HW3

Seraphym Ignacio

2025-02-13

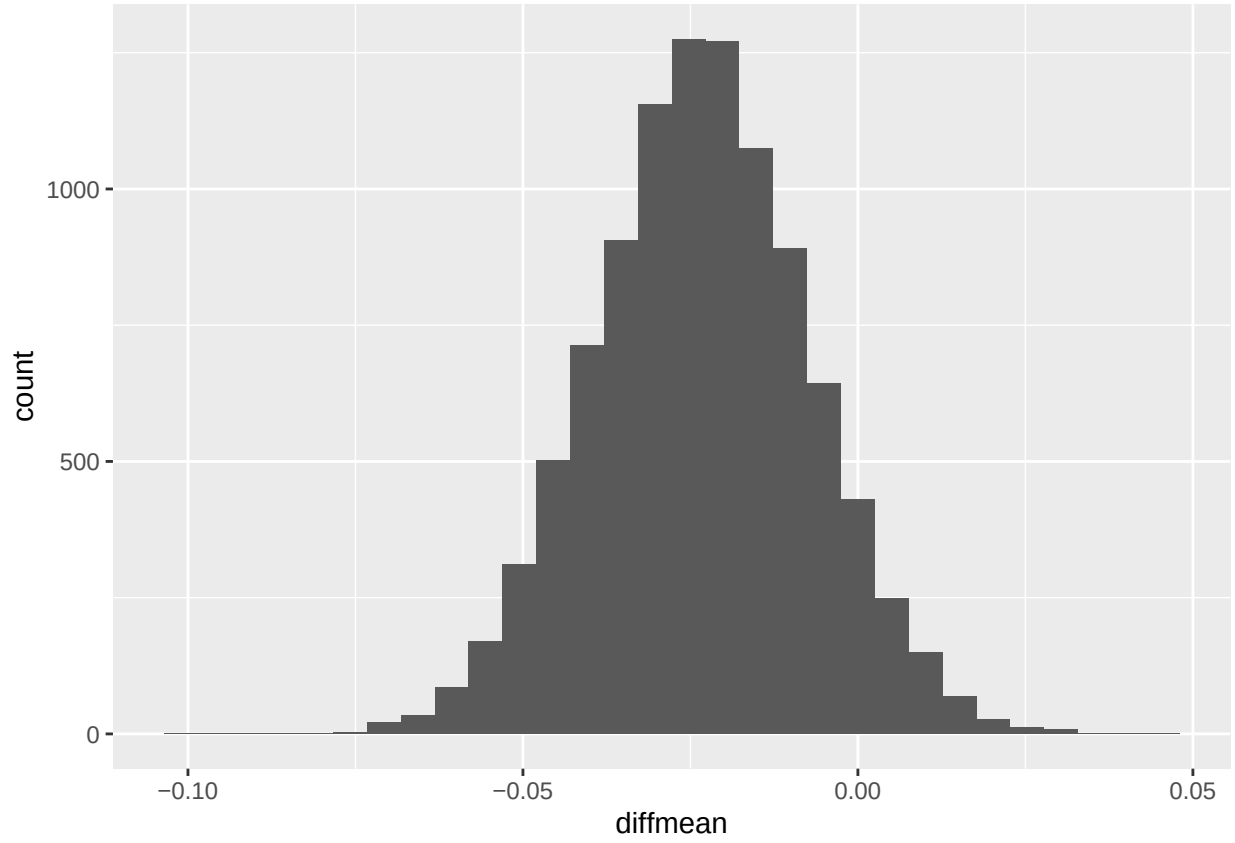
ID: sri346

GitHub Link:

Problem 1

```
##        N        Y
## 1.875882 1.852400
```

```
##    diffmean
## -0.02348235
```

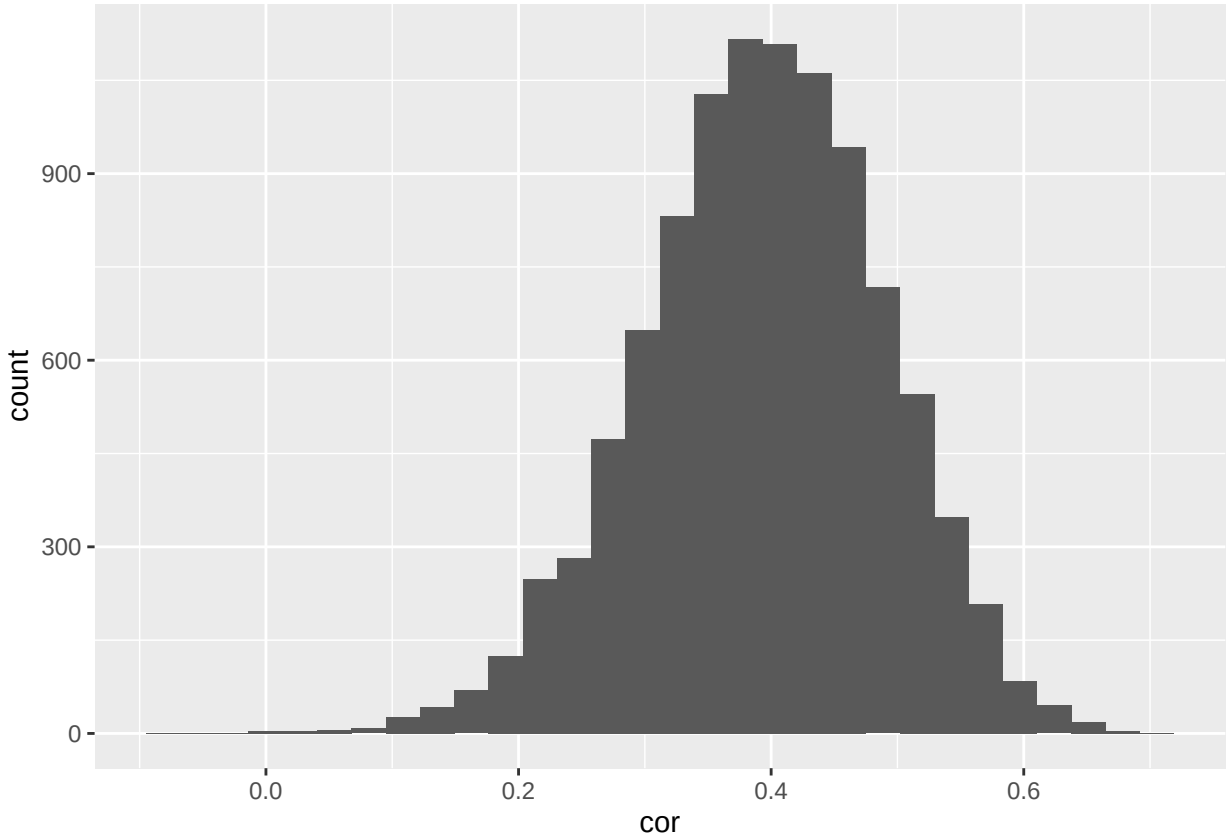


```
##       name       lower       upper level     method     estimate
## 1 diffmean -0.05467793 0.007911749  0.95 percentile -0.006611505
```

Theory A:

- Claim: Gas stations charge more if they lack direct competition in sight.
- Evidence: Gas stations with no competitors have prices $0.023 cents higher than gas stations with competitors. Compared to a wider population, this data was taken from 101 gas stations, so there is some uncertainty with these observations.
- Conclusion: With a 95% confidence, and ruling out a difference of zero, it can be stated that gas stations with no competitors do not have significantly higher prices than ones with competitors on average, with a difference of $0.007-$0.06 in mean values, which does not support Theory A.

```
## [1] 0.3961546
```



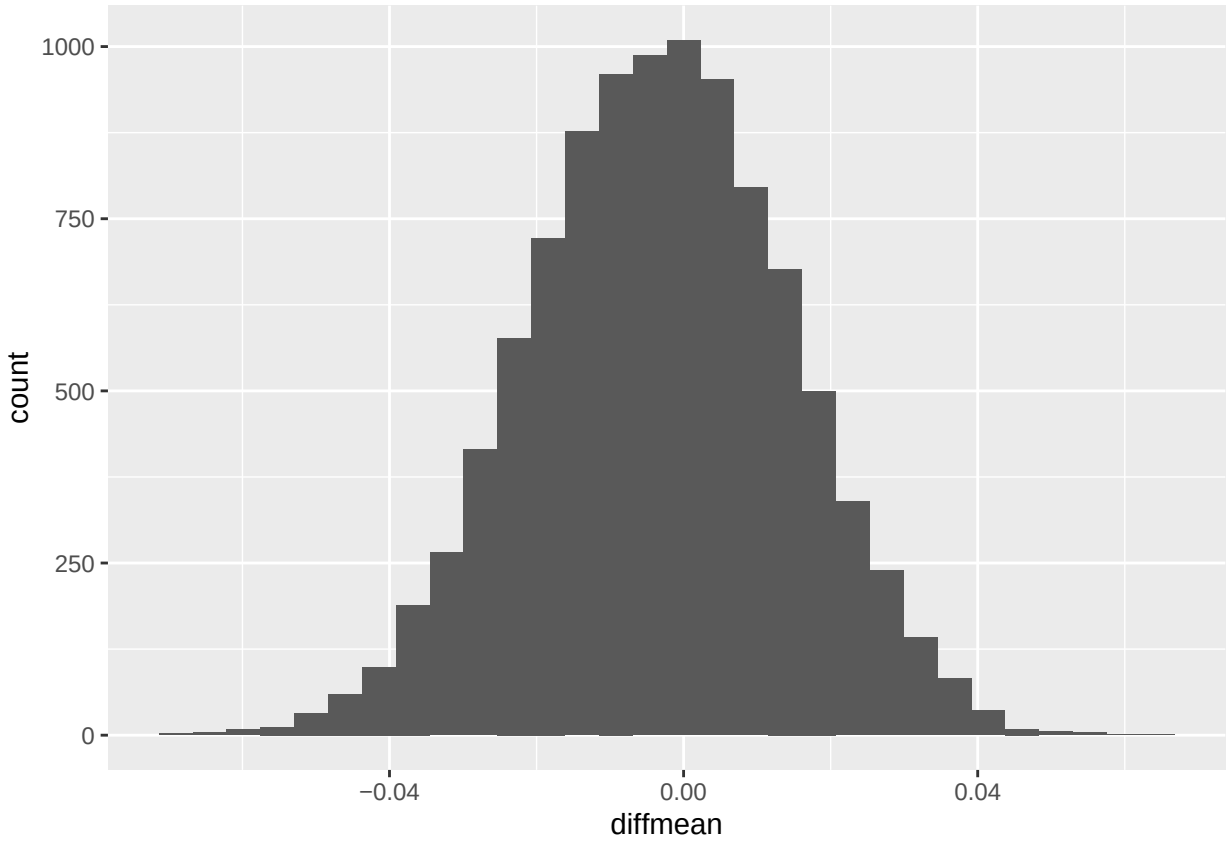
Theory B:

- Claim: The richer the area, the higher the gas prices.
- Evidence: The correlation between gas prices and income is 0.40, which represents a moderately positive relationship between the two observed variables. Again, all observations are said with some uncertainty as there were only 101 out of the many gas stations in the population.

- Conclusion: With 95% confidence and ruling out a difference of zero, there is a correlation of 0.57-0.20 between gas prices and area income. This positive relationship suggests that higher-income areas tend to have higher gas prices and supports Theory B.

```
##        N        Y
## 1.866316 1.863016
```

```
##     diffmean
## -0.003299916
```




```
##       name       lower      upper level     method   estimate
## 1 diffmean -0.03815223 0.03074049  0.95 percentile 0.01315447
```

Theory C:

- Claim: Gas stations at stoplights charge more.
- Evidence: Gas stations with no stoplights have prices of $0.003 higher than gas stations with stoplights. This observation comes with some uncertainty due to taking a small part of the population.
- Conclusion: This data does not rule out a difference of zero, and so it can be verified that the data supports Theory A and the average price between these types of gas stations are not statistically significant.

Theory D:

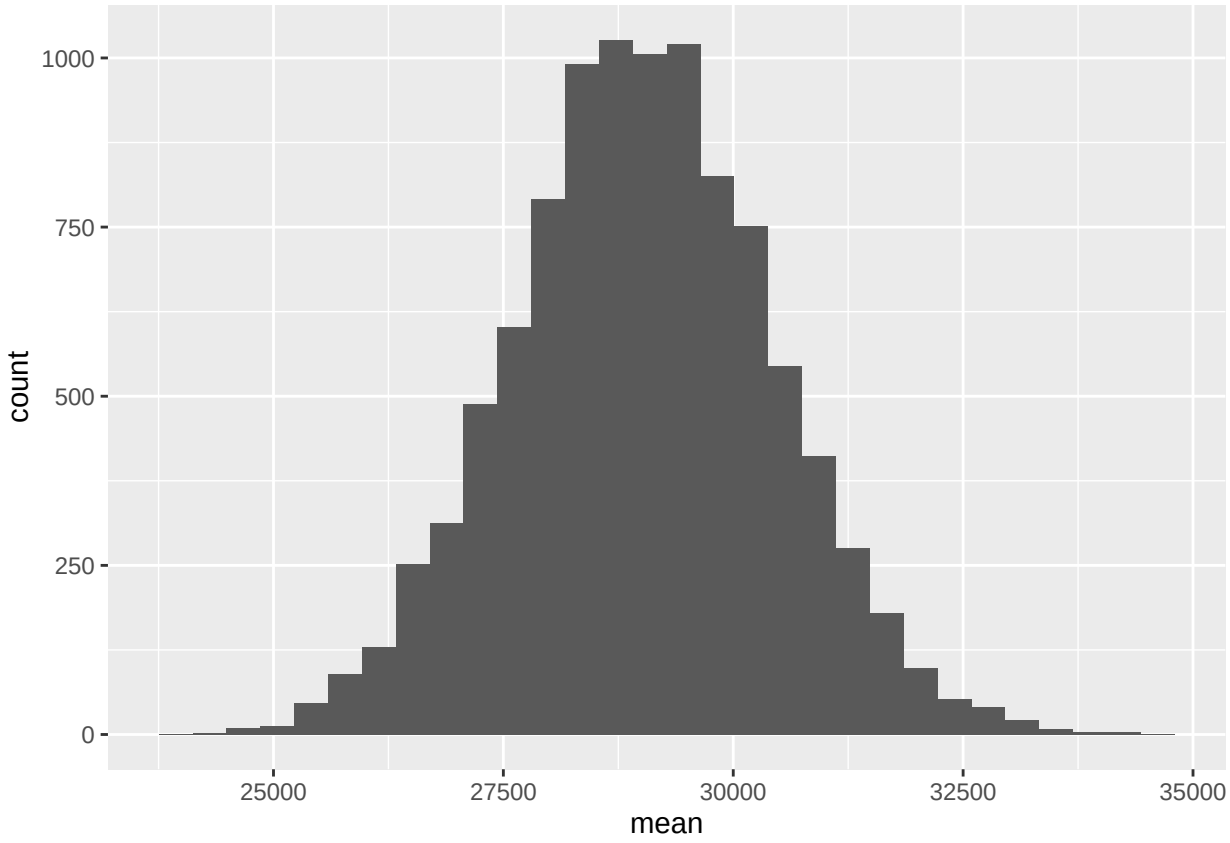
- Claim: Gas stations with direct highway access charge more.
- Evidence:
- Conclusion:

Theory E:

- Claim: Shell charges more than all other non-Shell brands.
- Evidence:
- Conclusion:

Problem 2

```
##   std_err_
## 1 1418.602
```

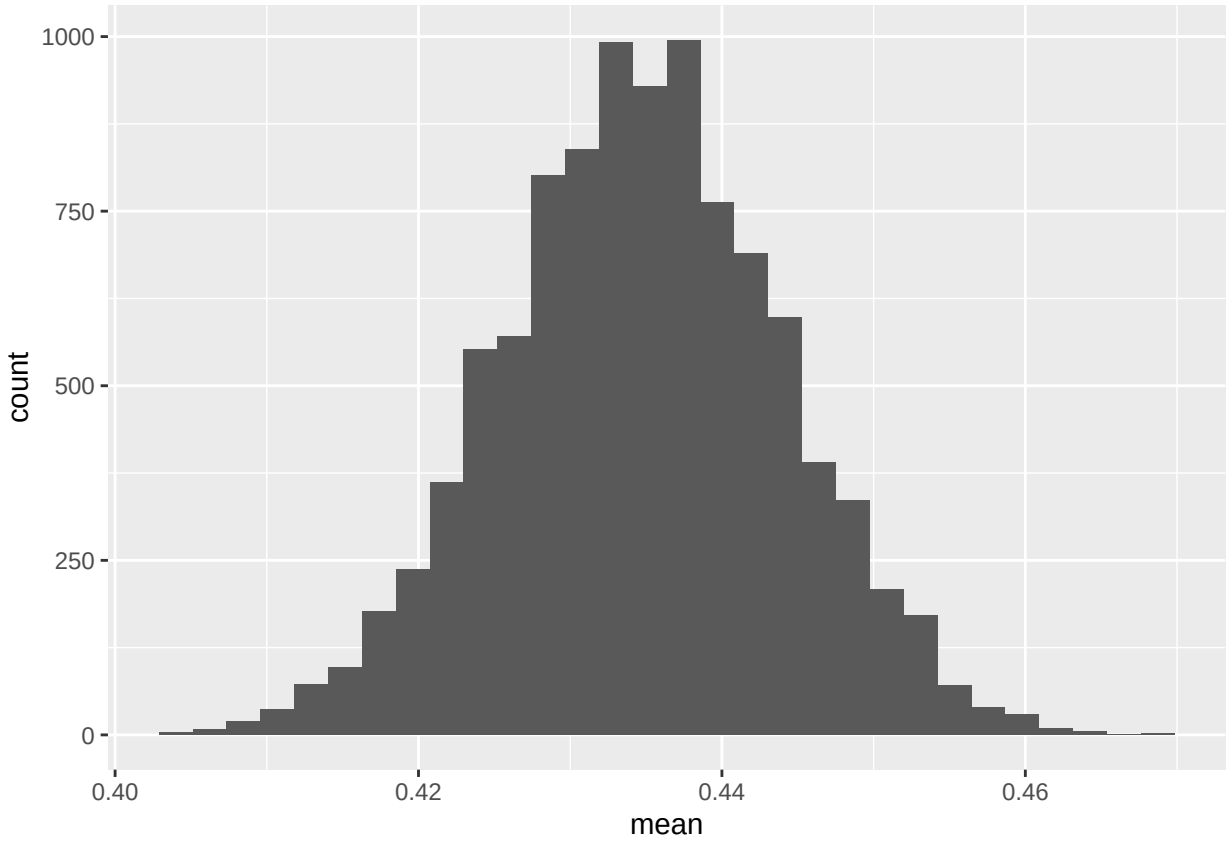



```
##   name    lower    upper level     method estimate
## 1 mean 26218.97 31789.13  0.95 percentile 27605.85
```

Part A

Based on these 116 cars from a 30,000-car dataset, it can be stated with 95% confidence that the average mileage of the 2011 S-Class 63 AMGs is 26.5 thousand. The true average mileage would be expected to fall between approximately 26 thousand and 32 thousand.

```
##      std_err_
## 1 0.009251666
```



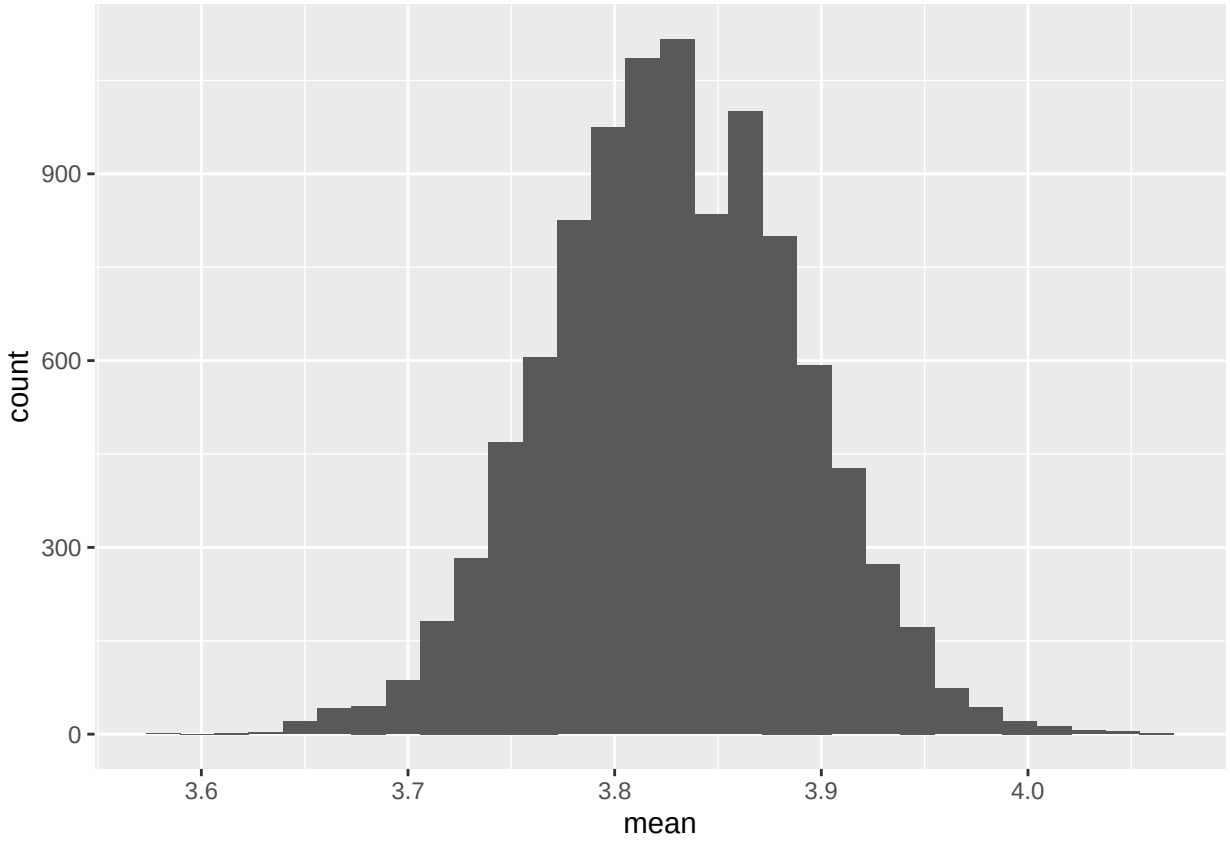

```
##   name     lower     upper level     method estimate
## 1 mean 0.4164071 0.4527518  0.95 percentile 0.423676
```

Part B

Based on these 2889 cars from a 30,000-car dataset, it can be stated with 95% confidence that the average amount of cars painted black of the 2014 S-Class 550s is 0.43. The true average amount would be expected to fall between approximately 0.42 and 0.45.

Problem 3

```
##     std_err_
## 1 0.06033505
```

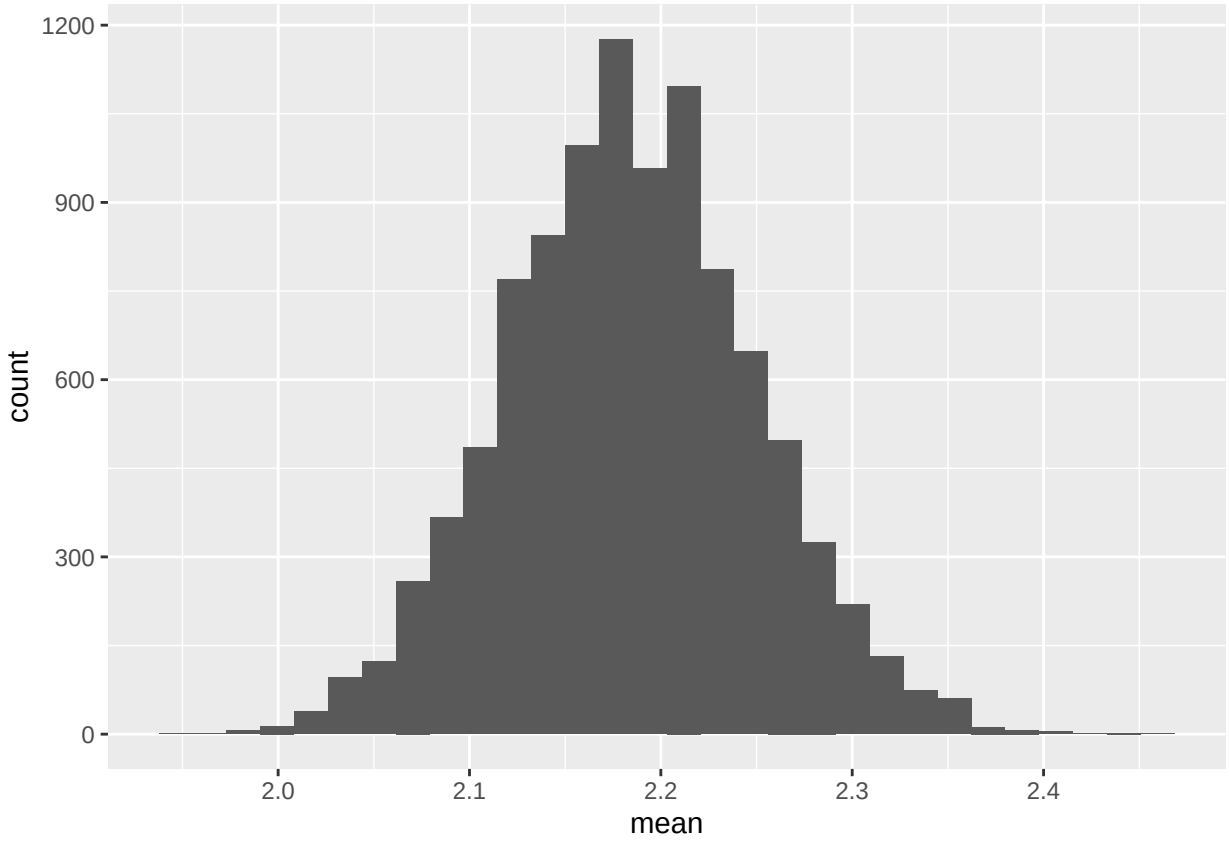



```
##   name    lower    upper level     method estimate
## 1 mean 3.710638 3.944681  0.95 percentile 3.787234
```

Part A

1. Question: Who makes people happier: Ed or earl?
2. Approach: To answer this question, I used viewer responses from the dataset for two shows called "Living with Ed" and "My Name is Earl". Then,
3. Results:
4. Conclusion:

```
##     std_err_
## 1 0.06502972
```

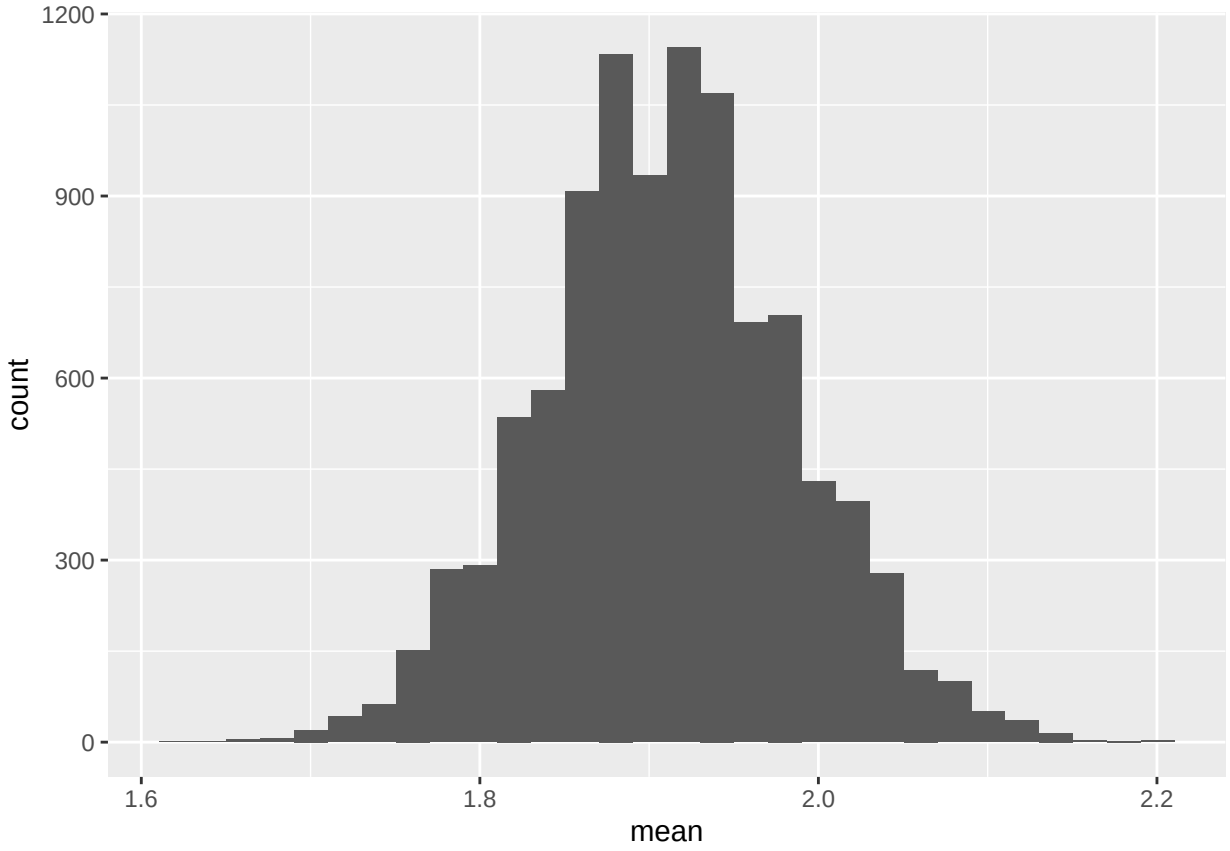



```
##   name    lower    upper level     method estimate
## 1 mean 2.059211 2.315789  0.95 percentile 2.233553
```

Part B

1. Question: Which reality contest made people feel more annoyed: "The Biggest Loser" or "The Apprentice: Los Angeles"
2. Approach: To answer this question, I used viewer responses from the dataset for two shows called "The Biggest Loser" and "The Apprentice: Los Angeles". Then,
3. Results:
4. Conclusion:

```
##     std_err_
## 1 0.07527388
```




```
##   name    lower    upper level     method estimate
## 1 mean 1.767956 2.060773  0.95 percentile 1.933702
```

Part C

1. Question: Is the premise of "Dancing with the Stars" confusing?
2. Approach: To answer this question, I used viewer responses from the dataset for the show "Dancing with the Stars". Then,
3. Results:
4. Conclusion:

Problem 4

1. Question:
2. Approach:
3. Results:
4. Conclusion: