ECON 470 HW1

Ben Yang

2023-01-25

Instructions

This assignment is just about GitHub and data management. The goal is to give you a chance to practice wrangling and tidying data. We do this very early in the class because we will start doing some empirical analysis using real data soon. The faster you are comfortable with the datasets, the better. For more detailed instructions on how to submit your homework answers, please see the overview page here.

Building the data

The purpose of this part of the assignment is essentially to practice database management. Most of your professional lives will likely involve managing data. It can be tedious but also extremely rewarding when you finally get to find out what's going on in the analysis stage. Anyway, let's get to work! All of these questions require you to use the Medicare Advantage GitHub Repo.

Enrollment Data

Run the R code to organize the Monthly Plan Enrollment Data. Once you've created your final dataset (it's called *full_ma_data* in my code), answer the following:

1. How many observations exist in your current dataset?

The current dataset have 1.913e+07 rows of observations.

2. How many different *plan_types* exist in the data?

There are 27 different *plan_types* exist in the data.

Table 1: Plan Types by Year

Plan Type	2007	2008	2009	2010	2011	2012	2013	2014	2015
Medicare Prescription Drug Plan	920,058	963,478	945,794	893,609	771,694	815,223	826,907	1,122,209	991,457
PFFS	364,285	630,756	683,361	385,733	45,781	36,423	31,919	24,905	13,658
HMO/HMOPOS	60,012	70,176	479,978	506,802	528,473	507,272	530,909	523,304	479,275
Employer/Union Only Direct Contract PDP	32,358	29,113	25,860	28,700	28,697	28,669	25,526	25,528	25,630
Regional PPO	26,402	27,990	25,943	24,442	22,773	21,602	19,970	19,773	17,578
Local PPO	17,427	38,470	405,197	417,551	515,700	636,701	633,884	664,716	704,993
1876 Cost	5,855	5,459	5,825	6,035	6,851	7,633	7,731	7,069	7,157
MSA	4,422	16,515	12,267	135	6,421	6,416	6,431	6,449	6,518
MSA Demo	3,274	NA	NA	NA	NA	NA	NA	NA	NA
Employer Direct PFFS	3,247	NA	NA	NA	NA	NA	NA	NA	NA
SHMO	1,125	NA	NA	NA	NA	NA	NA	NA	NA
MN Senior Health Options	968	NA	NA	NA	NA	NA	NA	NA	NA
PSO (State License)	421	535	87	123	176	171	NA	NA	NA
National PACE	405	548	616	717	781	858	953	1,118	1,216
PSO (Federal Waiver of State License)	162	NA	NA	NA	NA	NA	NA	NA	NA
Continuing Care Retirement Community	95	122	158	142	NA	NA	NA	NA	NA
ESRD I	75	122	123	117	NA	NA	NA	NA	NA
MA Health Senior Care Options	73	NA	NA	NA	NA	NA	NA	NA	NA
WI Partnership Program	42	NA	NA	NA	NA	NA	NA	NA	NA
MN Disability Health Options	21	NA	NA	NA	NA	NA	NA	NA	NA
Pilot	15	12	201	53	3	3	2	2	2
HCPP - 1833 Cost	13	13	3,938	3,604	11	11	10	9	9
ESRD II	12	12	7	8	NA	NA	NA	NA	NA
Employer/Union Only Direct Contract PFFS	NA	3,332	3,335	3,332	3,329	3,323	NA	NA	NA
RFB PFFS	NA	NA	3,006	NA	NA	NA	NA	NA	NA
Medicare-Medicaid Plan HMO/HMOPOS	NA	NA	NA	NA	NA	NA	265	1,319	4,130

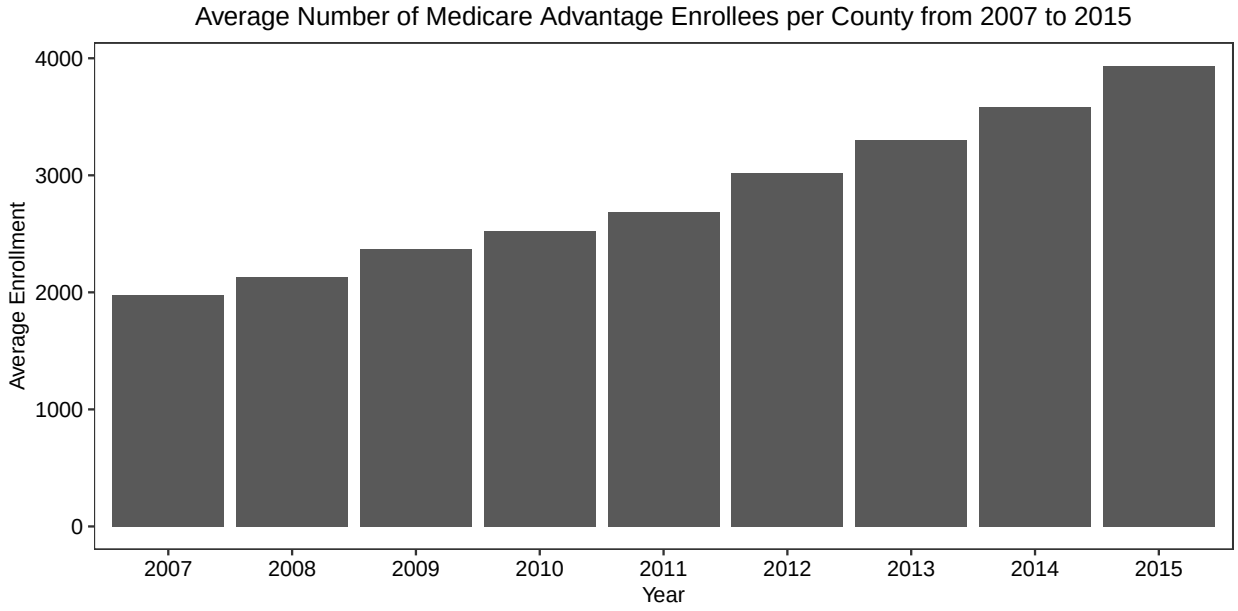
Table 2: Plan Types by Year

Plan Type	2007	2008	2009	2010	2011	2012	2013	2014	2015
Medicare Prescription Drug Plan	398,167	428,936	415,027	391,205	295,458	289,044	278,091	301,082	269,153
PFFS	51,987	105,859	89,586	54,119	22,038	17,449	12,945	6,053	4,232
HMO/HMOPOS	30,670	34,545	36,166	34,460	33,931	37,551	37,179	38,893	36,588
Regional PPO	7,254	7,794	8,470	10,659	10,995	11,279	9,660	10,420	8,531
Local PPO	6,116	7,612	9,929	11,652	13,874	17,030	17,089	17,169	16,728
1876 Cost	5,048	4,577	4,781	4,923	5,829	6,647	6,759	6,207	6,329
MSA	2,177	3,303	2,459	68	131	132	145	163	232
SHMO	458	NA	NA	NA	NA	NA	NA	NA	NA
National PACE	395	548	615	717	781	858	953	1,118	1,216
PSO (State License)	376	394	75	97	141	143	NA	NA	NA
MSA Demo	129	NA	NA	NA	NA	NA	NA	NA	NA
PSO (Federal Waiver of State License)	110	NA	NA	NA	NA	NA	NA	NA	NA
ESRD I	75	122	123	117	NA	NA	NA	NA	NA
Continuing Care Retirement Community	68	66	60	64	NA	NA	NA	NA	NA
RFB PFFS	NA	NA	3,006	NA	NA	NA	NA	NA	NA
Medicare-Medicaid Plan HMO/HMOPOS	NA	NA	NA	NA	NA	NA	265	1,319	4,130

3. Provide a table of the count of plans under each plan type in each year. Your table should look something like Table 1.

4. Remove all special needs plans (SNP), employer group plans (eghp), and all "800-series" plans. Provide an updated version of Table 1 after making these exclusions.

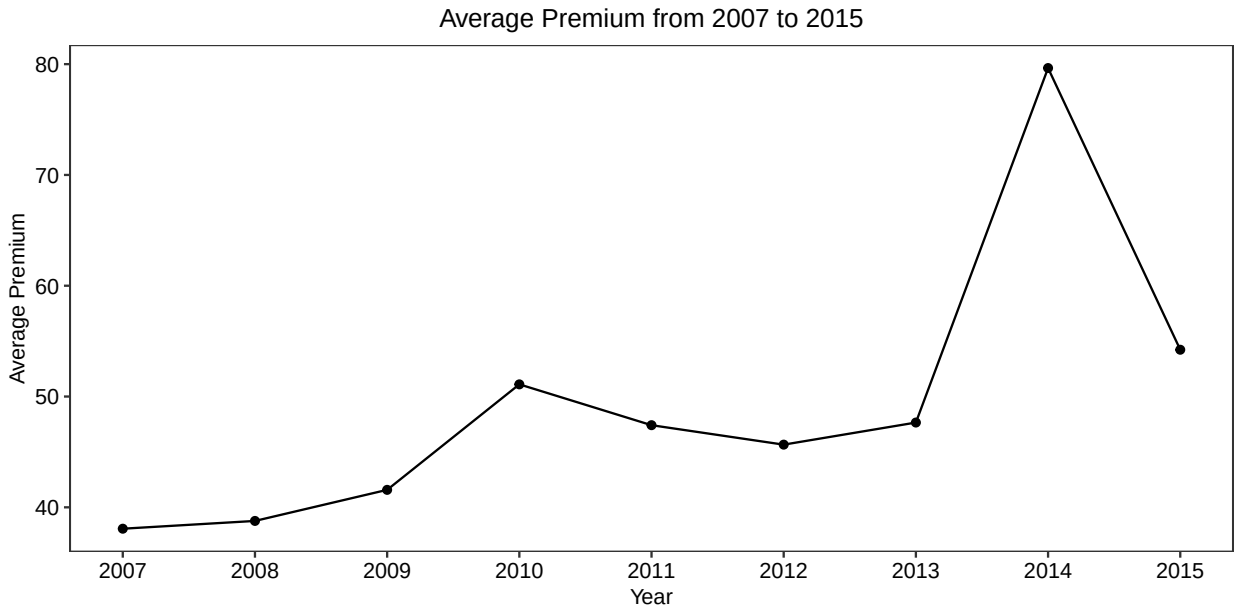
5. Merge the contract service area data to the enrollment data, and restrict the data only to contracts that are approved in their respective counties. The R script to create the service area dataset is here: Contract Service Area. And you can follow the _BuildFinalData.R script to see where/how I join the datasets. Limiting your dataset only to plans with non-missing enrollment data, provide a graph showing the average number of Medicare Advantage enrollees per county from 2008 to 2015. Be sure to format your graph in a meaningful way.



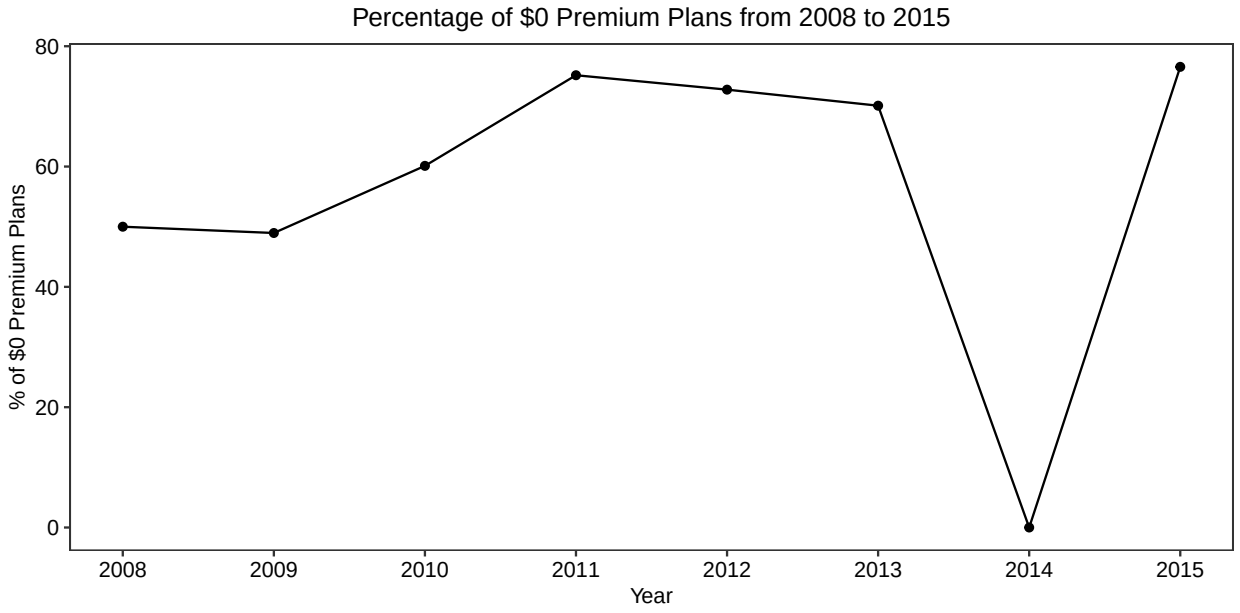
Premium Data

Now we're going to incorporate the plan premium information. This is part of the "Plan Characteristics" data, and the underlying R scripts for these files can be found here: Plan Characteristics.

6. Merge the plan characteristics data to the dataset you created in Step 6 above. Note that you'll need to join the Market Penetration Data in order to get the information you need to merge the plan characteristics. This is because the plan characteristics data only have state name and county (not FIPS codes). The penetration files have both FIPS codes and state/county names, so that dataset serves as a good crosswalk file. Provide a graph showing the average premium over time. Don't forget about formatting!



7. Provide a graph showing the percentage of $0 premium plans over time. Also. . . remember to format things.



Summary Questions

With all of this data work and these great summaries, let's take a step back and think about what all this means.

8. Why did we drop the "800-series" plans?

9. Why do so many plans charge a $0 premium? What does that really mean to a beneficiary?

The beneficiary of a plan that charges $0 premium does not need to pay for the insurance coverage directly. Instead, the federal government contracts with private insurance companies to administer the Medicare Advantage plans. If the Medicare Advantage plan spend less than the flat fee given by the government, it may choose to offer more benefits. As a result, private insurers are incentivized to compete for more business with low premiums and additional benefits, such as dental, vision and/or prescription coverage. Many Medicare Advantage plans also offer preventive care and disease management programs to help people better manage their health, so that the plan may have healthier enrollees with lower expected healthcare costs.

However, the beneficiary may still need to pay deductibles and co-pays for the services covered by the Medicare Advantage plan, despite paying $0 premium. Besides, the beneficiary still have to pay the Medicare Part B premium. Moreover, Medicare Advantage plans usually enter into contracts with a network of doctors and hospitals for cost management purposes, which means that the beneficiary needs to pay more money out of pocket for visiting a healthcare provider outside the plan's network.

10. Briefly describe your experience working with these data (just a few sentences). Tell me one thing you learned and one thing that really aggravated you.

Working with these data was intimidating at first due to the size of the data set and the unfamiliar columns. After understanding the meanings of the data entries, the assignment becomes much more manageable. The provided data code of merging data sets to build the final data set is very helpful. One thing I learned is that the Medicare Advantage plans seem to effectively make healthcare insurance more affordable for people, as showed by the percentage of $0 premium plans.