

Who is Smarter Preliminary Analysis

Peter Williams

4/21/2020

The Data

The survey data used in this report was gathered from students during the Spring semester of 2019 at Utah Valley University.

Cleaning the Data

Exporting data from qualtrics results in a lot of useless information (such as geographical location) that needs to be removed. Also, because the survey questions are not “forced response”, a lot of elements are missing from the data set. The unnecessary variables were removed as well as the NAs with the following scripts:

- “survey_tidying_script.R”
- “removing_missing_values_SP2019.R”

Analyzing the data

Because we don’t have data about the GPA yet, I ran a chi-squared (two-variable), to determine if there was a relationship between gender and perceived intelligence (measured with a factored question “who is smarter, you or your group member?”. The results of the chi-squared and the fisher’s exactness tests are as follows:

```
##
##      Cell Contents
## |-----|
## |              Count |
## |      Expected Values |
## |      Row Percent |
## |      Column Percent |
## |      Std Residual |
## |-----|
##
## Total Observations in Table:  524
##
##      | df$Who_is_smarter.
## df$gender2 |      1 |      2 | Row Total |
## -----|-----|-----|-----|
##      F |      79 |      153 |      232 |
##      |  95.191 |  136.809 |      |
##      |  34.052% |  65.948% |  44.275% |
##      |  36.744% |  49.515% |      |
##      |   -1.659 |    1.384 |      |
## -----|-----|-----|-----|
##      M |      136 |      156 |      292 |
##      |  119.809 |  172.191 |      |
##      |  46.575% |  53.425% |  55.725% |
##      |  63.256% |  50.485% |      |
##      |    1.479 |   -1.234 |      |
## -----|-----|-----|-----|
## Column Total |      215 |      309 |      524 |
```

```

##           | 41.031% | 58.969% |           |
## -----|-----|-----|-----|
##
##
## Statistics for All Table Factors
##
##
## Pearson's Chi-squared test
## -----
## Chi^2 = 8.380402    d.f. = 1    p = 0.003792886
##
## Pearson's Chi-squared test with Yates' continuity correction
## -----
## Chi^2 = 7.870792    d.f. = 1    p = 0.005023963
##
##
## Fisher's Exact Test for Count Data
## -----
## Sample estimate odds ratio: 0.5928807
##
## Alternative hypothesis: true odds ratio is not equal to 1
## p = 0.004225989
## 95% confidence interval: 0.4081039 0.8582259
##
## Alternative hypothesis: true odds ratio is less than 1
## p = 0.002450615
## 95% confidence interval: 0 0.8110263
##
## Alternative hypothesis: true odds ratio is greater than 1
## p = 0.9986221
## 95% confidence interval: 0.4323979 Inf
##
##
##
## Minimum expected frequency: 95.19084

```

After examining the expected values in each column, we can see that the first assumption of a chi-squared test was met (expected values > 5 , our lowest was ~95). With the assumption satisfied, the next point to examine is the chi-squared p-value (0.003792886), which indicates a p-value < 0.05 , allowing us to assume there is some kind of relationship between gender and perceived intelligence. The Fisher's exact test confirms this, with a $p < 0.001$. It would appear from the contingency table that females (34%) are more likely to perceive themselves as less smart than males (46%).

Next Step

Now that a relationship has been established between gender and perceived intelligence, I would assume the next step would be to compare the genders against the grade in the class and overall GPA. This would give us more insight into whether the perceptions were validated or if they possibly had some affect on the outcome of their grades. I am not sure what the actual article examined, I still haven't been able to find it.