Comparing Rising Healthcare Costs in New England and the American Midwest

Data Watch Paper

MILI 6963 - Section 70

Danny Moncada

CARLSON SCHOOL
OF MANAGEMENT
UNIVERSITY OF MINNESOTA

Contents

1. Introduction & Background

Healthcare costs for both providers and patients are rising across the continental United States; here in Minnesota, healthcare spending is projected to hit **$94.2 billion** [3, pg. 28] by 2026, which represents an increase of **78%** compared to spending in 2018 [3, pg. 28]. As an analyst working for Cigna Healthcare in the Connecticut/New York metro area for over six years, I witnessed firsthand not only the escalating costs associated with maintaining one's health, but also the difficulty for many American citizens to understand and comprehend their claims information. With benefits summaries reaching a level of complexity only seen on the balance sheets of Fortune 500 companies, it is no wonder so many people "don't have an idea of what health care costs." [1, pg. 1] As a graduate student (*Carlson School of Management, M.S. in Business Analytics*) and employee of the University of Minnesota now living in Minneapolis, MN, I wanted to investigate whether there were any parallels between my old life in Connecticut and my new life in the Midwest. Thanks to the skills developed in my coursework as part of the healthcare analytics course, I am better prepared to answer this question.

While there is much literature on the subject, including daily news articles exploring the explosion in medical expenditure in the United States, it was not until reviewing a report developed by the *Minnesota Department of Health* when I realized the desperate situation facing my new home; "health care spending growth is expected to outpace economic progress, resulting in a greater share of the Minnesota economy devoted to health care (18.6 percent by 2026, up from 13.9 percent in 2016)." [3, pg. 28] Per the article, *demographic* shifts will contribute to increased Medicare enrollment and greater spending to treat multiple chronic diseases, and the number of Minnesotans eligible for Medicare is projected to increase 3.3 percent per year, on average, between 2017 and 2026 [3, pg. 29]. This confirmed a few suspicions that I had before diving into the analysis of claims data available to me: the

population in Minnesota is getting older at a much faster rate than in previous generations, and this shift **will** (whether it is directly or indirectly can be debated) contribute to additional costs to keep our population healthy.

2. Data Engineering & Analysis

Here is a brief overview of the analysis that was done in prep for this paper. The source of the data were two SAS programs that were developed as apart of the MILI 6963 Healthcare Analytics course; the SAS programs mined through claims data made available in the SynUSA Synthetic Health Insurance Analytic File System (SHIAF) [4, pg 1]. The training database used during the course allowed analysts in our class to generate claims-based analytics. The data available in this system includes two years worth of insurance claims data, with the ability to segment claims and patient information into different age categories and these five major insurance types:

- **Employer Sponsored Insurance (ESI)**
- **Medicaid (under 65)**
- **Medicare Fee For Service**
- **Medicare Advantage**
- **Non-Group (both Affordable Care Act exchange and non-exchange)**

I generated two sets of files needed for the analysis; the first set of Excel files were simply demographic related information for the two areas I was investigating, Connecticut/New York (combined) and Minnesota/Wisconsin (combined). The other set of files generated were related to patient and provider out of pocket expenses. This file was what would be used to determine what, if any, differences there were in the patient out-of-pocket expenses between the two regions.

The first proto SAS program subsetted all of the available claim information to only include patients in two specific beneficiary states; for brevity, I won't constantly call out these out; I will use **Group A** for Connecticut & New York, and **Group B** for Minnesota &

Wisconsin. Subsequently, the data was mined even further so that each patient was attributed to a "Age" segment; here is a breakdown of the different segments:

- **0 to 18 years old**
- **19 to 34 years old**
- **35 to 44 years old**
- **45 to 64 years old**
- **65 and older**

With this we could determine how many patients of each age belonged to each of the major insurance types we were investigating throughout the course. Finally, to build the visualizations, we removed all of the unnecessary rows from the Excel output and simply plotted the distributions of patients for each insurance type, grouped by age.

The second proto SAS program performed a similar subsetting of available claims information for Group A and Group B. However, instead of mining for patient demographic information, this program instead analyzed claim information to look at patient and provider "burden"; patient burden being the sum of the patient's coinsurance plus copayments plus deductibles. The purpose of this program was to measure how much patients were paying out of pocket, and similarly patients were grouped into different segments:

- **Under $1000**
- **$1,000 to $4,999**
- **$5,000 to $9,999**
- **$10,000 to $29,999**
- **More than $30,000**
- **Unknown patient cost****

The last segment was not part of the original SAS program; I took the extra step to include patient claims that did not fit in any of the first five segments by subtracting all patients belong to these segments from the overall population; the remainder were bucketed into this "Unknown patient cost" segment. To generate the visualization, I then grouped all of the major insurance types together in order to stack them into the spending buckets.

To generate the plot with projected healthcare spending in Minnesota (from 2012 - 2026), I used the information provided in the *Minnesota Department of Health*'s report that was referenced in the introduction. I took healthcare spending that has been observed for the last few years in Minnesota, and utilized a very simple function in Python to generate a linear model that filled in values from 2016 through 2026. I created another column to show the projected growth if spending followed this linear model to demonstrate how much this changes year over year as we reach that date.

3. Results, Insights & Outstanding Questions

Results

The first observation made was that projected growth for spending in healthcare in Minnesota (based on the linear model) would never drop below **5.26%**. This right away raises alarm bells, for the simple reason being: *who has to foot that cost*? In my experience working at Cigna, providers are reluctant to foot additional cost or overhead, and these costs will roll onto patients. That growth, combined with "...the historical pattern of moderate growth in direct spending by individuals" [3, pg. 19] mean that many Minnesotans "...painfully feel the rising burden of increasing out-of-pocket costs." [3, pg. 19] This is not a sustainable level of growth; even the S&P 500 won't ever reach this level of growth!

This is not just a problem that is affecting only Minnesotans however; similar struggles are being faced by patients in Connecticut. According to Vicki Veltri, executive director of the Office of Health Strategy, "We know from our insurance bill how much we're paying in premiums, how much our copay is, but we don't know … what the carriers are actually paying." [1, pg. 1]. Due to this uncertainty, the Office of Health Strategy developed a new cost estimator tool to reveal how much medical treatment costs vary across Connecticut's

health care facilities; procedures like knee replacements can cost nearly **three** times as much at certain hospitals than they do at others [1, pg. 1].

With this context, let's take a look at what the analyses generated for this paper tell us about the populations in Group A and Group B. When we look at the distribution of the five major insurance groups by age, we see that the age splits are very similar; the largest groups belong to the **45 to 64** age range. When combining the *Employer ESI*, *Medicaid*, and *Under 65 Non-Group* insurance groups, there's a total of ~6.57 million patients in Group A, and ~3.46 million patients in Group B. What is really telling is that the *Medicare FFS* and *Medicaid* groups making up the **65 and over** age come nowhere near that range; there are a total of ~3.12 million patients in Group A **and** Group B, respectively. We don't know the exact age breakdown of these groups, but we can presume that the majority of the patients in the **45 to 64** group will eventually age and join the **65 and over** age population. *Why is an aging medical population important?* Let's review the outcome from the other analysis to see the impact firsthand.

I looked at the patient burden (which was calculated using coinsurance combined with copayments combined with deductibles). What was incredibly striking was that in Group A, almost 80% of the claims where the patient had to pay an out-of-pocket cost **greater than $30,000** came from the *Medicare FFS* group (the remaining 20% was attributed to the *Employer ESI* group). In Group B, this was even **more** stark, 100% of claims where the cost to the patient was greater than $30,000 came from the *Medicare FFS* group. This was extremely jarring to see, because this means that a population made up of almost entirely of people expected to be in their retirement or twilight years were suddenly being tasked with out-of-pocket expenses totaling more than $30,000. In addition to that, 90% of the claims where the out-of-pocket cost to the patient was between **$10,000 and $29,999**. If trends stay similar to what was observed in 2014 and 2015, that means that many of the population entering that

retirement age will more than likely face a costly medical procedure (and maybe even multiple), during a time when they will not have an employer provided insurance plan, and may not have access to as much money as in their earlier years.

What I would like to do as a result of this analysis and project is to revisit this subject when we have more claims data to review; I'd like to see if spending stayed the same for the various groups we analyzed. In addition, I would like to review where healthcare spending for the state of Minnesota is in five years, to see how close to the projected spend of $94.2 billion we actually reach.

Insights

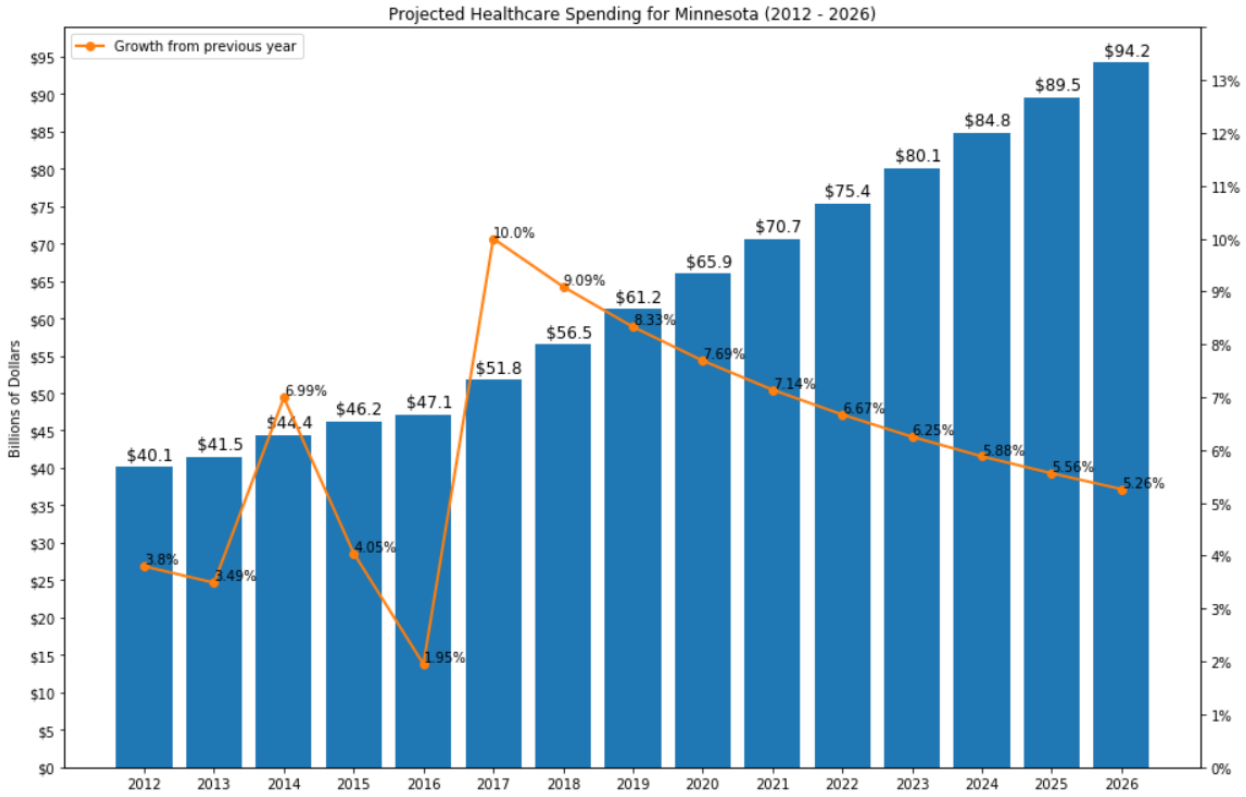
- Based on data available from an existing report from the Minnesota Department of Health, projected healthcare spending is expected to reach **$94.2** billion by 2026.
- At this rate, that means the growth from the previous year (how much spending grows year over year) will never drop below **5.26%**, and the next three years will see growth of **7.69%**, **7.14%**, and **6.67%** respectively.
- There are ~3.46 million patients that belong in the **45 to 64** age range in Minnesota & Wisconsin; out of the entire medical population, that accounts for **~26%** of the entire medical population that was queried (total of ~13.4 million patients).
- The second largest group belongs to the **65 and older** age range; there are ~3.11 million patients, accounting for 23% of the entire medical population queried.
- 100% of claims where the out-of-pocket cost to the patient was **greater than $30,000** came from the *Medicare FFS* insurance group.

Outstanding Questions

- *With more of the medical population moving into retirement (as seen in the later analyses) and off of employer sponsored plans, will patients start to feel the additional burden related to rising healthcare costs?*
- *Simple math tells us that ~49% of the population is in or will be entering retirement age; how will these patients pay for the medical care they will need as they enter their twilight years?*
- *More and more patients will be moving from the 45 to 64 age band into the 65 and older group; will these patients be on the hook for greater out-of-pocket expenses?*

3. Exhibits

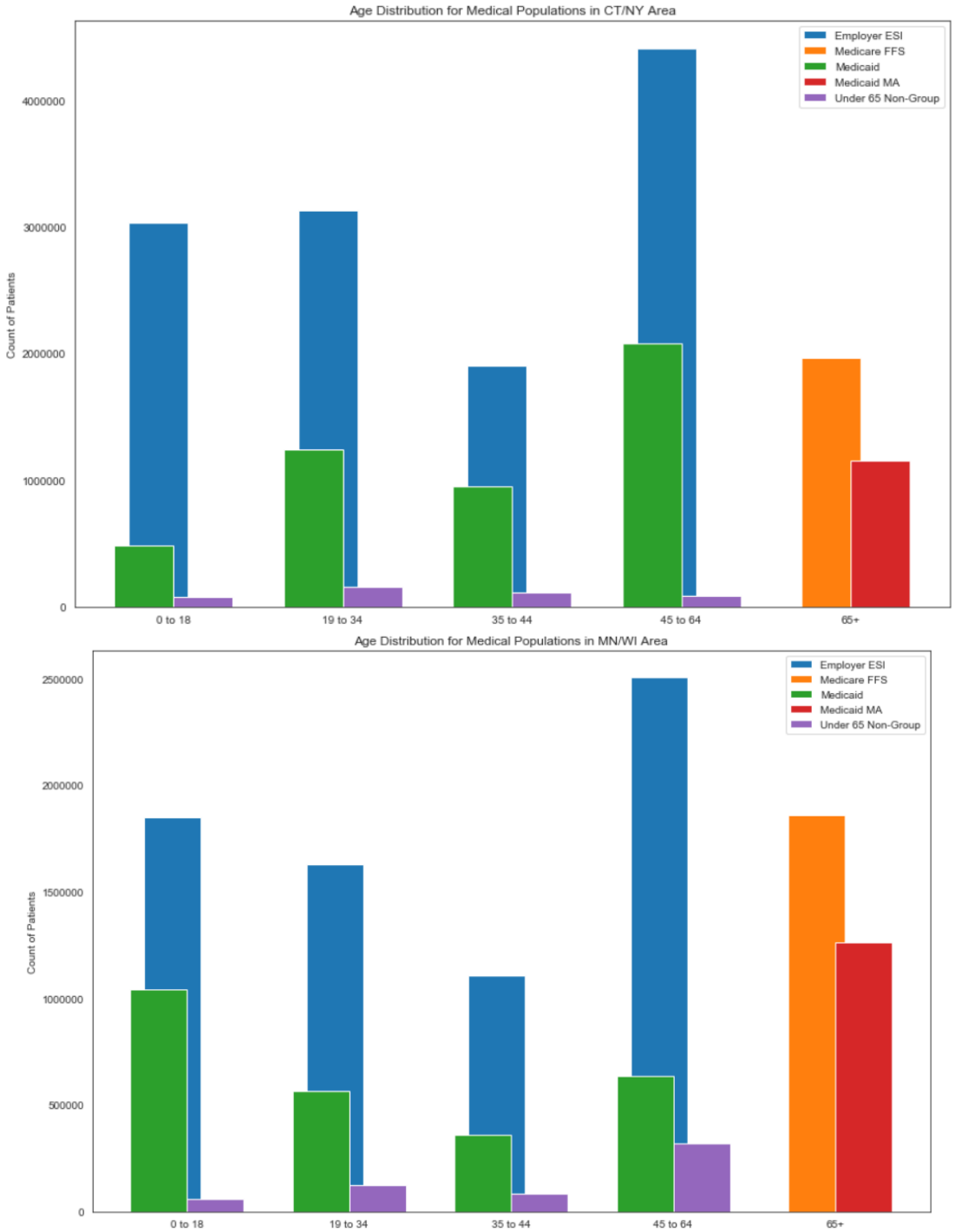
Exhibit 1 - Projected Healthcare Spending for Minnesota (2012 - 2026)



Source: Minnesota Department of Health

Notes: Author's computations. The projected spend calculations were based on building a simple linear model and using a function provided in Python to populate the estimated spend.

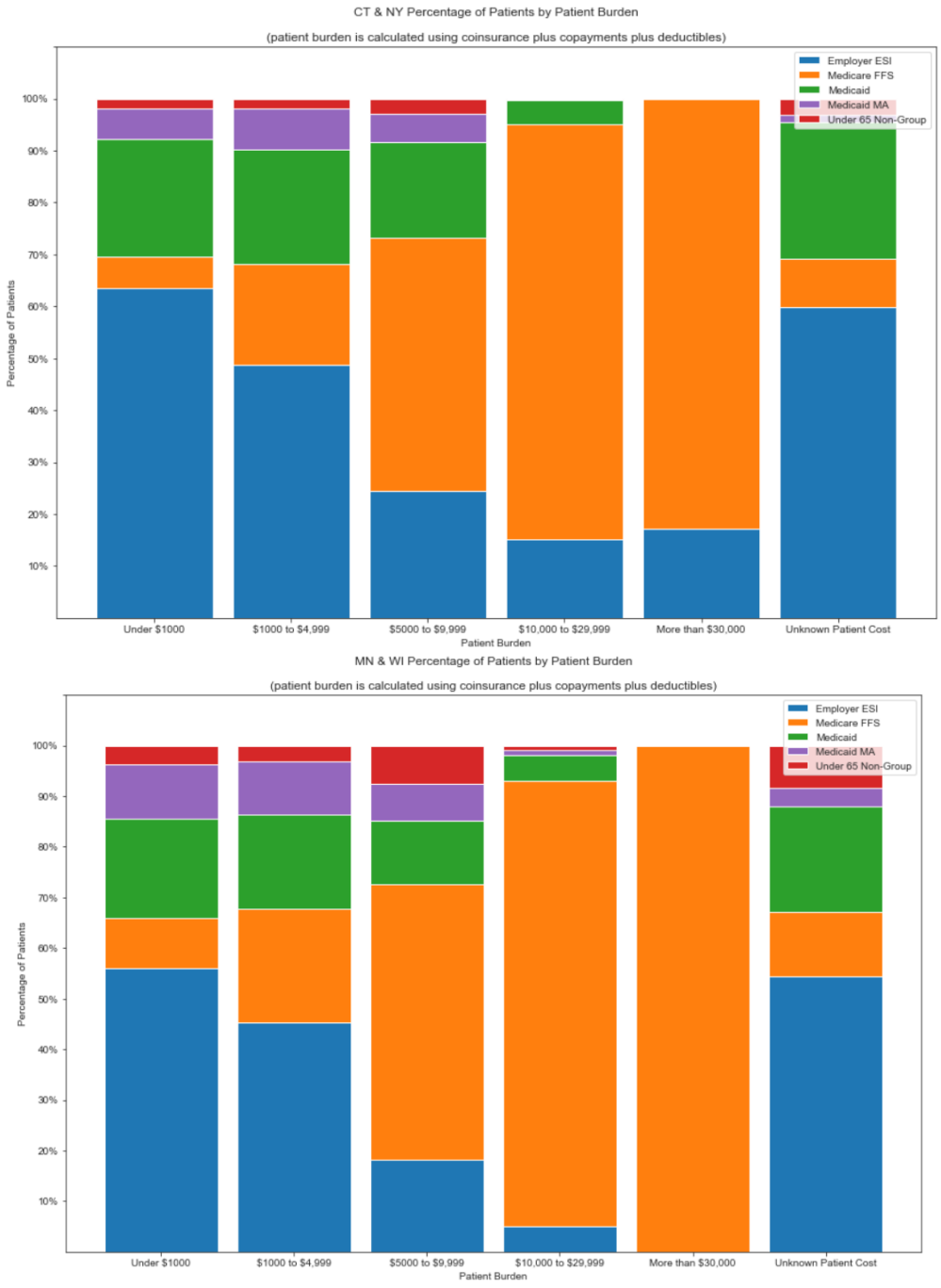
Exhibit 2 - Age Distributions for Medical Populations in Connecticut & New York compared to Minnesota & Wisconsin



Source: SHIAF claims database

Notes: Author's computations. Two programs were run separately, one to subset the Connecticut & New York populations, the other for Minnesota & Wisconsin populations.

Exhibit 3 - Percentage of Patient Burden for Medical Populations in Connecticut & New York compared to Minnesota & Wisconsin



Source: SHIAF claims database

Notes: Author's computations. Two programs were run separately, one to subset the Connecticut & New York populations, the other for Minnesota & Wisconsin populations. Data aggregation and manipulation to generate the plots is explained in an earlier section.

4. Appendix

Python Environment Details

```
Here are the environment details...

C:\Python\envs\MSBA2020\python.exe
3.7.1 (default, Oct 28 2018, 08:39:03) [MSC v.1912 64 bit (AMD64)]
sys.version_info(major=3, minor=7, micro=1, releaselevel='final', serial=0)

This notebook is using pandas version: 0.25.3.
This notebook is using numpy version: 1.16.4.
This notebook is using seaborn version: 0.9.0.
This notebook is using matplotlib version: 3.1.2.
```

Data Sources

Data was generated using proto SAS programs developed during the course. Here are the Excel files that were used:

```
## Load the CT/NY files

ctny_week3_file = "WEEK3D_HCAlytics_2018_DM_CTNY.XLSX"
ctny_week5_file = "WEEK5_HCAlytics_2019_DM_CTNY.XLSX"
```

```
## Load the MN/WI files

mnwi_week3_file = "WEEK3D_HCAlytics_2018_DM_MNWI.XLSX"
mnwi_week5_file = "WEEK5_HCAlytics_2019_DM_MNWI.XLSX"
```

Raw Data Table (for Projected Healthcare Spending in Minnesota visualization):

```
## Show the final dataframe!
minnesota_healthcare_df
```

	Year	Spend (Billions of Dollars)	Growth
0	2012	40.10	3.80
1	2013	41.50	3.49
2	2014	44.40	6.99
3	2015	46.20	4.05
4	2016	47.10	1.95
5	2017	51.81	10.00
6	2018	56.52	9.09
7	2019	61.23	8.33
8	2020	65.94	7.69
9	2021	70.65	7.14
10	2022	75.36	6.67
11	2023	80.07	6.25
12	2024	84.78	5.88
13	2025	89.49	5.56
14	2026	94.20	5.26

Raw Data Tables (for Age Distribution visualizations):

Connecticut & New York

Group	Label	All Patients	Employer ESI	Medicare FFS	Medicaid	Medicaid MA	Under 65 Non-Group
s_age0018	NaN	3586115	3027996	0	482875	0	75243
s_age1934	NaN	4528449	3131478	0	1242824	0	154145
s_age3544	NaN	2964619	1905848	0	946584	0	112186
s_age4564	NaN	6575748	4404113	0	2082266	0	89368
s_age65P	NaN	3120963	0	1968622	0	1152341	0

Minnesota & Wisconsin

Group	Label	All Patients	Employer ESI	Medicare FFS	Medicaid	Medicaid MA	Under 65 Non-Group
s_age0018	NaN	2952767	1850926	0	1040804	0	61036
s_age1934	NaN	2321521	1631029	0	565755	0	124736
s_age3544	NaN	1552200	1108645	0	359212	0	84342
s_age4564	NaN	3463297	2504347	0	637757	0	321192
s_age65P	NaN	3119659	0	1857308	0	1262351	0

Raw Data Tables (for Percentage of Patient Burden for Medical Populations visualizations):

Connecticut & New York

Group	Label	All Patients	Employer ESI	Medicare FFS	Medicaid	Medicaid MA	Under 65 Non-Group
S_BS_01	Total Patient Burden: < $1,000 =1, ELSE 0	14016654	8905450	855909	3154823	843558	256912
S_BS1_5	Total Patient Burden: $1,000 to <5,000 =1, E...	3229429	1571868	630511	711383	251561	64103
S_BS510	Total Patient Burden: $5,000 to < 10,000 =1, ...	244726	59803	119371	45178	13358	7015
S_BS1030	Total Patient Burden: $10,000 to < 30,000 =1,...	71663	10849	57278	3276	0	259
S_BSGT30	Total Patient Burden: $30,000+ =1, ELSE 0	7017	1202	5815	0	0	0
OTHER	unknown	3206405	1920263	299737	839889	43863	102653

Minnesota & Wisconsin

Group	Label	All Patients	Employer ESI	Medicare FFS	Medicaid	Medicaid MA	Under 65 Non-Group
S_BS_01	Total Patient Burden: < $1,000 =1, ELSE 0	8626989	4825441	870615	1688025	923984	318923
S_BS1_5	Total Patient Burden: $1,000 to <5,000 =1, E...	2319466	1049857	520934	432852	242587	73233
S_BS510	Total Patient Burden: $5,000 to < 10,000 =1, ...	246298	44787	134279	30492	18220	18518
S_BS1030	Total Patient Burden: $10,000 to < 30,000 =1,...	61162	3129	53729	3108	667	528
S_BSGT30	Total Patient Burden: $30,000+ =1, ELSE 0	2682	0	2682	0	0	0
OTHER	unknown	2152845	1171733	275066	449051	76890	180104

This notebook contains all of the code used to generate the plots for all of the analyses described in the paper:

- **MILI6963_HealthcareAnalytics_DannyMoncada_Visualizations.ipynb**

6. References

[1] Blanco, A. (2019, October 8). New tool aims to remove the mystery of health care costs at Connecticut hospitals. *Hartford Courant*. Retrieved from https://www.courant.com/news/connecticut/hc-news-connecticut-healthcare-tool-20191008-4xq4deajnjgqzhnz424d4mwz7y-story.html

[2] Smith, S. (2019, February 28). Minnesota saw lower health care spending growth in 2016. *Minnesota Department of Health*. Retrieved from https://www.health.state.mn.us/news/pressrel/2019/costs022819.html

[3] Malcolm, Jan K. (2019, February 1). Minnesota Health Care Spending: 2015 and 2016 Estimates and Ten-Year Projections. *Minnesota Department of Health*. Retreived from https://www.health.state.mn.us/data/economics/docs/costs/healthspending2019.pdf

[4] Parente, Stephen T. The Healthcare Marketplace. *Carlson School of Management, Medical Industry Leadership Institute*. Retrieved from http://ehealthecon.hsinetwork.com/

[5] Gaudard, Olivier. (2017, December 1). #13 Percent Stacked Barplot. *The Python Graph Gallery*. Retrieved from https://python-graph-gallery.com/13-percent-stacked-barplot/