

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 19,126,783 rows of observations.

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

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

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

Table 1: Plan Count by Year

plan_type	2014	2015	2008	2009	2007	2010	2013	2012	2011
Medicare Prescription Drug Plan	1122209	991457	963478	945794	920058	893609	826907	815223	771694

plan_type	2014	2015	2008	2009	2007	2010	2013	2012	2011
Local PPO	664716	704993	38470	405197	17427	417551	633884	636701	515700
PFFS	24905	13658	630756	683361	364285	385733	31919	36423	45781
HMO/HMOPOS	523304	479275	70176	479978	60012	506802	530909	507272	528473
NA	NA	NA	NA	27505	NA	277533	NA	NA	NA
Employer/Union Only Direct Contract PDP	25528	25630	29113	25860	32358	28700	25526	28669	28697
Regional PPO	19773	17578	27990	25943	26402	24442	19970	21602	22773
MSA	6449	6518	16515	12267	4422	135	6431	6416	6421
1876 Cost	7069	7157	5459	5825	5855	6035	7731	7633	6851
Medicare-Medicaid Plan HMO/HMOPOS	1319	4130	NA	NA	NA	NA	265	NA	NA
HCPP - 1833 Cost	9	9	13	3938	13	3604	10	11	11
Employer/Union Only Direct Contract PFFS	NA	NA	3332	3335	NA	3332	NA	3323	3329
MSA Demo	NA	NA	NA	NA	3274	NA	NA	NA	NA
Employer Direct PFFS	NA	NA	NA	NA	3247	NA	NA	NA	NA
RFB PFFS	NA	NA	NA	3006	NA	NA	NA	NA	NA
National PACE	1118	1216	548	616	405	717	953	858	781
SHMO	NA	NA	NA	NA	1125	NA	NA	NA	NA
MN Senior Health Options	NA	NA	NA	NA	968	NA	NA	NA	NA
PSO (State License)	NA	NA	535	87	421	123	NA	171	176
Pilot	2	2	12	201	15	53	2	3	3
PSO (Federal Waiver of State License)	NA	NA	NA	NA	162	NA	NA	NA	NA
Continuing Care Retirement Community	NA	NA	122	158	95	142	NA	NA	NA
ESRD I	NA	NA	122	123	75	117	NA	NA	NA
MA Health Senior Care Options	NA	NA	NA	NA	73	NA	NA	NA	NA
WI Partnership Program	NA	NA	NA	NA	42	NA	NA	NA	NA
MN Disability Health Options	NA	NA	NA	NA	21	NA	NA	NA	NA
ESRD II	NA	NA	12	7	12	8	NA	NA	NA

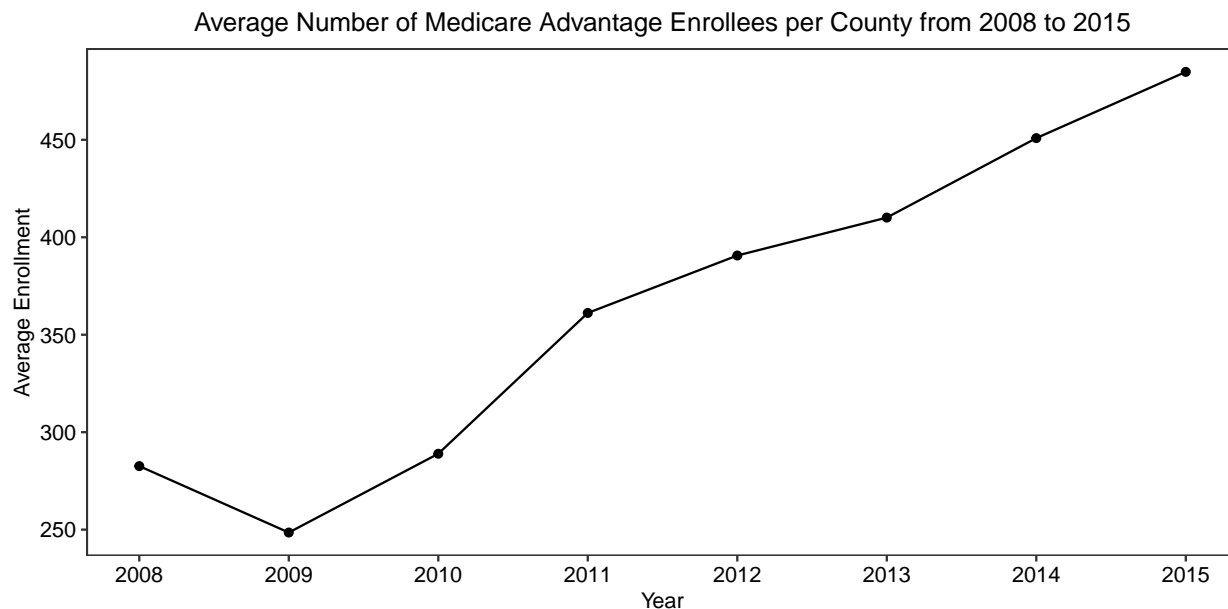
4. Remove all special needs plans (SNP), employer group plans (eghp), and all “800-series” plans. Provide an updated version of Table 1 after making these exclusions.

Table 2: Plan Count by Year (excl. snp, eghp, 800-series)

plan_type	2008	2009	2007	2010	2014	2011	2012	2013	2015
Medicare Prescription Drug Plan	428936	415027	398167	391205	301082	295458	289044	278091	269153
PFFS	105859	89586	51987	54119	6053	22038	17449	12945	4232
HMO/HMOPOS	34545	36166	30670	34460	38893	33931	37551	37179	36588
NA	NA	13619	NA	29733	NA	NA	NA	NA	NA
Local PPO	7612	9929	6116	11652	17169	13874	17030	17089	16728
Regional PPO	7794	8470	7254	10659	10420	10995	11279	9660	8531
1876 Cost	4577	4781	5048	4923	6207	5829	6647	6759	6329
Medicare-Medicaid Plan HMO/HMOPOS	NA	NA	NA	NA	1319	NA	NA	265	4130
MSA	3303	2459	2177	68	163	131	132	145	232
RFB PFFS	NA	3006	NA	NA	NA	NA	NA	NA	NA

plan_type	2008	2009	2007	2010	2014	2011	2012	2013	2015
National PACE	548	615	395	717	1118	781	858	953	1216
SHMO	NA	NA	458	NA	NA	NA	NA	NA	NA
PSO (State License)	394	75	376	97	NA	141	143	NA	NA
MSA Demo	NA	NA	129	NA	NA	NA	NA	NA	NA
ESRD I	122	123	75	117	NA	NA	NA	NA	NA
PSO (Federal Waiver of State License)	NA	NA	110	NA	NA	NA	NA	NA	NA
Continuing Care Retirement Community	66	60	68	64	NA	NA	NA	NA	NA

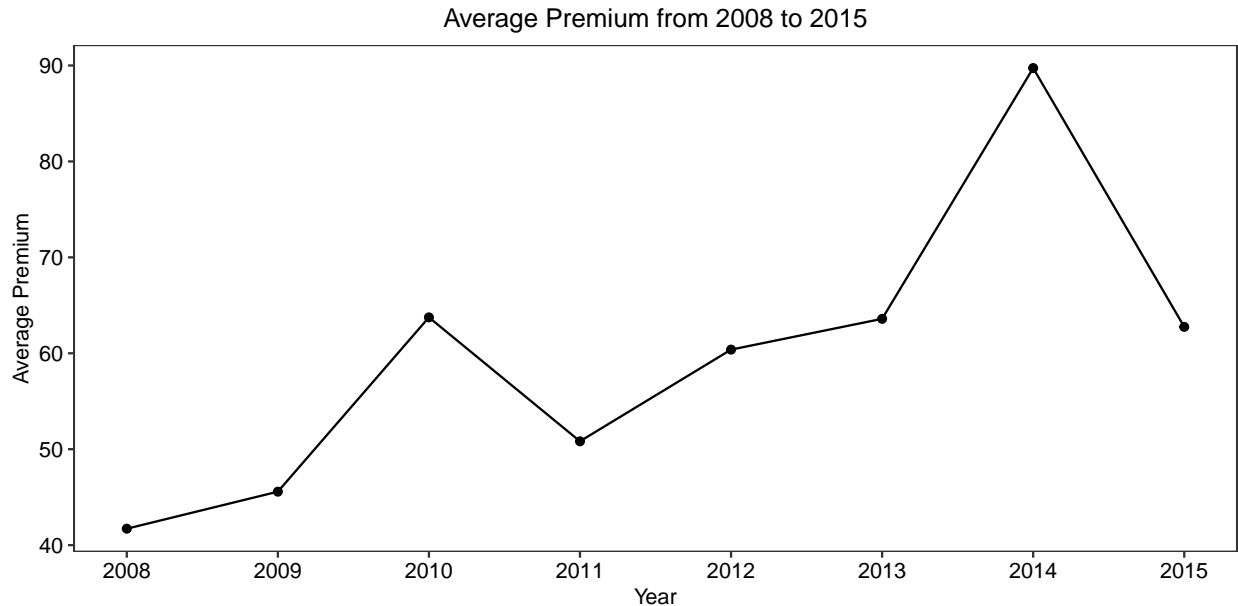
5. Merge the contract service area data to the enrollment data, and restrict the data only to contracts that are approved in their respective counties. The R script to create the service area dataset is here: [Contract Service Area](#). And you can follow the [__BuildFinalData.R](#) script to see where/how 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.



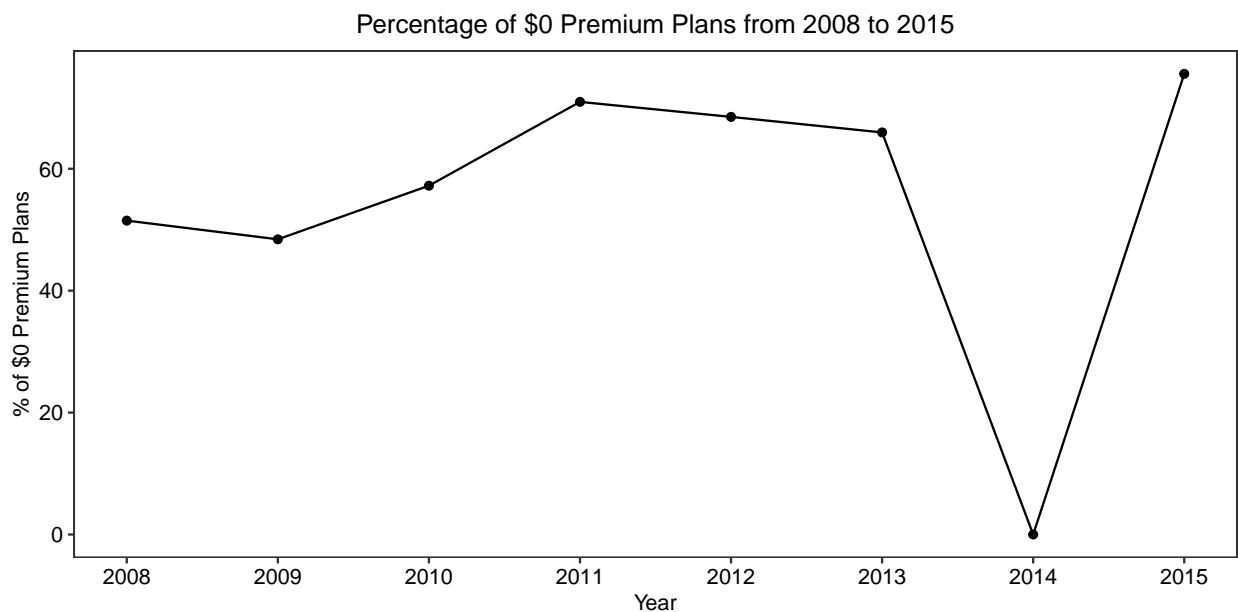
Premium Data

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 unfamiliar columns. After understanding the meanings of the data entry, the assignment becomes much more manageable. The code of merging data sets to build the final data set provided is very helpful. One thing I learned is that the Medicare Advantage plans effectively make healthcare services more affordable as showed by the percentage of \$0 premium plans.