

Other Formats

- [PDF](#)

Homework One Submission 3

Author

Safia Read

This is my third submission of the first homework for Econ 470.

[Link to Github](#)

<https://github.com/safiaread/homework-1>

Enrollment Data

1. How many observations exist in your current dataset?

By using the `nrow()` function, I found 13,276,162 total observations in the full dataset.

2. How many different `plan_types` exist in the data? By using the `unique()` function, I found 17 plans in the dataset, not including "NA".
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 types by year

Plan Type	2010	2011	2012	2013	2014	2015
1876 Cost	6,035	6,851	7,633	7,731	7,069	7,157
Employer/Union Only Direct Contract PDP	28,700	28,697	28,669	25,526	25,528	25,630
HCPP - 1833 Cost	3,604	11	11	10	9	9
HMO/HMOPOS	506,802	528,473	507,272	530,909	523,304	479,275
Local PPO	417,551	515,700	636,701	633,884	664,716	704,993
MSA	135	6,421	6,416	6,431	6,449	6,518
Medicare Prescription Drug Plan	893,609	771,694	815,223	826,907	1,122,209	991,457
National PACE	717	781	858	953	1,118	1,216
PFFS	385,733	45,781	36,423	31,919	24,905	13,658
Pilot	53	3	3	2	2	2
Regional PPO	24,442	22,773	21,602	19,970	19,773	17,578

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: Revised plan types by year

Plan Type	2010	2011	2012	2013	2014	2015
1876 Cost	4,923	5,829	6,647	6,759	6,207	6,329
HMO/HMOPOS	34,460	33,931	37,551	37,179	38,893	36,588
Local PPO	11,652	13,874	17,030	17,089	17,169	16,728
MSA	68	131	132	145	163	232
Medicare Prescription Drug Plan	391,205	295,458	289,044	278,091	301,082	269,153
National PACE	717	781	858	953	1,118	1,216
PFFS	54,119	22,038	17,449	12,945	6,053	4,232
Regional PPO	10,659	10,995	11,279	9,660	10,420	8,531

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 2010 to 2015. Be sure to format your graph in a meaningful way.

## Average Enrollments Per County Over Time

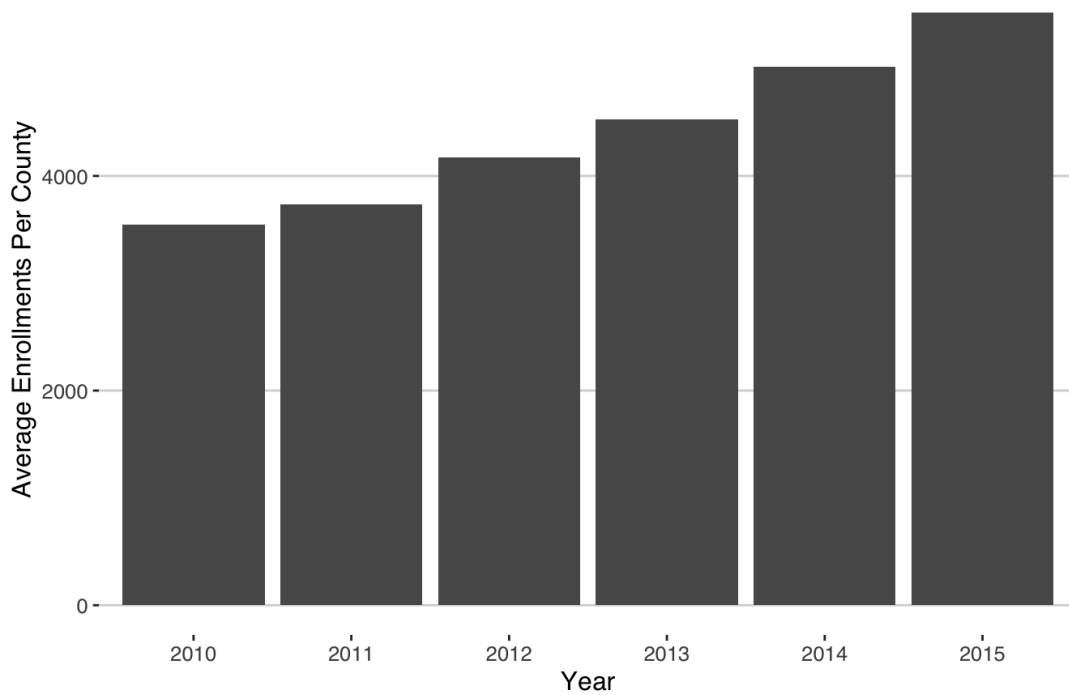


Figure 1: Average enrollment per county from 2010-2015.

## Premium Data

6. Merge the plan characteristics data to the dataset you created in Step 5 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!

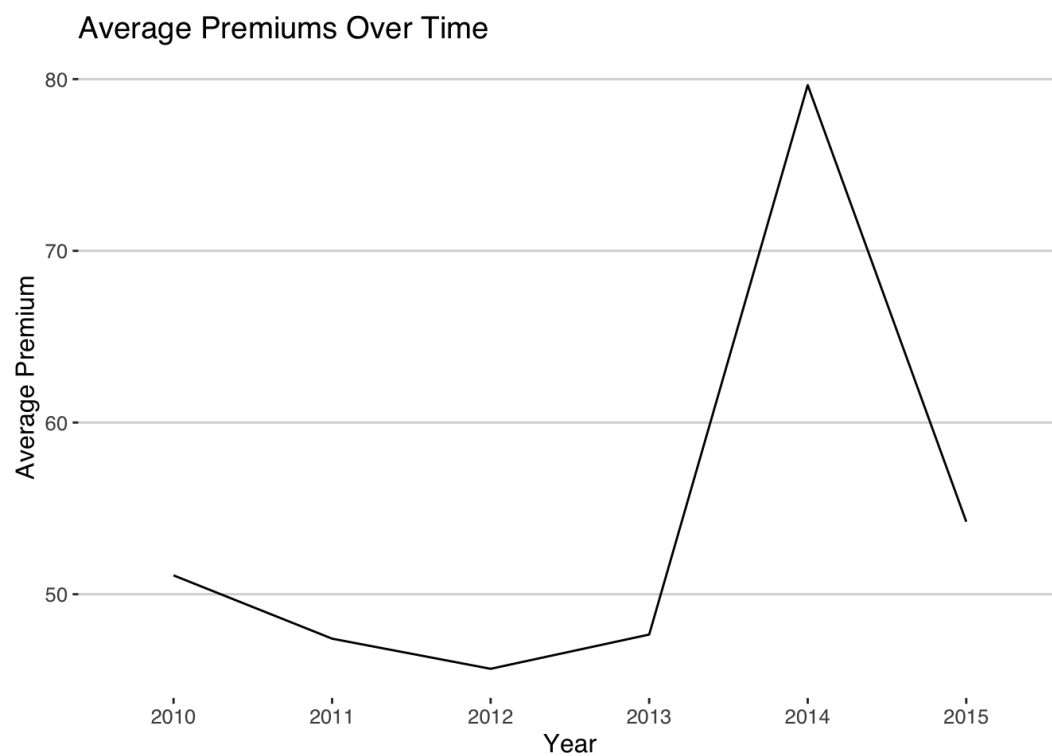


Figure 2: Average plan premium per county from 2010-2015.

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

I need the merged dataset from above to run this code but this is what I assume I need to do once it is merged.

## Percentage of Zero Premium Plans Over Time

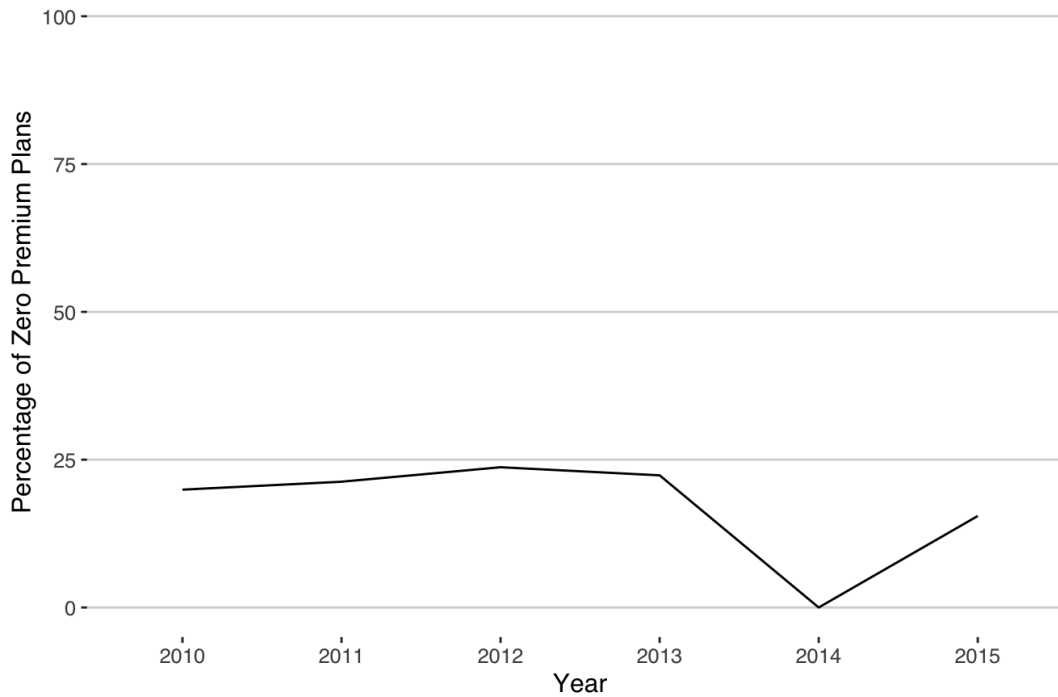


Figure 3: Percentage of \$0 premium plans over time from 2010-2015.

### Summary Questions

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

800 series plans are offered directly to employers and unions, and so are not available to an average person. This means they are a special group of observations that aren't reflective of the average person's Medicare Advantage experience.

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

The \$0 monthly premium is in addition to a premium for Part B of Medicare Advantage, which all enrollees have to pay. Medicare Plan B charges premiums for an extended variety of services, however the tradeoff is a small network of providers with high out of pocket costs. Traditional medicare has no 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.

With this data, it was definitely an adjustment to manipulate it in VSCode. I have previously used RStudio, and I think you were right that this is a less intuitive interface. I also have never worked with a dataset so large before, so it took some time to figure out how to upload it to my computer and organize everything so my computer would run. However, I think I learned a lot about Github and creating qmd files. The hardest part was the merging of datasets for question 6, which would have been difficult to figure out on my own at the level I am at now.