Mini Project #2

Timothy Elbert
telbert6@gatech.edu

1 BACKGROUND AND SIGNIFICANCE

The US healthcare system is the most expensive in the world by pretty much every metric. It's not that Americans have a tendency to go to the doctor or be hospitalized more than the rest of the world. It's really just that the prices are so much higher. Another known phenomenon of the US healthcare system is how much prices vary, sometimes by orders of magnitude between two hospitals within the same city (Hankin, 2020). The average US citizen is not better off financially than those living in other Western Democracies and definitely has fewer safety nets. They aren't healthier either and require a similar level of medical care. And yet, they are responsible for paying extremely exorbitant healthcare costs. 41% of working-aged people (72 million) have problems paying their medical bills, and the trend is increasing (Commonwealth Fund, 2020). Americans often cannot afford their own healthcare system without government assistance, such as it is. Clearly, based on the 41% statistic, this support is often inadequate. Meanwhile, Medicare alone is costing the government 14% of its budget ($644 billion) every year, which is also trending up (Peterson, 2020).

Chronic conditions are a big portion of US spending, as discussed in the course. Infact, the industry of treating chronic conditions costs $1.1 trillion a year, which is 6% of the gdp (Waters & Graf, 2020). Chronic conditions by necessity result in repeat encounters with medical professionals directly related to treatment. Chronic conditions often also cause acute conditions either directly or indirectly that result in further encounters, like broken bones as a result of osteoporosis.

Another issue is bias in how people are treated in the healthcare system, be it based on sex (Schopen, 2017) or race. For example, black people consistently are rated less on the pain scale by physicians for similar conditions to white people (Hall et al., 2015). In the case of the medical field, these slight biases can and do cause serious damage and death.

2 PROBLEM

Chronic conditions are not always random or evenly distributed. There are environmental factors like pollution and hours of daylight, as well as social norm differences based on geography and history of a particular area. Certain areas are more prone to certain conditions, but it's not always obvious to the people in the area that what they are seeing in terms of caseload is out of the ordinary. They might not even think to look for patterns or trends, especially if the numbers are low, as they are in a certain area. However, early detection of an anomalous number of people in the area with a chronic condition might allow for the identification and resolution of an environmental factor new or unique to that area.

In a similar vein, it is often hard to pinpoint inequality as it happens, especially considering doctor patient confidentiality, how hard it is to identify inequality in a highly specialized field like medicine, and the fact that most people don't even know themselves that they are biased. There is nothing inherently suspicious about a doctor evaluating a black woman's pain at a certain number. It's a judgement call. Only with bigger numbers can a trend be detected and mitigated.

3 PROPOSED SOLUTION OR IDEA

This dashboard is a proposed tool to help mitigate these problems. It is based on the Synpuf synthetic data, specifically the summary of the beneficiaries, which lists demographic information, chronic condition status, and also how much money the patient and their provider had paid over the course of 2008 and 2009.

The first figure is a county level heatmap of the percent of patients that have at least one chronic health condition. This allows for the immediate detection of hotspots that merit investigation.

The second component is the two rows of pie charts, each row representing a different demographic domain. This is here to make it easy to detect inequality. The first two charts in a row might point to environmental factors or bias. The first shows the demographic breakdown by population and the second shows the demographic breakdown by who suffers or at least has been diagnosed as suffering from a chronic condition. The third pie chart is a more direct indicator of a systemic problem as it is a calculation of how much money the average person suffering from this condition has had to pay over the timespan of the visualization.

This is a tool that allows administrators to ask questions. It facilitates their ability to manipulate the data quickly, without technical skill. It doesn't serve up what is worth investigating up front. Hopefully, this means that it is more scalable, responsive, and multiuse.

All the code, including the app itself and the data cleaning notebook can be found at this git-repo. Only after completing the project did I realize that I had been pushing to it as my personal git account Badunoff, not the Georgia Tech git account telbert6. Please let me know if this is an issue or if I can answer any questions.

4 COMPLEXITY OR EFFORT

This effort was not terribly complex. The most complicated portion was getting county-level information to display on the map, which only knows countries and states. In order to do that, I had to import a json from plotly that contained the shape information for the counties, as well as a lookup table pulled from nber.org to convert from the SSA code provided by the Synpuf data to the FIPS code of the shape file. The unfortunate part is that I was only able to get the lookup table for 2011, which is slightly beyond the scope of my data. This is my primary suspect as to why Alaska does not appear in the visualization.

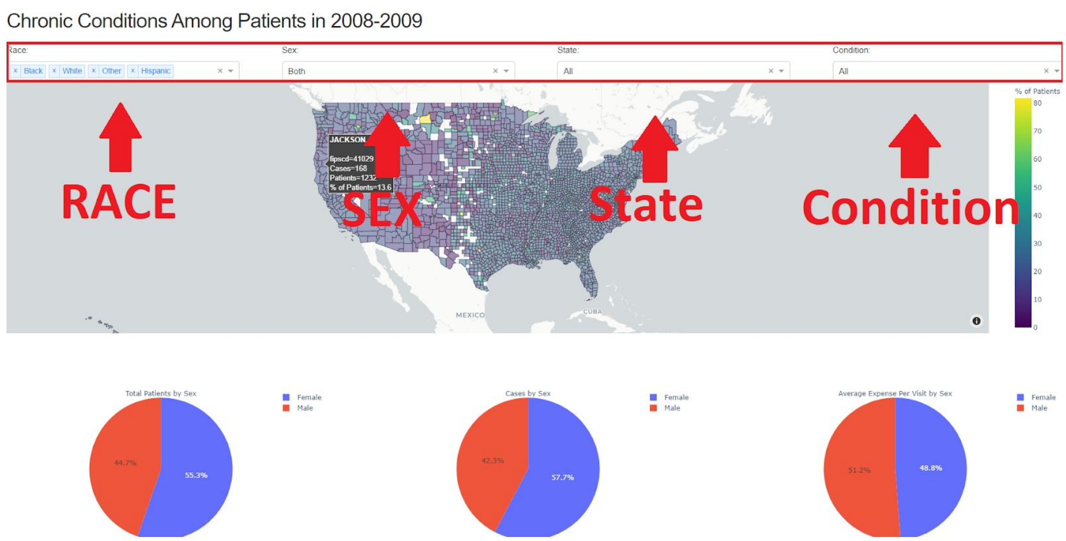
Learning plotly was frustrating, at times, but provided a far more controllable environment than in Tableau, where it seemed like if I wanted one kind of visualization, I could not have another. The complexity of managing multiple data sources and having filters work across them made what should have been a simple processing incredibly complicated.

One design decision I made was to filter out counties that had less than 5 patients in order to at least somewhat mitigate the fact that small data makes large rates of chronic disease meaningless.

Future iterations of this, if I were to build upon it, would include getting demographic information for individual counties, not just putting them in context of the rest of the state or country by chronic case rates. It is possible to get this information by playing with the filters, but it requires memory and several updates. I would also like to somehow add more context based on individual claims. The current visualization is based on the summary file aggregated over 2 years.

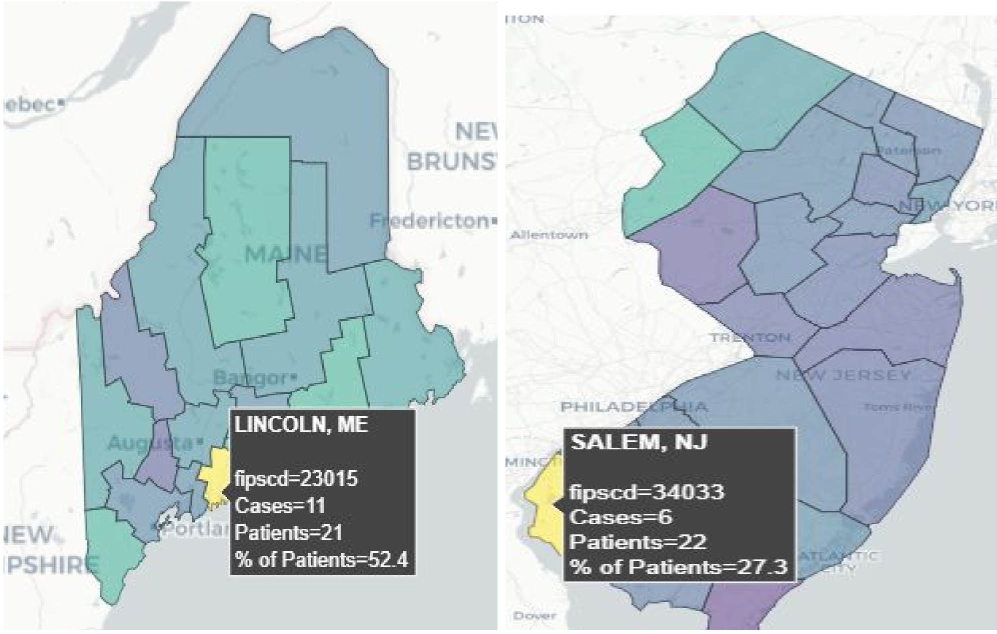
It would also be nice to do this on real data and not synthetic, because as smart as the model generating the synthetic data might be, I feel like certain real world information might not be accounted for.

5 TUTORIAL AND DEMONSTRATION



The filters on top allow to further explore anomalies by drilling further down by condition, sex, and race. These are all common dropdowns. Race is multi select, in case there are specific questions about minorities in any combination that the user might ask.

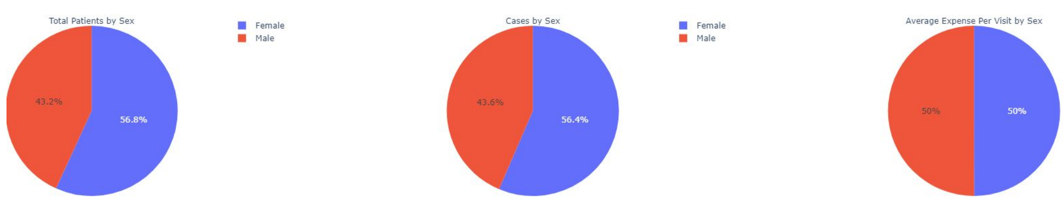
This map clearly shows when a county over-indexes in chronic conditions in general or even specific chronic conditions relative to both the national and state averages (if filtering by state).



LEFT: Maine depression rates. RIGHT: New Jersey cancer rates

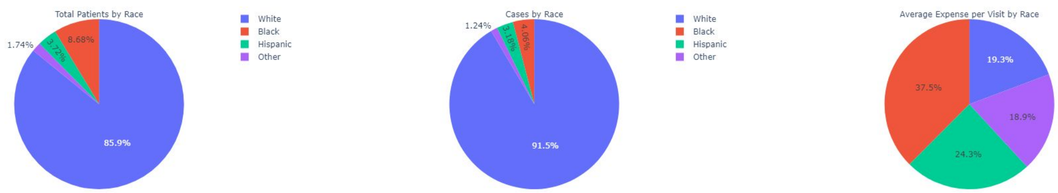
For example, if you select just depression and filter down to just Maine, you see that an alarming 52% of patients in that Lincoln county are known to have chronic depression, when the surrounding counties have no more that 30%. Salem, NJ has a 27% cancer rate across beneficiaries compared with single digits nearly everywhere else in the state. Because it's a percentage of the total patient population, it controls for more and less populated areas.

The map helps identify problematic trends in the chronic conditions themselves, across various combinations of demographics. However, that does not necessarily help detect if there is a bias issue for that condition in that area. That is what the 6 pie charts in 2 rows. The first row is Sex and the second row is based on Race.



Califnornia rates of chronic kindney disease by sex

The first chart shows the demographic breakdown of patients in general. This is intended to be compared with the second chart which shows the breakdown of that demographic based on patients who have the chronic condition. If there is a significant difference between the ratios between the two charts, then this is an issue that should be further investigated. It’s an indication of inequality and that work needs to be done to figure out the reason. The example above is a perfect demonstration of equality based on sex. The rate of kidney disease detection in the second pie chart matches almost perfectly with the patient demographic breakdown by sex. Kidney disease is detected at the same rate for men and women. Furthermore, the third chart shows that the treatment is extremely equitable, because both men and women with kidney disease pay the same average amount in medical bills, within literally $11, with over a thousand patients. This is incredibly consistent and the best example that I’ve found.



Florida rates of chronic cancer by race

The second row, though, shows some inequity based on race. Cancer is detected in white patients (blue) 6% more and in black patients (red) about 4% less relative to demographic differences based on the first two pie charts. This might be because of a cultural phenomenon of the fact that the white population is more elederly due to an old fad of retiring in Florida. As a result, rates are higher. This may or may not be the reason, but the visualization definitely shows that something is going on. The inequality is perhaps more visible in the third chart that points to the medical costs per person suffering from cancer. For whatever reason, even though Black cancer patients (Red) end up over-indexing on what their medical treatment costs by a full 12%. Hispanic cancer patients (Green) pay what seems to be an even share, but it’s still significantly more than what White cancer patients (Blue) and Other races (Purple) pay. This would seem to actually cast doubt on the guess that the reason cancer is detected in more White patients is because of an elderly white population. Elderly people tend to have more medical problems and thus, I would expect, would cost more to treat over a given period of time. It would be worth digging deeper into individual claims of these patients to see where the difference lies. This is beyond the scope of this visualization, though. It did its job and showed that there is something that needs to be examined.

It takes several seconds for the visualization to update (the more filters, the faster). You know when the process is completed when you see that the text in the browser tab switches back to “Dash” from “Updating”

7 REFERENCES

1. Commonwealth Fund. (2020). Survey: 79 Million Americans Have Problems with Medical Bills or Debt. Retrieved September 20, 2020, from https://www.commonwealthfund.org/publications/newsletter-article/survey-79-million-americans-have-problems-medical-bills-or-debt
2. Hall, W. J., Chapman, M. V., Lee, K. M., Merino, Y. M., Thomas, T. W., Payne, B. K., Eng, E., Day, S. H., & Coyne-Beasley, T. (2015). Implicit Racial/Ethnic Bias Among Health Care Professionals and Its Influence on Health Care Outcomes: A Systematic Review. American journal of public health, 105(12), e60–e76. https://doi.org/10.2105/AJPH.2015.302903
3. Hankin, A. (2020, August 29). How U.S. Healthcare Costs Compare to Other Countries. Retrieved September 20, 2020, from https://www.investopedia.com/articles/personal-finance/072116/us-healthcare-costs-compared-other-countries.asp
4. Schopen, F. (2017, November 20). The healthcare gender bias: Do men get better medical treatment? Retrieved September 20, 2020, from https://www.theguardian.com/lifeandstyle/2017/nov/20/healthcare-gender-bias-women-pain
5. Waters, H., & Graf, M. (2020, September 2). The Costs of Chronic Disease in the U.S. Retrieved September 20, 2020, from https://milkeninstitute.org/reports/costs-chronic-disease-us