A glance into US Healthcare Expenditures & Health Outcomes

the Outliers

Sam Sims | Nurmaa Dashzeveg | Kathryn McAtee | Fran Ruiz

Meet The Outliers

the Outliers

Kathryn McAtee

- Project Manager
- Data Analyst & Developer
- Background: CPG and Pharma IT

Fran Ruiz

- Presentation Development
- Data Analyst & Developer
- Background: Instructional design & systems engineering

Nurmaa Dashzeveg

- Data Analyst & Developer
- Background: Data & Cancer Research

Sam Sims

- Data Analyst & Lead Developer
- Background: Technology End User Operations

Agenda

- Meet the Team
- Introduction to the Project & the Question to be Answered
- Our Hypothesis
- The Outliers at Work: Analysis
- How our Analysis Measured Up
- Lessons Learned

Intro & Questions to be Answered

US Healthcare spend continues to increase year over year[1] and projections are not slowing down.

Is the personal healthcare investment making a positive impact on health? And does that answer change based on where you live, how much you make, if you receive government assistance, or where all of that spend is being allocated?

Is the US personal healthcare investment connected to a positive mortality outcome?

Is the investment worth it?

Hypothesis

The Outliers believe that
the average American is paying more for healthcare than they were in previous years, but the value of the investment is declining year over year.

We believe the rate of increase does NOT correlate to mortality outcomes.

[1] https://www.healthsystemtracker.org/chart-collection/u-s-spending-healthcare-changed-time/
https://www.statista.com/statistics/184955/us-national-health-expenditures-per-capita-since-1960/

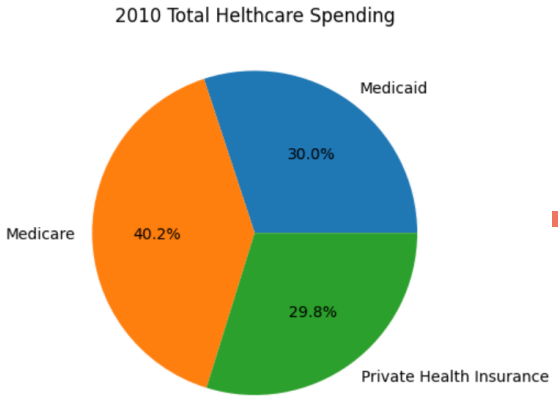
Areas to Analyze

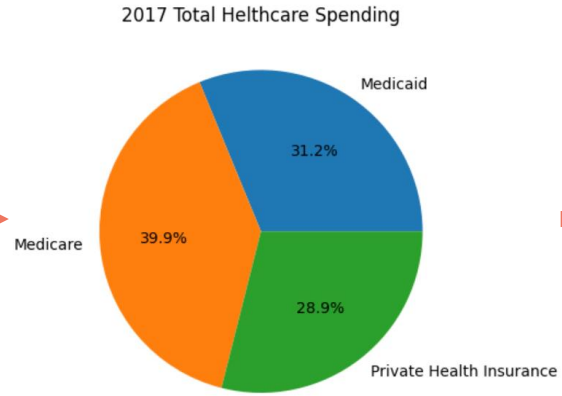
	Presenter:
1. Demonstrate the healthcare expenditure over the year (2010-2017)	Sam
2. The total healthcare spend in government programs.	Kathryn
3. The relation between total population and enrollment in Medicare between 2010 to 2017 in the US.	Fran
4. The relation of high-income earners population and healthcare spend.	Nurmaa
5. The statistics of mortality rates and healthcare spend.	Sam

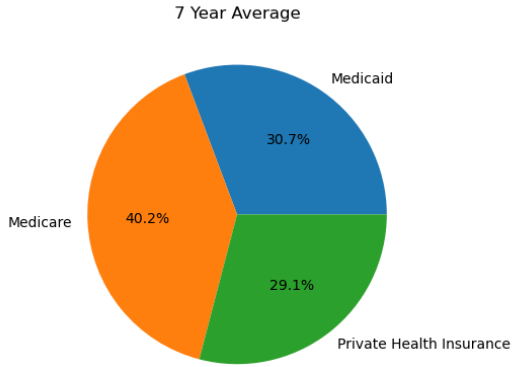
1. Demonstrate the healthcare expenditure over the year (2010-2017)

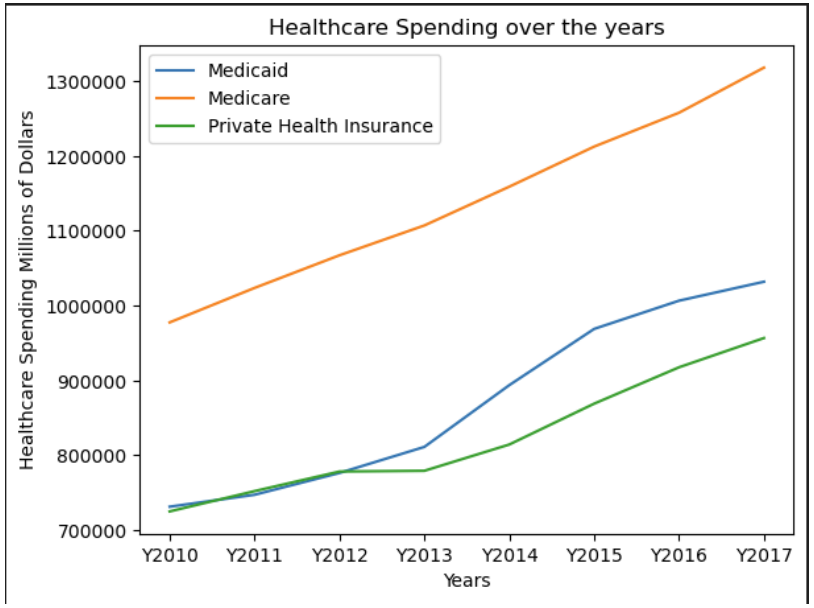
Analyst Sam Sims

Overview of total US spending 2010-2017: Medicare, Medicaid, and Private Insurance

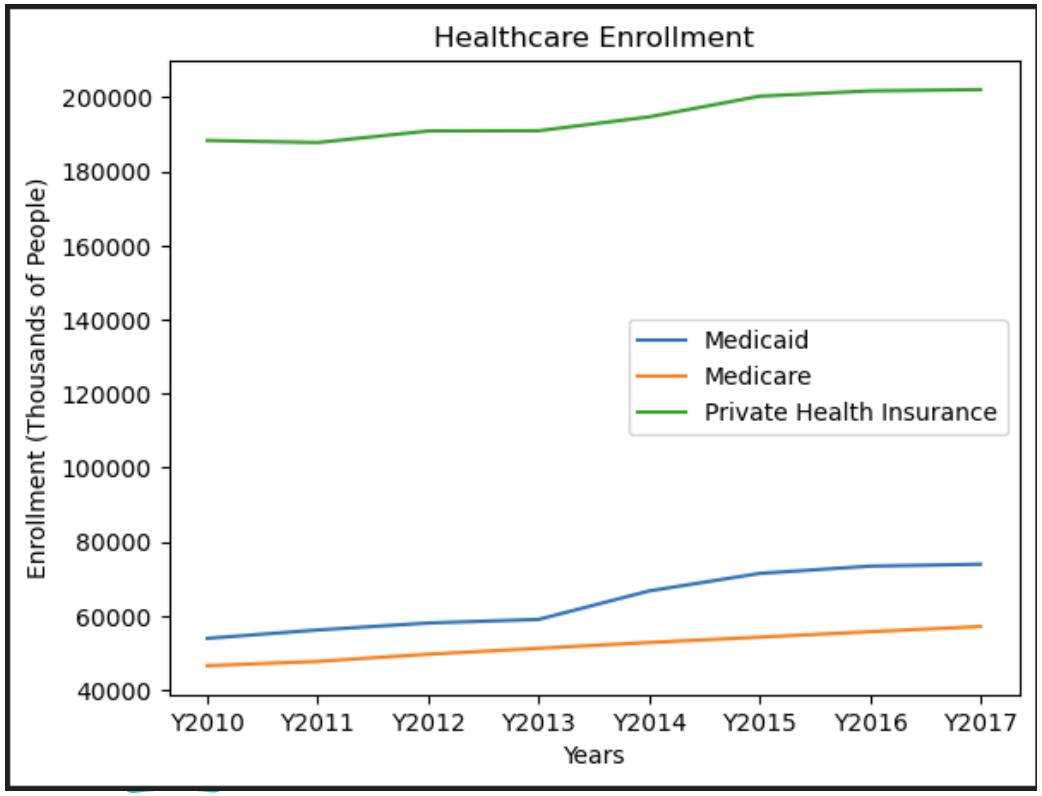


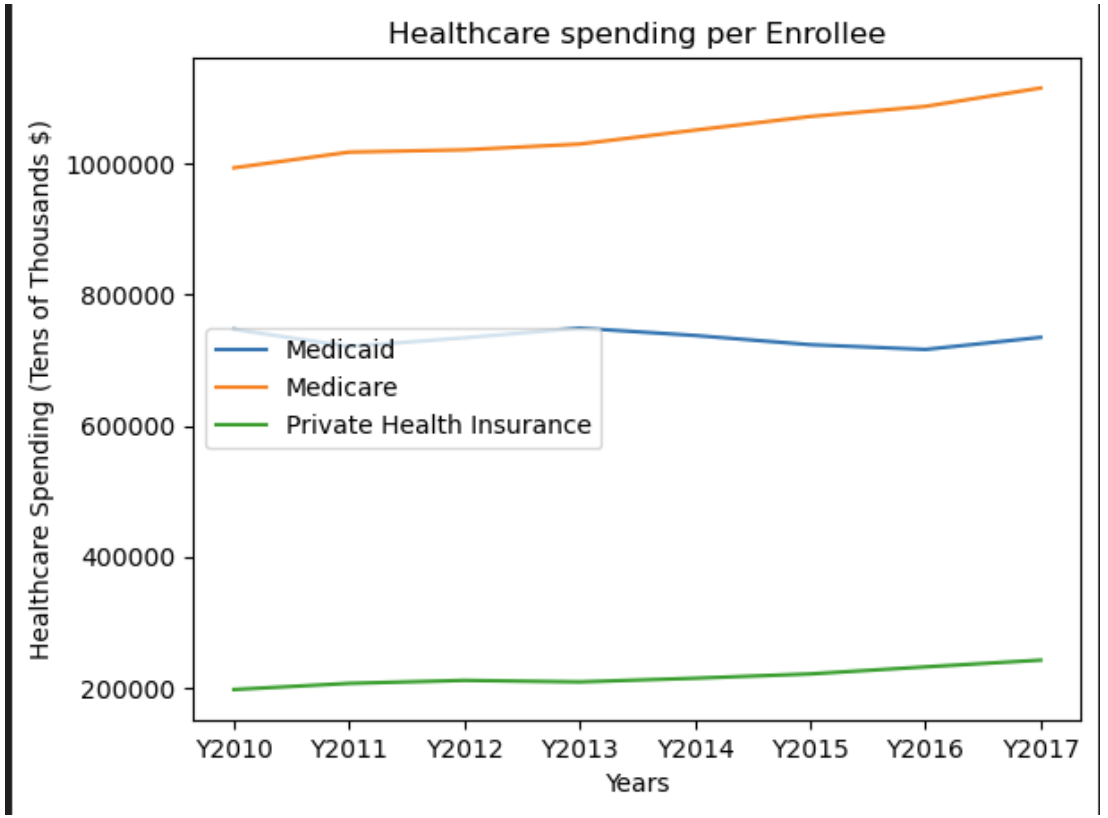






Comparison of Enrollment and spending per Enrollee

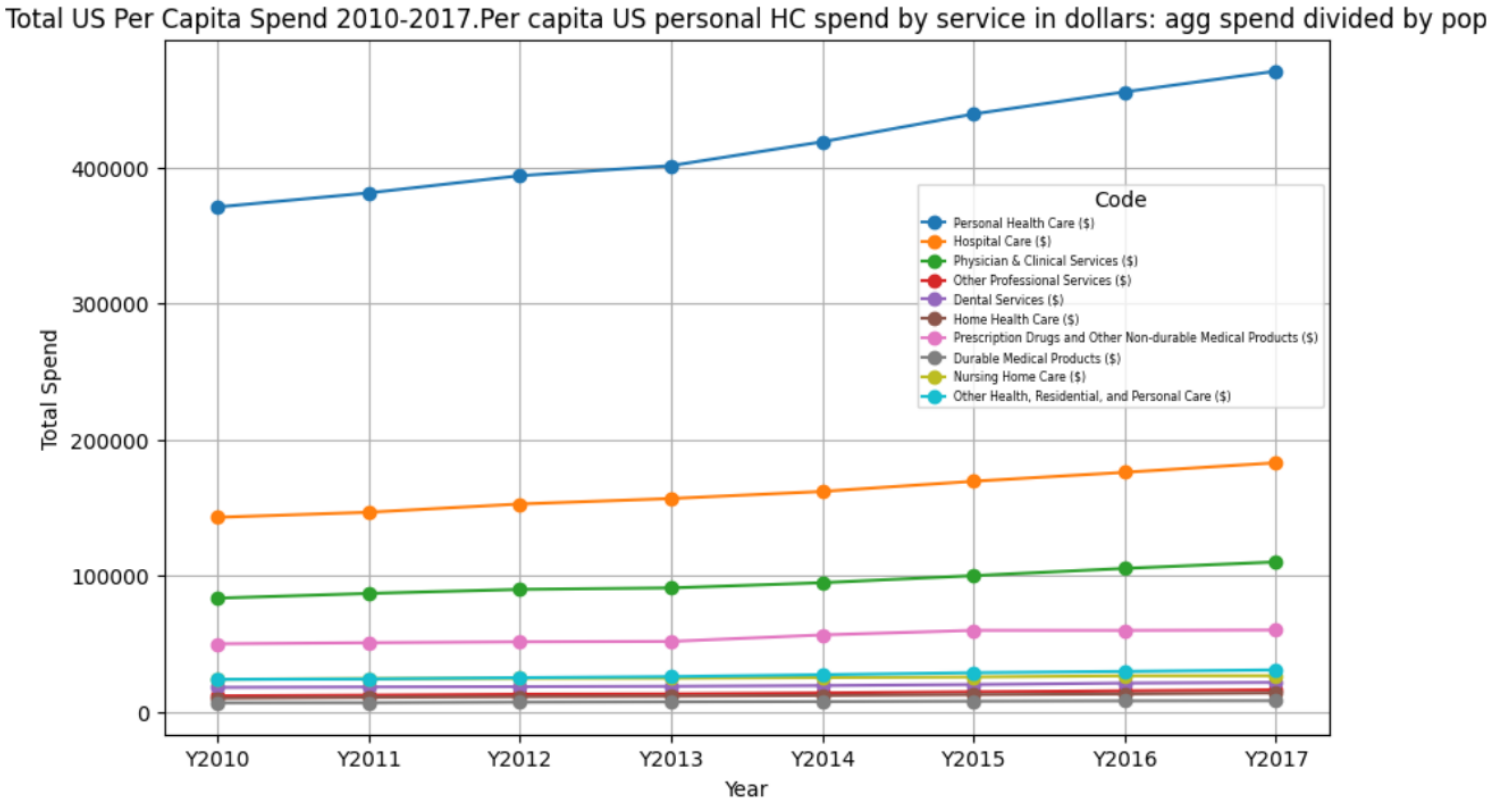




A look into government healthcare program personal spend.

Analyst Kathryn McAtee

Healthcare Spend is Increasing



Personal Health Care

Hospital Care

Physician & Clinical Services

Other Professional Services

Dental Services

Home Health Care

Prescription Drugs

Durable Medical Products

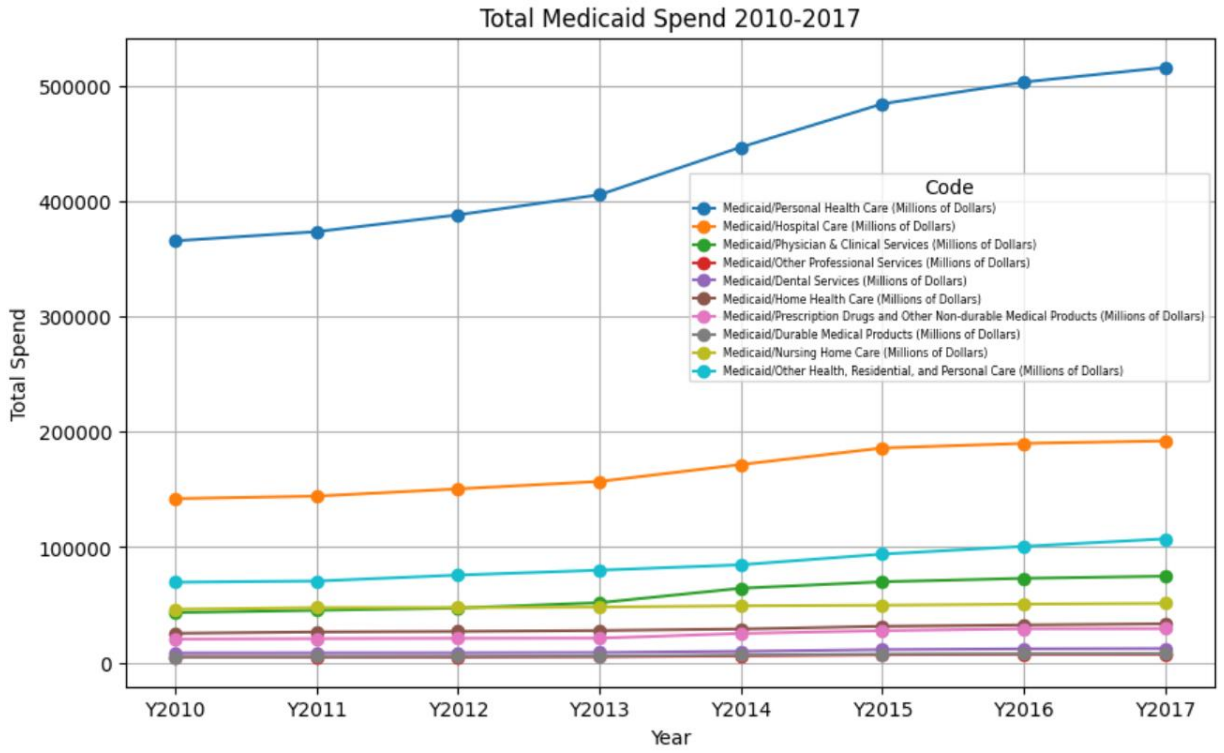
Nursing Home Care

Other Health Care

Population/ Enrollment

1 2 3 4 5 6 7 8 9 10 11

Medicaid Increases by Service and Year



Personal Health Care

Hospital Care

Physician & Clinical Services

Other Professional Services

Dental Services

Home Health Care

Prescription Drugs

Durable Medical Products

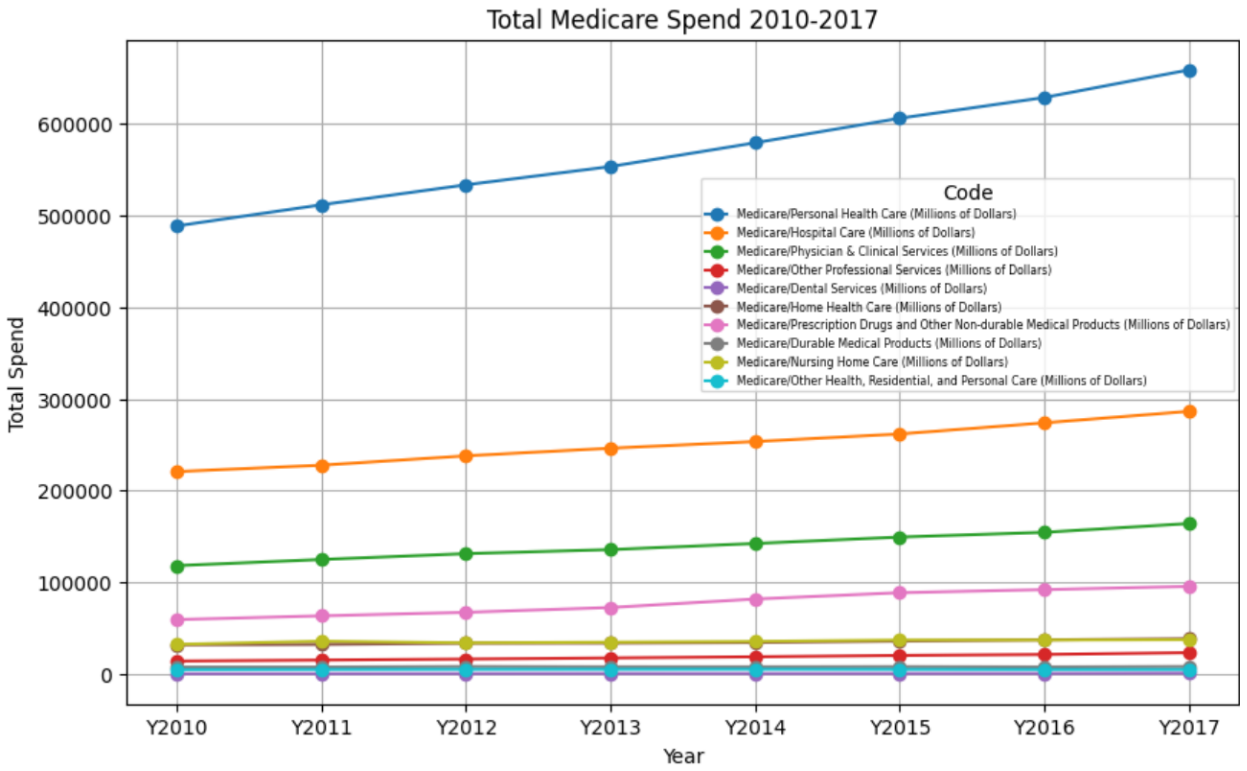
Nursing Home Care

Other Health Care

Population/ Enrollment

1 2 3 4 5 6 7 8 9 10 11

Medicare Increases by Service and Year



Personal Health Care

Hospital Care

Physician & Clinical Services

Other Professional Services

Dental Services

Home Health Care

Prescription Drugs

Durable Medical Products

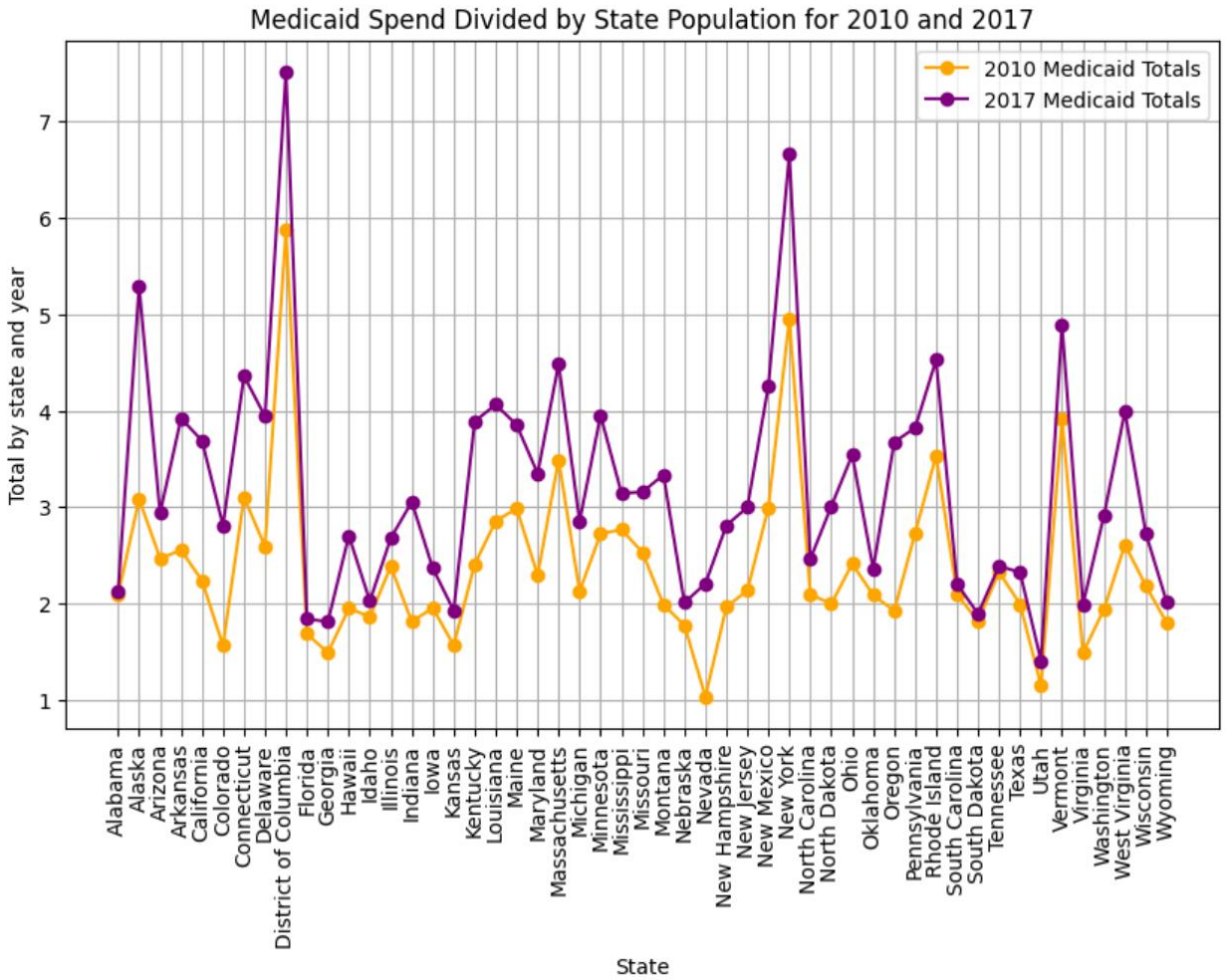
Nursing Home Care

Other Health Care

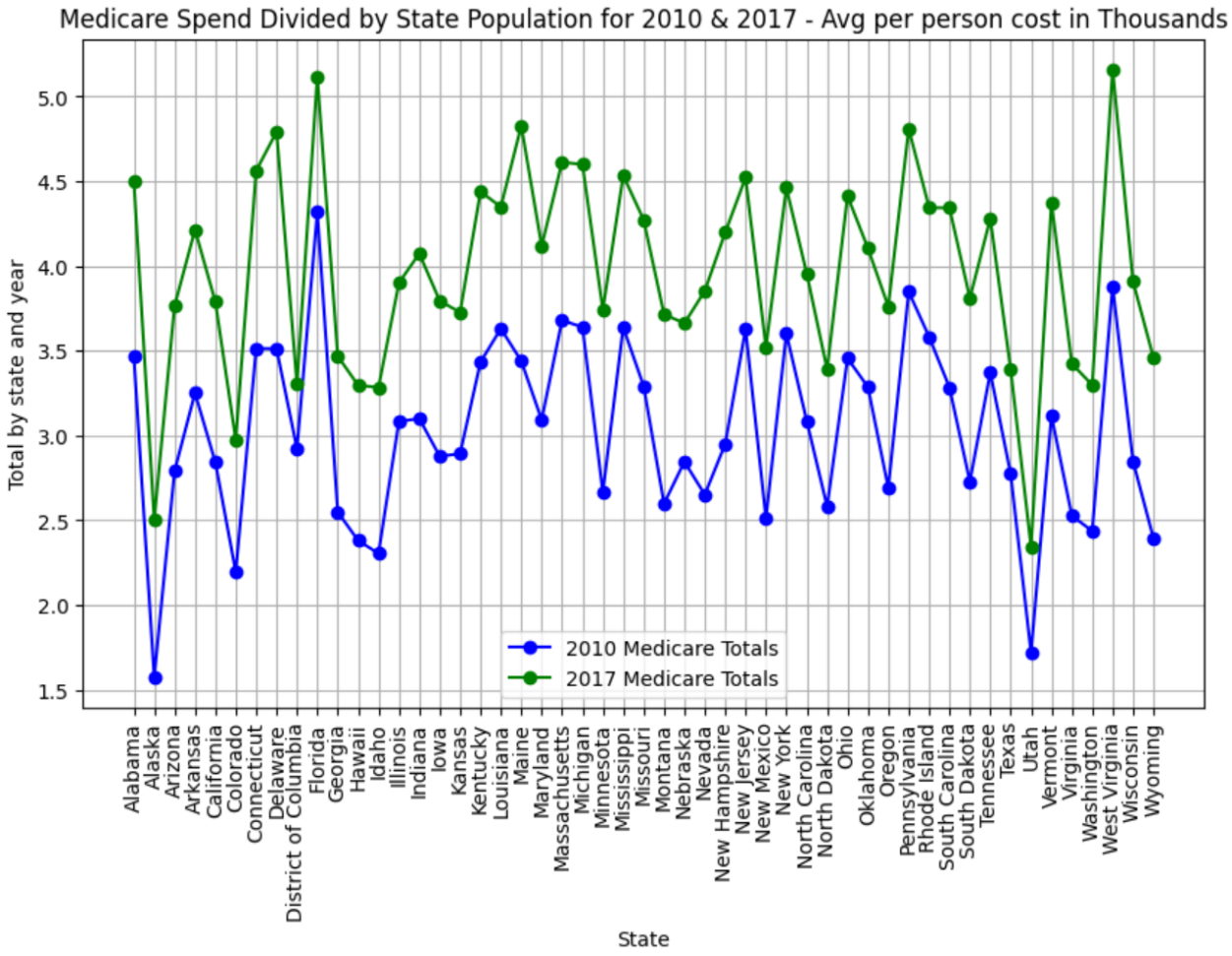
Population/ Enrollment

1 2 3 4 5 6 7 8 9 10 11

Medicaid Spend Divided by State Population



Medicare Spend Divided by State Population



3. The relation between total population and enrollment in Medicare between 2010 to 2017 in the US.

Analyst Fran Ruiz

Data Set 1

new_census_data

	Year	State	Total Population	Age Categories 65 Years And Over
0	2010	Alabama	4712651.0	636208.0
1	2010	Alaska	691189.0	50457.0
2	2010	Arizona	6246816.0	830827.0
3	2010	Arkansas	2872684.0	405048.0
4	2010	California	36637290.0	4066739.0
...	...	...	...	...
407	2017	Virginia	8365952.0	303515.0
408	2017	Washington	7169967.0	253125.0
409	2017	West Virginia	1836843.0	61792.0
410	2017	Wisconsin	5763217.0	205743.0
411	2017	Wyoming	583200.0	22300.0

412 rows × 4 columns

count_states.count()

```
Year                                  412
State                                 412
Total Population                      408
Age Categories 65 Years And Over      408
dtype: int64
```

new_census_data.dropna()

```
Year                                  408
State                                 408
Total Population                      408
Age Categories 65 Years And Over      408
dtype: int64
```

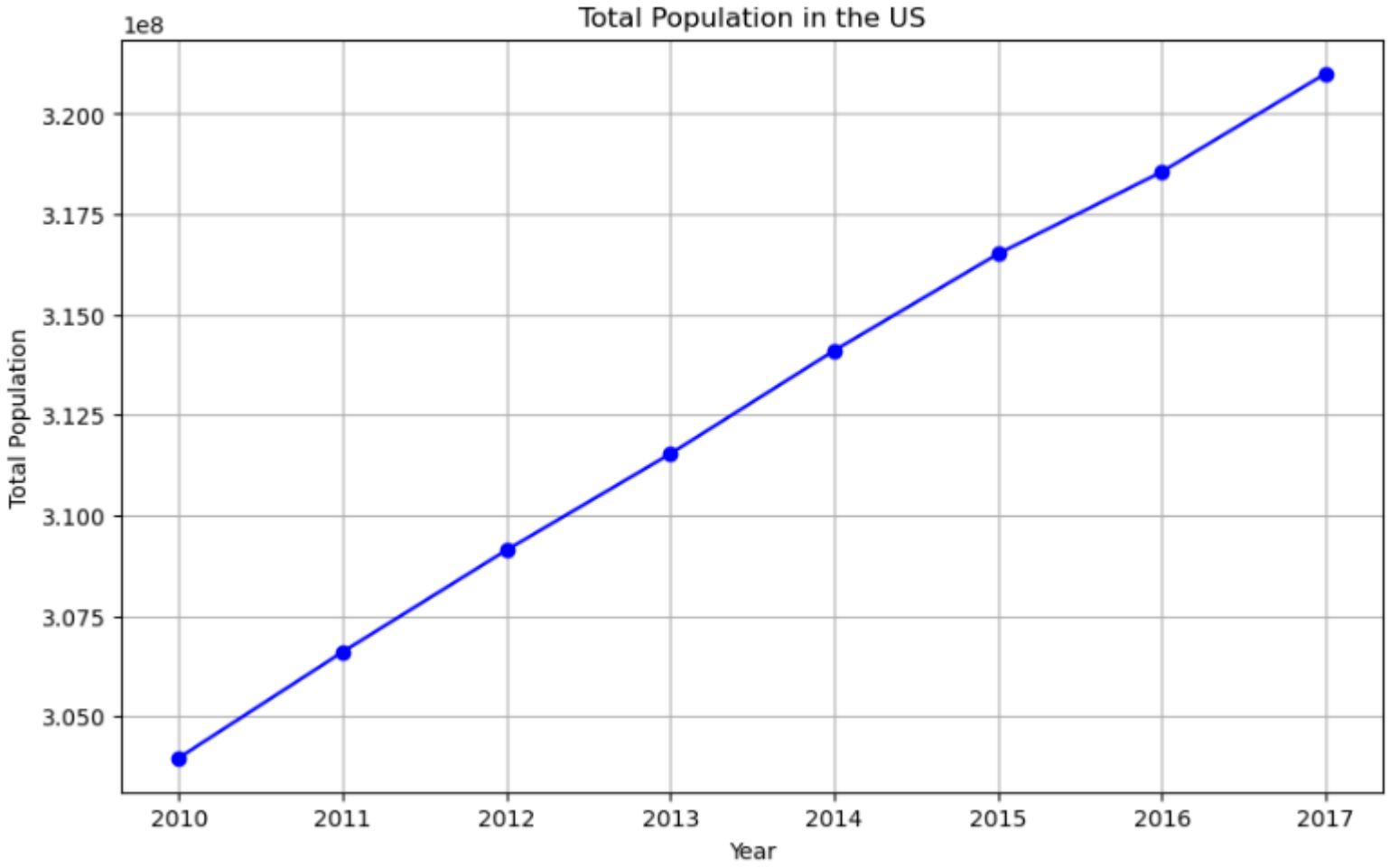
	Year	Total States
0	2010	51
1	2011	51
2	2012	51
3	2013	51
4	2014	51
5	2015	51
6	2016	51
7	2017	51

Data Set 2

medicare_data.head(10)

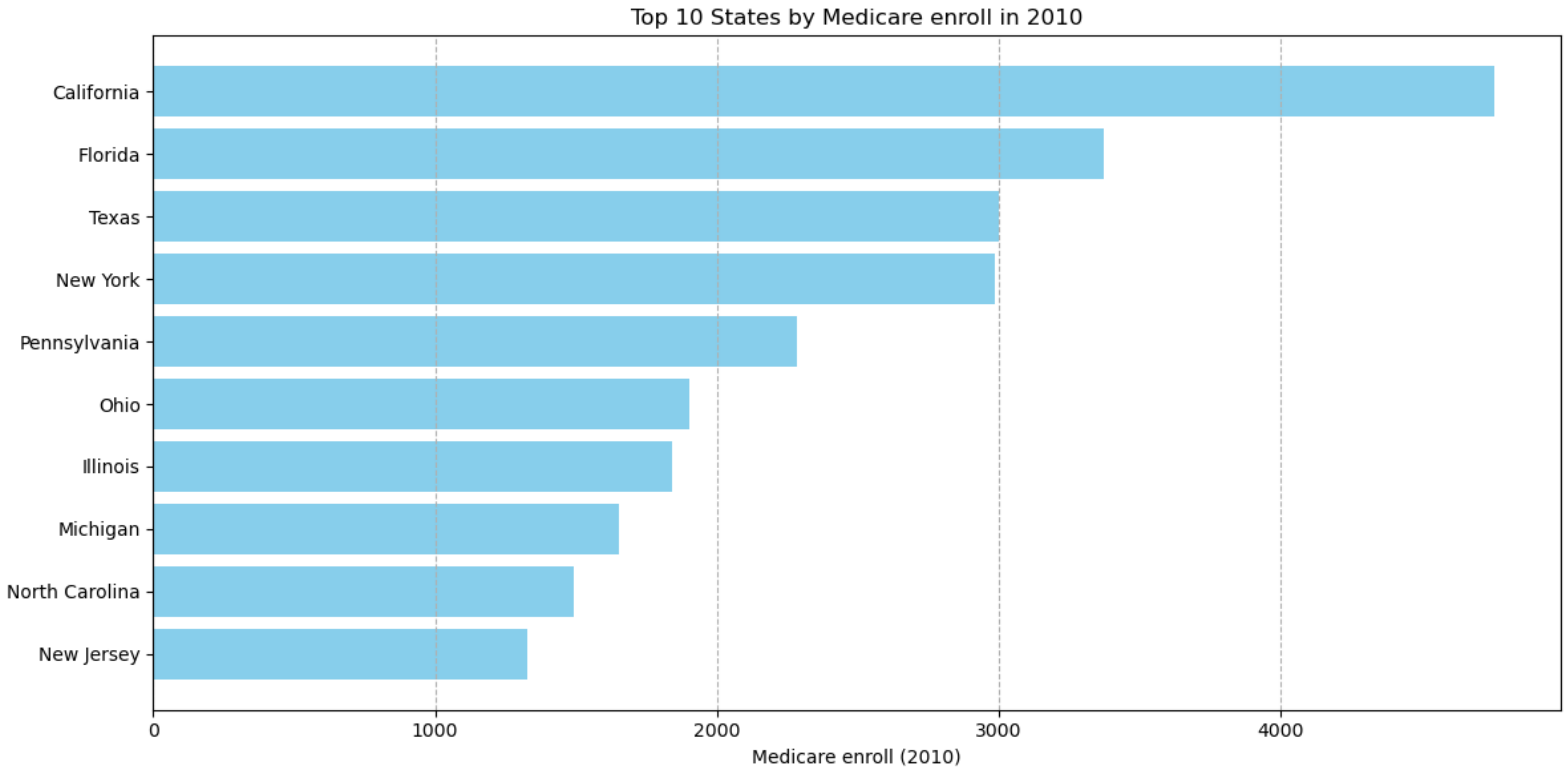
	State	Medicare 2010	Medicare 2011	Medicare 2012	Medicare 2013	Medicare 2014	Medicare 2015	Medicare 2016	Medicare 2017	Code	Item
0	Alabama	845	864	896	922	947	968	988	1006	11	Medicare Enrollee Population (Thousands)
1	Alaska	66	69	73	76	80	84	88	92	11	Medicare Enrollee Population (Thousands)
2	Arizona	930	962	1009	1051	1096	1140	1180	1224	11	Medicare Enrollee Population (Thousands)
3	Arkansas	531	542	559	572	584	594	607	617	11	Medicare Enrollee Population (Thousands)
4	California	4757	4901	5111	5294	5476	5645	5814	5966	11	Medicare Enrollee Population (Thousands)
5	Colorado	625	650	688	722	755	786	816	840	11	Medicare Enrollee Population (Thousands)
6	Connecticut	568	577	595	608	619	630	642	654	11	Medicare Enrollee Population (Thousands)
7	Delaware	149	154	161	168	175	181	187	193	11	Medicare Enrollee Population (Thousands)
8	District of Columbia	78	80	82	84	87	88	90	91	11	Medicare Enrollee Population (Thousands)
9	Florida	3375	3472	3621	3757	3900	4040	4164	4289	11	Medicare Enrollee Population (Thousands)

Total Estimate Population in The US from 2010 to 2017



Top ten states by Medicare enroll in 2010

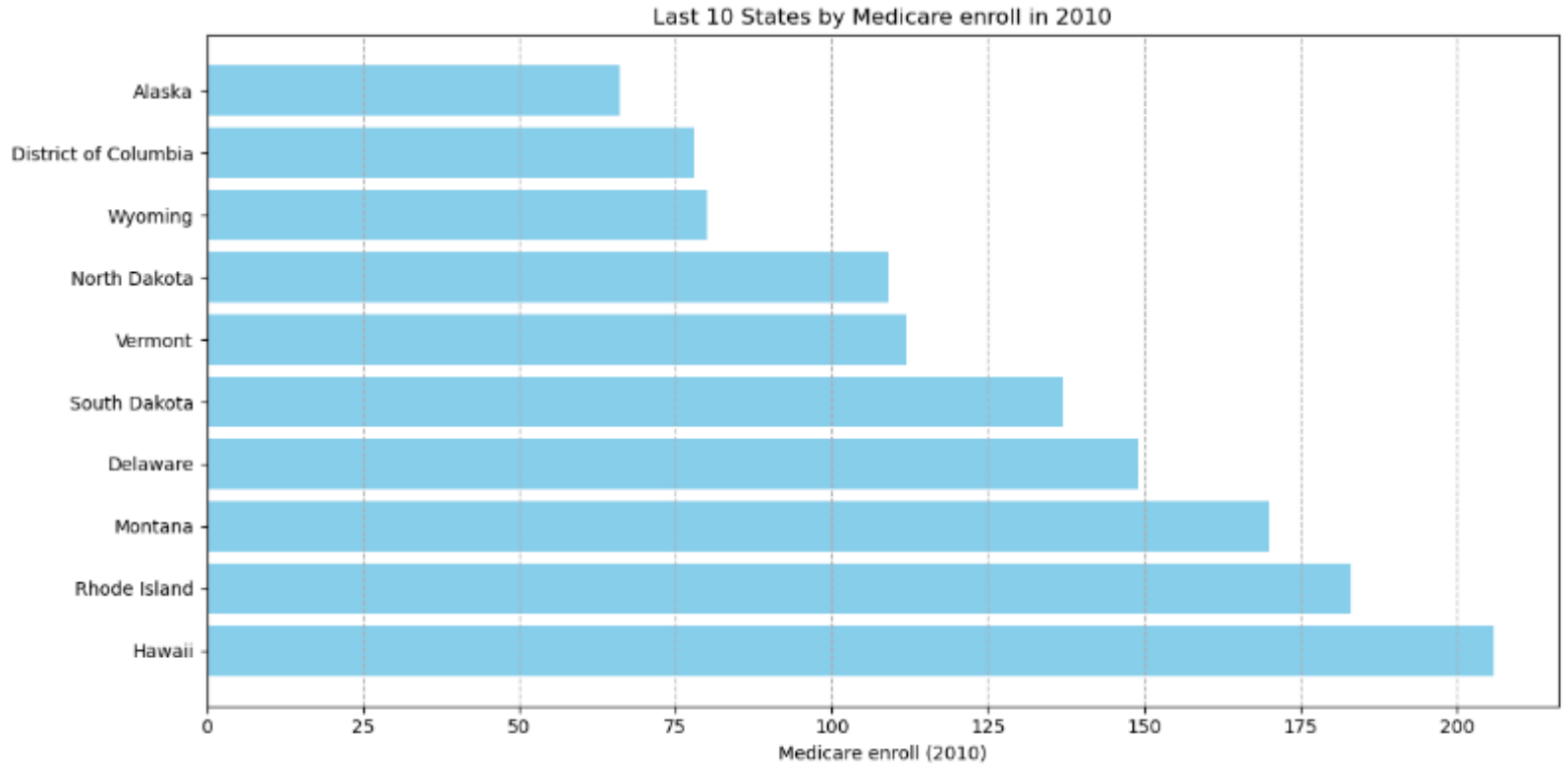
nlargest()



	State	Medicare 2010
4	California	4757
9	Florida	3375
43	Texas	3001
32	New York	2988
38	Pennsylvania	2283
35	Ohio	1901
13	Illinois	1839
22	Michigan	1651
33	North Carolina	1490
30	New Jersey	1327

The last ten states by Medicare enroll in 2010

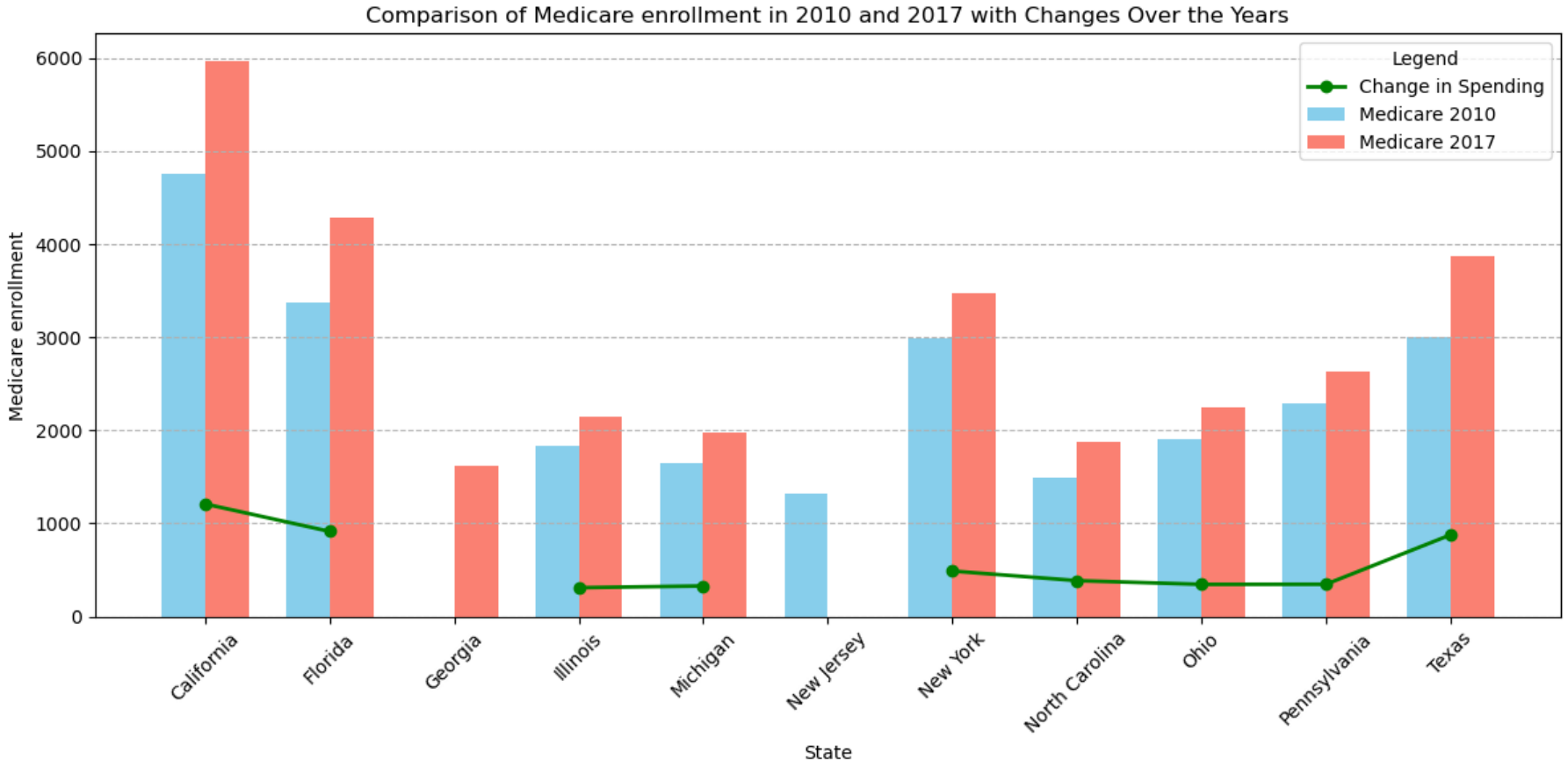
Nsmallest()



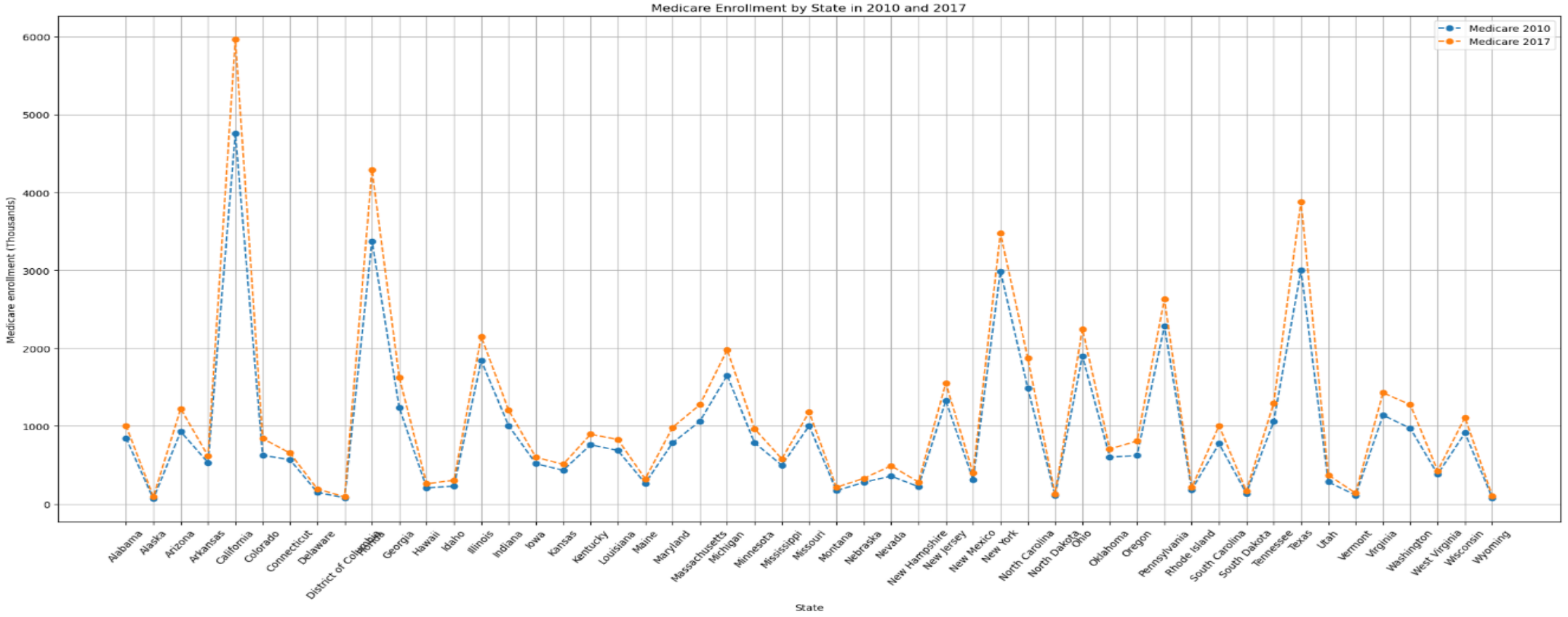
	State	Medicare 2010
1	Alaska	66
8	District of Columbia	78
50	Wyoming	80
34	North Dakota	109
45	Vermont	112
41	South Dakota	137
7	Delaware	149
26	Montana	170
39	Rhode Island	183
11	Hawaii	206

Top ten states by Medicare enroll in 2010 and 2017

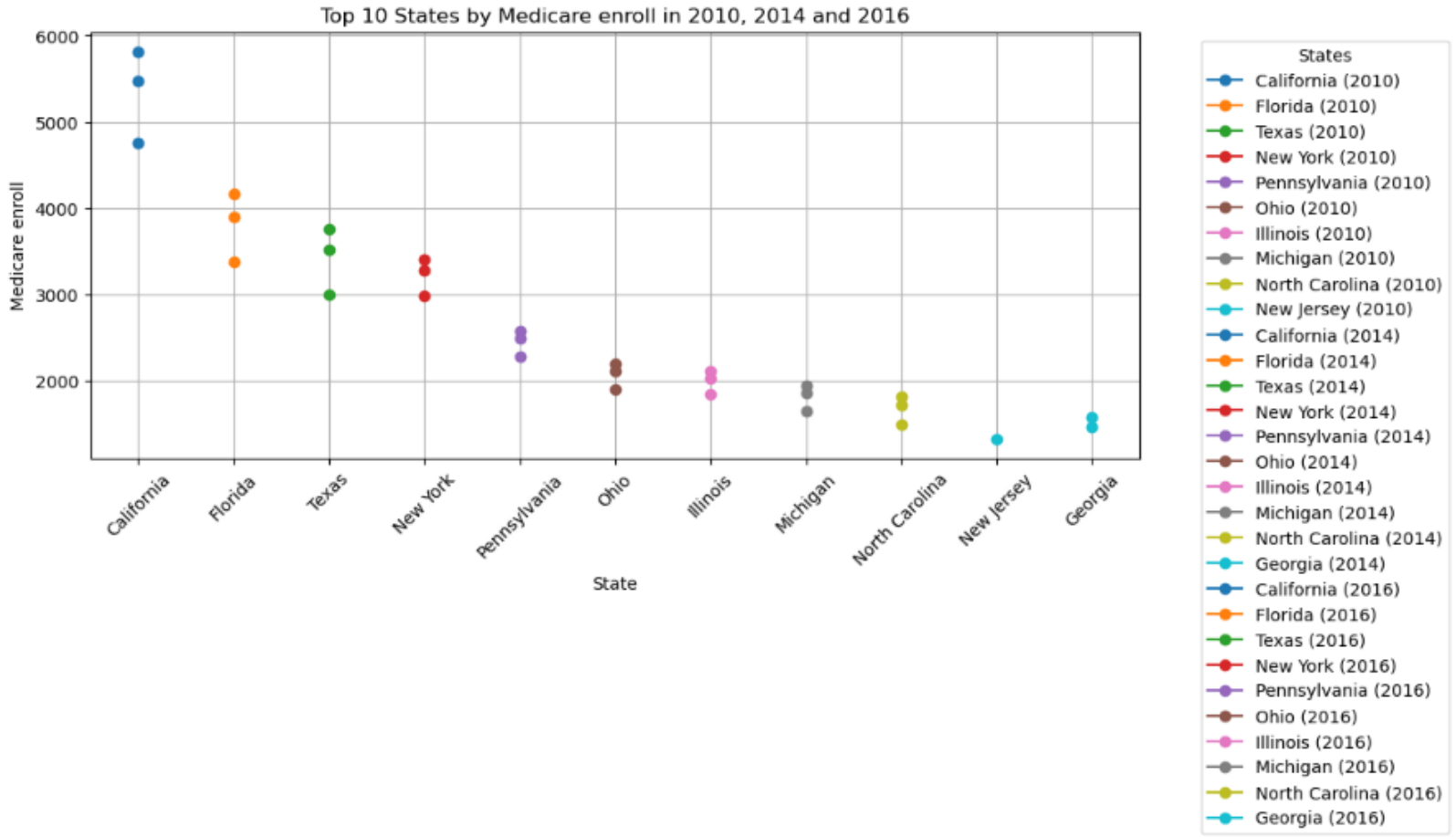
nlargest()



Medicare Enrollment in the US per State



Top Ten States by Medicare Enrollment in 2010, 2014 and 2017



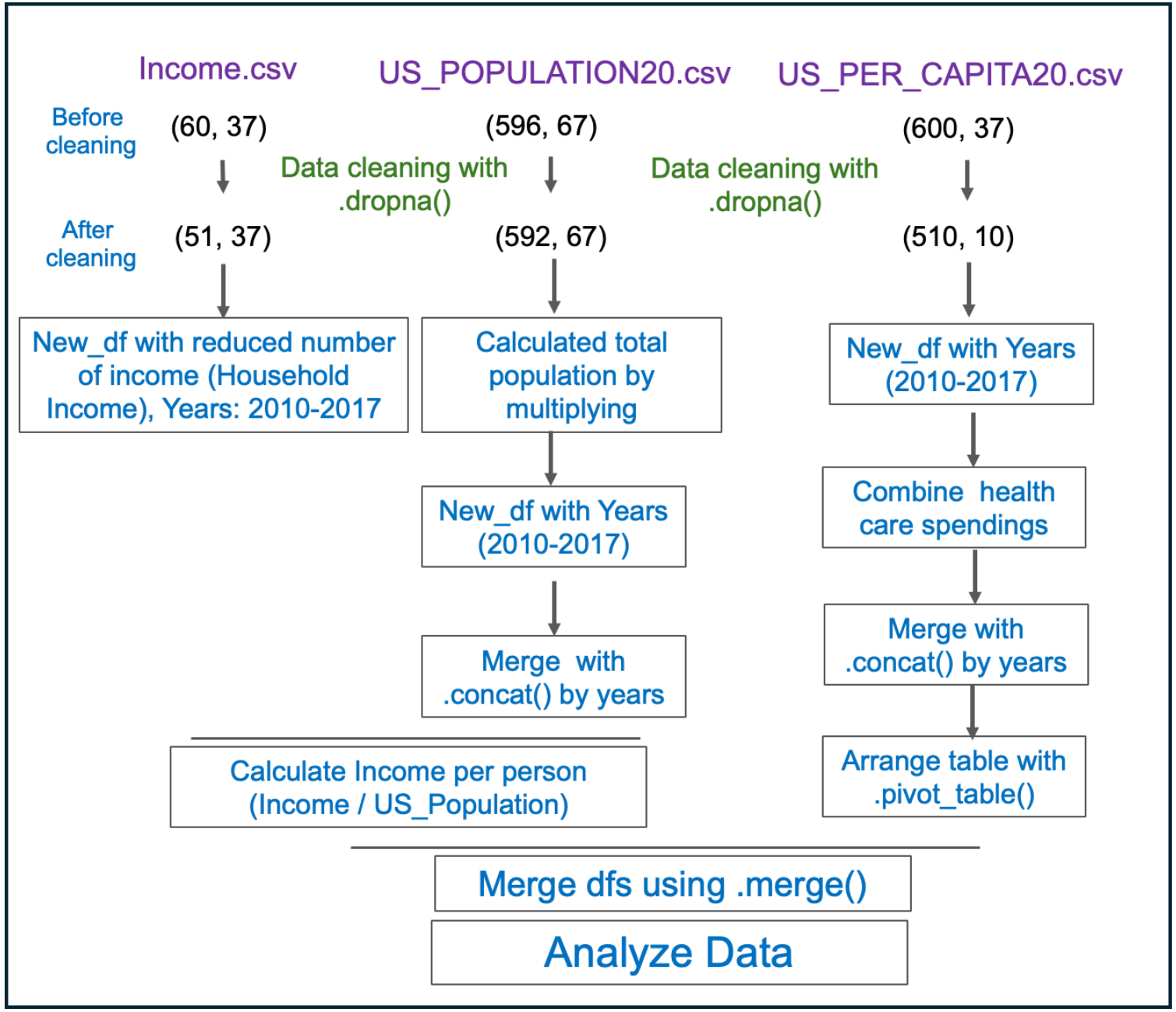
4. The relation of high-income earners population and healthcare spend.

Analyst Nurmaa Dashzeveg

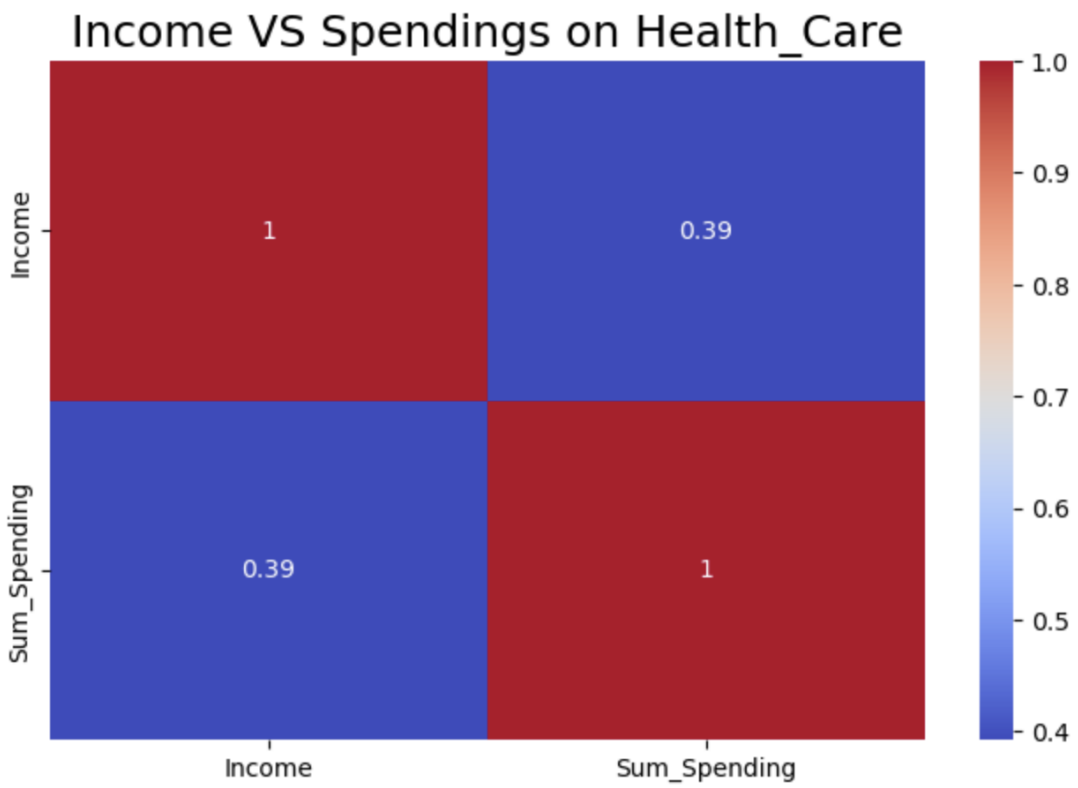
Hypothesis:

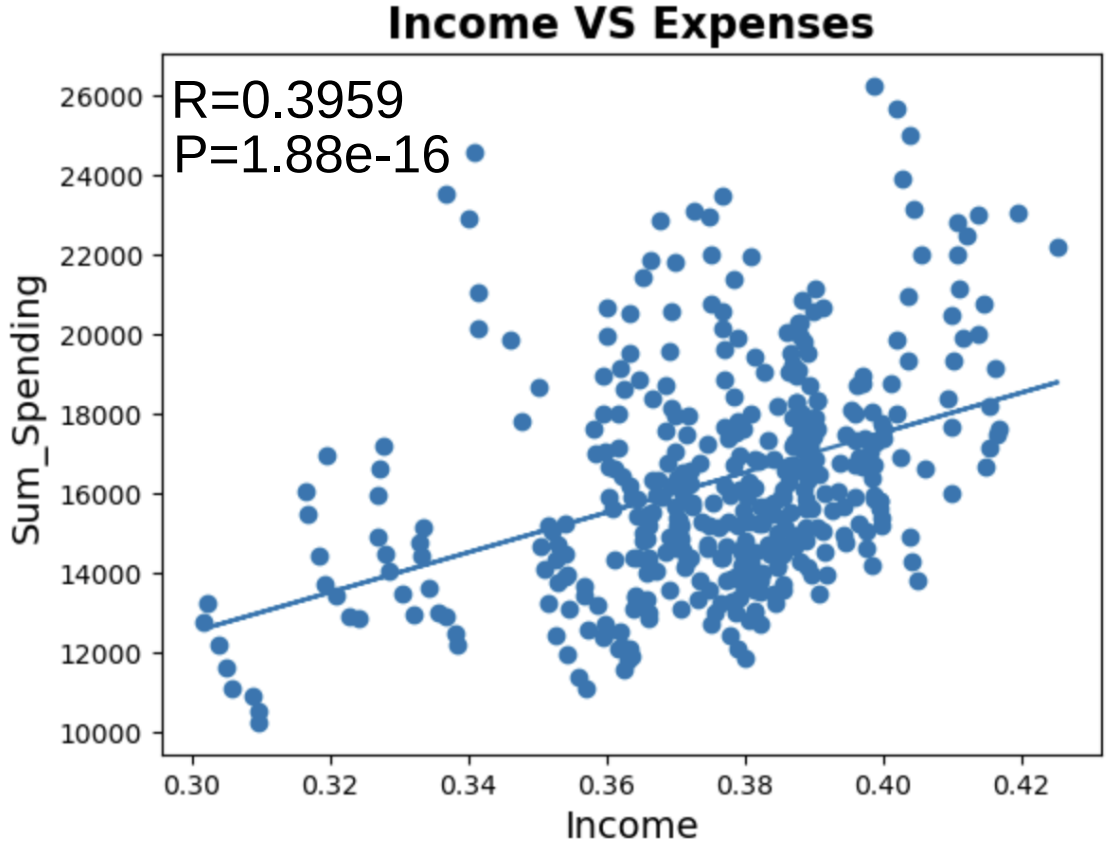
Do wealthier populations (by state) invest more in health care?

Strategy of data analysis



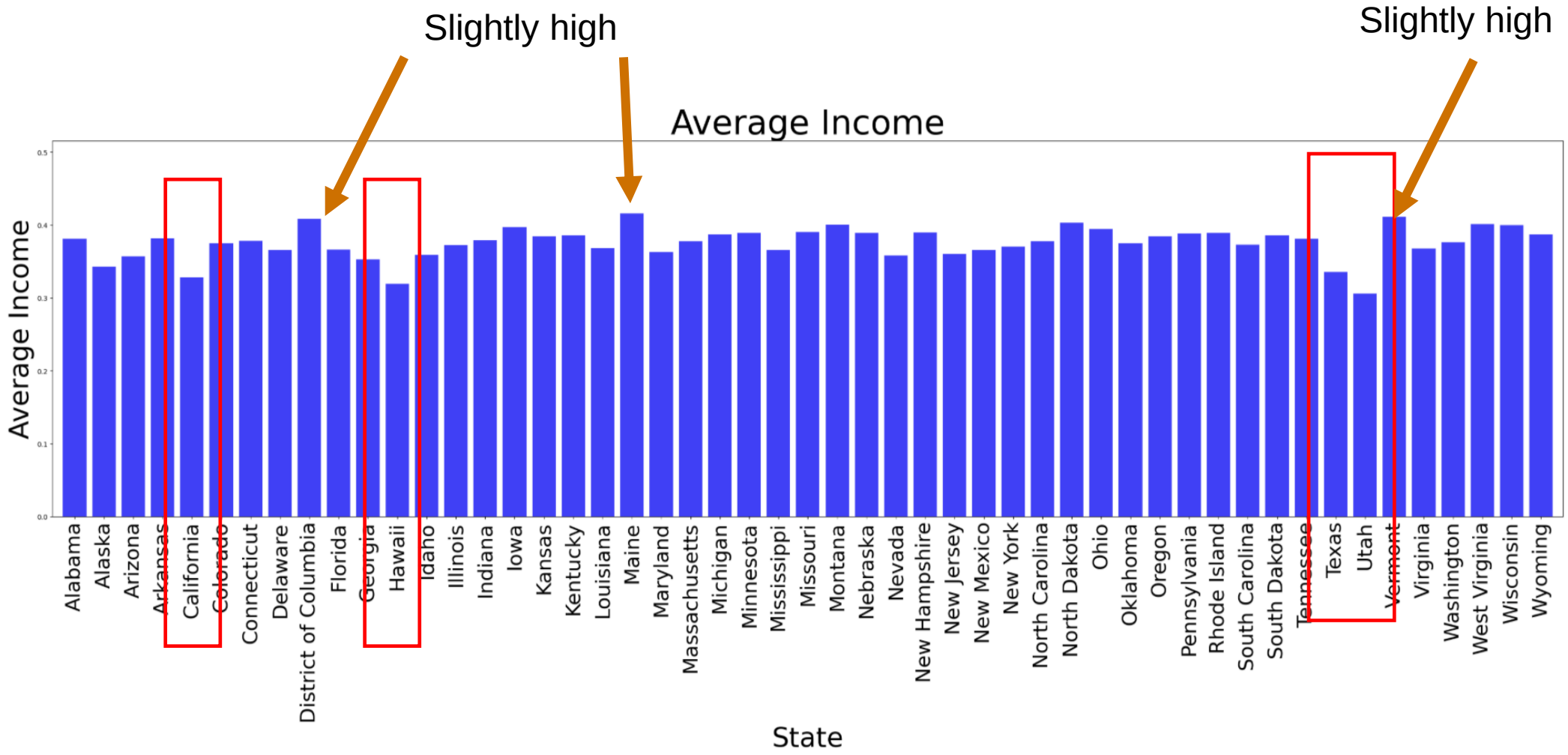
Correlation in between Income and Health Care Expenses





Income and health care expenses have mild positive correlation.

Average Income by States



It might there are very big variations in California

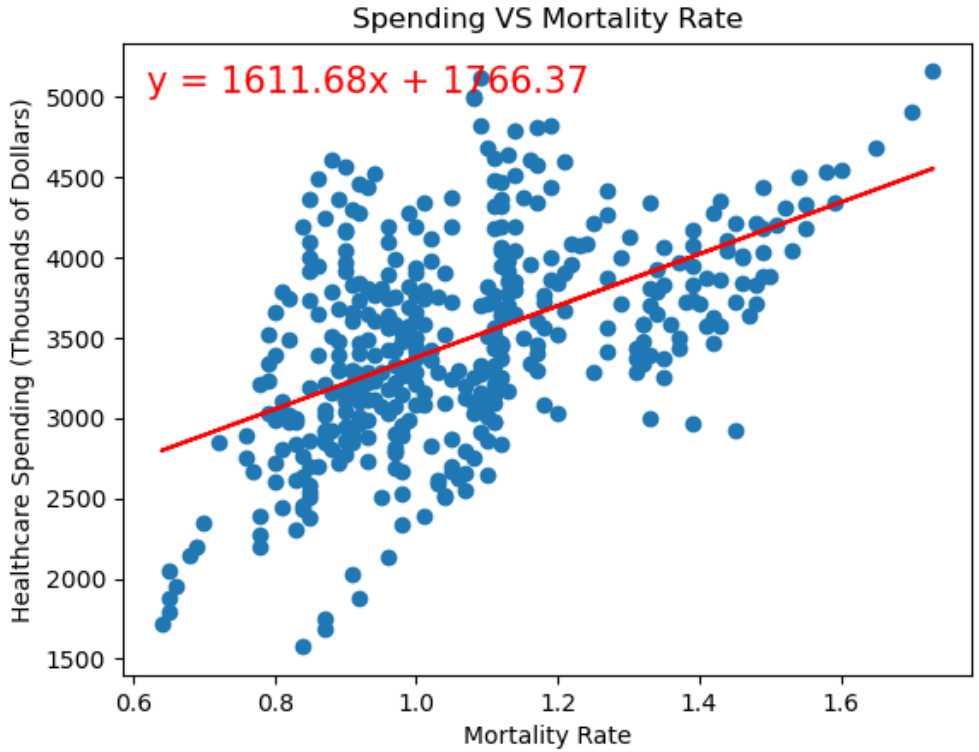
Lowest Income

5. The statistics of mortality rates and healthcare spend.

Analyst Sam Sims

Mortality VS Per Capita Spending

R^2 .267



Revisiting Our Key Questions

Is the US personal healthcare investment connected to a positive mortality outcome?

Is the investment worth it?

The data shows that increased healthcare investment does not lead to better mortality outcomes – so, in the context of this analysis, our team was accurate in our prediction that higher health spend is not positively correlated to a better mortality rate.

Though the Outliers Hypothesis turned out to be true, we *do* believe the investment is worth it, though not in the sole context of mortality rate.

Lessons Learned

- Project Learnings
  - We took on too much of a challenge for our first project; need to remember less is more.
  - It all starts with the data. We quickly realized the importance of finding a good, solid data set would make all of our lives easier. The quality of the data supersedes the quality of the idea itself.
  - Without proper data cleaning, the analysis could have led to inaccurate insights, such as misrepresenting population or Medicare enrollment trends due to faulty or missing data.
  - Where libraries exist – use them over manipulating CSVs!
  - Do more pushes and pulls more often to keep things in sync and avoid duplicating work.
  - Understand the question/hypothesis well before starting data collection and analysis and think logically to analyze data, there are thousands of ways to analyze data.
  - Be Friends with Google!
- Group Learnings
  - With such a short timeframe to deliver, we quickly realized we needed to lean on each other's strongest skills to deliver.
  - We saw the benefit of quickly selecting an idea with a 'good enough' data set, despite a few challenges along the way. Staying organized and having good communication as a team contributes just as much to the success of the project as the code itself.

Thank you

The Outliers Team

Project Requirements

Readme File

- ❑ Must have Complete information of your project.
- ❑ Team members.
- ❑ Introduction of the project.
- ❑ What does it do?
- ❑ Who is the audience?
- ❑ What problem does it solve
- ❑ Where did the data come from?
- ❑ instructions on how to use your software
- ❑ If anything needs to be installed (instructions on packages/libraries, API-Keys etc)

Jupyter Notebook

*1 comprehensive & consolidated notebook for the group

- ❑ Make sure the cells are clean and clear with lots of comments.
- ❑ Data joins and data cleanup should be properly explained and the findings be documented. Also do
- ❑ After doing an EDA add a markdown cell to explain what you found out from the analysis.

Final Report

*between 12-15 pages

- ❑ Team members.
- ❑ Introduction of the project.
- ❑ What problem does it solve
- ❑ What issues did you have while dealing with it
- ❑ Did you join the data with another external source? Explain what you had to do.
- ❑ Complete EDA with explanation.
- ❑ Problems faced
- ❑ Lessons learnt.
- ❑ Any other wishlist.

Presentation & Slides

Total time duration 5 min x # of team members (20~25 minutes) Followed by 15 minutes of Q/A

- ❑ 1-3 introductory slides
- ❑ Introduce the project
- ❑ Who is the audience?
- ❑ What problem does it solve?
- ❑ All the questions. What is your Hypothesis?
- ❑ Who are the team-members and what is their role
- ❑ 2-3 slides for each question. Supported by analysis and Visuals that are clear and well labeled
- ❑ Lessons learnt
- ❑ 1 slide for Conclusion

NOTE: The main repo should have a directory called **FinalProject1** with only these 4 files and a **data** directory with your csv file(s). If you are using your own .py files, please document them in README, the notebook and the final document.

There can be other directories for each team member with their individual work in the git repo.