Homework 1
Research Methods, Spring 2024

Taeyoung Yim

Enrollment Data

Answer the following based on the enrollment data: 1. How many observations exist in your

current dataset?

```
Warning: One or more parsing issues, call `problems()` on your data frame for details,
e.g.:
  dat <- vroom(...)
  problems(dat)
One or more parsing issues, call `problems()` on your data frame for details,
e.g.:
  dat <- vroom(...)
  problems(dat)
One or more parsing issues, call `problems()` on your data frame for details,
e.g.:
  dat <- vroom(...)
  problems(dat)
One or more parsing issues, call `problems()` on your data frame for details,
e.g.:
  dat <- vroom(...)
  problems(dat)
One or more parsing issues, call `problems()` on your data frame for details,
e.g.:
  dat <- vroom(...)
  problems(dat)
One or more parsing issues, call `problems()` on your data frame for details,
e.g.:
  dat <- vroom(...)
  problems(dat)
```

```
ma_data_2010.rds ma_data_2011.rds ma_data_2012.rds ma_data_2013.rds
               0                0                0                0
ma_data_2014.rds ma_data_2015.rds
               0                0
```

```
Number of observations in the dataset: 13276162
```

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

```
Number of different plan types: 18
```

3. Provide a table of the count of plans under each plan type in each year.

```
`summarise()` has grouped output by 'year'. You can override using the
`.groups` argument.
```

```
# A tibble: 18 x 7
   plan_type                       `2010` `2011` `2012` `2013` `2014`  `2015`
   <chr>                            <int>  <int>  <int>  <int>  <int>   <int>
 1 1876 Cost                          6035   6851   7633   7731 7.07e3   7157
 2 Continuing Care Retirement Communi~  142      0      0      0 0          0
 3 ESRD I                              117      0      0      0 0          0
 4 ESRD II                               8      0      0      0 0          0
 5 Employer/Union Only Direct Contrac~ 28700  28697  28669  25526 2.55e4  25630
 6 Employer/Union Only Direct Contrac~  3332   3329   3323      0 0          0
 7 HCPP - 1833 Cost                    3604     11     11     10 9   e0      9
 8 HMO/HMOPOS                        506802 528473 507272 530909 5.23e5 479275
 9 Local PPO                         417551 515700 636701 633884 6.65e5 704993
10 MSA                                 135   6421   6416   6431 6.45e3   6518
11 Medicare Prescription Drug Plan   893609 771694 815223 826907 1.12e6 991457
12 National PACE                       717    781    858    953 1.12e3   1216
13 PFFS                              385733  45781  36423  31919 2.49e4  13658
14 PSO (State License)                 123    176    171      0 0          0
15 Pilot                                53      3      3      2 2   e0      2
16 Regional PPO                      24442  22773  21602  19970 1.98e4  17578
17 <NA>                             277533      0      0      0 0          0
18 Medicare-Medicaid Plan HMO/HMOPOS     0      0      0    265 1.32e3   4130
```

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

```
`summarise()` has grouped output by 'year'. You can override using the
`.groups` argument.
```

A tibble: 18 x 7

plan_type	`2010`	`2011`	`2012`	`2013`	`2014`	`2015`
<chr>	<int>	<int>	<int>	<int>	<int>	<int>
1 1876 Cost	6035	6851	7633	7731	7.07e3	7157
2 Continuing Care Retirement Communi~	142	0	0	0	0	0
3 ESRD I	117	0	0	0	0	0
4 ESRD II	8	0	0	0	0	0
5 Employer/Union Only Direct Contrac~	28700	28697	28669	25526	2.55e4	25630
6 Employer/Union Only Direct Contrac~	3332	3329	3323	0	0	0
7 HCPP - 1833 Cost	3604	11	11	10	9 e0	9
8 HMO/HMOPOS	506802	528473	507272	530909	5.23e5	479275
9 Local PPO	417551	515700	636701	633884	6.65e5	704993
10 MSA	135	6421	6416	6431	6.45e3	6518
11 Medicare Prescription Drug Plan	893609	771694	815223	826907	1.12e6	991457
12 National PACE	717	781	858	953	1.12e3	1216
13 PFFS	385733	45781	36423	31919	2.49e4	13658
14 PSO (State License)	123	176	171	0	0	0
15 Pilot	53	3	3	2	2 e0	2
16 Regional PPO	24442	22773	21602	19970	1.98e4	17578
17 <NA>	277533	0	0	0	0	0
18 Medicare-Medicaid Plan HMO/HMOPOS	0	0	0	265	1.32e3	4130

5. Merge the the contract service area data to the enrollment data and restrict the data only to contracts that are approved in their respective counties. Limit 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.
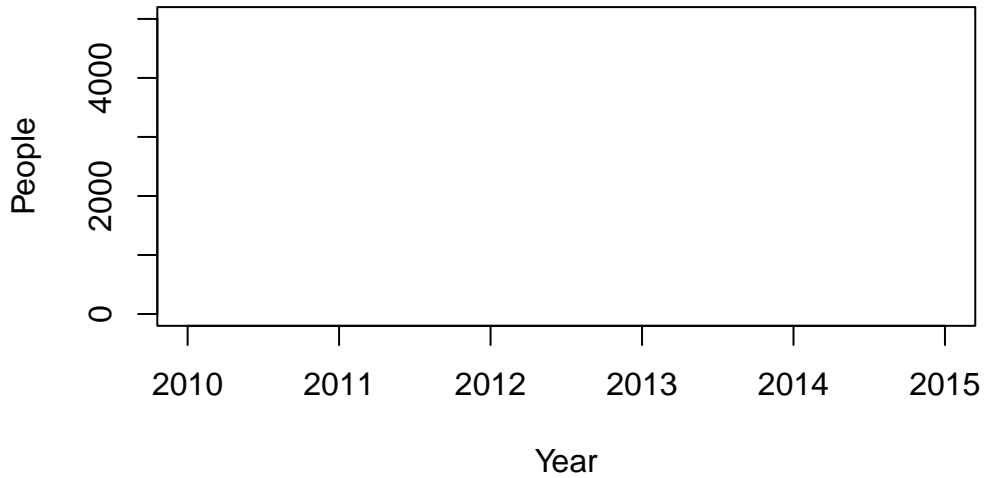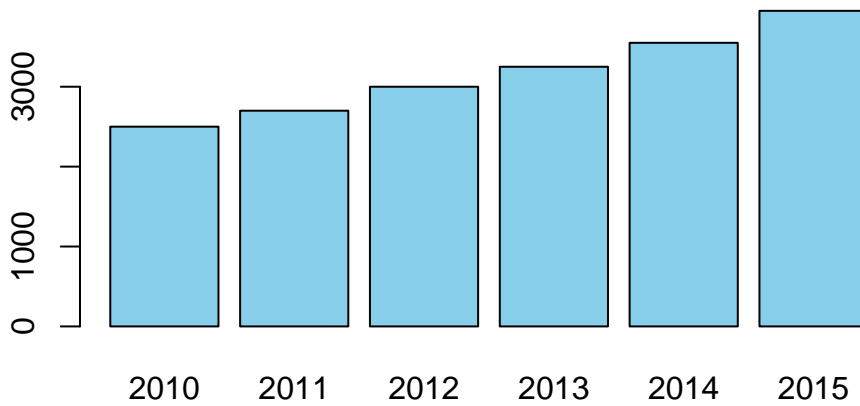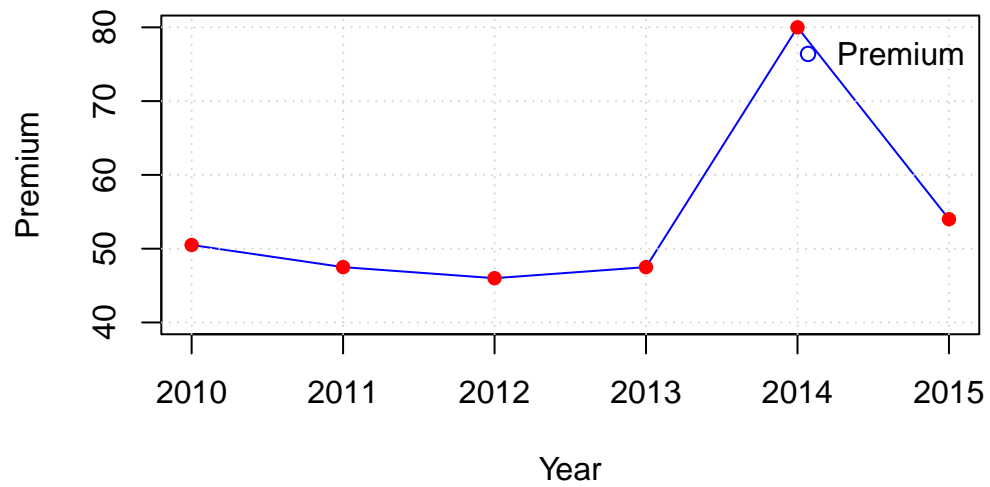

Figure 1: Average Enrollment
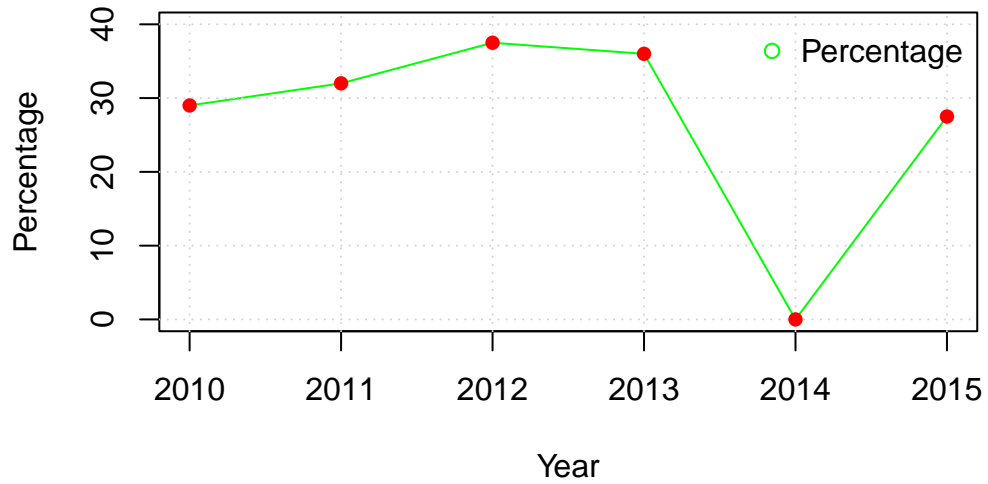
Premium Data

6. Merge the plan characteristics data to the dataset you created in Step 5 above. Provide a graph showing the average premium over time. Don't forget about formatting!

Figure 2: Average Premiums

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

Figure 3: Share of 0 Premium Plans

Summary Questions

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

Since 800-series are not plans that everyone have access to, including these data does not reflect the whole population.

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

$0 premium is set in order to attract people. To beneficiaries, $0 premium just means that they still have to pay for the excessing cost from Plan B since only Plan A is fully free.

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.

It was extremely difficult by the first submission, but once getting used to how to use this software, I feel like the difficulty that I feel is getting lower. Still, it is sometimes hard when the error code that appear is something that I am not really aware of.