level in a state is an important consideration as it indicates the purchasing power of consumers, which can greatly impact business success.

2. Rate of Total Convictions: This factor provides insights into the level of integrity and stability within a state's business environment. A lower rate of convictions suggests a more reliable and trustworthy setting for business operations.

3. Competitor Landscape: The presence or absence of competitors in a state can influence business opportunities. States with fewer competitors may offer greater potential for market penetration and growth.

refocus

Conclusion

While thoroughly studying the link between mean income and per capita corruption conviction rates across various states, a minor negative correlation is detected. This suggests that states demonstrating higher average incomes generally experience fewer cases of corruption convictions, indicating an inverse connection between wealth and corruption. However, this correlation is relatively weak, with a correlation coefficient standing at -0.23, meaning that while there's a certain trend, there could also be other influencing elements at play.

Significantly, it's intriguing to see that wealthier states, such as the ones in the top 5 bracket averaging an income of $413K, show lower corruption conviction rates. In contrast, states on the lower income end of the spectrum, including those in the bottom 5 averaging an income of $241K, record higher corruption conviction occurrences. This pattern further emphasizes the inverse relationship between wealth levels and corruption convictions.

refocus

Conclusion

The research also investigated the geographical variance in conviction rates across America. The findings revealed that the southern region had a higher average of corruption conviction rates, while the Midwest region exhibited fewer. The Northeast and West showed a blend of results, insinuating that other factors or unique regional characteristics may be influencing these statistics.

Besides, a minor positive correlation of roughly 0.17 was identified between healthcare expenditures and corruption conviction rates. This suggests that an increase in healthcare spending slightly elevates corruption convictions. While this correlation is weak, it might be indicative of the potential for corruption in situations of larger healthcare budgets.

refocus

Conclusion

The derived findings imply that socio-economic variables like income levels and healthcare expenditure contribute to corruption conviction rates. Furthermore, these findings emphasize the need for establishing an environment promoting transparency and accountability, particularly in states with lower average income and higher health expenditures.

refocus

Conclusion

Answers to Business Questions

The goal of this analysis is to determine which state can we expand our smartwatch health tracker business. Our findings suggest that states New Hampshire and Nort Dakota are the best choices.

To explain our findings let's first go back to what factors we need to consider while running the Financial Viability Metric:

1. Average Income

2. Rate of Total Convictions

3. Competitor Landscape

refocus

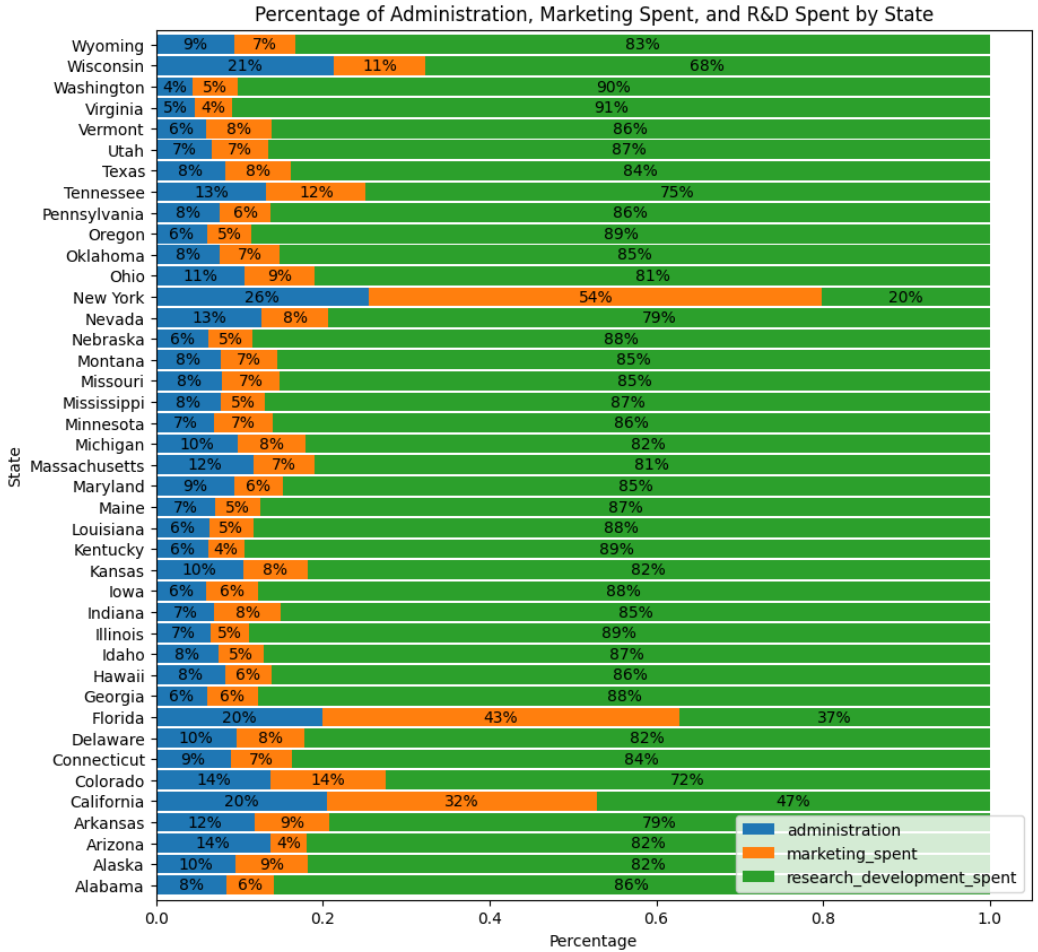
Conclusion

Summing up all the expenses by the competitors, we found out that they spend more on R&D.

	total_administration	total_marketing_spent	total_research_development_spent	\
0	$ 27,487,722.23	$ 33,435,379.40	$ 216,110,664.77	

	total_profit
0	$ 317,693,365.25

refocus



Conclusion

Running multiple linear regression analysis showed that 'research_development_spent' and 'marketing_spent' have significant impacts on profit.

R & D showed a positive coefficient which means that as R & D increases the profit also increases. While Marketing had a negative coefficient which implies that an increase is associated with a decrease in profit. Both assumes that all other variables remain constant. This tells us that it would be best to focus on R&D to maximize profits

OLS Regression Results

Dep. Variable:	profit	R-squared:	0.295
Model:	OLS	Adj. R-squared:	0.287
Method:	Least Squares	F-statistic:	37.23
Date:	Fri, 30 Jun 2023	Prob (F-statistic):	3.92e-20
Time:	15:13:30	Log-Likelihood:	-4068.7
No. Observations:	271	AIC:	8145.
Df Residuals:	267	BIC:	8160.
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.444e+06	2.18e+05	6.631	0.000	1.02e+06	1.87e+06
research_development_spent	0.4361	0.086	5.060	0.000	0.266	0.606
administration	-2.7200	1.808	-1.504	0.134	-6.281	0.841
marketing_spent	-2.7843	0.548	-5.081	0.000	-3.863	-1.705

Omnibus:	10.658	Durbin-Watson:	1.664
Prob(Omnibus):	0.005	Jarque-Bera (JB):	6.433
Skew:	0.210	Prob(JB):	0.0401
Kurtosis:	2.373	Cond. No.	4.58e+06

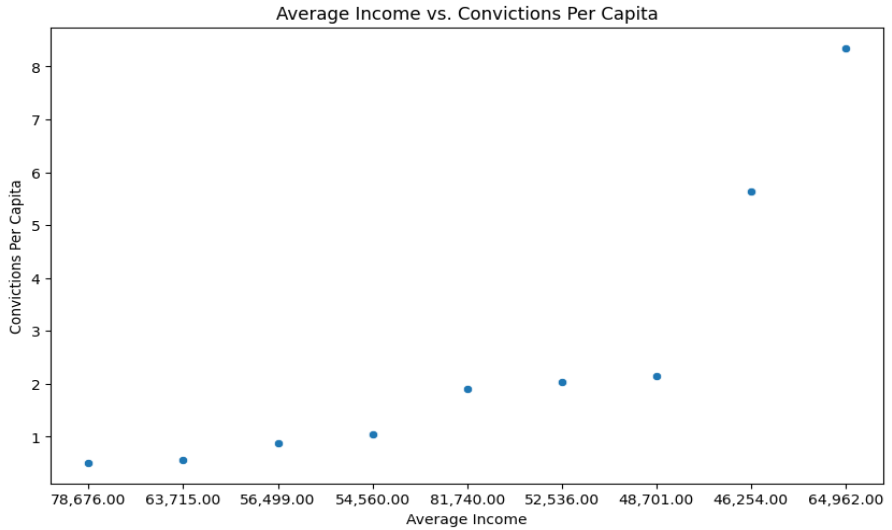
refocus

Conclusion

To answer our business question on which state should we expand it would be better to enter in the Untapped market and determine which one fits our criteria (high income and low rate of corruption).

This norrows us to two states, New Hampshire and North Dakota.

	US States	Avg_Income	Rate of Avg. income	Convictions Per Capita	Rate of Total Convictions
0	New Hampshire	78,676.00	2.49%	0.51	0.60%
1	North Dakota	63,715.00	2.02%	0.57	0.67%
2	South Dakota	56,499.00	1.79%	0.87	1.03%
3	North Carolina	54,560.00	1.73%	1.05	1.24%
4	New Jersey	81,740.00	2.59%	1.90	2.24%
5	South Carolina	52,536.00	1.66%	2.04	2.41%
6	New Mexico	48,701.00	1.54%	2.14	2.53%
7	West Virginia	46,254.00	1.46%	5.64	6.66%
8	Rhode Island	64,962.00	2.06%	8.35	9.85%



refocus

Conclusion

Challenges Encountered During the Analysis

- Due to the limited sample size, the accuracy of the findings may have been affected.
- The analysis only examined the relationship between state income and corruption conviction index. Additional population and health spending datasets was not enough. Other relevant variables such as market demand, cost of doing business, and specific industry conditions should also be considered for a comprehensive business expansion strategy.

refocus

Conclusion

Recommendations for Future Research

To expand on this analysis, future research could explore the following carefully when doing expansions. These factors are broadly classified into internal factors pertaining to the business and external factors relating to the market and overall environment. Here are some of them:

1. Market Demand: Understanding the demand for your product or service in the potential location is crucial. This could involve analyzing demographic data, consumer behavior, and trends in the industry.

2. Cost of Doing Business: This includes costs related to rent or property prices, utilities, taxes, and labor. Locations with lower costs can increase the profitability of the business.

3. Labor Market: The availability of skilled labor in the potential location is another important factor. This could involve analyzing data on education levels, unemployment rates, and wage levels.

4. Infrastructure: The quality of infrastructure, including transportation, utilities, and internet connectivity, can impact the operations of the business.

5. Regulatory Environment: The ease of doing business, regulations related to your industry, and the political stability of the location can also impact the success of the business.

6. Competition: Understanding the competitive landscape in the potential location is also important. This could involve analyzing the number of competitors, their market share, and their strategies.

7. Cultural Factors: Cultural compatibility can also impact the success of the business, especially if the business is expanding to a location with a different culture.

refocus

Conclusion

8. Economic Conditions: The overall economic health of the location, including factors like GDP growth rate, inflation rate, and economic stability, can also impact the business.

9. Access to Suppliers and Distribution Channels: The proximity to suppliers and distribution channels can impact the cost and efficiency of the supply chain.

10. Customer Accessibility: Locations that are easily accessible to customers can help in attracting more customers to the business.

In terms of data, businesses can leverage various sources such as census data, economic indicators, industry reports, and market research studies to gather information on these factors. It's also beneficial to engage with local business agencies, chambers of commerce, and industry associations to gain insights about the business environment in the potential location.

refocus