

Homework Four Submission 2

Safia Read

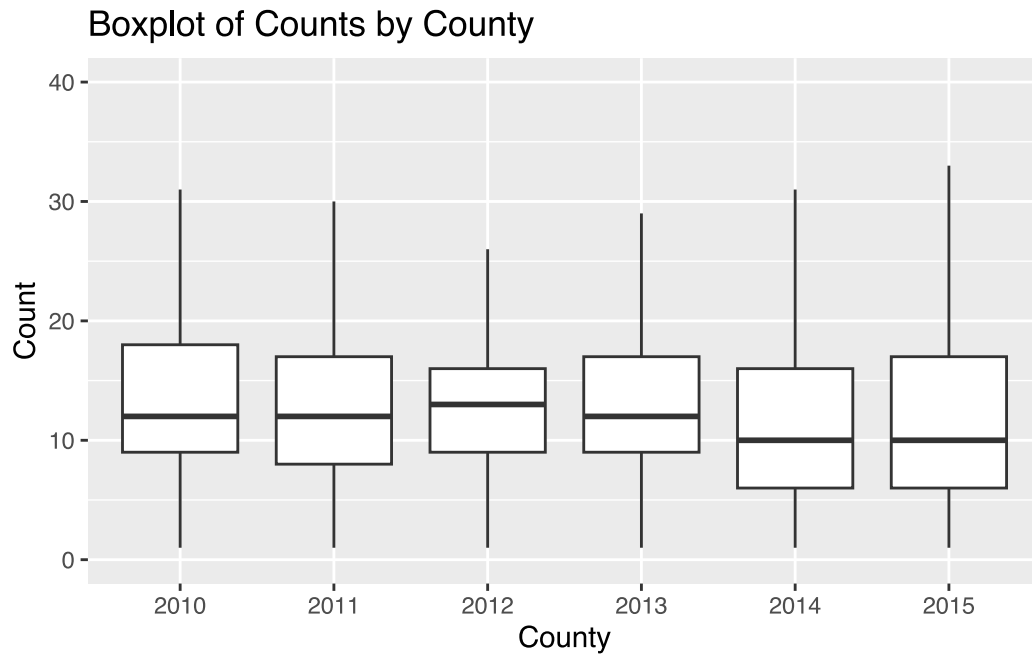
This is my second submission of the fourth homework for Econ 470.

Link to Github

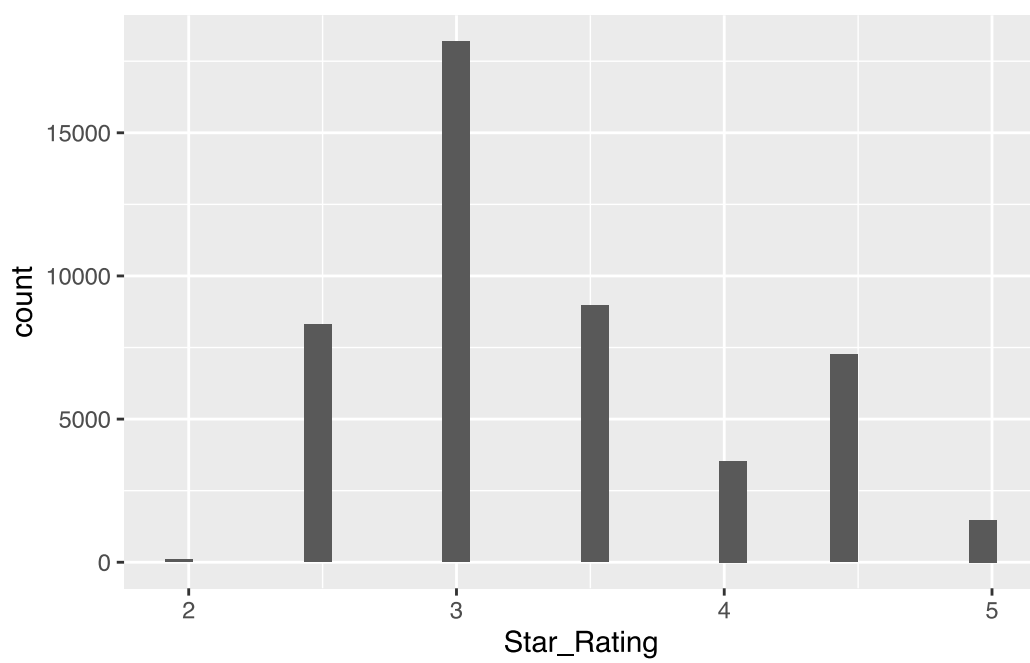
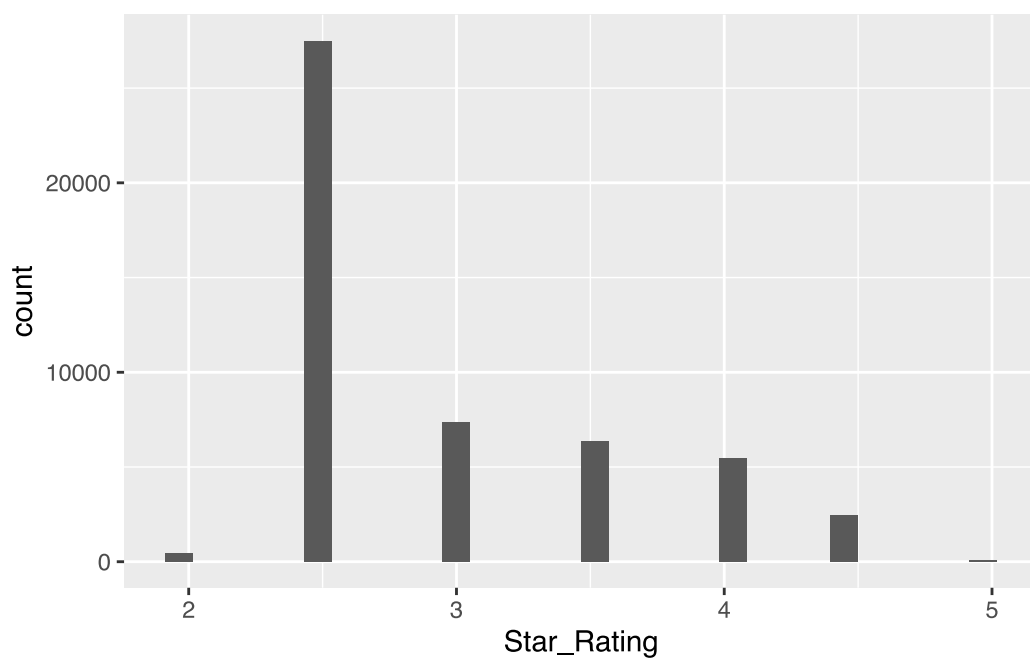
https://github.com/safiaread/homework_4

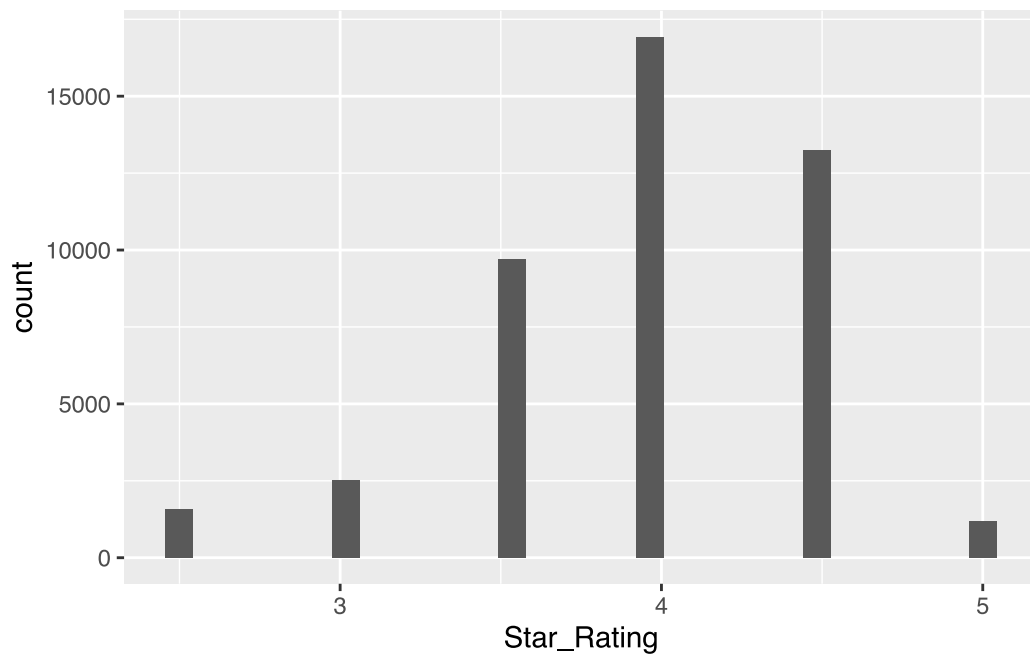
Question 1 It looks like the number of plans is pretty low.

Warning: Removed 672 rows containing non-finite values (``stat_boxplot()``).



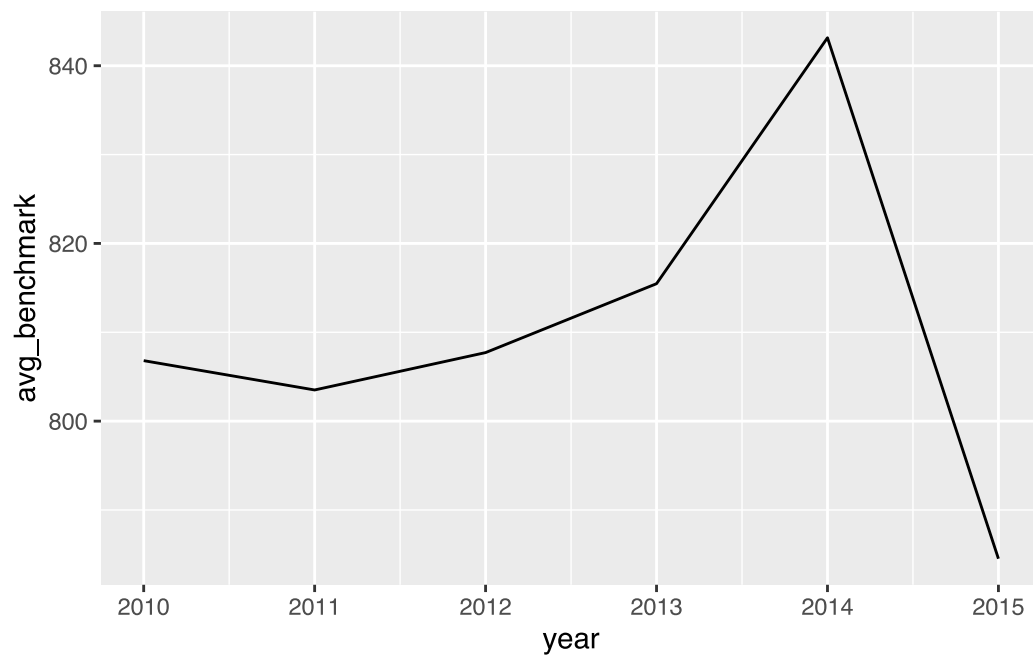
Question 2 The star ratings have increased generally.





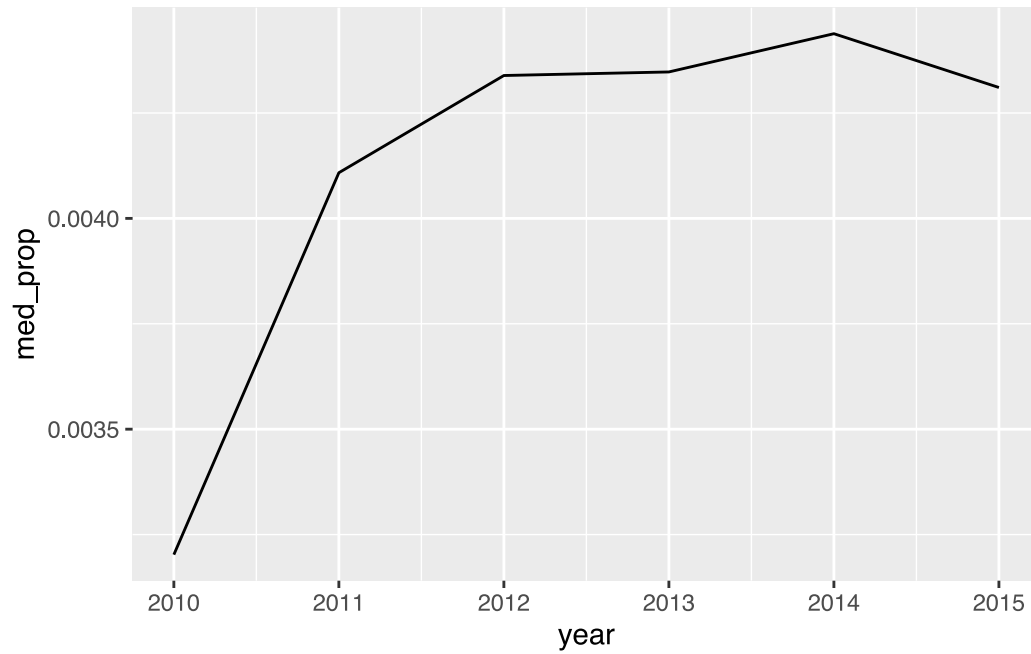
Question 3

The average benchmark has fallen according to this graph, but I think the variable calculation is off.



Question 4

Medicare Advantage has increased in popularity over time. I would expect that to correlate with increased benchmark payments.



Question 5

```
# A tibble: 1 × 5
  three three_five four four_five five
  <dbl>   <dbl> <dbl>   <dbl> <dbl>
1 26010   31771 18626   6636    0
```

Question 6

```
[1] 0.004178597
```

```
[1] 0.0004189378
```

Question 7

The data is sensitive to the bandwidth and changes the estimate.

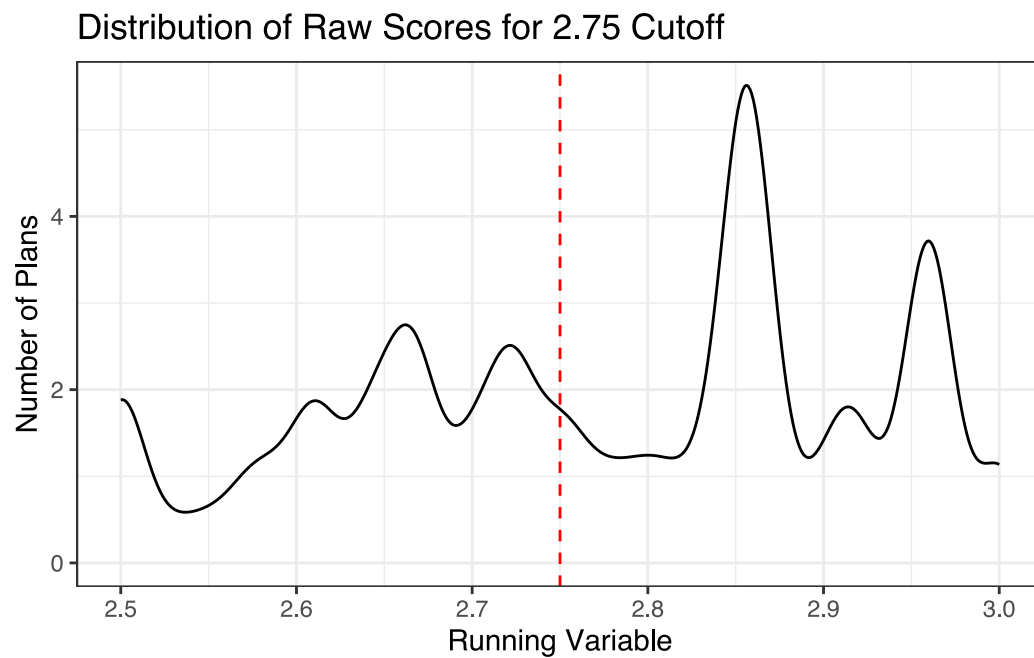
```
bandwidth estimates
1      0.1 0.002737363
2     0.12 0.003809948
3     0.13 0.005797619
4     0.14 0.005015187
5     0.15 0.004379916
```

	bandwidth	estimates
1	0.1	-0.0081092277
2	0.12	-0.0033249992
3	0.13	0.0002628067
4	0.14	-0.0007741078
5	0.15	-0.0007621881

Question 8

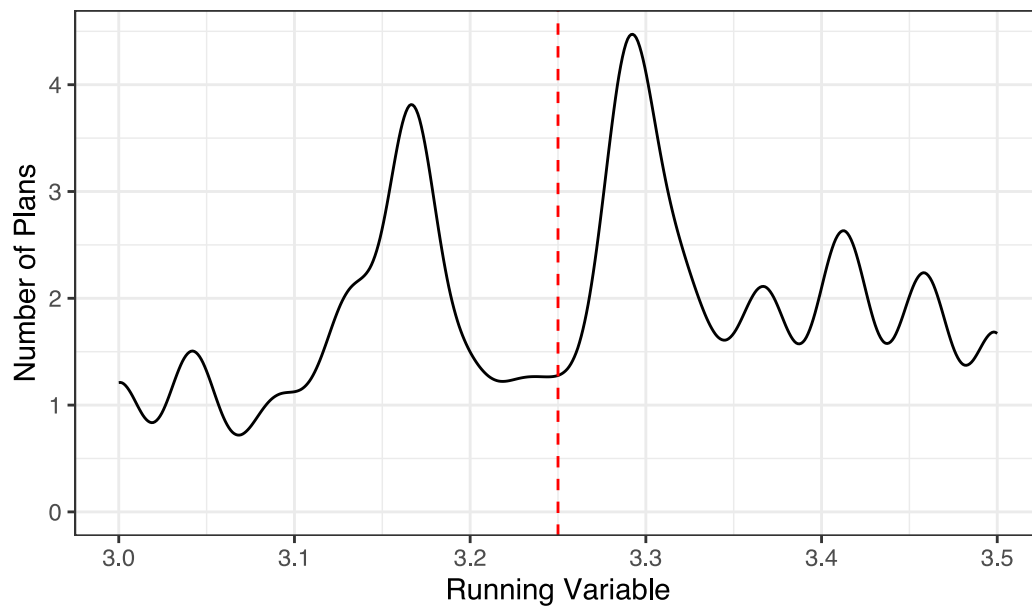
It looks like the data is clustered aaround the threshold.

Warning: Removed 202358 rows containing non-finite values (`stat_density()`).



Warning: Removed 207873 rows containing non-finite values (`stat_density()`).

Distribution of Raw Scores for 3.25 Cutoff



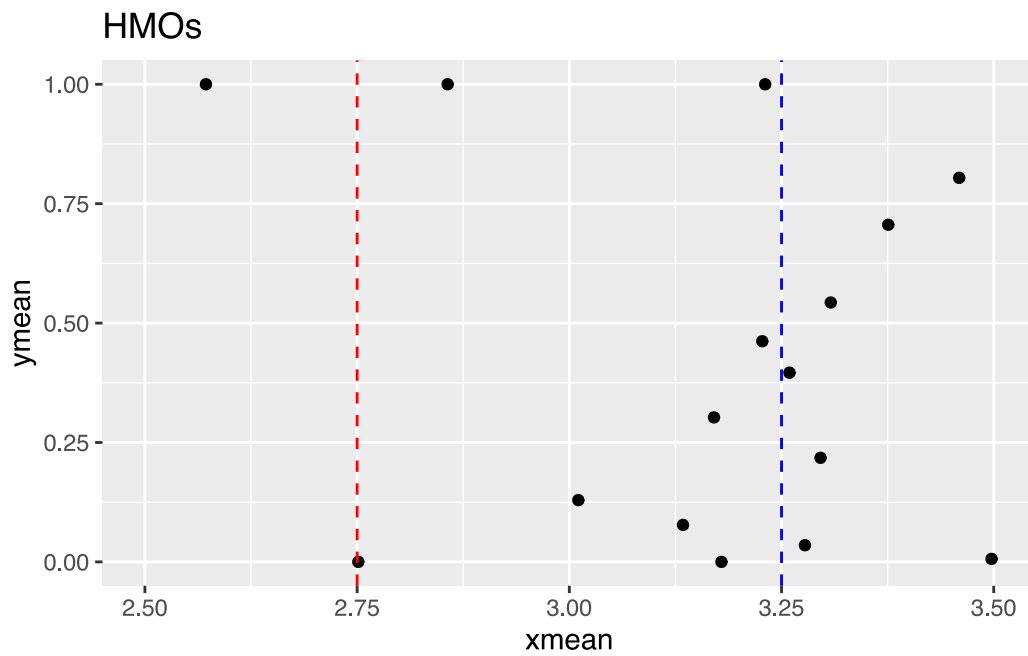
Question 9

From my binned scatterplots, we can see some covariate imbalance around the proportion of HMO plans, with a higher proportion of HMO's being above the 3.5 threshold. For the Part D plans, there seems to be a lower proportion of them past the 3.5 threshold. There isn't a lot of data around the 3 star threshold.

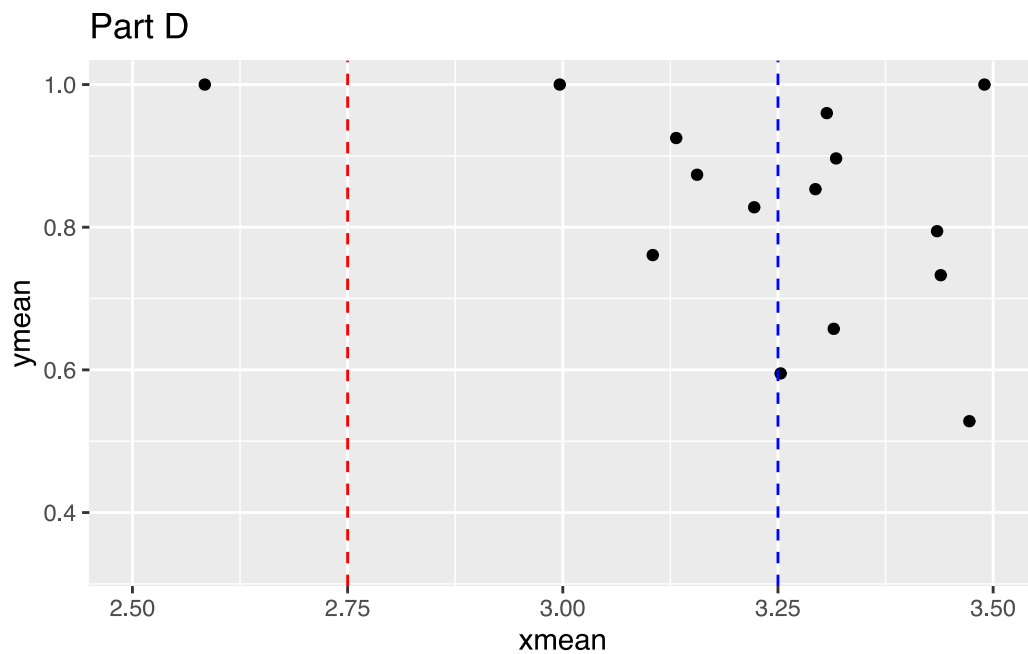
```
Warning: Unknown or uninitialised column: `dummy_partd`.
```

```
NULL
```

```
Warning: Removed 14 rows containing missing values (`geom_point()`).
```



Warning: Removed 5 rows containing missing values (``geom_point()``).



Question 10

Increasing star rating caused enrollments to decrease slightly over estimates with small bandwidths. I'm not sure why, maybe my estimates are off.