Homework 3 - https://github.com/Henryblake2777/Homework3

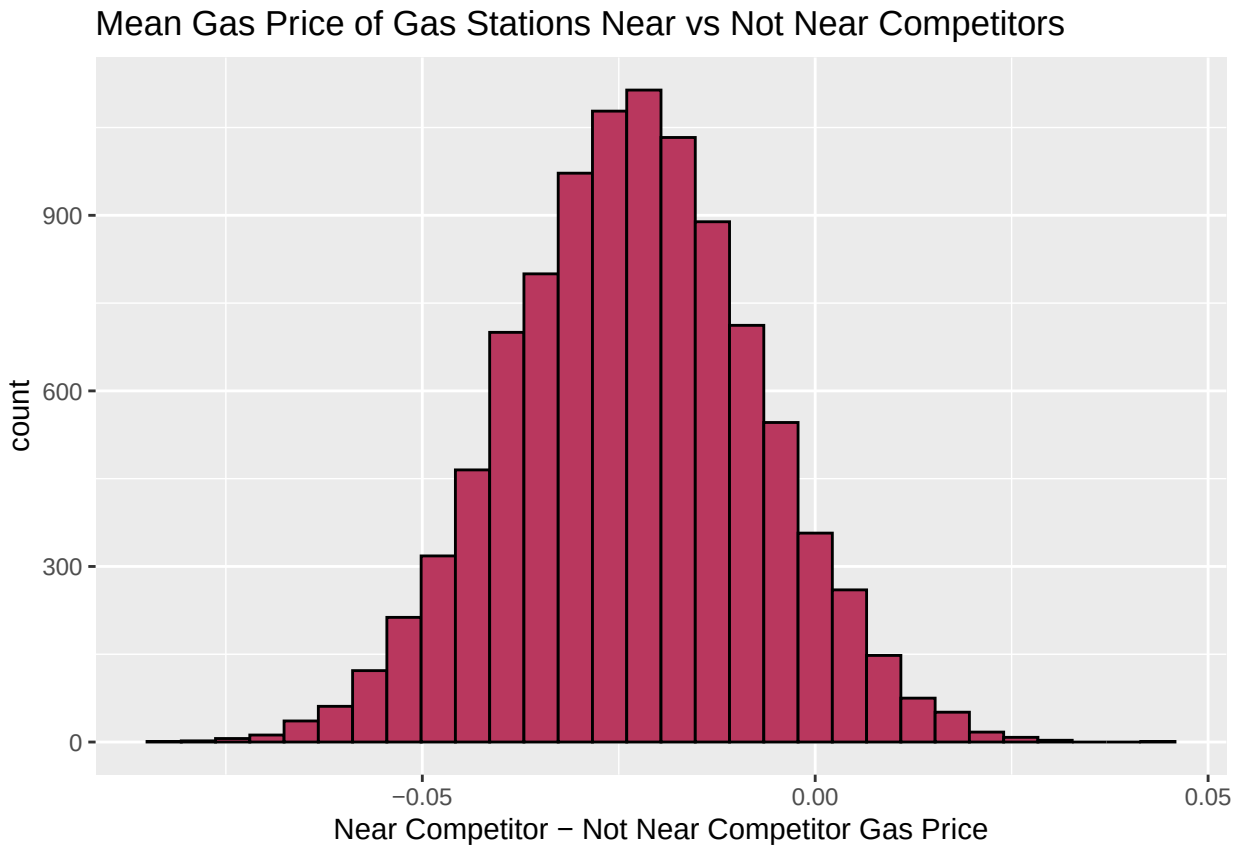
Henry Blake - hrb987

2025-02-07

Contents

Problem 1 - Gas

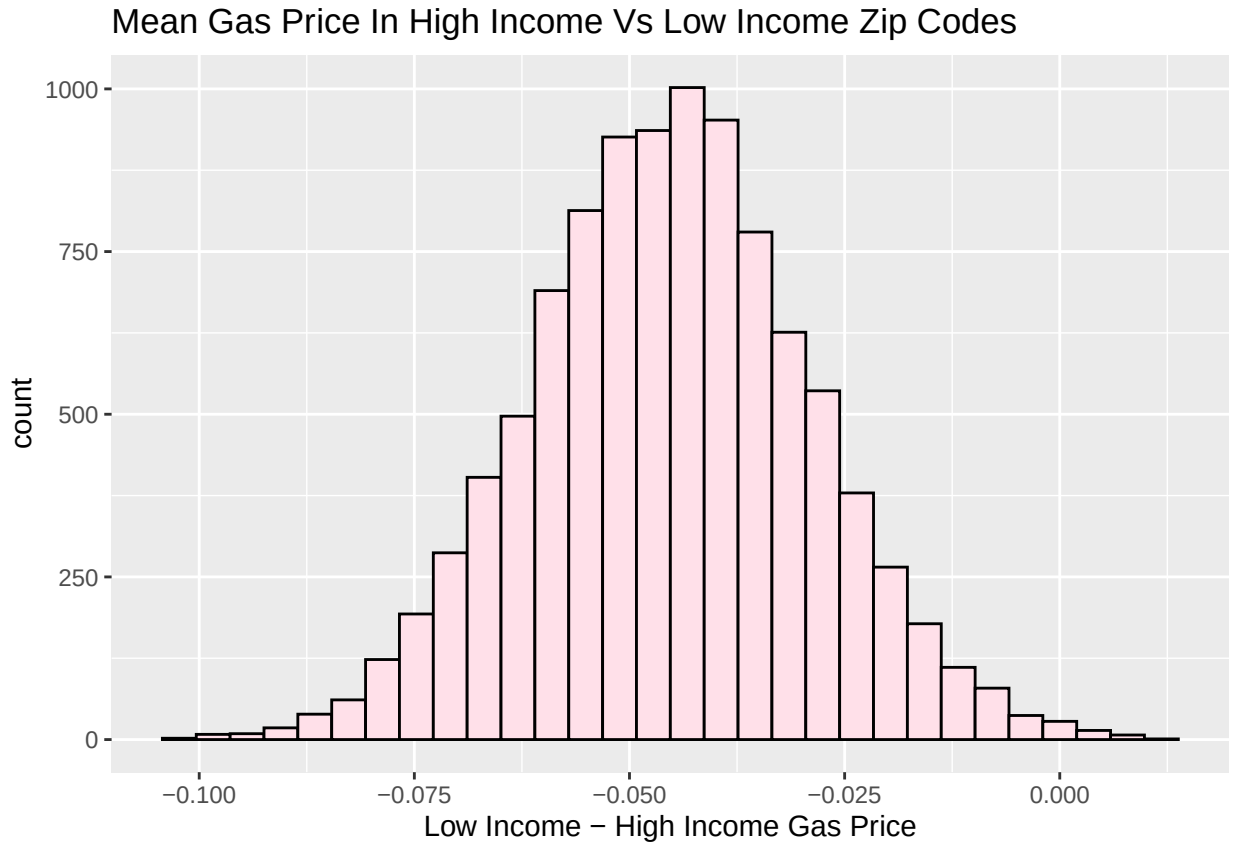
Theory A - Competition




```
##       name       lower       upper level     method    estimate
## 1 diffmean -0.05427803 0.007861687  0.95 percentile -0.02348235
```

I am 95 percent confident that the mean difference of gas price between gas stations with competitors and those without, is between -5.53 cents to 0.77 cents. Because zero is included in this range, I do not have significant evidence to say that the presence of competitors lowers gas prices.

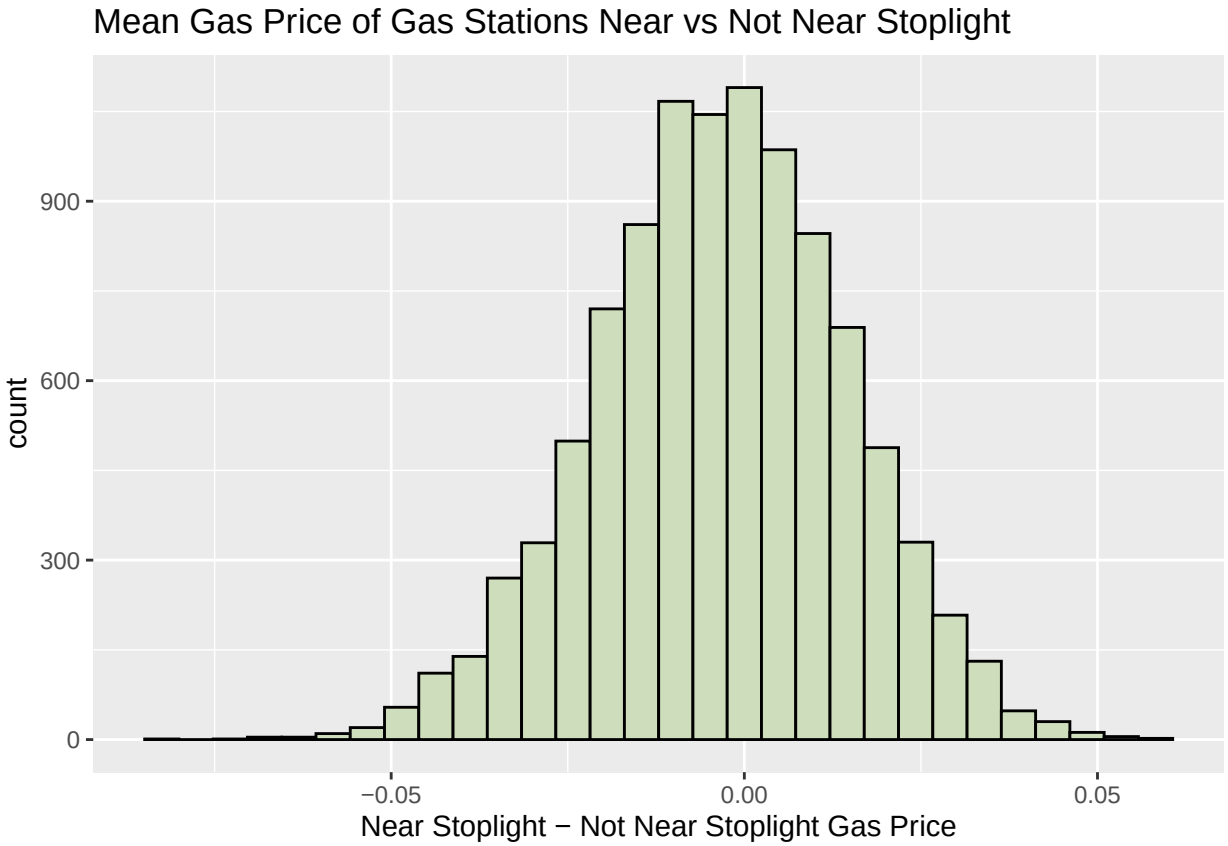
Theory B - Income




```
##       name       lower      upper level     method    estimate
## 1 diffmean -0.07700234 -0.0132654  0.95 percentile -0.04552049
```

I am 95 percent confident that the mean difference of gas price between low income and high income zip codes is between -7.72 cents to -1.43 cents. This is intriguing evidence to show that higher income zip codes charge more for gas. However, even more data will need to be found to further prove this conclusion.

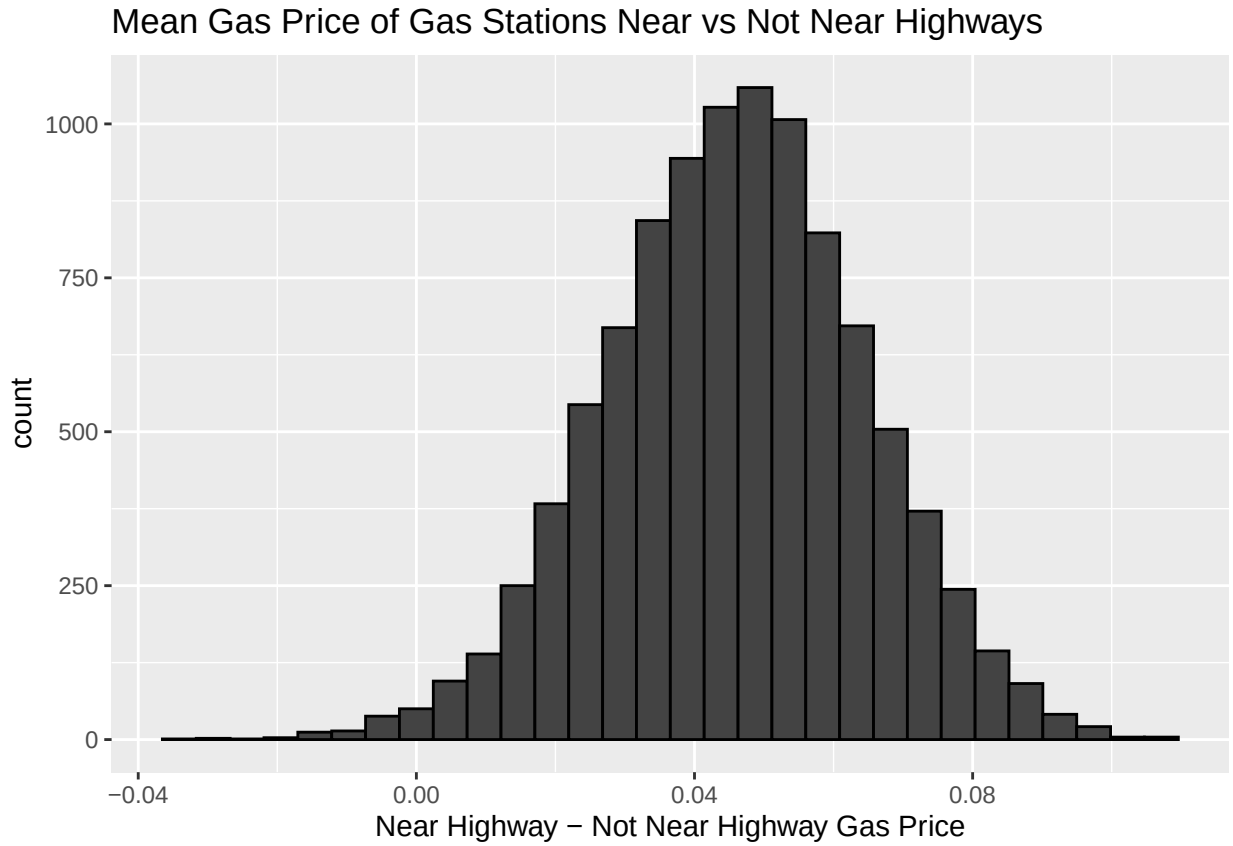
Theory C - Stoplights




```
##       name       lower      upper level     method     estimate
## 1 diffmean -0.03905899 0.03093117  0.95 percentile -0.003299916
```

I am 95 percent confident that the mean difference of gas price between gas stations with stoplights and those without, is between -3.84 cents to 3.04 cents. Because zero is included in this range, I do not have significant evidence to say that the presence of stoplights increases gas prices.

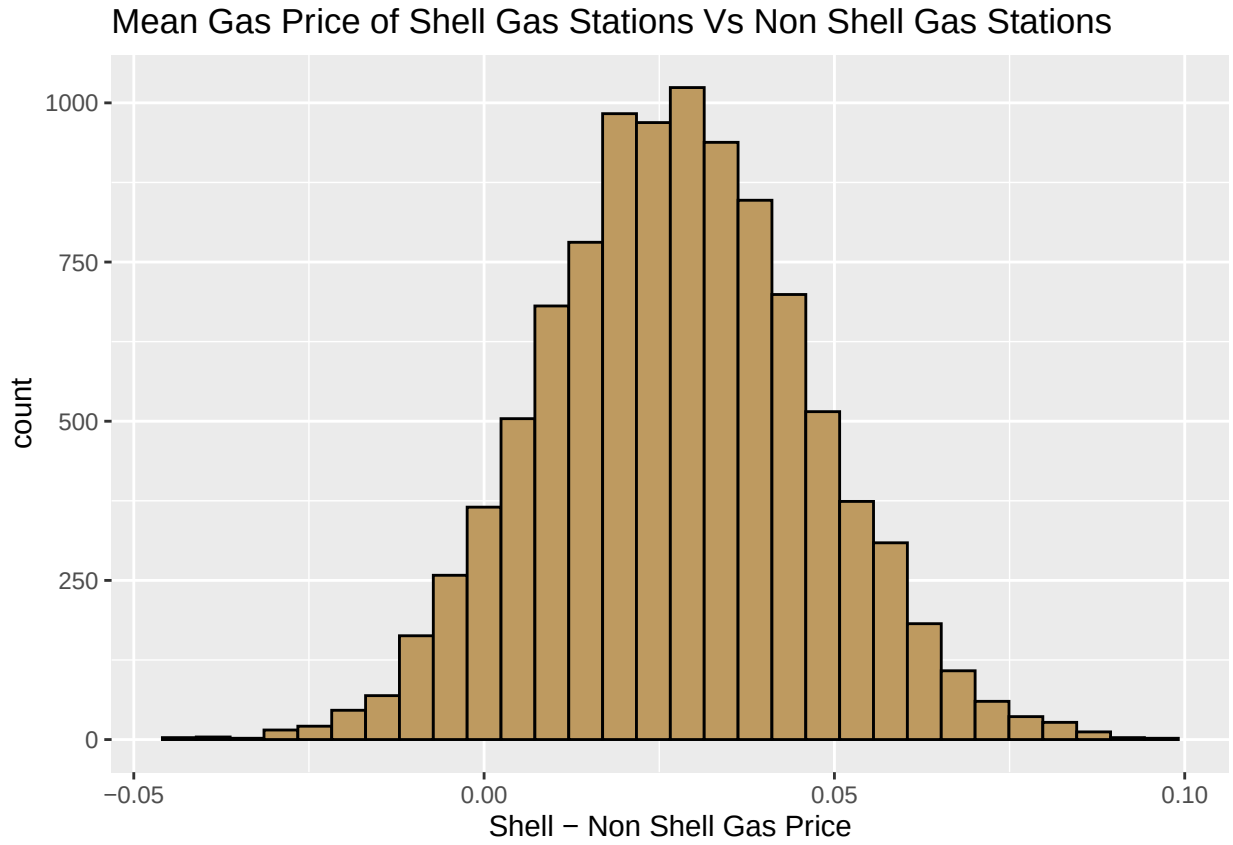
Theory D - Highways




```
##       name      lower      upper level     method  estimate
## 1 diffmean 0.00880108 0.08200244  0.95 percentile 0.0456962
```

I am 95 percent confident that the mean difference of gas price between gas stations with highway access to that without is between .79 cents to 8.11 cents. This is intriguing evidence to show that gas stations with highway access charge more for gas. However, even more data will need to be found to further prove this conclusion.

Theory E - Brands

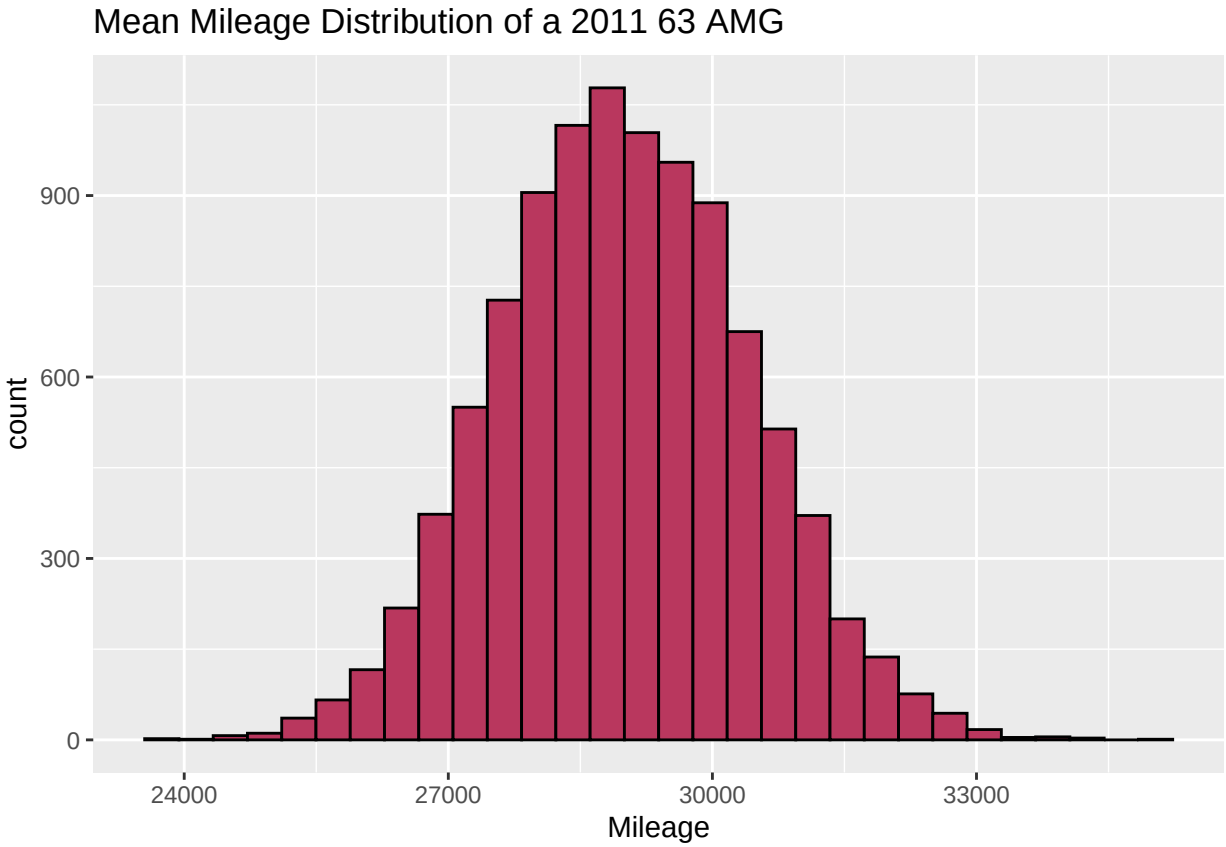



```
##       name        lower      upper level     method   estimate
## 1 diffmean -0.009297456 0.06506158  0.95 percentile 0.02740421
```

I am 95 percent confident that the mean difference of gas price between Shell stations and not Shell gas stations is between -1.02 cents to 6.55 cents. Because zero is included in this range, I do not have significant evidence to say that the presence of stoplights increases gas prices.

Problem 2 - Mercedes

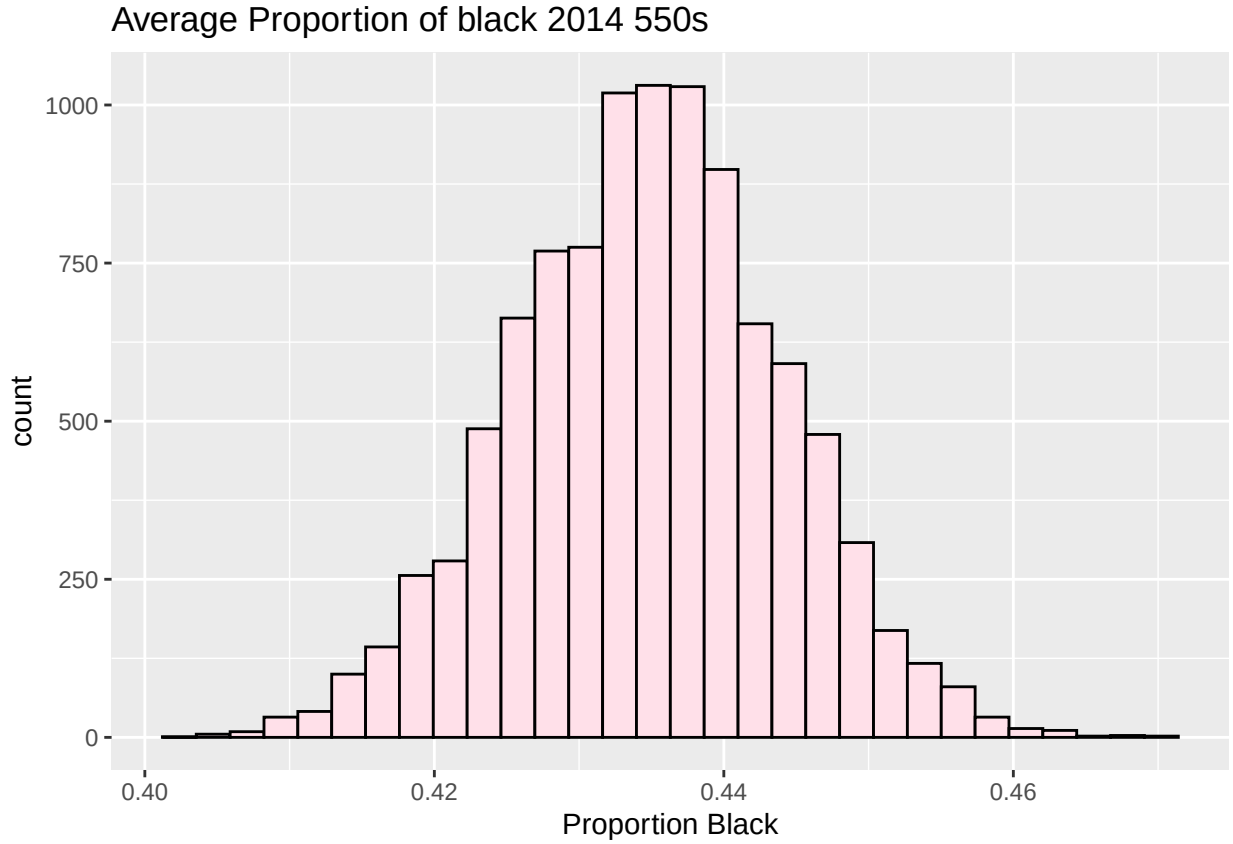
Part A - 63 AMG




```
##   name    lower    upper level     method estimate
## 1 mean 26301.58 31837.14  0.95 percentile 28997.34
```

I am 95 percent confident that the mean mileage of a 2011 63 AMG is between 26273 and 31814 miles.

Part B - 550




```
##         name     lower     upper level     method  estimate
## 1 prop_Black 0.4164071 0.4527518  0.95 percentile 0.4347525
```

I am 95 percent confident that the average proportion of 2014 550s that are black is between 41.64 percent and 45.31 percent

Problem 3 - TV

Part A - Ed or Earl?

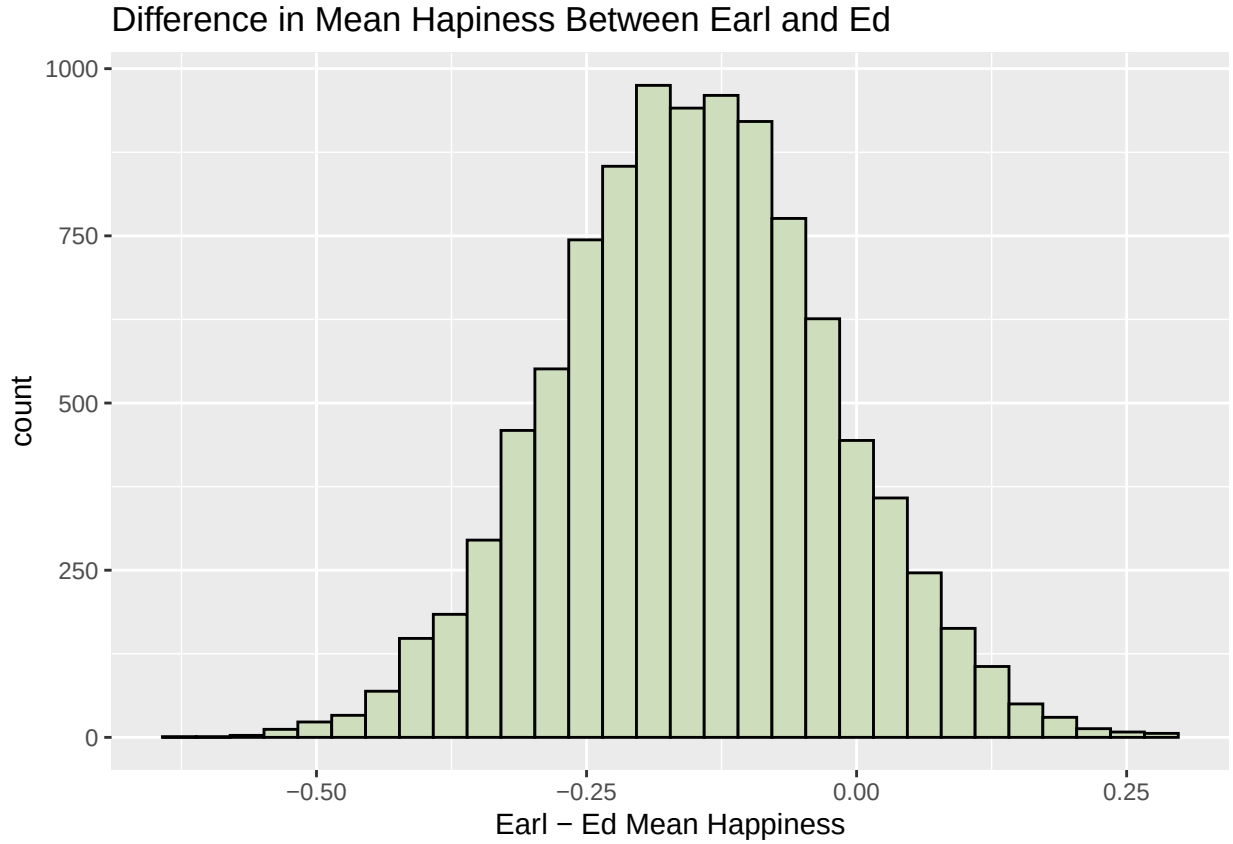
Question

Does “My Name is Earl” or “Living With Ed” make people happier?

Approach

I will use bootstrapping and a difference of means confidence interval

Results




```
##       name      lower     upper level     method   estimate
## 1 diffmean -0.3978529 0.1016288  0.95 percentile -0.1490515
```

Conclusion

I am 95 percent confident that they difference of mean happiness rating between "My Name is Earl" and "Living with Ed" is between -.398 and .095. Since 0 is in the interval I do not have significant evidence to say one show makes audience's happier than the other, however the data leans slightly towards Living with Ed and might be worth looking more into.

Part B - Loser or LA?

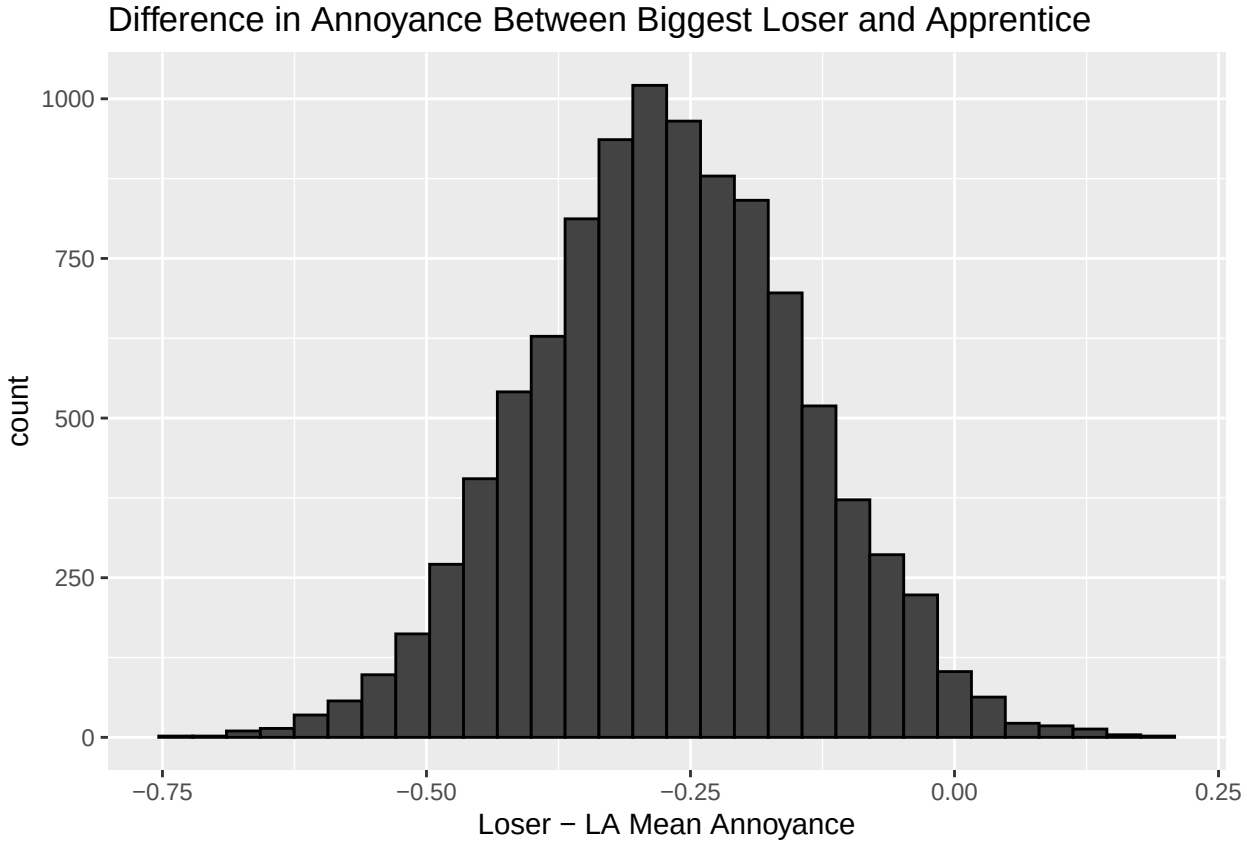
Question

Does "The Biggest Loser" or "My Name is Earl" make people more annoyed?

Approach

I will use bootstrapping and a difference of means confidence interval

Results




```
##       name      lower       upper level     method  estimate
## 1 diffmean -0.5224352 -0.02161089  0.95 percentile -0.270997
```

Conclusion

I am 95 percent confident that the difference in mean annoyance between "The Biggest Loser" and "The Apprentice: Los Angeles" is between -.523 and -.023. There is intriguing evidence to show that people find "The Apprentice: Los Angeles" more annoying, but more data may be needed.

Part C - Dancing with Confusion?

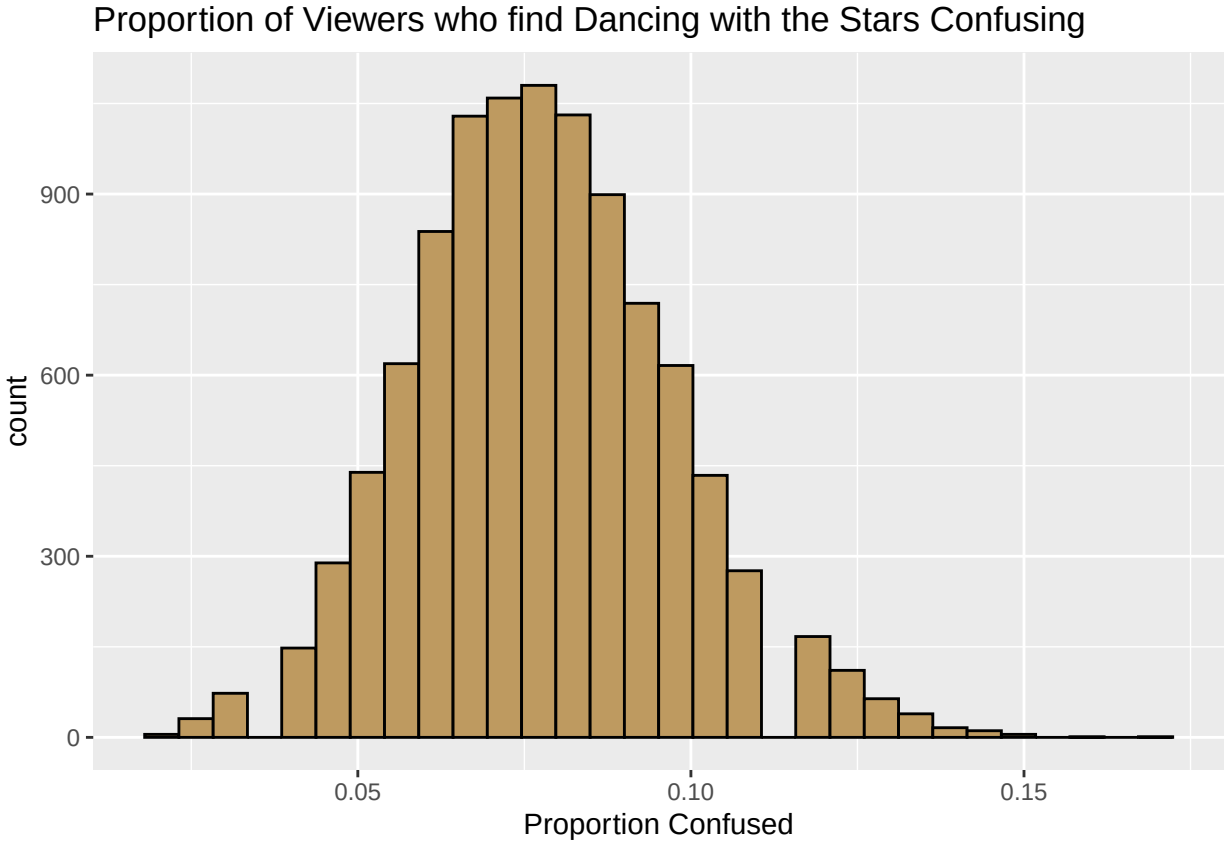
Question

What proportion of viewers find "Dancing with the Stars" confusing?

Approach

I will use bootstrapping and a proportion confidence interval

Results




```
##             name      lower     upper level     method   estimate
## 1 prop_Confusing 0.03867403 0.1160221  0.95 percentile 0.07734807
```

Conclusion

I am 95 percent confident that the average proportion of viewers who find "Dancing with the Stars" confusing is between 3.87 and 11.60 percent. This seems to be a pretty low percentage of people and indicates that the show isn't confusing.

Problem 4 - EBay

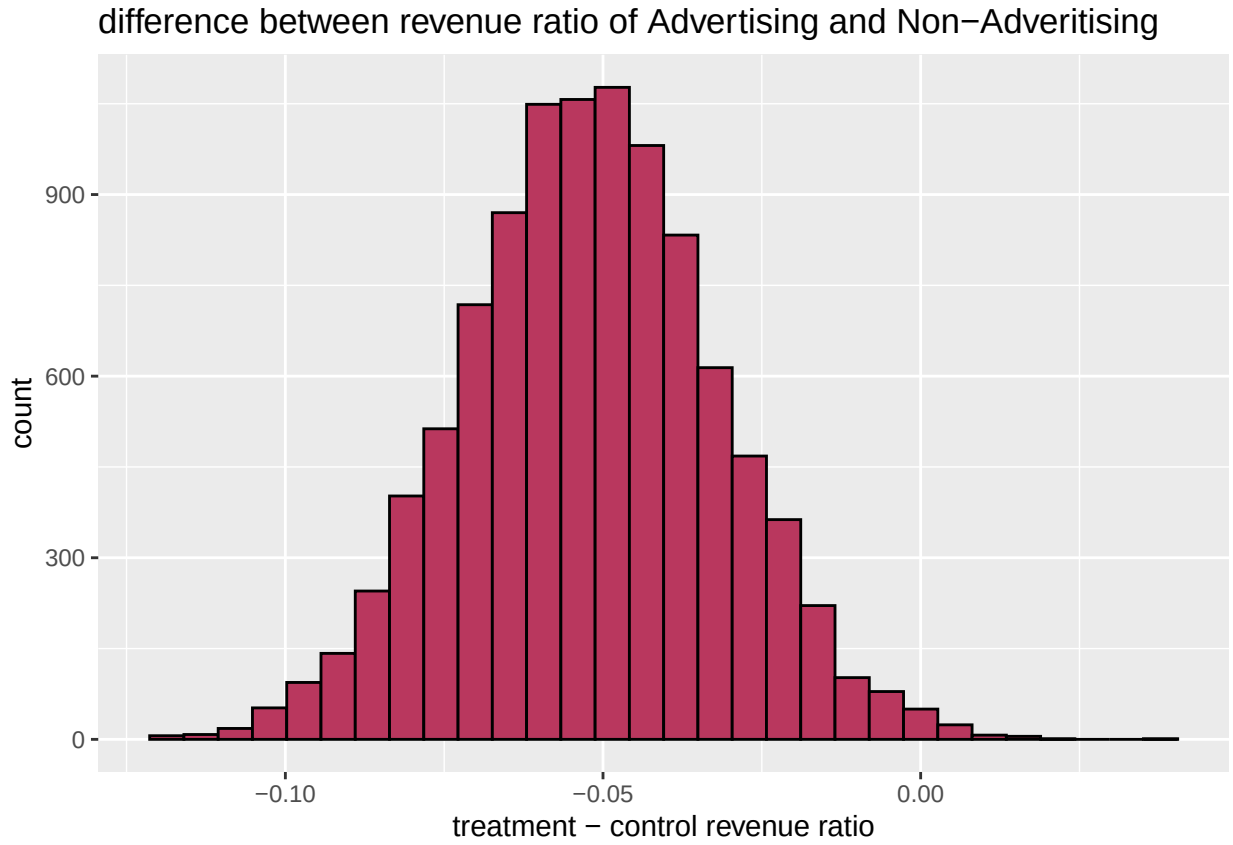
Question

Does Paid Advertisements on Google increase Ebay's revenue?

Approach

I will bootstrap and construct a difference of means confidence interval

Results




```
##       name       lower       upper level     method    estimate
## 1 diffmean -0.09074116 -0.01291888  0.95 percentile -0.05228145
```

Conclusion

I am 95 percent confident that the difference between the mean revenue ratio between the treatment group and the control group is between -.0921 and -.0144. There is intriguing evidence to show that paid advertisements increase Ebay's revenue, but more data will need to be collected to be even more certain