Levi Watson

STAT 480

02/23/2023

Wear What You Like – Statistical Analysis

1. Introduction

The goal of this study was to investigate whether a teacher's interaction with students varied based on the clothing choice of the student.

1.1 Study Design:

The sample used in this experiment was a sample of 231 students that were taught by a certain professor over the course of a few classes. These students were classified into three groups based on clothing type: Unisex, Standard, and Other. The researcher counted the number of positive and negative interactions each student had with the professor, their clothing type, as well as recording the student's gender.

1.2 Variables:

Quantitative Variables:

- count: (Number of Interactions), range of 0-14

Categorical Variables:

- clothing: Unisex/Std/Other
- gender: female/male
- itype: PosInt/Reproof

2. Methodology

The analyses for this data were performed using R, implementing various packages. Three main groups of tests were performed: Exploratory Data Analysis (EDA), Chi-Squared Goodness-of-Fit Tests, and Chi-Squared Tests of Independence.

2.1 Exploratory Analysis:

EDA was performed on each of the variables. A histogram was used to visualize the numeric variable (scale), while bar charts were used to visualize the frequencies of the categorical variables (clothing, gender, and itype).

2.2 Goodness of Fit Tests:

Goodness of Fit tests were used to see if the total number of interactions had with the professor differed from the number we'd expect based on the 3 categorical variables separately. The number of interactions (based on one of the categorical variables at a time) were compared against the number of interactions that would be given if there were no difference on number of interactions between the groups.

2.3 Tests of Independence:

Tests of independence were used to explore if any of the categorical variables changed the proportion of interactions with regard to another variable. We tested if interaction type is affected by gender or clothing,

as well as see if different clothing types are worn by different genders more.

3. Results

The tables and graphs indicating the results of all the tests are included in section 5 – Tables/Graphs.

The EDA frequency histograms/graphs (Section 5.1) show that we have slightly more males than females in our sample, and that Other is the most common clothing type, followed by Standard, and then Unisex.

In the Goodness-of-Fit tests, it was found that, first, gender has a significant effect on total number of interactions. We compared the actual counts to the hypothesized counts (that we would have if there was no effect), and found that the difference in distribution is statistically significant (see 5.2.1). From the bar chart, it seems that males get significantly more interactions than we'd expect if both genders got the same number of interactions, while females have significantly less than expected.

It was also found that total number of interactions is different based on clothing type as well. The same sort of test that was run for gender was run for clothing, and there was a significant difference in distribution as well. The plot suggests that the 'Other' clothing type decreases the number of interactions, while 'Unisex' increases it and 'Std' has no difference (Section 5.2.2).

Additionally, it was found that the professor gave significantly more positive than negative interactions. The number of positive and negative

interactions were tested for equal proportions, and the test indicated the proportions were significantly different from equal. The plot of actual vs expected (5.2.3) indicates that the professor gave many more positive interactions than negative interactions, while the expected count shows the same amount for each interaction type.

In the tests for independence, it was found that first, gender does not significantly affect the number of positive and negative interactions. According to the chart (Section 5.3.1), the proportion of positive interactions is slightly higher for females than males, but not by any significant amount.

It was also found that clothing does significantly affect the proportion of positive and negative interactions with the professor. It seems that the 'Unisex' group gets a larger percent of positive interactions compared to the other two groups (See Section 5.3.2).

Finally, the relationship between gender and clothing type was tested, and the test results indicate, though there is some difference in the types of clothing worn by males and females, that the difference is not statistically significant (Section 5.3.3).

4. Conclusions

Does clothing affect the number of interactions with the professor? According to this sample, yes, it does. It appears that the professor interacted with students wearing unisex clothing more often and more positively than the other two clothing groups. Gender was not found to be a significant factor in interaction type, though males

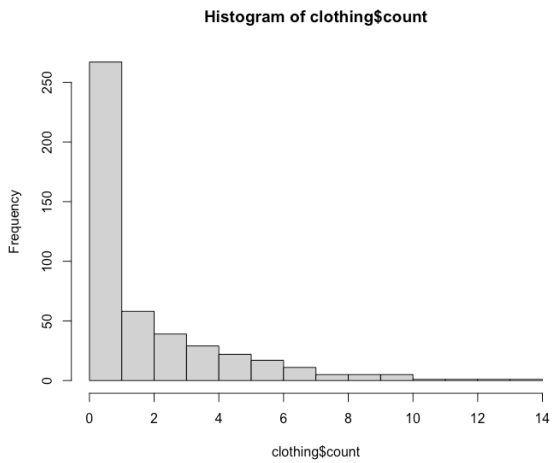
received more interactions total. The professor did give out many more positive interactions than negative ones as well. The question that follows: Is the professor inadvertently playing favorites with students wearing unisex clothing? Perhaps further study is needed; certain demographics of students could dress similarly and have similar soft skills or academic skills. Could those kinds of factors influence the interactions with the professor more than clothing type alone?

5. Tables/Graphs

The results of all the statistical tests and the corresponding graphs can be found in the following pages.

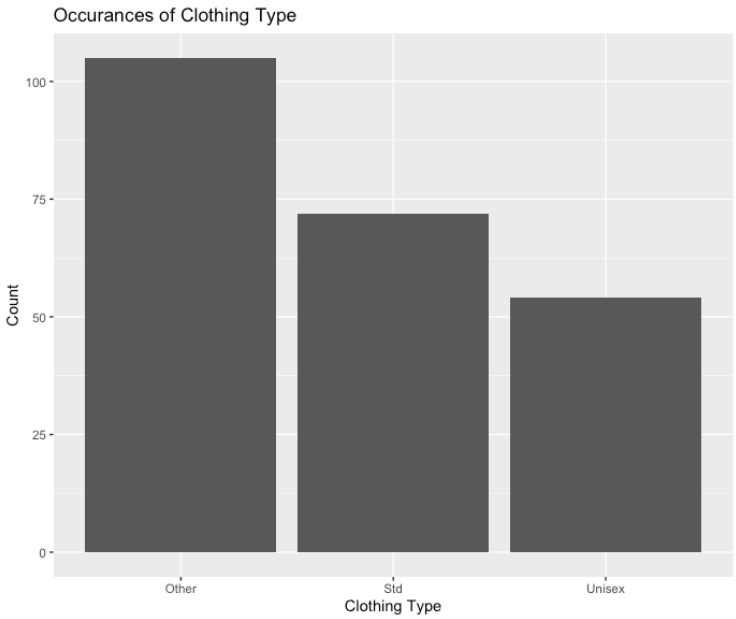
5.1. **Distribution of Variables:**

Distribution of Count of Interactions:



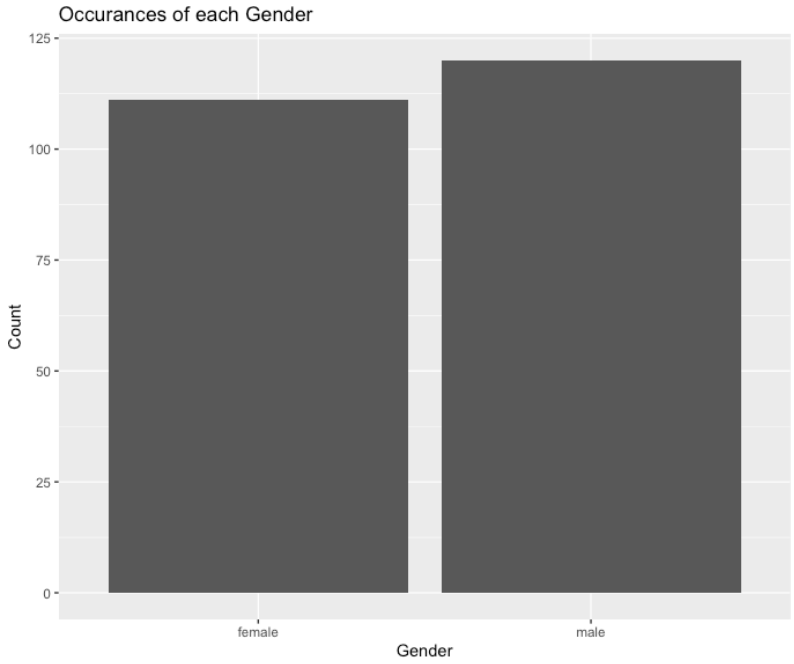
Minimum:	1st Quartile:	Median:	Mean:	3rd Quartile:	Maximum:
0.0	0.0	1.0	1.944	3.0	14.0

Distribution of Clothing Type:



Clothing Type:	Frequency:	Percent:
Other	105	45.5%
Standard	72	31.2%
Unisex	54	32.3%

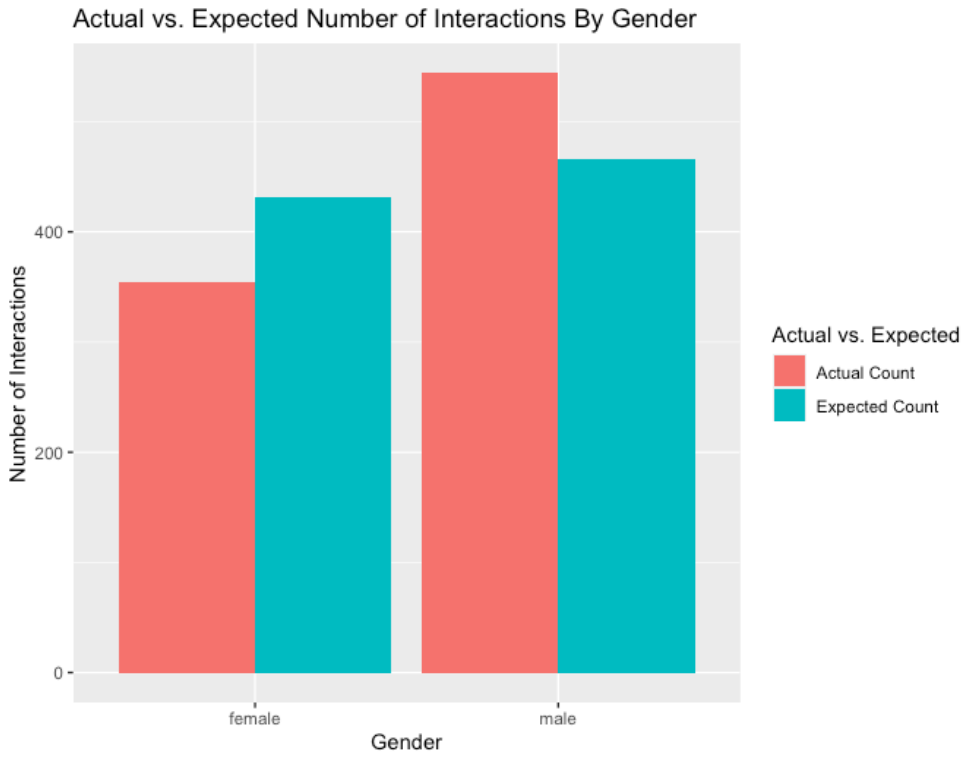
Distribution of Gender:



Gender	Frequency:	Percent:
Male	120	51.9%
Female	111	48.1%

5.2. **Goodness of Fit Tests:**

5.2.1. Total Interactions by Gender

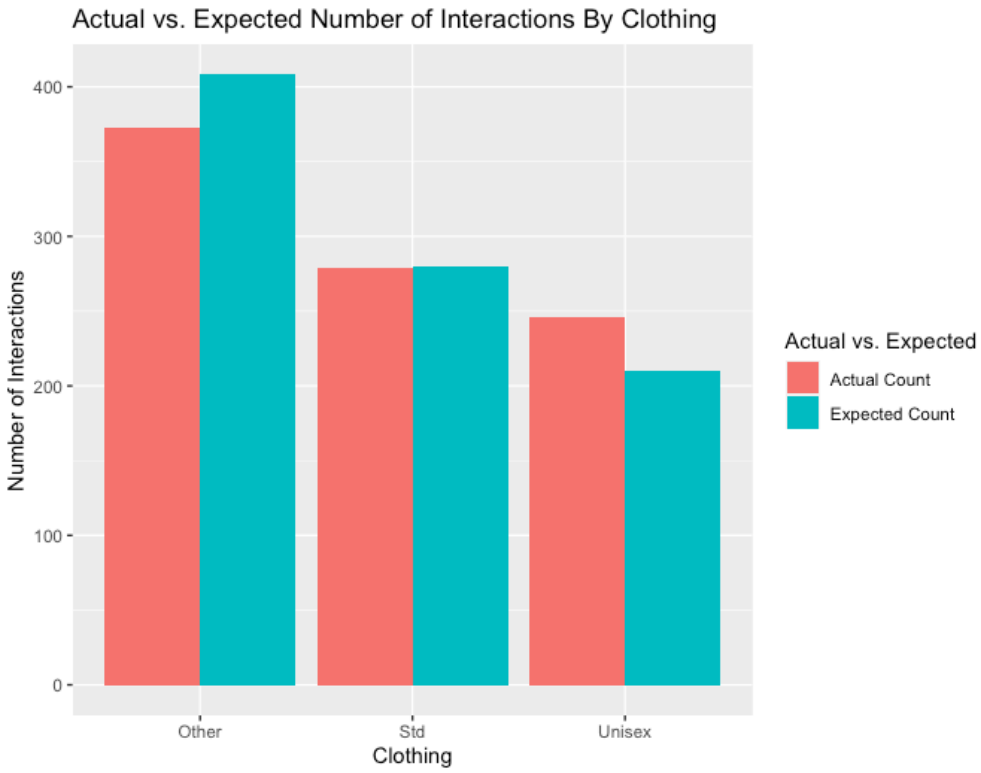


Males have a higher Actual Count than expected, while females have a lower Actual count than expected.

```
[1] "Is number of interactions different based on gender?"

        Chi-squared test for given probabilities

data:  genderprop$totalcount
X-squared = 26.799, df = 1, p-value = 2.257e-07
```

The p-value of 2.257e-07 indicates that there is a significant difference between the proportion of actual counts by gender and the expected proportion.

5.2.2. Total Interactions by Clothing

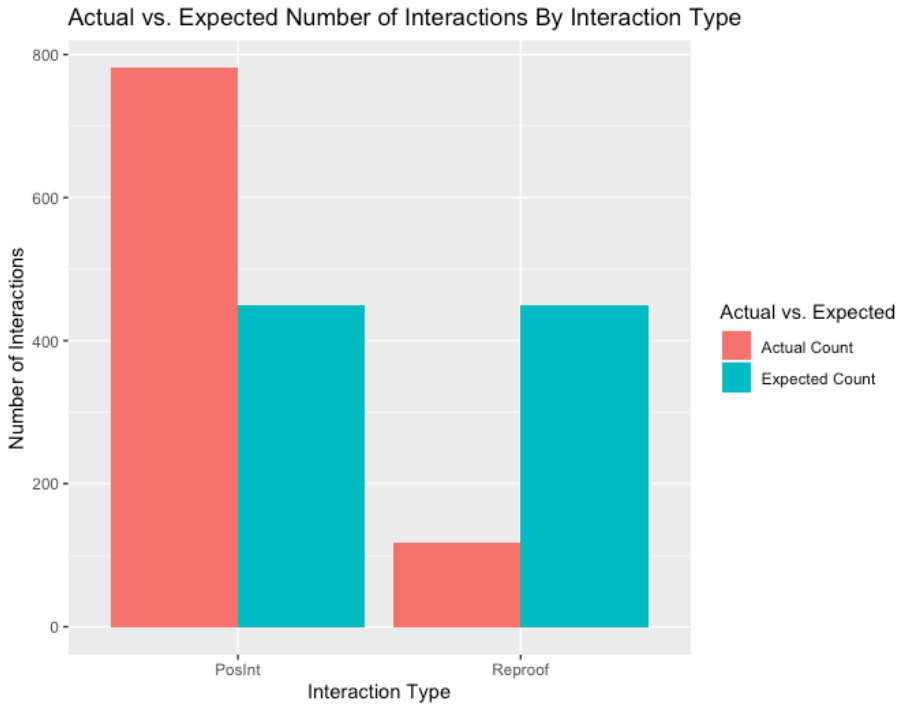


If we expect that each clothing will receive the same number of interactions from the professor, Unisex has a higher Actual Count than expected, Std has the Actual Count that we'd expect, and Other has a lower count than expected.

```
[1] "Is number of interactions different based on clothing?"

        Chi-squared test for given probabilities

data:  clothingprop$`Actual Count`
X-squared = 9.2357, df = 2, p-value = 0.009874
```

The p-value of 0.009874 indicates that the proportion of counts across clothing differs from the expected proportion.

5.2.3. Total Interactions by Type



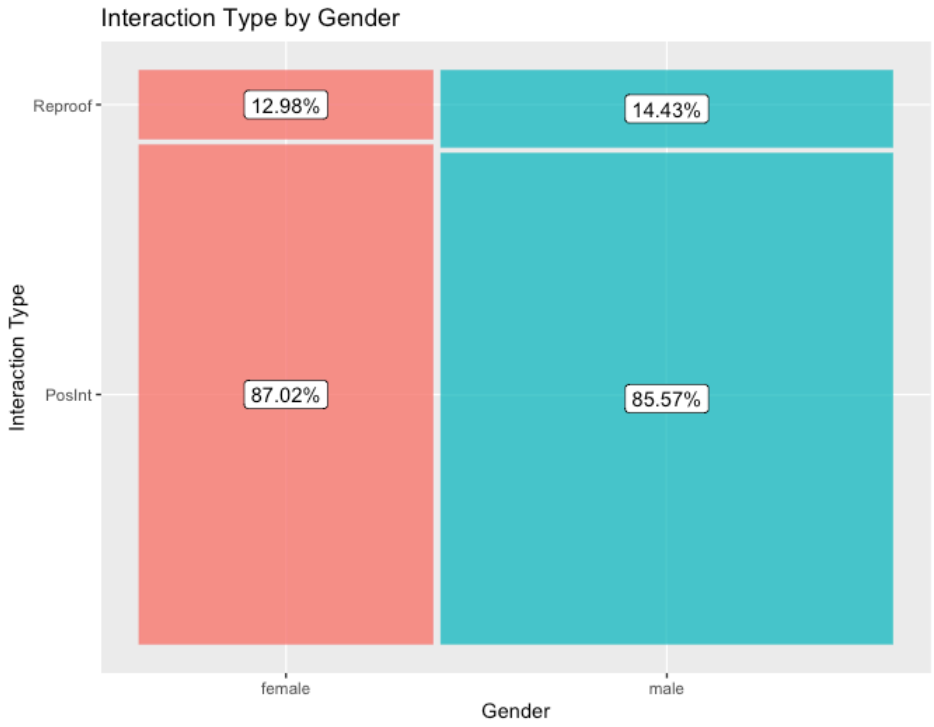
There are many more Positive interactions than negative interactions where we'd hypothesize that the counts would be equal.

```
[1] "Is number of interactions different based on interaction type?"

        Chi-squared test for given probabilities

data:  itypeprop$`Actual Count`
X-squared = 490.98, df = 1, p-value < 2.2e-16
```

The p-value of <2.2e-16 indicates that the counts based on interaction type significantly differ from the expected proportions.

5.3. **Tests of Independence:**

5.3.1. Interaction Type by Gender

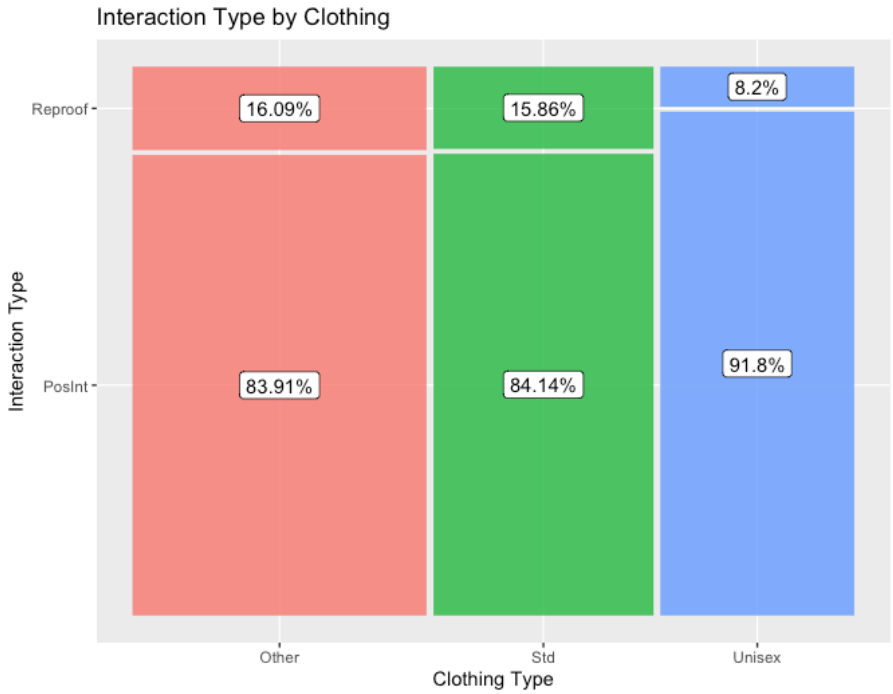


The percentage of positive and negative interactions change slightly from female to male students, but not by much.

```
"Does Gender affect Interaction Type?"

        Pearson's Chi-squared test with Yates' continuity correction

data:  genderitypematrix
X-squared = 0.28302, df = 1, p-value = 0.5947
```

The high p-value of 0.5947 indicates that the two variables are independent. This means that the number of positive vs. negative interactions isn't significantly affected by gender.

5.3.2. Interaction Type by Clothing

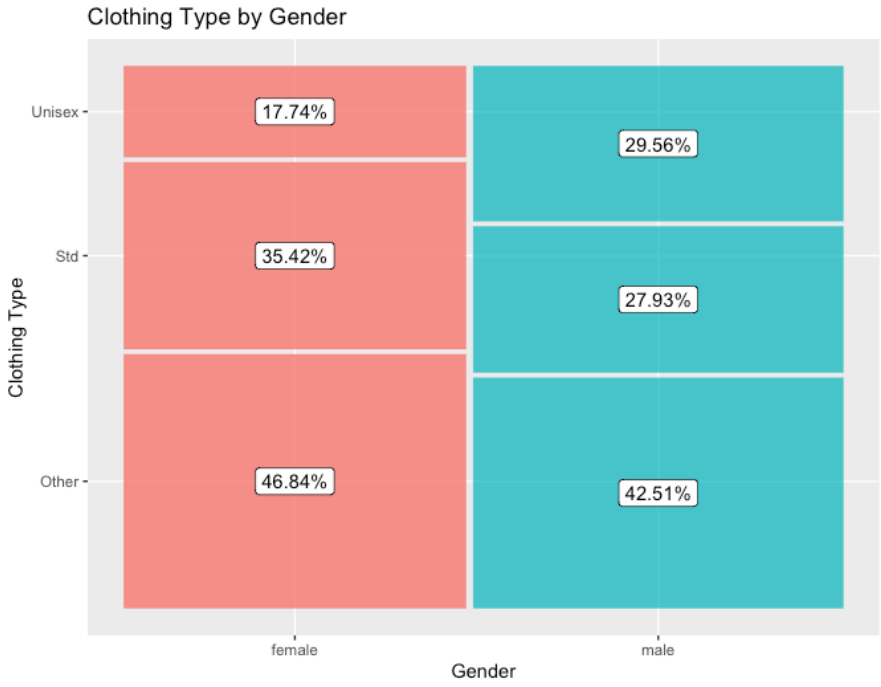


We can see that the percent of positive interactions increases slightly from Other to Std, and then it increases by a bit more from Std to Unisex.

```
"Does Clothing affect Interaction Type?"

        Pearson's Chi-squared test

data:  clothingitypematrix
X-squared = 9.7625, df = 2, p-value = 0.007587
```

The p-value of 0.007587 indicates that clothing and interaction type are not independent. This means that the number of positive vs. negative interactions is affected by clothing type.

5.3.3. Clothing by Gender



We see that different genders wear certain clothing types more than others. Males seem to wear Unisex clothing more, while females wear Std and Other more than males. Are these differences significant?

```
[1] "Does Gender affect Clothing Choice?"

        Pearson's Chi-squared test

data:  clothing2$clothing and clothing2$gender
X-squared = 4.9071, df = 2, p-value = 0.08599
```

The p-value of 0.08599 indicates that clothing type and gender are independent. This means gender doesn't significantly affect the clothing choice of a student.