

### **Pre-Analysis Steps**

After collecting the raw data from Google Sheets, we downloaded it as an Excel file. Then, we performed a by-eye search to remove any responses that did not answer all of the questions in the survey. A by-eye search was performed due to the small size of the results. By coincidence, this removed all of the **Graduate** students' responses, leaving only **Undergraduate** students.

Next, we checked the **Rationale** column to make sure that the answers provided were valid, and deleted the responses that weren't. Once we completed our cleaning process, the original 92 rows were reduced to 77. Finally, we converted the data to a .csv file to make our analysis easier.

### **Analysis Methods**

Our analysis methods for the cleaned data was to begin with exploratory charts in which we could visualize the total number of **Yes** and **No** answers to whether or not somebody considered a hot dog a sandwich. Continuing our analysis, we wanted to consider if different class year's had the same proportionality of Yes/No answers as the total number of respondents. Another analysis that we used was the Chi Squared test.

### **Evaluation of Success**

In order to evaluate the success of our analyses, we would look for a clear correlation (positive or negative) in our plots. If the plots have no correlation, we will not be able to draw a conclusion from the results. The result of a simulated chi squared test is that  $p \approx 0.9$ , indicating that year and response are not correlated. This exploratory analysis was successful because we were able to determine that UVA students respond similarly to this question regardless of their year. Below are graphs that display the total number of responses across years at UVA. It is swayed in the direction of third and fourth years so a larger sample would be necessary to produce more definitive results.





