Pre-analysis steps

We uploaded the dataset and imported libraries (pandas and numpy). We also performed data cleaning by renaming columns, removing graduate students, and dropped columns (Time stamp, grad status). We also conducted some brief exploratory data analysis, looking at the value counts of the variables. We found that ‘rationale’ would be a difficult variable to work with as the free-response nature of the variable means that there are many unique string values.

Analysis Methods

We plotted the data through matplotlib and made a barplot where the x axis if the student thought that hotdogs were a sandwich or not (yes, no) and the counts. Since we filtered out the graduate students, we have data pertaining to undergraduate students only.

Evaluation of success

Based on the results, we concluded that there are more UVA undergraduate students that believe that a hotdog is not a sandwich than UVA students who believe a hotdog is a sandwich. Because of the limited data, we were not able to conduct an analysis on the “rationale” variable, thereby not having the opportunity to explore the causality.