

## DS 4002: Analysis Plan

Kendall Freese, Blake Richardson, & Jeremy Grossman

### Pre-Analysis Steps

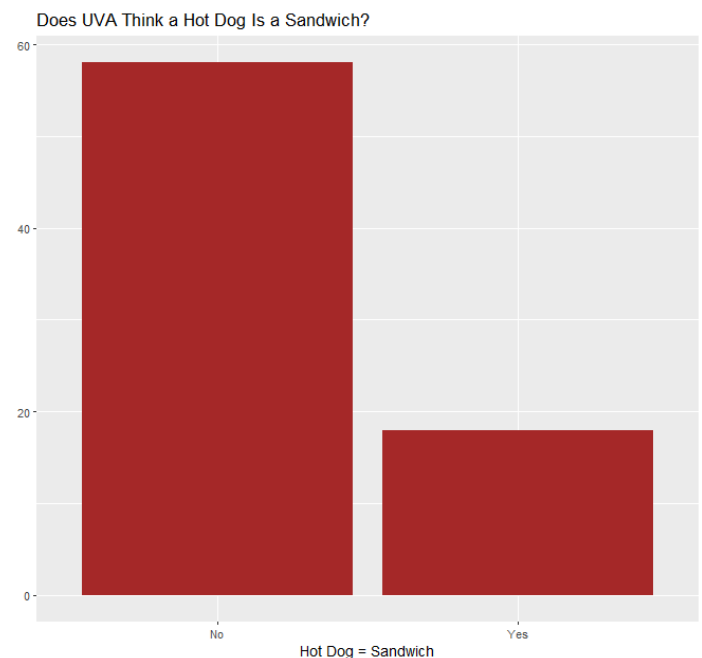
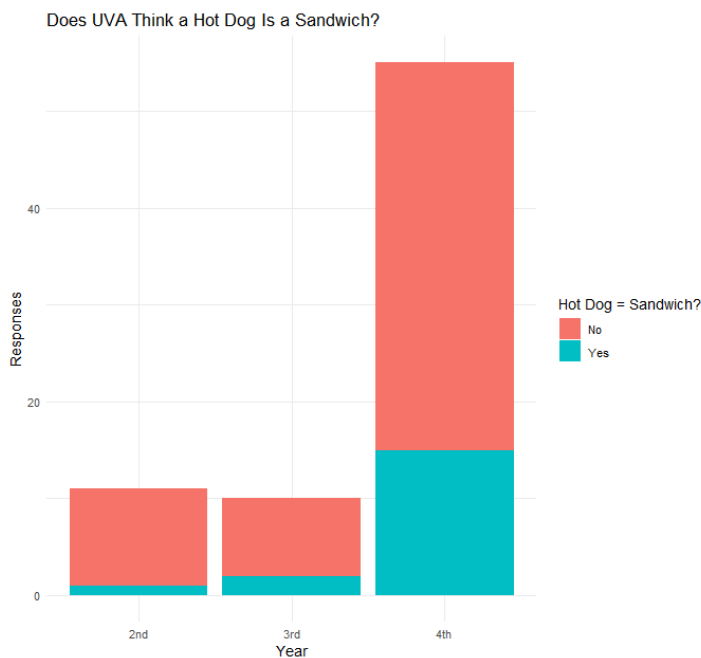
First, we needed to rename the columns, since initially, some column names were “Do you believe a hotdog is a sandwich?” and “What year are you?” We ended up setting the column names as Time, Student, Year, and Sandwich. Then, we needed to filter out all rows where the Student column is a “No”, since our hypothesis was about UVA students.

### Analysis Methods

We chose to analyze the data through a visualization of the frequency of particular survey responses and a stacked v. simple bar plot. These visualizations allowed us to draw conclusions regarding what the most common survey response was and how answers varied amongst particular student years. Additionally, we ran a one-sample proportion test to evaluate whether the proportion of students who consider hot dogs to be sandwiches differs from 50%.

### Evaluation of Success

To evaluate the success of our work we look at the statistical significance of our responses. The p-value is negligible (approximately 0), thus we can conclude that the results are statistically significant and can go on to draw conclusions. It’s important to note some potential limitations including that this sample may not be fully representative of the UVA student population due to selection bias from the survey, accessibility issues, and other unobserved variables.



95% Confidence Interval: We are 95% confident that the true proportion of UVA undergraduate students that believe a hot dog is a sandwich is between 15.0% and 35.1%. Therefore, we do have enough evidence to reject the null hypothesis and say that less than 50% of UVA students believe a hot dog is a sandwich

We are happy with our final output graphs and data analysis of undergraduates perspectives on if a hotdog is a sandwich.