The Teeth, the Brain, and the Socioeconomic Inequalities?: DATA STORY

Natalie Smirnoff

Let’s explore the data through my favorite medium: pictures. Follow along (if the code allows) with this [link](https://jupyter.org/try-jupyter/lab/index.html?path=notebooks%2FDATA205FINALPROJECT+%28copy%29.ipynb) or with the repo documents.

First and foremost, let’s recap: we’re going to be exploring a few elements relating to relationship between mental health (MH) and dental visits (DV). In my initial data exploration I found a strong relationship between the two and wanted to explore it further.

QUICK NOTES BEFORE WE GET INTO IT: data for mental health, general health, and health insurance is reported by poor performance.

MENTAL HEALTH IS: Mental health not good for >=14 days among adults aged >=18 years

GENERAL HEALTH IS: Fair or poor self-rated health status among adults aged >=18 years

HEALTH INSURANCE IS: Current lack of health insurance among adults aged 18-64 years

So the higher the number for these fields, the higher the number of people struggling with mental health, general health, or lack of insurance.

See [here](https://www.cdc.gov/places/measure-definitions/index.html) for additional descriptions of each value in the PLACES dataset.

1. DATA STRUCTURE

So the first thing we’re going to do is a quick sanity check: does the data look correct to what we know to be true of the world.

A graph of population across area types

Description automatically generated

*Fig 1. How much of the population falls into each population type. 1 is most urban all the way to 6 being most rural.*

A graph of a number of area types

Description automatically generated

*Fig 2. How many of the FIPS codes falls into each population type. 1 is most urban all the way to 6 being most rural.*

More or less, yes. The data does look kind of like what we would expect. Fig. 1 shows that population is heavily skewed to urban areas, much more than it really is, and that goes with the typical problem of studies underrepresenting rural areas. Fig. 2 shows the distribution of FIPS codes (not land in terms of Acres) and fits the reality that I know to be true of most counties falling into that more rural category.

A graph of different types of health

Description automatically generated with medium confidence

*Fig 3. Data distribution graphs of the selected variables to explore.*

Fig. 3 exists to check data distribution. It’s not pretty but it gets the job done for a quick look. You can see in the first to second graph how standardizing dentists to “per 1000” helps spread the distribution a little a give more context to the story of each number. The other distributions have their heads and tails but are for the most part pretty bell shaped.

1. CORRELATIONS

Next lets explore how all if our data looks when it plays with the other data. Looking at Fig. 4, for the most part, everything more or less falls into what you would expect it to, expect for our area of interest. I made a lot of different correlation charts because I had a hard time wrapping my head around what I was actually looking at and each one helped me understand a little better.

A graph of a diagram

Description automatically generated with medium confidence

*Fig 4. Faceted graph of correlations of interest.*

A screenshot of a table

Description automatically generated

*Fig 5. Correlations in numbers.*

A collage of different health insurance graphs

Description automatically generated with medium confidence

*Fig. 6 Close up of faceted graph of correlations of interest, focusing on some areas of interest.*

Fig. 6 Shows a more comprehensive view of some of the elements we’re looking at. Here you can clearly see how as MH goes down, DV also goes down. You would expect to see the same with MH and annual checkups (AC), but it’s Actually the opposite, as MH goes down, AC goes up. I also took a look at MH and DV and how it related to health insurance (HI) since I know health insurance plays a big part in people’s health choices.

A graph of different colored squares

Description automatically generated

*Fig. 7 Correlation heat map. Everything below the absolute value of 0.2 is greyed out since those relationships aren’t significant and to remove distractions.*

Fig. 7 is the best visual for what we have going on. In addition to everything we have already discussed, we can see that there is also a significantly stronger correlation between HI and DV than between HI and AC. Another element of the DV vs. AC WWE smackdown that is very different.

1. EXPLORATIONS

Here we’re going to explore these relationships in greater detail through graphs and bring in another dimension: Population type. The NCHS breaks population type into six categories: 1 being most urban and 6 being most rural. Feel free to consult documentation [here](https://www.cdc.gov/nchs/data/series/sr_02/sr02_166.pdf) to learn more about their classification.

I wanted to see if different population types had different attitudes about AC and DV by seeing if the attendance levels were different.

For some of these graphs I’ve condensed classification 1 and 2 into Urban, 3 and 4 into Suburban, and 5 and 6 into Rural.

A graph of a number of people

Description automatically generated with medium confidence

*Fig. 8 Graph of number of people per 1000 going to Annual Checkups and Dental Visits across population types.*

I made Fig. 8 to see if there is a significant different between number of people going to the dentist vs doctor, and how it looks across population types. I expected DV to be much lower than what they actually are and I expected the more rural areas to see lower dentist and doctor visits. However, in reality DV wasn’t too much lower than AC and it was pretty even across the board.

A green and red line graph

Description automatically generated

*Fig. 9 Mental health and Dental Visits correlations across population types.*

I remade my cornerstone correlation chart and broke it down into population types to see if there was a significant difference. There wasn’t. And then I made 6 more because I really liked making these for some reason.

A graph with green dots and red line

Description automatically generated

*Fig. 10 Mental health and number of dentists per 1000 correlations across population types.*

Fig. 10 has a slightly stronger correlation in the urban sector so I wanted to see if there was anything else I could see by breaking it back into the original six classification types. See Fig. 11 below.

A green and red dots

Description automatically generated

*Fig. 11 Mental health and number of dentists per 1000 correlations across the 6 population types.*

Population code 2 is the one skewing the data. I’m not sure why but I have a mini hypothesis: If this code represents an area like Bethesda where people are discouraged from having a car (either by poor parking or some other element) but the area doesn’t quite provide the walkability needed to really cover all needs, then maybe the lack of dental professionals in the area puts people into a position that they can’t make it to the dentist as often as they like? Who knows?

A graph with a red line and a green line

Description automatically generated

*Fig. 12 Mental health and Health insurance correlations across population types.*

A green and red line graph

Description automatically generated

*Fig. 13 Dental visits and Health insurance correlations across population types.*

Fig. 13 is interesting because the overall correlation chart he this weird little triangle shape and you can see here that most of this shape comes from the rural sector (obviously because the rural sector has more counties and more of the data comes from there). But it’s interesting to see how for the beginning of the graph there does seem to be a very strong correlation between having more health insurance and people going to the dentist.

A green and red line graph

Description automatically generated

*Fig. 14 Dental visits and General health across population types.*

A graph with green and red lines

Description automatically generated

*Fig. 15 Dental visits and number of dentists per 1000 correlations across population types.*

I think I made so many of these because I had the bias where I really expected there to be more of a difference across the population types. But if anything it showed me that people care about equally about making their way to the doctor regardless of the cultural differences across the population types.

A screenshot of a computer

Description automatically generated

*Fig. 16 Screenshot of the two t-tests.*

I ran two T-Tests to check significance. These numbers felt off to me so I tried a number of fixes, group-by’s and pivots in the data so try to “fix” them but everything returned something in this realm. I spend too much time on it so I said I’d revisit the issue later but by the time I remembered them it was a little late to explore the issue further.







*Fig. 17, 18, 19, 20 Screenshots of the chi-squared tests. Grouped together. The variables compared are listed above the results in the #description.*

I also transformed some of the quantitative variables into categorical variables by placing them into “grades” to see if there was a grater significant relationship once you grouped them. The grade beginning and endpoints are more arbitrary than I would like but I did consult on them so I do have some confidence in them.

I consulted with a friend who works on data for the government, and we decided that the best way to divide everything up is this: (depending on data structure of each variable) mean +/- 1 standard deviation. Everything above/ below that number is either and A or an F (remember that in MH, DV, and HI, the higher the number the higher the “failure”) and everything else is divided equally into the remaining grades. I did minor rounding For the most part I feel like this captured what I was going for pretty well.

While the above figures didn’t show much of a relationship between MH/DV and population type, Fig. 17 and 18 state that there is a significant relationship and they are dependent. Fig. 19 confirms that there is a relationship between dental visits and number of dentists per 1000 people. And Fig. 20 states that there is a relationship between MH and number of dentists per 1000 people.

Since I don’t think dentist availability would be too impacted by the overall mental health of the county (though maybe there’s something to be explored with the dental profession having some of the highest suicide rates amongst all professions), I think the data is telling me that just maybe, counties that have a smaller dentist to client ratio have more mental health issues. So just maybe, regularly going to the dentist has positive impacts on your mental health?

Lets not forget that correlation is not causation and move on.

I made some maps to visualize the data further. They mostly speak for themselves and only exist because I’m a visual person, but I do have some interesting thoughts/observations about them.

A map of the united states

Description automatically generated

*Fig. 21 Map showing mental health grade by county.*

Fig. 21 for the most part looks like what I would expect (don’t look for Florida, there’s no data on it) except for N Dakota, S Dakota, and Nebraska. I originally thought that since the data reports on negative MH, these areas just had 0 values for MH and went into the data. But no, these states have values and they are just legitimately low. So good for them!

A map of the united states

Description automatically generated

*Fig. 22 Map showing dental visit grade by county.*

Fig. 22 is about as expected. And visually matches our exploration or MH relating to DV when compared with above map.

A map of the united states

Description automatically generated

*Fig. 23 Map showing dentist per 1000 grade by county.*

Fig. 23 is rather messy looking but if you squint and move back a little bit you can still see the pattern around the South to South-Eastern areas seeming to have more areas of failure than some of the other areas. Quick note that all of this data is excluding Chattahoochee County, GA because it had more than double the next highest number of dentist ratio.

A map of the united states

Description automatically generated

*Fig. 24 Map showing mental health grade by county in Maryland.*

A map of the united states

Description automatically generated

*Fig. 25 Map showing dental visit grade by county in Maryland.*

A map of the united states

Description automatically generated

*Fig. 26 Map showing dentist per 1000 grade by county in Maryland.*

For those of you curious about how Maryland looks, see Fig. 24, 25, and 26.

1. CONCLUSIONS

Let’s do a quick recap of what we’ve discovered with all the above work.

* Mental health and Dental visits: RELATED. The poorer the mental health, lower the dental visits.
* Mental health and dentist availability: NOT RELATED. UNLESS! You group them into categorical “grades”. The said grades are subjective though so be mindful of bias.
* Mental health, Dental visits, and population type: NOT RELATED. UNLESS! You group them in “grades”. Again, subjective. So some kind of cultural differences between population types maybe does something?
* Dental visits and dentist availability: SLIGHTLY RELATED. So maybe lower dentist availability impacts DV. Or maybe dentists won’t bother moving to an area where people have no interest in going to the dentist. Which would create a vicious cycle.
* Health insurance and everything: RELATED
* Dental Visits and general health: STRONGLY RELATED. AND LET’S DISCUSS THIS ONE FURTHER!!

I’m not comfortable stating that there are clear causations here. But I am comfortable sharing my thoughts and ideas. I stated earlier “So just maybe, regularly going to the dentist has positive impacts on your mental health?” but I stated it as a question for a reason. I don’t know if I believe that. Or at least, I don’t know if I believe that alone. I think there are greater things to discuss before we can get into that.

* Cultural Priorities: As a culture, we just don’t put as much emphasis on dental care as we do on annual checkups. For whatever reason we’ve decided teeth are an aesthetic choice and it reflects in everything from many health insurance plans not including dental to me rescheduling my dentist appointments 4 out of 5 times because I have “more important things to do”. I had 11 cavities filled in one go when I was 17.
* Physical/Mental battery: When you’re not feeling well physically or mentally, you go into triage mode. And if culturally we put more emphasis on annual checkups than dental visits, then you know what to cut. You’ll still go to the doctor, and plan for the dentist when you’re “feeling better” (and if you have a chronic illness and are feeling better, then you’re probably going to spend that time doing something a little more fun than *the dentist*).
* Professional Availability: (This one does tie into what we’ve looked at a bit). If you do decide to be responsible when feeling better, and there is no professional within g reasonable distance from you (what is reasonable will vary person by person), then you require more of said batter to go to that professional and suddenly the task is a lot more daunting.
* Health Insurance: Medical procedures and now even common checkups are expensive. Without insurance, people simply may not be able to afford them. And going back to priorities, if you have to allocate your funds, it will be into what your culture (and by extension you) deem the priority.

All of this to say, maybe going to the dentist has positive impacts on your mental health, and that better teeth give you more confidence to smile (or whatever all those ads say). But I think it’s also likely that you’re just not going to prioritize the “less important” health checkup when you’re already not feeling your best.

1. DATA PRODUCT

So finally, let’s look at my data product. Again, as long as everything works, you should be able to access it [here](https://jupyter.org/try-jupyter/lab/index.html?path=notebooks%2FDATA205FINALPROJECT+%28copy%29.ipynb) or through the code in the repository. If something doesn’t work, then I will describe it. It’s a scatter map of the US counties where the size of the bubble represents the number of uninsured individuals (so the larger the bubble, the more uninsured) and the color represents the number of dentists per 1000 people per county. There is also a hover function where hovering over each bubble states (in order): the county name, the number of people per 1000 without health insurance, latitude, longitude, the number of people per 1000 with poor mental health, the number of people per 1000 who have been to the dentist in the last year, mental health grade, dental visit grade, and the number of dentists per 1000 people.

A map of the united states

Description automatically generated

*Fig. 27 “Interactive” map showing dentist per 1000 (bubble color) and health insurance (bubble size).*

I selected to group by health insurance and dentists per 1000, instead of mental health and dental visits which is what the majority of my project looked at. And the reason for that is: you can’t make people go to the dentist, and you can’t make people feel better, but you can improve/increase their resources so they can make easier choices for themselves. So if you’re a member of a group trying to lobby for increased health insurance coverage, you can pinpoint which areas are struggling the most (see that outer ridge of Texas?) and concentrate your efforts there first. If you’re trying to start a roving dentist group, you can find counties in the red (most of them) and direct people there first. Or you can set up programs to incentivize newly licensed dental to help areas of grater need.

There are lots of applications I haven’t even thought of but the point is, the relationship between dental visits and mental health is fascinating, but I think at the end of the day the most important thing it does is highlight failures in our social and cultural safety nets. So my final product exists to shed more light on those failures and maybe help people address them. As long as plotly decides to work.