Homework 4

Submission 1

Conor Mulligan

First submission of homework 4.

Link to Github

Summarize The Data

1. Remove all SNPs, 800-series plans, and prescription drug only plans (i.e., plans that do not offer Part C benefits). Provide a box and whisker plot showing the distribution of plan counts by county over time. Do you think that the number of plans is sufficient, too few, or too many?
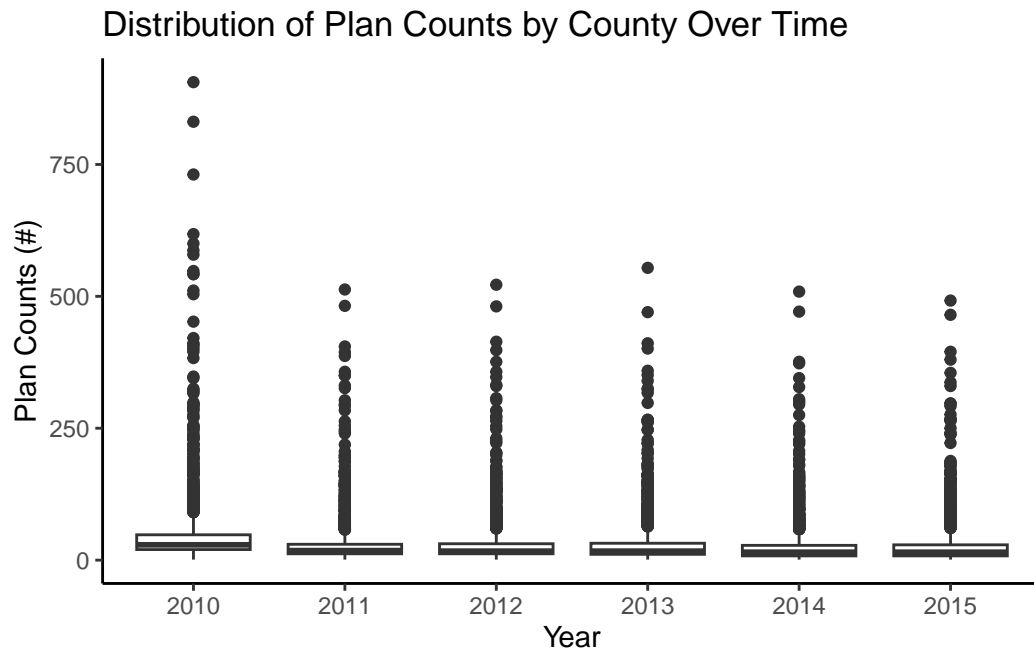

Figure 1: Distribution of Plan Counts Over Time (By County)

This box and whisker plot shows the distribution of plan counts by county. Some outliers can be seen in the data, particularly around years like 2010. It may be wise to remove these, but otherwise the number of plans seems sufficient.

2. Provide bar graphs showing the distribution of star ratings in 2010, 2012, and 2015. How has this distribution changed over time?
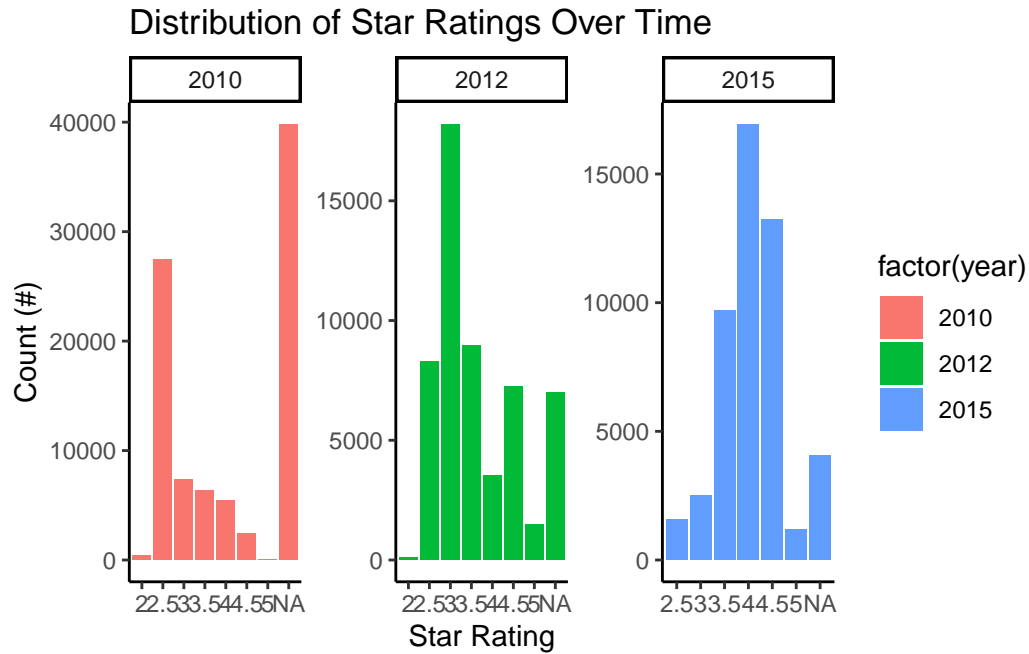

Figure 2: Distribution of Star Ratings: 2010, 2012 and 2015

This graph displays a shift towards higher star ratings over time. The distribution seems fairly even across all star ratings in 2010. In the 2015 area of the graph, there are far more ratings for 4 stars and 5 stars; subsequent years also show a decrease in lower star ratings.

3. Plot the average benchmark payment over time from 2010 through 2015. How much has the average benchmark payment risen over the years?
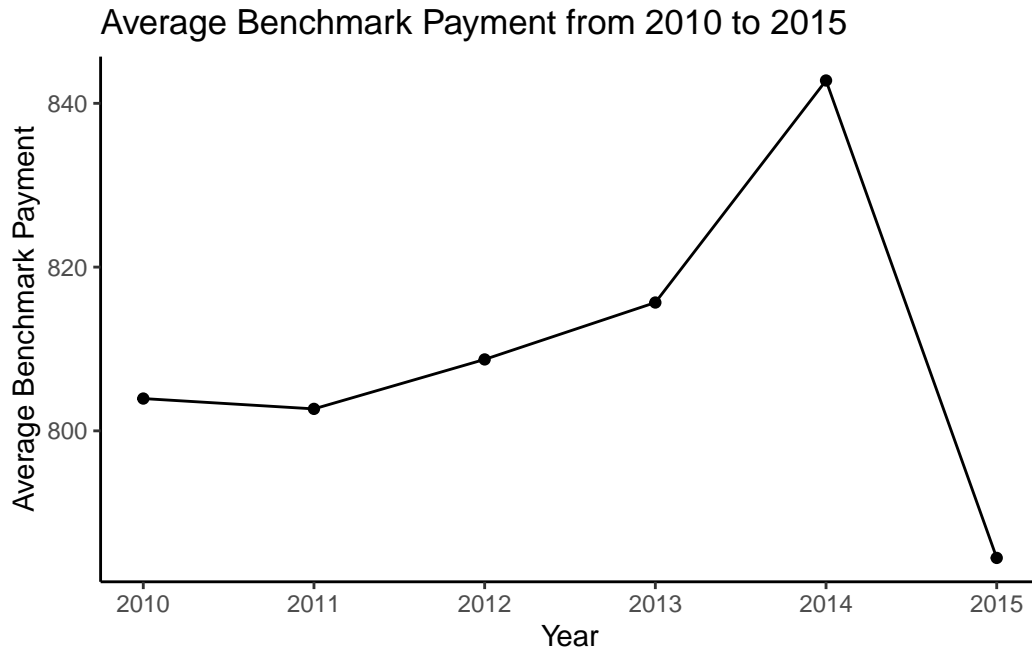

Figure 3: Average Benchmark Payment From 2010 to 2015

This graph shows the average benchmark payment rising before falling again over time.

4. Plot the average share of Medicare Advantage (relative to all Medicare eligibles) over time from 2010 through 2015. Has Medicare Advantage increased or decreased in popularity? How does this share correlate with benchmark payments?
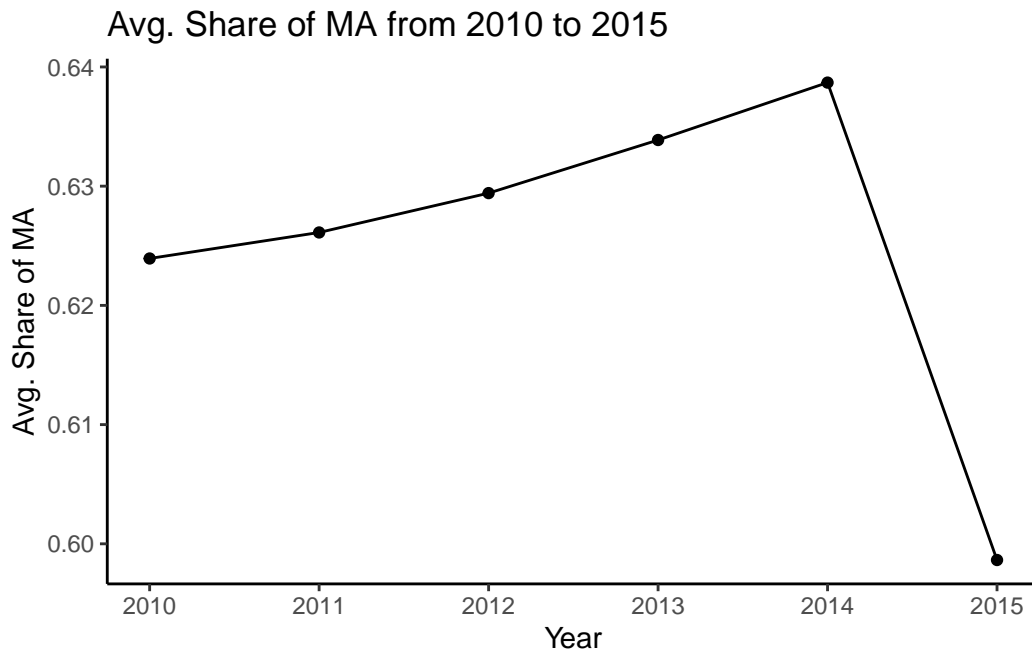
Figure 4: Average Share of Medicare Advantage From 2010 to 2015

This graph shows that the average share of Medicare Advantage enrollees increased from 2010 to 2015. It follows a similar trend to the previous graph, suggesting a possible correlation between two.

Estimate ATEs

5.Calculate the running variable underlying the star rating. Provide a table showing the number of plans that are rounded up into a 3-star, 3.5-star, 4-star, 4.5-star, and 5-star rating.

```
Loading required package: knitr
```

Table 1: Count of Plans Rounded into Nearest Half Ratings

star_rating	number_of_plans
1.5	8
2.0	712
2.5	5059
3.0	4962
3.5	3611
4.0	1935
4.5	50

6. Using the RD estimator with a bandwidth of 0.125, provide an estimate of the effect of receiving a 3-star versus a 2.5 star rating on enrollments. Repeat the exercise to estimate the effects at 3.5 stars, and summarize your results in a table.

```
            Length Class  Mode
Estimate    4      -none- numeric
bws         4      -none- numeric
coef        3      -none- numeric
se          3      -none- numeric
z           3      -none- numeric
pv          3      -none- numeric
ci          6      -none- numeric
beta_Y_p_l  2      -none- numeric
beta_Y_p_r  2      -none- numeric
V_cl_l      4      -none- numeric
V_cl_r      4      -none- numeric
V_rb_l      4      -none- numeric
V_rb_r      4      -none- numeric
N           2      -none- numeric
N_h         2      -none- numeric
N_b         2      -none- numeric
M           2      -none- numeric
tau_cl      2      -none- numeric
tau_bc      2      -none- numeric
c           1      -none- numeric
p           1      -none- numeric
q           1      -none- numeric
bias        2      -none- numeric
kernel      1      -none- character
all         0      -none- NULL
vce         1      -none- character
bwselect    1      -none- character
level       1      -none- numeric
masspoints  1      -none- character
rdmodel     1      -none- character
beta_covs   0      -none- NULL
call        9      -none- call
```

Used simple summary function for now it seems that it cannot be put into a knitr function; gives error "cannot coerce." Will clean up for subsequent submissions.

	Length	Class	Mode
Estimate	4	-none-	numeric
bws	4	-none-	numeric
coef	3	-none-	numeric
se	3	-none-	numeric
z	3	-none-	numeric
pv	3	-none-	numeric
ci	6	-none-	numeric
beta_Y_p_l	2	-none-	numeric
beta_Y_p_r	2	-none-	numeric
V_cl_l	4	-none-	numeric
V_cl_r	4	-none-	numeric
V_rb_l	4	-none-	numeric
V_rb_r	4	-none-	numeric
N	2	-none-	numeric
N_h	2	-none-	numeric
N_b	2	-none-	numeric
M	2	-none-	numeric
tau_cl	2	-none-	numeric
tau_bc	2	-none-	numeric
c	1	-none-	numeric
p	1	-none-	numeric
q	1	-none-	numeric
bias	2	-none-	numeric
kernel	1	-none-	character
all	0	-none-	NULL
vce	1	-none-	character
bwselect	1	-none-	character
level	1	-none-	numeric
masspoints	1	-none-	character
rdmodel	1	-none-	character
beta_covs	0	-none-	NULL
call	9	-none-	call

7. Repeat your results for bandwidths of 0.1, 0.12, 0.13, 0.14, and 0.15 (again for 3 and 3.5 stars). Show all of the results in a graph. How sensitive are your findings to the choice of bandwidth?
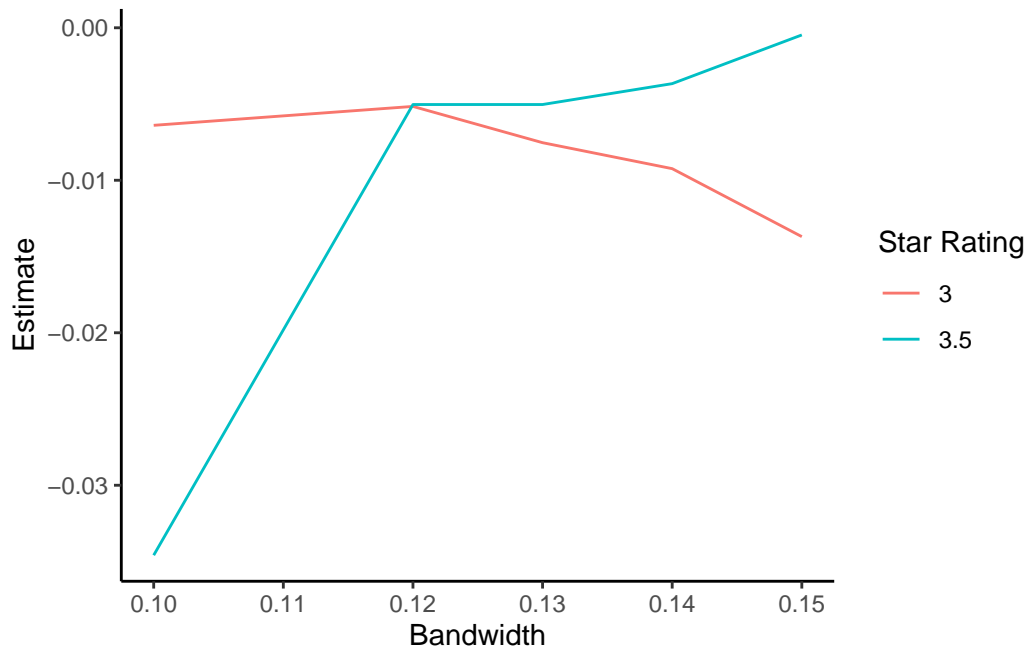

Figure 5: Star Ratings with New Bandwidths

This graph shows the RD estimate for the effect of a 3-star rating compared to a 2.5-star rating as positive; the same is true for the 3.5 vs. 3 star ratings. This would indicate higher ratings are correlated with higher enrollment. The flatness of the lines suggests relatively low sensitivity.

8. Examine (graphically) whether contracts appear to manipulate the running variable. In other words, look at the distribution of the running variable before and after the relevent threshold values. What do you find?

```
          Length Class  Mode
hat         3      -none- list
sd_asy      3      -none- list
sd_jk       3      -none- list
test        4      -none- list
hat_p       3      -none- list
sd_asy_p    3      -none- list
sd_jk_p     3      -none- list
test_p      4      -none- list
N           5      -none- list
h           2      -none- list
opt        20      -none- list
X_min       2      -none- list
X_max       2      -none- list
bino        5      -none- list


          Length Class  Mode
hat         3      -none- list
sd_asy      3      -none- list
sd_jk       3      -none- list
test        4      -none- list
hat_p       3      -none- list
sd_asy_p    3      -none- list
sd_jk_p     3      -none- list
test_p      4      -none- list
N           5      -none- list
h           2      -none- list
opt        20      -none- list
X_min       2      -none- list
X_max       2      -none- list
bino        5      -none- list
```

These do not seem to be running properly for me; will need to edit code to get graph to work properly.

9. Similar to question 4, examine whether plans just above the threshold values have different characteristics than contracts just below the threshold values. Use HMO and Part D status as your plan characteristics.

For this question, I have attempted to use the matchit funtion to create a mahalonobis distribution with plan_type and part_d from the previously defined three star ma df I created. However, for now, I cannot seem to get plan_type to be recognized as part of the dataset, so I am unable to also get the HMO as it is a characteristic (or should be) of plan_type.

10. Summarize your findings from 5-9. What is the effect of increasing a star rating on enrollments? Briefly explain your results.

Overall, there seems to be a positive correlation between higher star ratings and increased enrollment in Medicare Advantage plans. High star ratings tend to have more enrollees versus plans with low star ratings. The two graphs showing trends of increasing star ratings over time and increased enrollments supports this notion.