

CSCI 5323 Final Report

Understanding US Federal Spending Through Establishing Correlation

Ahmed Al Hasani, Mohamed Al-Rasbi

Introduction

Throughout the semester, we constantly discussed the different ways we can utilize the tools we are were taught to solve various problems. We were truly fascinated by the many concepts we learned, pair this with the phrase “datacenter computing” we’ve been hearing previously, we were excited that what seemed to be an abstract and foreign term is finally a matter we are not only familiar with, but becoming adept at implementing and using as we aim to become problem solvers in computer science related fields.

We searched for several days endlessly for different problems that we are genuinely interested and motivated in solving using the tools we learned. We came across many datasets and held so many conversations if it is a unique problem, an interesting dataset, aligns with our curiosities, and a good investment to use our time and the tools we learned to approach and dissect to outputs and outcomes to share with the class.

We were tackling a classification/prediction project at another class and in our research together that we did not want to use this project as another machine/deep learning and prediction models project as well. Therefore, because we were exposed to various talks, we learned that data visualization, management, and aggregation and transformation are skills that some data scientists and engineers lack and we believed it would be the ideal opportunity to learn and apply those skills in this project.

Consequently, we wanted a dataset that is large, preferably hosted by AWS, that will force us to use Spark as our data aggregation tool. We found many datasets on AWS and Kaggle which include:

- Twitter hate speech
- Yelp Review
- World Bank Third World Development Information
- USA Spending

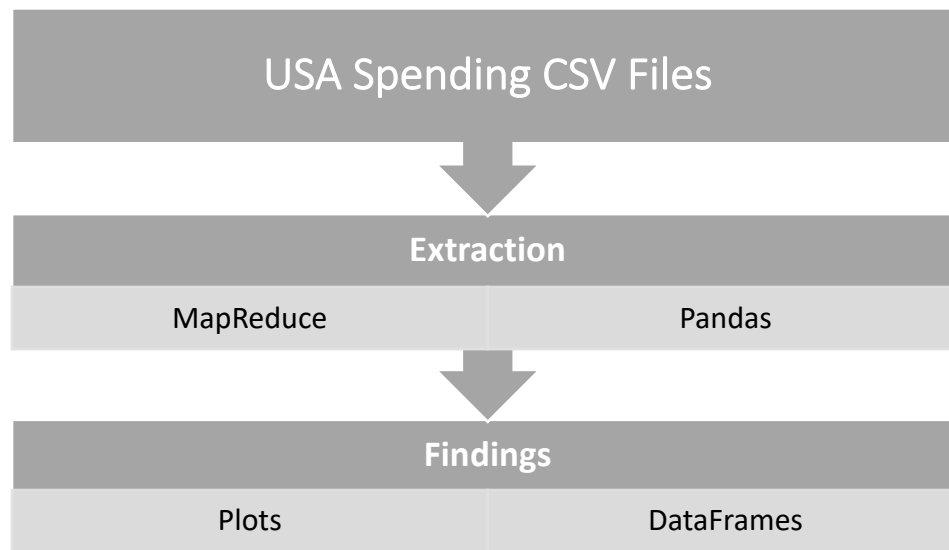
We did not want to use Tweets as our data, because our research already tackles a similar problem. Yelp reviews seemed to us to be a similar problem to the second project assigned in the semester. Finally, the World Bank dataset had many missing information for various data rows.

We finally settled on the USA Spending dataset, and we were excited to tackle the dataset because it was hosted on AWS RDS since its size exceeds 40 GB. Additionally, the dataset has many tables and various file types. Therefore, there are many aspects of this dataset that we can

tackle using various tools. Lastly, the dataset had many attributes that will allow us to extract and analyze different information and answer many different questions.

In this report, we will talk about our dataset in more detail and related work, our initial proposed design versus our actual implemented architecture, our approach, and finally evaluation of the results and conclusion.

System Design



a) USA Spending CSV Files

1. Get CSV files for 6 different states from 2010 to 2018
2. States: California, Oregon, Colorado, Arizona, Ohio, Pennsylvania

b) Extraction

1. MapReduce: use MapReduce for answering the question about political traits
2. Pandas: use Pandas library functions to answer the question about geographical traits

c) Findings

1. Show the final findings and conclusions using plots and dataframes

USA Spending Dataset

USA Spending dataset and database was initially hosted on AWS RDS

(<https://registry.opendata.aws/usaspending/>), however, the datasets were taken down earlier

this year and instead, they are now hosted and are available on USASpending.gov, a governmental website that contains all the information previously available on Amazon.

We are primarily interested in datasets that contain the types of financial aid/support/grants the federal government provides to states and recipients. As a result, we narrowed down our focus to specific datasets with specific attributes.

a) Datasets Used

There are various datasets provided through the website and hosted previously on AWS. We relied on two datasets provided by USASpending.gov which are:

1. Federal Accounts: Account Balances
2. Treasury Accounts: Account Balances

There are numerous and various datasets hosted and provided by USASpending.gov. For instance, there were some datasets stored as .h5 files. These datasets contained:

- An index
- A description of the transaction or account, written in a paragraph
- A URL or special notes

H5 datasets are not applicable to aggregate, nor can we communicate information that is standardized among all accounts and data lines. Hence, we focused only on CSV federal and treasury files.

b) Data Description & Attributes

Federal and Treasury datasets contain several columns which include:

- Award Description
- Cost
- Awarding Agency
- Sub Agency
- County
- Recipient Type

There are more attributes included in the dataset, but we found the above attributes to be the most meaningful ones for our project. Additionally, the earliest available dataset for each state starts 2010 and expands to 2018.

It is important to note and define what are awards. Awards are the money the federal government has promised to pay a recipient. Recipients can include a company, organization,

government entity. The award can be in the form of a contract, grant, loan, or direct payment (and other forms as well). ^[1]

c) Related Work

We did not find many related work online or similar ones either. The only project/research that we found that leverages the same dataset is a WIP codebase that is meant to download, clean and normalize spending data. ^[2]

Problem and Solution

To guide our efforts in extracting meaningful information and trends from the dataset, we had to think of a set of questions and problems that we can abide by as we explore the dataset. As a result, when we studied the different information contained within the dataset, we started seeing recurrences of certain information and we imagined that there are patterns and correlation between different attributes. As a result, if there is a trend and pattern, there is a relationship to be discovered in the dataset.

Since our datasets pertain only to states funding and awards, each state has certain characteristics and traits that will dictate how much funding they receive. While there are many traits we can focus on, we believed the geographical and political traits are the main ones that can help us discover these hidden relationships.

We divided the two traits amongst ourselves. Mohamed tackled the geographical traits, whereas Ahmed tackled the political ones.

Geographical

Geographical traits will encapsulate so many relationships. For example, the size of a state, its population, and the region it belongs to can dictate how much funding that state receives. Initially, we assumed the bigger the state, the more funding it receives. However, some states are very large but not as dense as others. Therefore, we believed focusing on a state's population and its regions is more accurate indicators when establishing correlation with funding and total awards given.

We compared all six states throughout 2010-2018 for two main reasons. First, they vary in population. Second, we chose those 6 states depending on their location: Oregon and California from the west coast, Colorado and Arizona from the mid-west, Ohio and Pennsylvania from the east. Hence, it will be clear if population and region have any effects on the trends.

Political traits can drive many of the relationships with a state's total funding. For instance, how officials prioritize certain issues versus others, the goals and aims they set out for their terms,

and the lobbying efforts in attempt to change policies while adhering to their constituents' needs and demands.

Political

For this project, we narrowed our focus only on a state's senators and their political agendas or political positions on two main issues which are public/government spending on health and education. We wanted to find whether a state is anti or pro public funding by determining a state senators' positions on these two matters. If a state is pro public funding then we assumed that funding for health and education would be high, and vice versa.

Ontheissues.org is a website that lists a political figure/leader's positions on various issues which also include health and education. They list both a political leader's rating in that area given by another institute and his/her position or vote for various policies.

For the education and health stance for each senate, the ratings (if applicable) are given by the NEA (National Education Association) and APHA (American Public Health Association). To narrow the scope and compare two states that are close in traits but are generally considered to follow or exhibit political characteristics that are different, I focused on Colorado (generally democrat/blue state) and Arizona (Red state/republican). It is important to note that Colorado's population is 5.6 million, whereas Arizona's is roughly around 7 million.

Colorado's senators, from 2010 to 2018, and their rating in public are:

- Mark Udall (2009 – 2015): (%80, %75) ^[3]
- Michael Bennet: against Affordable Care repeal
- Cory Gardner (2015 – Present): Against government health and anti public health funding

Arizona's senators, from 2010 to 2018, and their rating in public are:

- Jon Kyl (2018 – Present): (18%, 0%)
- Jeff Flake (2013 – Present): (%17, %33)
- John McCain (- 2018): (%45, %25)

Overall, Colorado is considered to be pro-public health and educational public funding based on our created metric whereas Arizona is considered to be anti-public health and educational public funding.

Findings

Political

To narrow the most helpful attributes I believed that will aid me in communicating meaningful results with the readers and the class during the presentations from the dataset are:

- Awarding Agency
- Total Funding
- CFDA Title (Catalog of Federal Domestic Assistance), which contains a description of the award given.

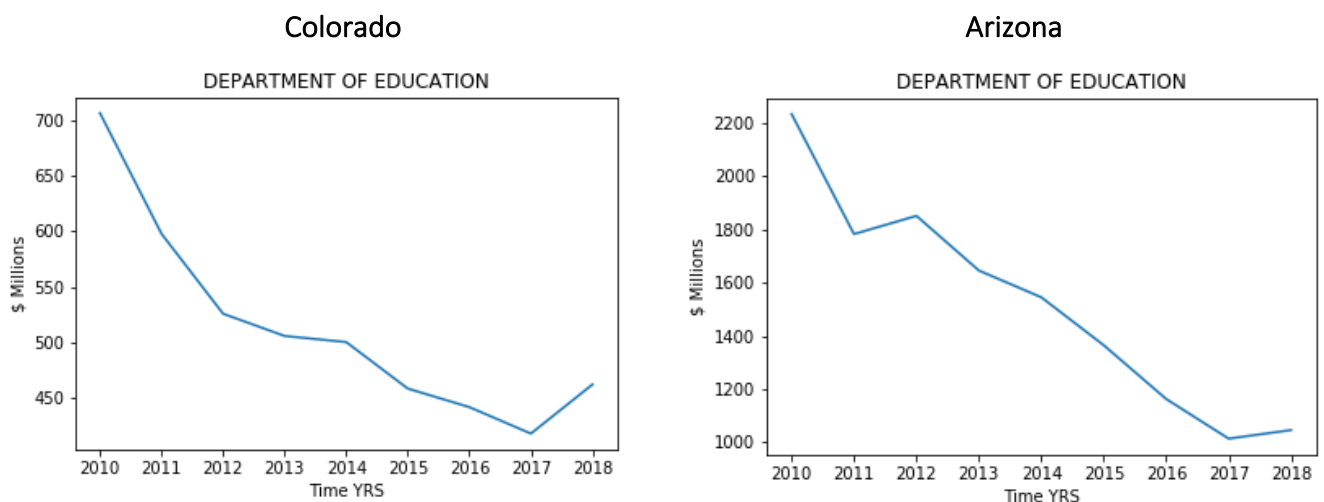
These three attributes communicate agency spending, how much each award costs, and the types of awards each agency is funding.

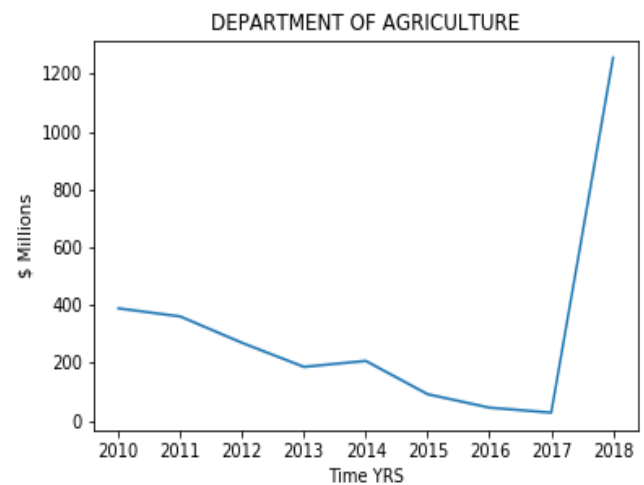
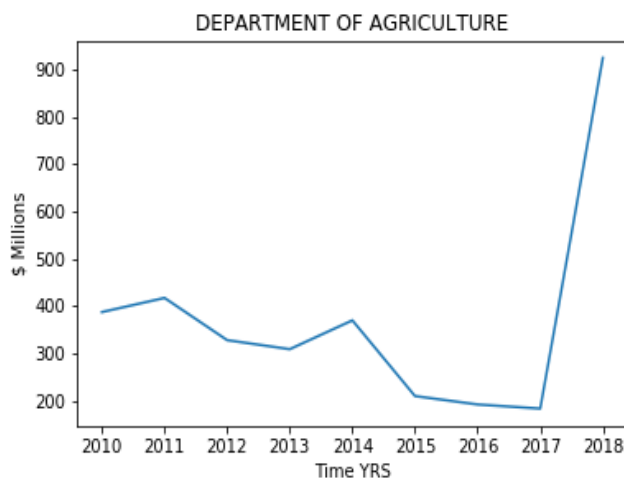
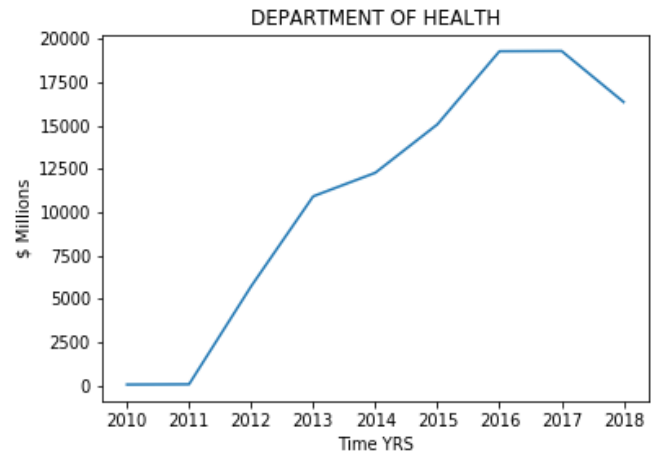
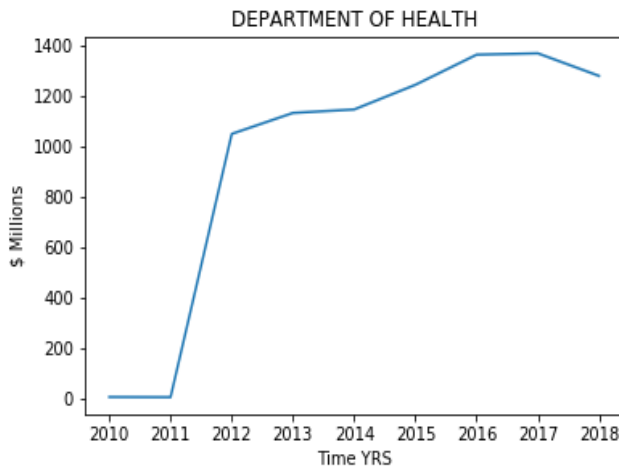
Initially, to get a sense of the data, I wanted to see the agency that provides the most funding (based on how many data rows are included for that agency) as well as the most frequent awards. Both states had the same result. The most awarding agency is the Department of Education, whereas the most frequent award is the Pell Grant, a higher education grant given to students. I was initially surprised that the two states shared the same results.

From there, I dove deeper in each attribute. I wanted to know the most active departments in each year. As a result, I believed Map Reduce would be essential in tackling this question. I wanted to map each department to frequency and then cost. Then, I wanted to plot the top 4 agencies in a year and repeat the steps for each year and for each state.

I noticed interesting trends, both states from 2010 to 2012 have the DOE as the highest awarding agency, then in 2012 to 2016, the highest awarding agency is DOH, which is then overtaken by Social Security Administration in 2017, however, this only happens in 2018 for Colorado, contrary to our initial assumptions given how anti public spending Arizona is compared to Colorado.

Summary findings for both states for awarding agencies:



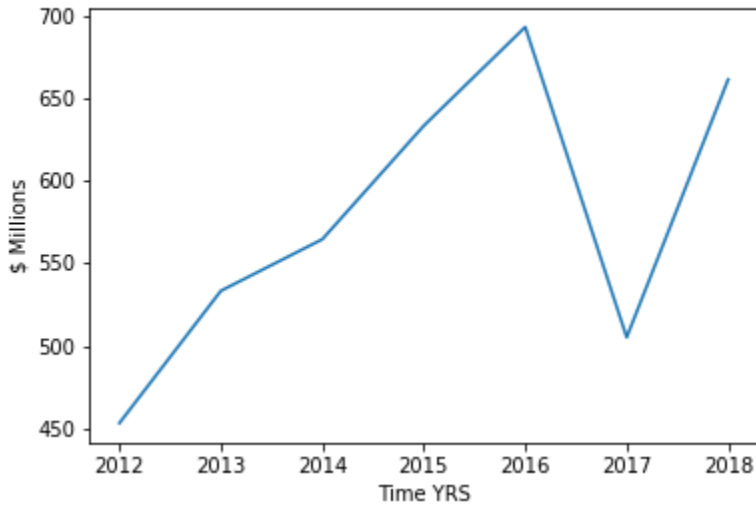


I wanted to apply the same method to awards and understand the highest awards by funding in each year for each state. For Colorado, from 2010 to 2016, the Pell Grant almost dominated the top spot each year followed by Medicare grants related to medical insurance, and hospital insurance. In 2017 and 2018, Colorado's highest funding awards were related to Social Security, such as retirement insurance, disability insurance or survivors insurance.

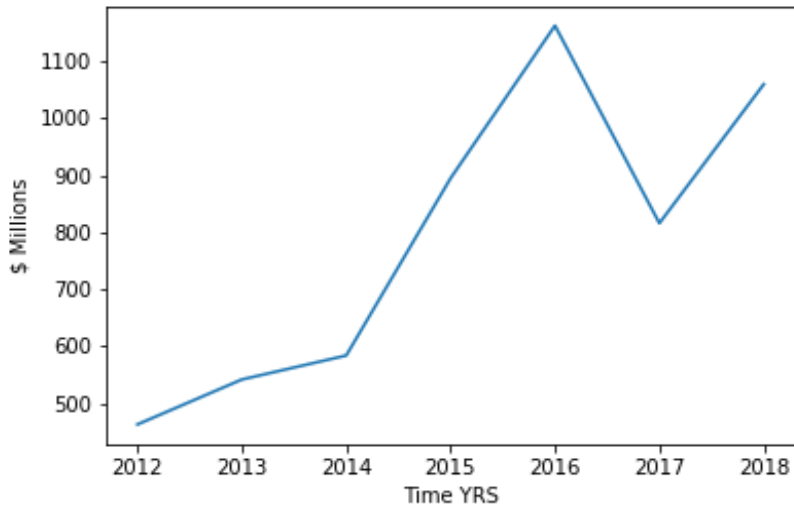
On the other hand, Arizona shared the similar funding behaviors as Colorado. The Pell Grant dominated the top spot from 2010 to 2015, when it was overtaken by Medicare Awards. In 2017 and 2018, both awards were overtaken by Social Security related awards. Again, both states follow similar trends.

Summary findings for both states in the award funding category:

Colorado: Medicare



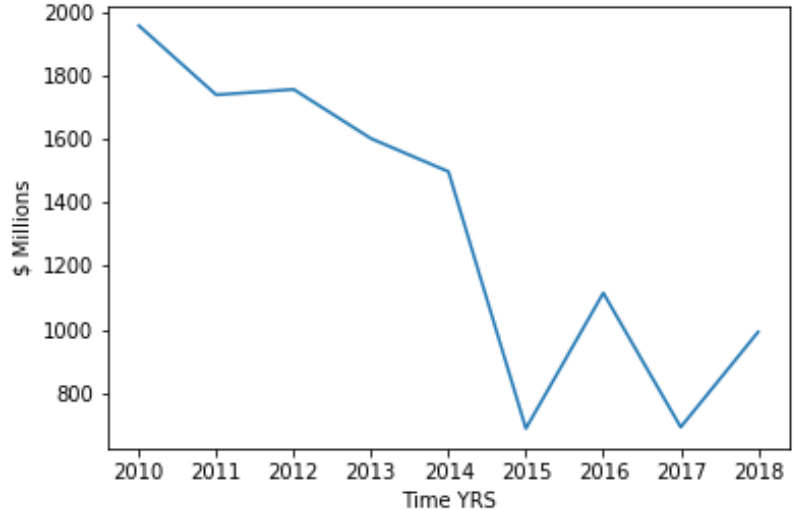
Arizona: Medicare



Colorado: Pell



Arizona: Pell



Lastly, I wanted to link the different awards the department of education and health fund for each state by executing map reduce functions that will count the number of occurrences for each award for each agency, then return dataframes based on the desired departments.

Both states again exhibited similar behaviors for the two departments. It is worthy to note that health spending was minimal from 2010 to 2012, and since then it steadily increased especially in 2017 and 2018. On the other hand, educational related awards were steady throughout the years.

Health Awards and Agency Count

arizona 2017

	cfda_title	awarding_sub_agency_name	Count
1	AFFORDABLE CARE ACT – NATIONAL HEALTH SERVICE CORPS	HEALTH RESOURCES AND SERVICES ADMINISTRATION	6
19	DISADVANTAGED HEALTH PROFESSIONS FACULTY LOAN REPAYMENT (FLRP) AND MINORITY FACULTY FELLOWSHIP P...	HEALTH RESOURCES AND SERVICES ADMINISTRATION	1
20	DISADVANTAGED HEALTH PROFESSIONS FACULTY LOAN REPAYMENT PROGRAM (FLRP)	HEALTH RESOURCES AND SERVICES ADMINISTRATION	2
32	HEALTH PROFESSIONS PRE-GRADUATE SCHOLARSHIP PROGRAM FOR INDIANS	INDIAN HEALTH SERVICE	1
33	HEALTH PROFESSIONS PREPARATORY SCHOLARSHIP PROGRAM FOR INDIANS	INDIAN HEALTH SERVICE	1
34	HEALTH PROFESSIONS SCHOLARSHIP PROGRAM	INDIAN HEALTH SERVICE	9
44	INDIAN HEALTH SERVICE EDUCATIONAL LOAN REPAYMENT	INDIAN HEALTH SERVICE	10
70	NATIONAL HEALTH SERVICE CORPS	HEALTH RESOURCES AND SERVICES ADMINISTRATION	35
71	NATIONAL INSTITUTE ON MINORITY HEALTH AND HEALTH DISPARITIES (NIMHD) EXTRAMURAL LOAN REPAYMENT P...	NATIONAL INSTITUTES OF HEALTH	4
72	NATIONAL INSTITUTES OF HEALTH LOAN REPAYMENT PROGRAM FOR CLINICAL RESEARCHERS	NATIONAL INSTITUTES OF HEALTH	5
73	NATIONAL INSTITUTES OF HEALTH PEDIATRIC RESEARCH LOAN REPAYMENT PROGRAM	NATIONAL INSTITUTES OF HEALTH	2

arizona 2018

	cfda_title	awarding_sub_agency_name	Count
27	HEALTH PROFESSIONS PRE-GRADUATE SCHOLARSHIP PROGRAM FOR INDIANS	INDIAN HEALTH SERVICE	1
28	HEALTH PROFESSIONS PREPARATORY SCHOLARSHIP PROGRAM FOR INDIANS	INDIAN HEALTH SERVICE	3
29	HEALTH PROFESSIONS SCHOLARSHIP PROGRAM	INDIAN HEALTH SERVICE	9
40	INDIAN HEALTH SERVICE EDUCATIONAL LOAN REPAYMENT	INDIAN HEALTH SERVICE	19
59	NATIONAL HEALTH SERVICE CORPS	HEALTH RESOURCES AND SERVICES ADMINISTRATION	57
60	NATIONAL INSTITUTE ON MINORITY HEALTH AND HEALTH DISPARITIES (NIMHD) EXTRAMURAL LOAN REPAYMENT P...	NATIONAL INSTITUTES OF HEALTH	3
61	NATIONAL INSTITUTES OF HEALTH LOAN REPAYMENT PROGRAM FOR CLINICAL RESEARCHERS	NATIONAL INSTITUTES OF HEALTH	1
62	NATIONAL INSTITUTES OF HEALTH PEDIATRIC RESEARCH LOAN REPAYMENT PROGRAM	NATIONAL INSTITUTES OF HEALTH	2

colorado 2017

	cfda_title	awarding_sub_agency_name	Count
0	AFFORDABLE CARE ACT – NATIONAL HEALTH SERVICE CORPS	HEALTH RESOURCES AND SERVICES ADMINISTRATION	3
22	HEALTH PROFESSIONS PRE-GRADUATE SCHOLARSHIP PROGRAM FOR INDIANS	INDIAN HEALTH SERVICE	1
23	HEALTH PROFESSIONS SCHOLARSHIP PROGRAM	INDIAN HEALTH SERVICE	12
29	INDIAN HEALTH SERVICE EDUCATIONAL LOAN REPAYMENT	INDIAN HEALTH SERVICE	6
49	NATIONAL HEALTH SERVICE CORPS	HEALTH RESOURCES AND SERVICES ADMINISTRATION	27
50	NATIONAL INSTITUTE ON MINORITY HEALTH AND HEALTH DISPARITIES (NIMHD) EXTRAMURAL LOAN REPAYMENT P...	NATIONAL INSTITUTES OF HEALTH	4
51	NATIONAL INSTITUTES OF HEALTH LOAN REPAYMENT PROGRAM FOR CLINICAL RESEARCHERS	NATIONAL INSTITUTES OF HEALTH	24
52	NATIONAL INSTITUTES OF HEALTH PEDIATRIC RESEARCH LOAN REPAYMENT PROGRAM	NATIONAL INSTITUTES OF HEALTH	16

colorado 2018

	cfda_title	awarding_sub_agency_name	Count
11	DISADVANTAGED HEALTH PROFESSIONS FACULTY LOAN REPAYMENT PROGRAM (FLRP)	HEALTH RESOURCES AND SERVICES ADMINISTRATION	1
20	HEALTH PROFESSIONS PRE-GRADUATE SCHOLARSHIP PROGRAM FOR INDIANS	INDIAN HEALTH SERVICE	1
21	HEALTH PROFESSIONS SCHOLARSHIP PROGRAM	INDIAN HEALTH SERVICE	3
27	INDIAN HEALTH SERVICE EDUCATIONAL LOAN REPAYMENT	INDIAN HEALTH SERVICE	4
42	NATIONAL HEALTH SERVICE CORPS	HEALTH RESOURCES AND SERVICES ADMINISTRATION	48
43	NATIONAL INSTITUTE ON MINORITY HEALTH AND HEALTH DISPARITIES (NIMHD) EXTRAMURAL LOAN REPAYMENT P...	NATIONAL INSTITUTES OF HEALTH	1
44	NATIONAL INSTITUTES OF HEALTH LOAN REPAYMENT PROGRAM FOR CLINICAL RESEARCHERS	NATIONAL INSTITUTES OF HEALTH	7
45	NATIONAL INSTITUTES OF HEALTH PEDIATRIC RESEARCH LOAN REPAYMENT PROGRAM	NATIONAL INSTITUTES OF HEALTH	4

As can be seen, both states are very similar in the count of awards related to health agencies.

Educational Awards and Agency Count

colorado 2010

	cfda_title	awarding_sub_agency_name	Count
0	ACADEMIC COMPETITIVENESS GRANTS	DEPARTMENT OF EDUCATION	3274
1	CENTERS FOR INDEPENDENT LIVING, RECOVERY ACT.	DEPARTMENT OF EDUCATION	26
8	EDUCATION JOBS FUND	DEPARTMENT OF EDUCATION	2
16	FEDERAL PELL GRANT PROGRAM	DEPARTMENT OF EDUCATION	6275
18	FEDERAL SUPPLEMENTAL EDUCATIONAL OPPORTUNITY GRANTS	DEPARTMENT OF EDUCATION	60
20	FEDERAL WORK-STUDY PROGRAM	DEPARTMENT OF EDUCATION	55
34	NATIONAL SCIENCE AND MATHEMATICS ACCESS TO RETAIN TALENT (SMART) GRANTS	DEPARTMENT OF EDUCATION	1751
44	TEACHER EDUCATION ASSISTANCE FOR COLLEGE AND HIGHER EDUCATION GRANTS (TEACH GRANTS)	DEPARTMENT OF EDUCATION	571

arizona 2010

	cfda_title	awarding_sub_agency_name	Count
0	ACADEMIC COMPETITIVENESS GRANTS	DEPARTMENT OF EDUCATION	1737
2	CENTERS FOR INDEPENDENT LIVING, RECOVERY ACT.	DEPARTMENT OF EDUCATION	18
9	EDUCATION JOBS FUND	DEPARTMENT OF EDUCATION	2
15	FEDERAL PELL GRANT PROGRAM	DEPARTMENT OF EDUCATION	6493
17	FEDERAL SUPPLEMENTAL EDUCATIONAL OPPORTUNITY GRANTS	DEPARTMENT OF EDUCATION	54
19	FEDERAL WORK-STUDY PROGRAM	DEPARTMENT OF EDUCATION	50
31	NATIONAL SCIENCE AND MATHEMATICS ACCESS TO RETAIN TALENT (SMART) GRANTS	DEPARTMENT OF EDUCATION	777
39	TEACHER EDUCATION ASSISTANCE FOR COLLEGE AND HIGHER EDUCATION GRANTS (TEACH GRANTS)	DEPARTMENT OF EDUCATION	925

colorado 2014

	cfda_title	awarding_sub_agency_name	Count
0	ACADEMIC COMPETITIVENESS GRANTS	DEPARTMENT OF EDUCATION	2
11	FEDERAL PELL GRANT PROGRAM	DEPARTMENT OF EDUCATION	10340
12	FEDERAL SUPPLEMENTAL EDUCATIONAL OPPORTUNITY GRANTS	DEPARTMENT OF EDUCATION	95
13	FEDERAL WORK-STUDY PROGRAM	DEPARTMENT OF EDUCATION	71
28	NATIONAL SCIENCE AND MATHEMATICS ACCESS TO RETAIN TALENT (SMART) GRANTS	DEPARTMENT OF EDUCATION	1
35	POSTSECONDARY EDUCATION SCHOLARSHIPS FOR VETERAN'S DEPENDENTS	DEPARTMENT OF EDUCATION	1
46	TEACHER EDUCATION ASSISTANCE FOR COLLEGE AND HIGHER EDUCATION GRANTS (TEACH GRANTS)	DEPARTMENT OF EDUCATION	161

arizona 2014

	cfda_title	awarding_sub_agency_name	Count
0	ACADEMIC COMPETITIVENESS GRANTS	DEPARTMENT OF EDUCATION	1
8	FEDERAL PELL GRANT PROGRAM	DEPARTMENT OF EDUCATION	12326
9	FEDERAL SUPPLEMENTAL EDUCATIONAL OPPORTUNITY GRANTS	DEPARTMENT OF EDUCATION	95
10	FEDERAL WORK-STUDY PROGRAM	DEPARTMENT OF EDUCATION	89
27	POSTSECONDARY EDUCATION SCHOLARSHIPS FOR VETERAN'S DEPENDENTS	DEPARTMENT OF EDUCATION	4
35	TEACHER EDUCATION ASSISTANCE FOR COLLEGE AND HIGHER EDUCATION GRANTS (TEACH GRANTS)	DEPARTMENT OF EDUCATION	458

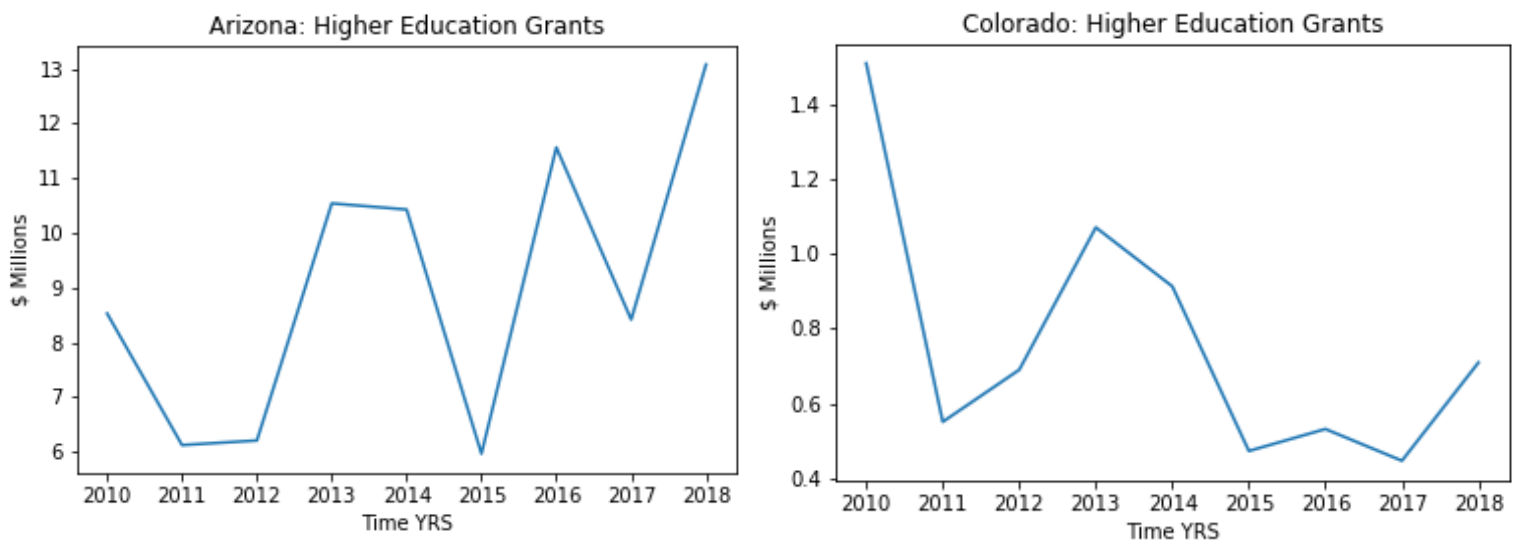
colorado 2017

	cfda_title	awarding_sub_agency_name	Count
17	FEDERAL PELL GRANT PROGRAM	DEPARTMENT OF EDUCATION	6778
18	FEDERAL SUPPLEMENTAL EDUCATIONAL OPPORTUNITY GRANTS	DEPARTMENT OF EDUCATION	183
19	FEDERAL WORK-STUDY PROGRAM	DEPARTMENT OF EDUCATION	175
58	PELL	DEPARTMENT OF EDUCATION	2510
62	POSTSECONDARY EDUCATION SCHOLARSHIPS FOR VETERAN'S DEPENDENTS	DEPARTMENT OF EDUCATION	1
85	TEACHER EDUCATION ASSISTANCE FOR COLLEGE AND HIGHER EDUCATION GRANTS (TEAC	DEPARTMENT OF EDUCATION	36
86	TEACHER EDUCATION ASSISTANCE FOR COLLEGE AND HIGHER EDUCATION GRANTS (TEACH GRANTS)	DEPARTMENT OF EDUCATION	67

arizona 2017

	cfda_title	awarding_sub_agency_name	Count
27	FEDERAL PELL GRANT PROGRAM	DEPARTMENT OF EDUCATION	8884
28	FEDERAL SUPPLEMENTAL EDUCATIONAL OPPORTUNITY GRANTS	DEPARTMENT OF EDUCATION	158
29	FEDERAL WORK-STUDY PROGRAM	DEPARTMENT OF EDUCATION	153
55	IRAQ AND AFGHANISTAN SERVICE GRANT	DEPARTMENT OF EDUCATION	1
78	PELL	DEPARTMENT OF EDUCATION	3184
83	POSTSECONDARY EDUCATION SCHOLARSHIPS FOR VETERAN'S DEPENDENTS	DEPARTMENT OF EDUCATION	3
110	TEACHER EDUCATION ASSISTANCE FOR COLLEGE AND HIGHER EDUCATION GRANTS (TEAC	DEPARTMENT OF EDUCATION	123
111	TEACHER EDUCATION ASSISTANCE FOR COLLEGE AND HIGHER EDUCATION GRANTS (TEACH GRANTS)	DEPARTMENT OF EDUCATION	364

Both states have similar awards and counts, but Arizona, unlike our assumption and hypothesis, has a higher award and funding frequency than Colorado. Additionally, there are various higher education grants, and we wanted to explore this award specifically between the two states.



While Arizona's senates overall exhibited a lower public education funding's rating than Colorado's, the state has a higher funding in that area.

Geographical

The question that I wanted to answer is, can we find correlation between population or location and the trends? To answer this question, I explored 4 trends for every state of the six states from 2010 to 2018.

Note:

Populations:

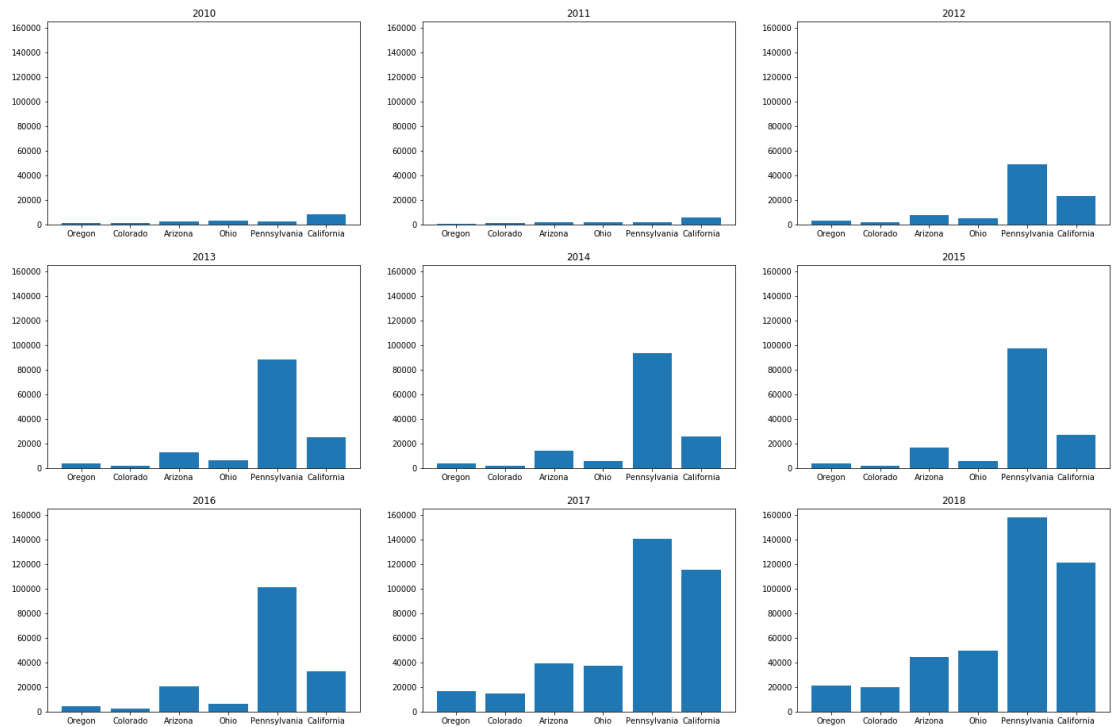
- Oregon: 4.143 million
- Colorado: 5.607 million
- Arizona: 7.016 million
- Ohio: 11.66 million
- Pennsylvania: 12.81 million
- California: 39.54 million

Regions:

- West coast: Oregon, California
- West: Arizona, Colorado
- East: Ohio, Pennsylvania

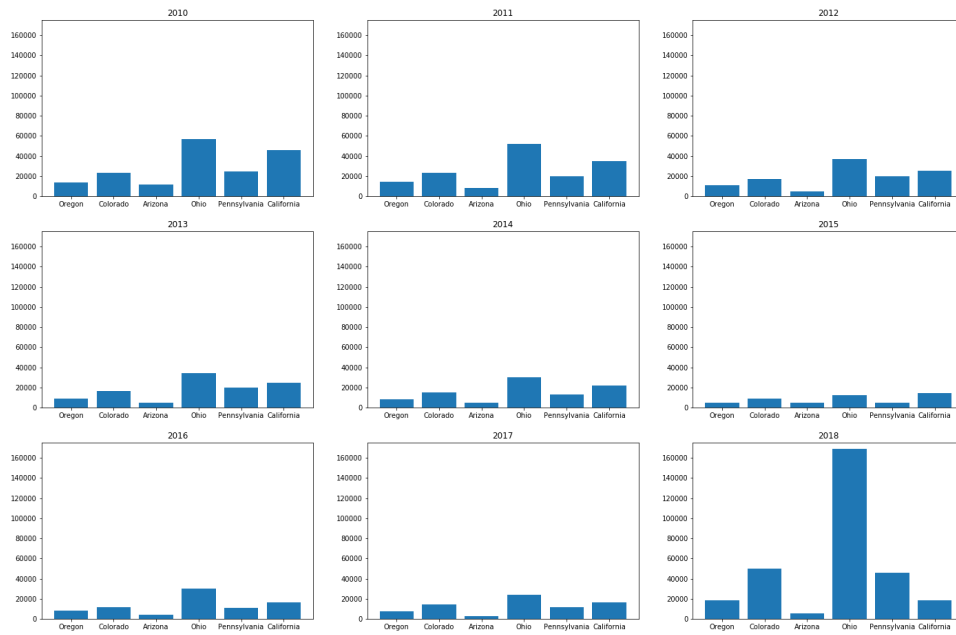
The first plot shows a comparison of the total funding every FY in the six states. We can see that there is a slight correlation between population and the total funding, however, that is not always the case. There could be some outliers. For example, California has more population than Pennsylvania, but we see that in almost all years, Pennsylvania has more total funding.

Total funding every FY in different states (in millions)

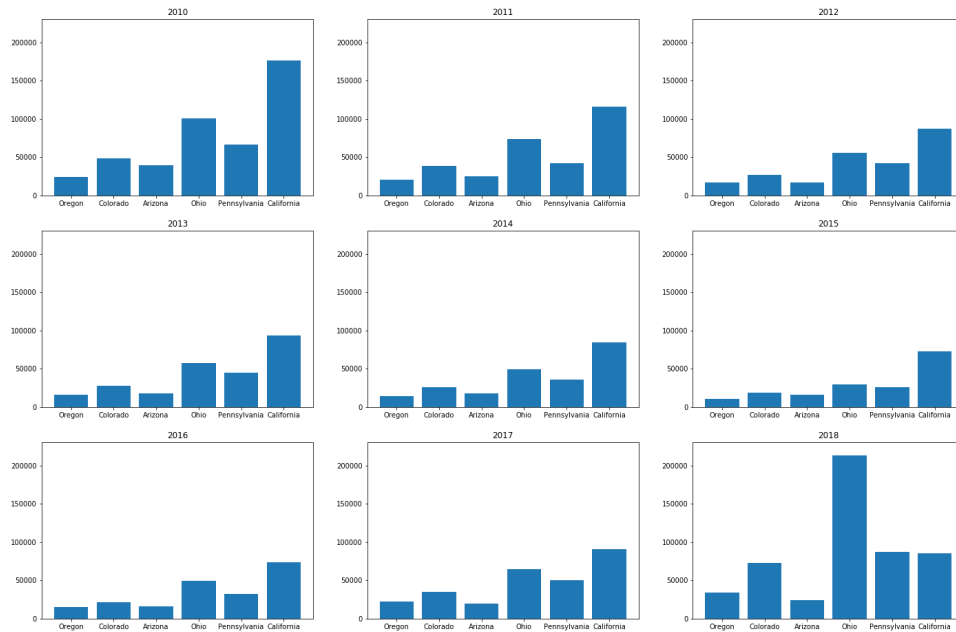


The second and the third plots show a comparison of the number of unique and total awards. The number of the total awards could be very big for some states, but when we take a look at the number of unique awards, we see there is a big difference. However, there is no correlation between these two trends and population or region.

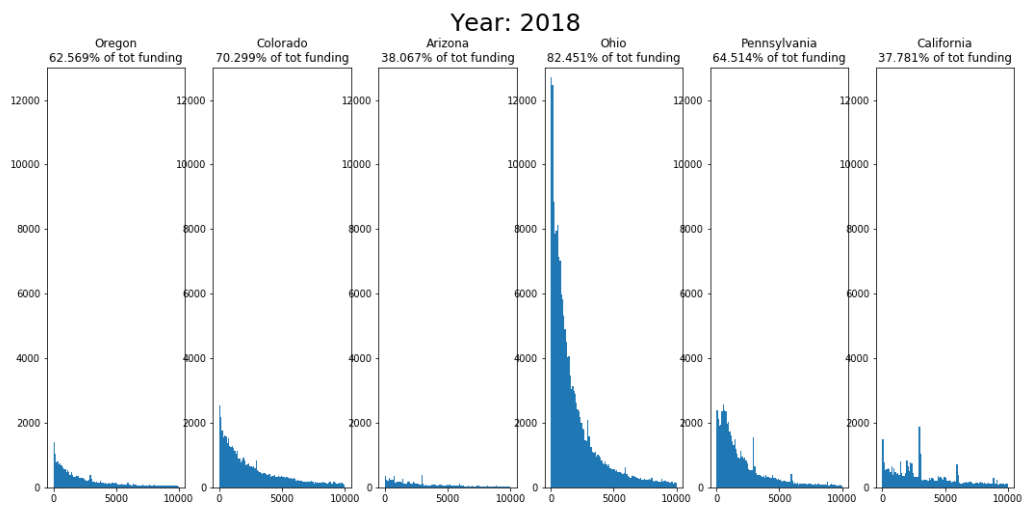
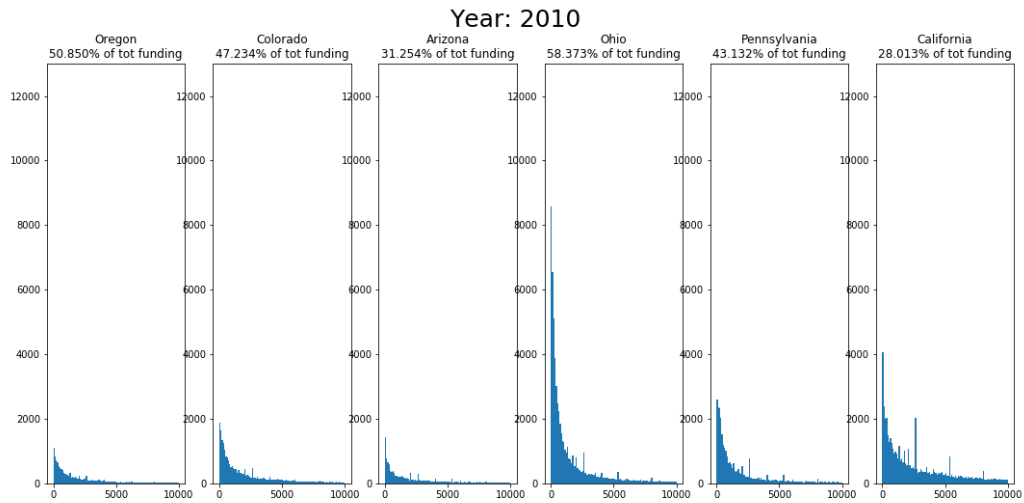
Num of unique awards given every FY in different states



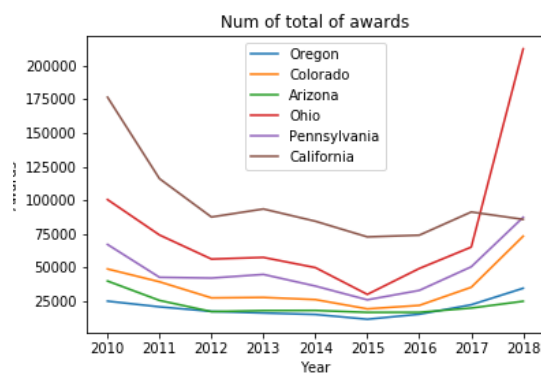
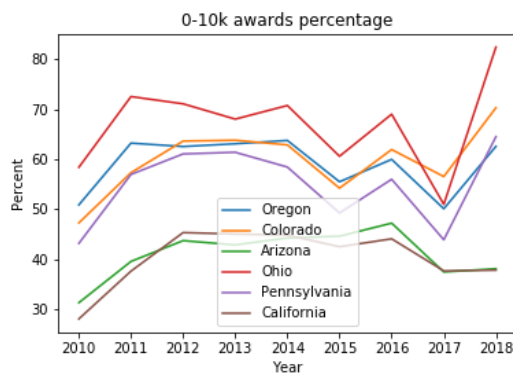
Num of total awards given every FY in different states

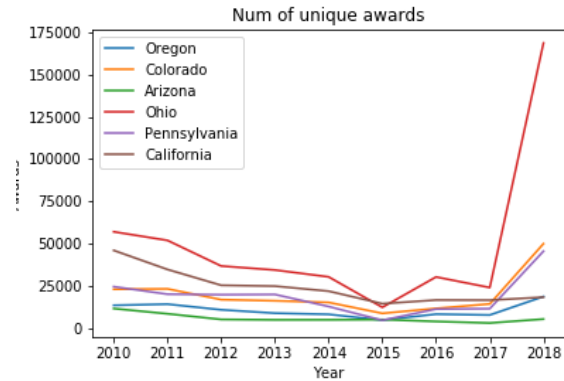
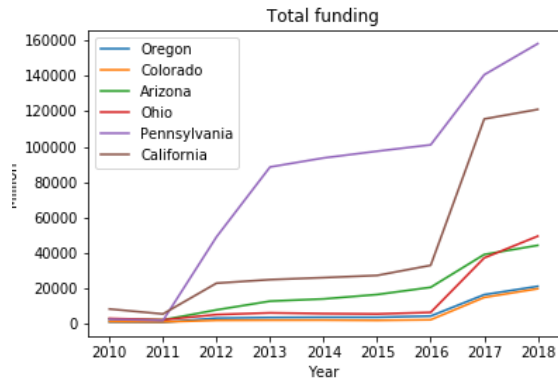


Finally, I thought that I should take a deeper look at each of these trends, so, I created histograms for the awards that have amounts between 0 and 10k, then calculated the percentage of those awards to the total number of awards.



Unfortunately, I was not able to see any correlation between the histograms and the geographic traits. However, by looking at all those trends, I saw similar patterns, which the plots below show.





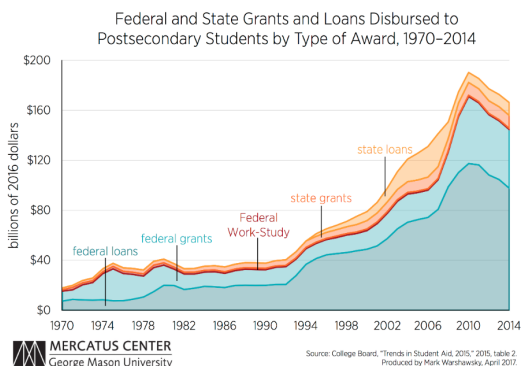
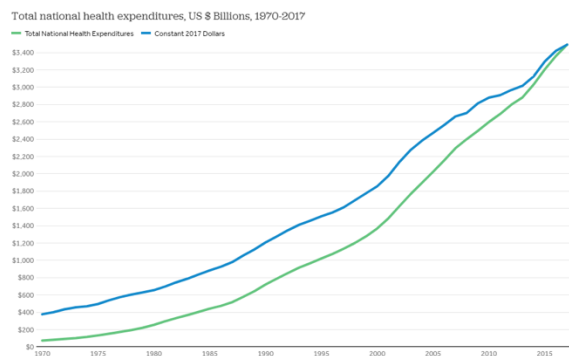
Conclusions

Political - Conclusion

As a result, I concluded that our assumption is invalid, as shown by the trends we extracted from our dataset. Both states follow very similar behaviors despite their overall stance in public health and educational funding. There are various reasons as to why this is the case, which can include:

- A senate's opinion and policy making efforts affects the nation as a whole, rather than a state only.
- There are other factors that we needed to consider, such as the administration's priorities for a term and the programs and policies it passes.
- A state's political figures policy making efforts does not take effect immediately but can be seen to have a slower and longer impact rather than an immediate one.

To verify that our data is accurate and reflects (closely) the national spending in both areas, I included the two graphs from Health System Checker Organization and George Mason University to verify education spending. ^{[4][5]}



Both graphs reflect accurately the trend we found in our dataset. As of 2010, the DOE's funding has dropped significantly year after year, whereas DOH's funding kept on increasing steadily yearly.

Geographical - Conclusion

We can see that there is a slight correlation between population and some of the trends, but that does not mean anything. Even though the conclusion that we were able to draw is that there is no correlation between the geographical traits and the trends we explored, there are similar patterns between the states.

Note: to see many more graphs and results, all bar graphs, plots and map reduce outputs for the political and geographical traits are listed Jupyter Notebooks in our project GitHub repository.

Review

Ahmed's review on Mohamed's work and performance

I honestly found Mohamed to be a pleasant and committed teammate. We both put the same effort and time, and we both shared the same dedication. In fact, to keep the work equal, we both decided to separate our work and for each to be assigned with a different problem. He would always contact me about the ideas he has, any questions, consistent updates about his work and approaches. He never missed a meeting, nor did he ever fall behind a deadline. He was always up to date with the project.

Mohamed's review on Ahmed's work and performance

I consider Ahmed the best teammate I have ever worked with. He is hardworking, committed and had a great vision of our project and how we should divide the work between us. He completed his part way before the deadline and was very helpful in answering my questions and completing the reports.

Reference

- 1) <https://www.usaspending.gov/#/>
- 2) https://github.com/unitedstates/federal_spending
- 3) <http://ontheissues.org/default.htm>
- 4) https://www.healthsystemtracker.org/chart-collection/u-s-spending-healthcare-changed-time/#item-total-health-expenditures-have-increased-substantially-over-the-past-several-decades_2017
- 5) <https://www.mercatus.org/publications/increased-federal-funding-higher-education-produces-adverse-effects>