Data Report

Edward Hernández

The College of William & Mary

Abstract

This paper addresses possible data analysis techniques that might be applied to an

survey dataset resulting from an MTurk immediately preceding the 2012 election.

Keywords: exploratory data analysis, R, MTurk

The MTurk experiment had four treatment conditions: a control, a stress condition, a

political stress condition, and a contentment condition. Participants were evenly distributed

among these conditions, as shown in Figure 1.[1]

Control	Stress	Politics	Content
448	445	440	446

Figure 1

The participants scored a mean of 48.3 on the Social Interaction Anxiety Scale,[2] and their

political polarization is illustrated in Figure 2.[3][4]

	n
True Independent	234
Leaning Independent	461
Weak Partisan	577
Strong Partisan	603

Figure 2

More data about the participants, like marital status, occupation, and whether or not they were

born in the US are included in the data as well.[5]

It is likely that a participant's anxiety strength at the thought of political discussion is

inversely correlated with their propensity to engage in political activity. Presumably, the anxiety

would be the independent variable that would influence the political engagement as a dependent

variable. The dataset includes both of these variables, which permits this analysis with relatively
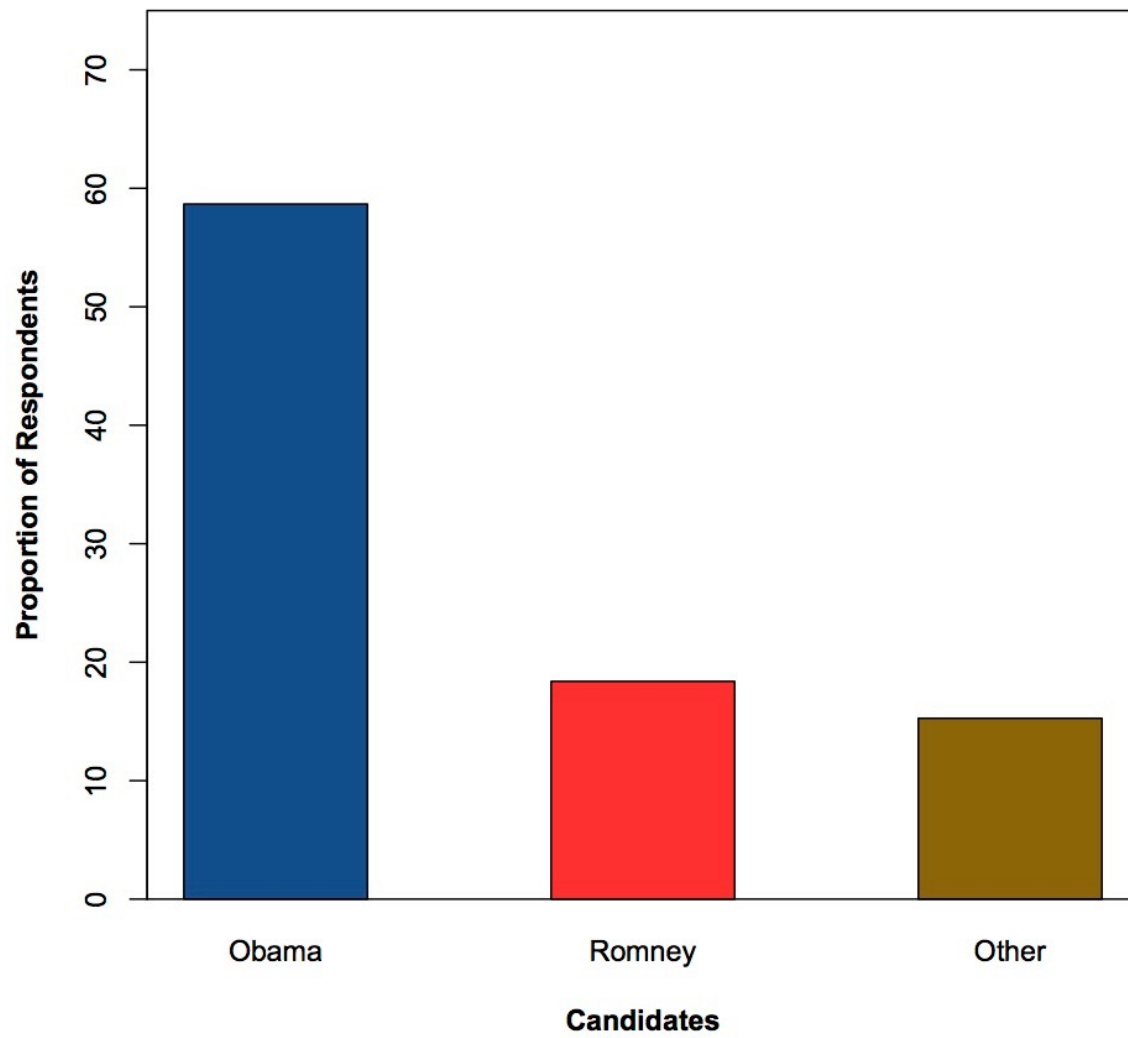
simple correlation and t tests.

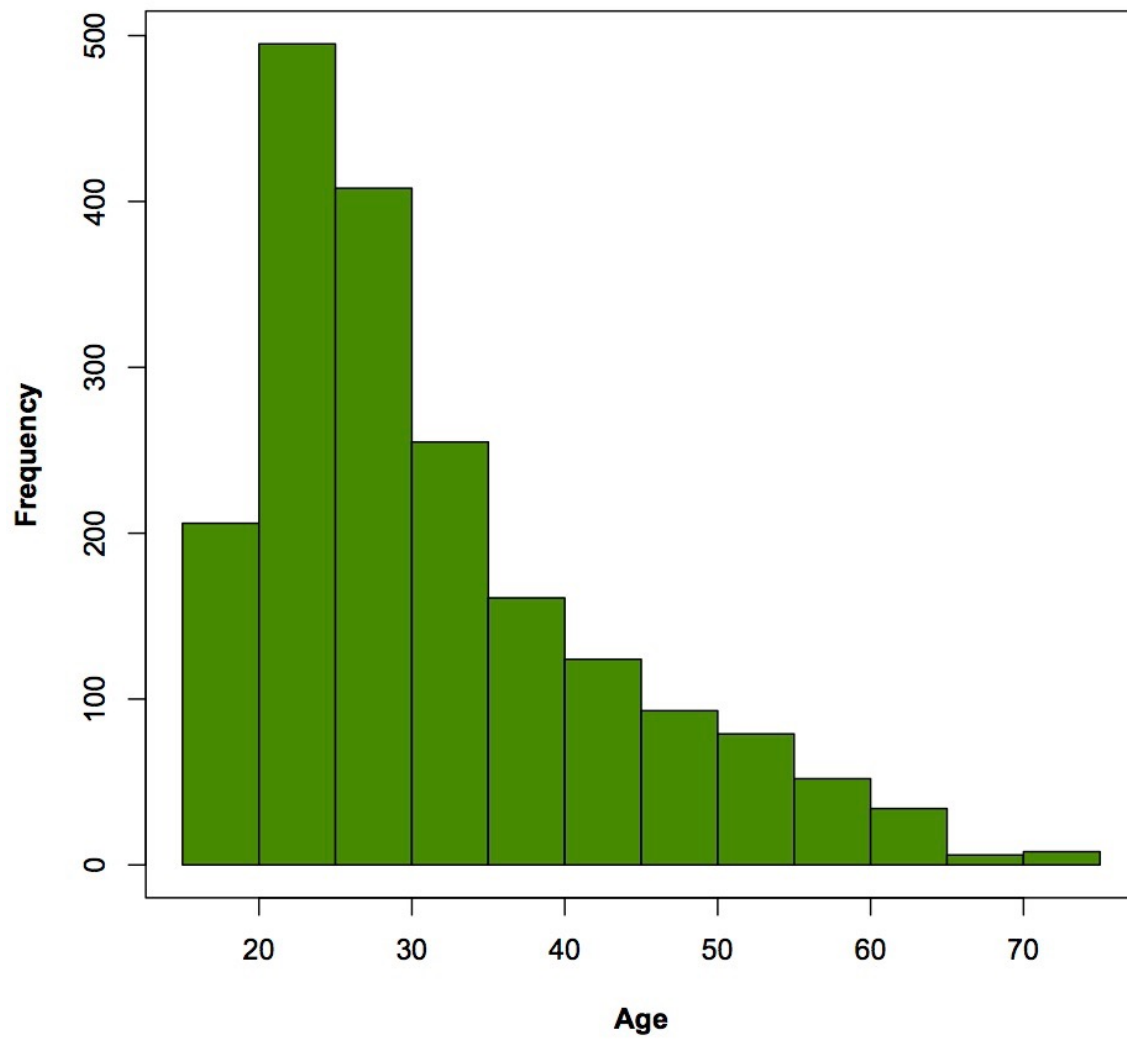[1] Superscript numbers correspond to numbered blocks of code in the accompanying code document.

[4] Their political engagement is also shown in the data, but not in this document.

The unit of analysis of this dataset in an individual respondent. 1926 respondents took the survey, 92% of whom finished it.[8] 38.2% of participants, and 31.3% of female participants were married.[9] The proportions of men and women are not significantly different from each other,[10] but the proportion of Democrats in the sample is significantly different from the proportion of Democrats in the national population.[11]

Interestingly, there was no statistically significant relationship between getting the political stress treatment and reporting likelihood to vote.[12] Though one might expect that a participant who is primed to experience political stress would be less likely to report an intention to vote in the upcoming election, the survey showed absolutely no relation between the two variables.

Some more salient relationships in the data are illustrated in the following figures.

Vote Choice Proportions in the MTurk Dataset

Histogram of MTurk Participant Age

Correlation Between Respondent Age and Response Length