

Trump Expecting to Win 53% of Popular Vote in 2020 US Presidential Election*

4% Confidence Interval Based on a Survey from July 2020

TBD

31 October 2020

Abstract

First sentence, second sentence, third sentence, fourth sentence

Keywords: forecasting, US 2020 Election, Trump, Biden, multilevel regression with post-stratification

1 Abstract

2 Introduction

3 Data

To train our model to predict the outcome of the 2020 US presidential election, we used Wave 49 of the Nationscape Dataset (results from the week of June 18-24, 2020). We will discuss how this data was collected, its key features, and what the data looks like in the section titled “Individual-level survey dataset”.

TODO: To make predictions on the outcome of the 2020 US presidential election, we used... We will discuss how this data was collected, its key features, and what the data looks like in the section titled “Post-stratification dataset”. The explanation of multilevel modelling with post-stratification can be found in the “Model” section.

3.1 Individual-level survey dataset

The Nationscape Project is 16-month-long voter study (from July 2019 to January 2021) that conducts weekly surveys regarding the 2020 US presidential election. We will solely discuss Wave 49 of the Nationscape Dataset for the remainder of this paper.

From June 18, 2020 to June 24, 2020, Nationscape collected data on public opinion about the 2020 presidential campaign and election by conducting online interviews. Their target is the American “population”. Unfortunately, the published information on their methodology is not more specific as to what constitutes a member of the American population (are they interested in the opinions of noncitizens?). Presumably, their target population is all individuals presently residing in the United States. Nationscape uses the audience of market research platform Lucid as its sampling frame, i.e., the survey respondents on Lucid are the frame for this dataset. Finally, a sample matching the demographics of the American population is selected from

*Code and data supporting this analysis are available at: https://github.com/JamesBond0014/sta304_ps4.

the frame. Unfortunately, further details on their sampling methodology were not provided: potentially they could have conducted simple random stratified sampling or picked random samples from the sampling frame (passing on individuals if quotas for their demographic were met). After being contacted by Lucid to take the survey, respondents are immediately redirected to Nationscape survey software where the questionnaire starts.

-methodology and approach processing the data -non response

-key features, -strengths -weaknesses -discuss variables you use -similar -combinining? -plot data -discuss plots

3.2 Post-stratification dataset

4 Model

$$Pr(\theta|y) = \frac{Pr(y|\theta)Pr(\theta)}{Pr(y)} \quad (1)$$

Equation (1) seems useful, eh?

Here's a dumb example of how to use some references: In paper we run our analysis in R (R Core Team 2020). We also use the `tidyverse` which was written by Wickham et al. (2019) If we were interested in baseball data then Friendly et al. (2020) could be useful. Tausanovitch and Vavreck (2019)

5 Results

Our data is of penguins (Figure 1).

Talk more about it.

Also bills and their average (Figure 2). (Notice how you can change the height and width so they don't take the whole page?)

Talk way more about it. # Discussion

5.1 Weaknesses and next steps

Weaknesses and next steps should also be included.

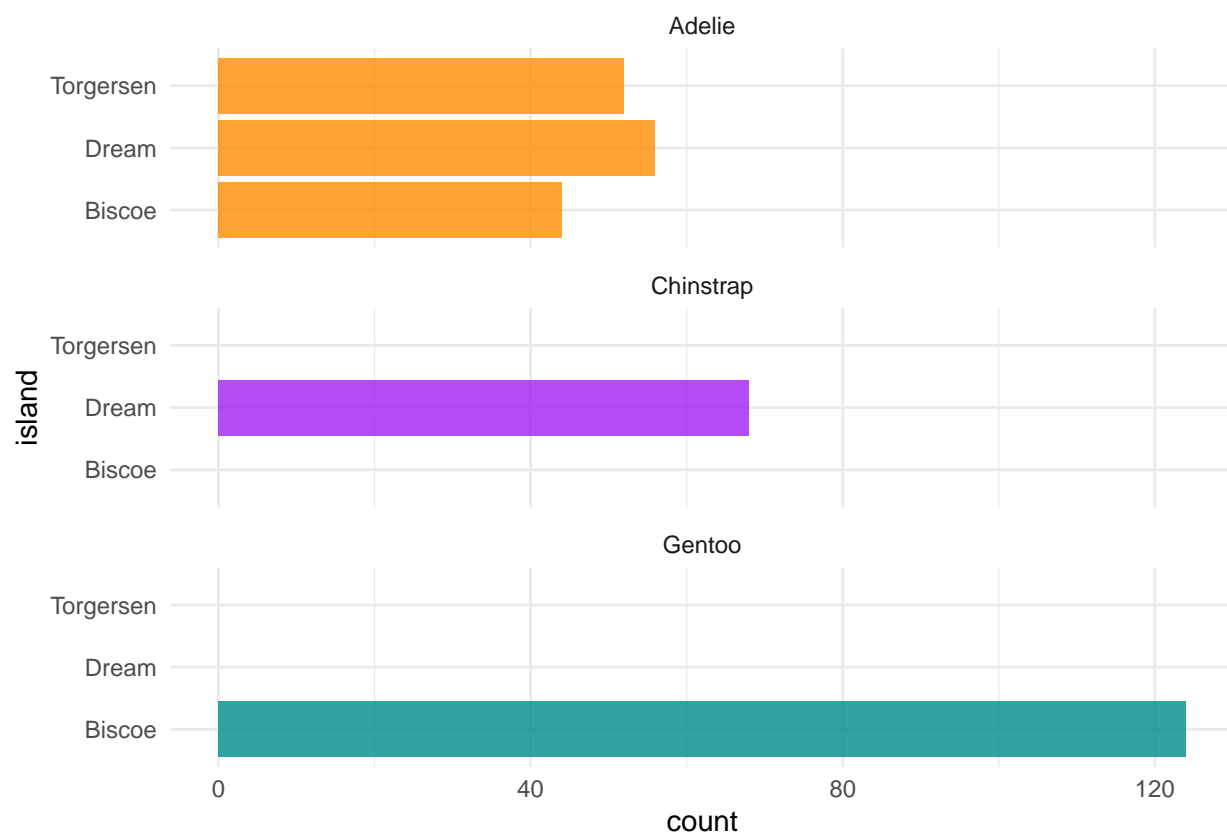


Figure 1: Bills of penguins

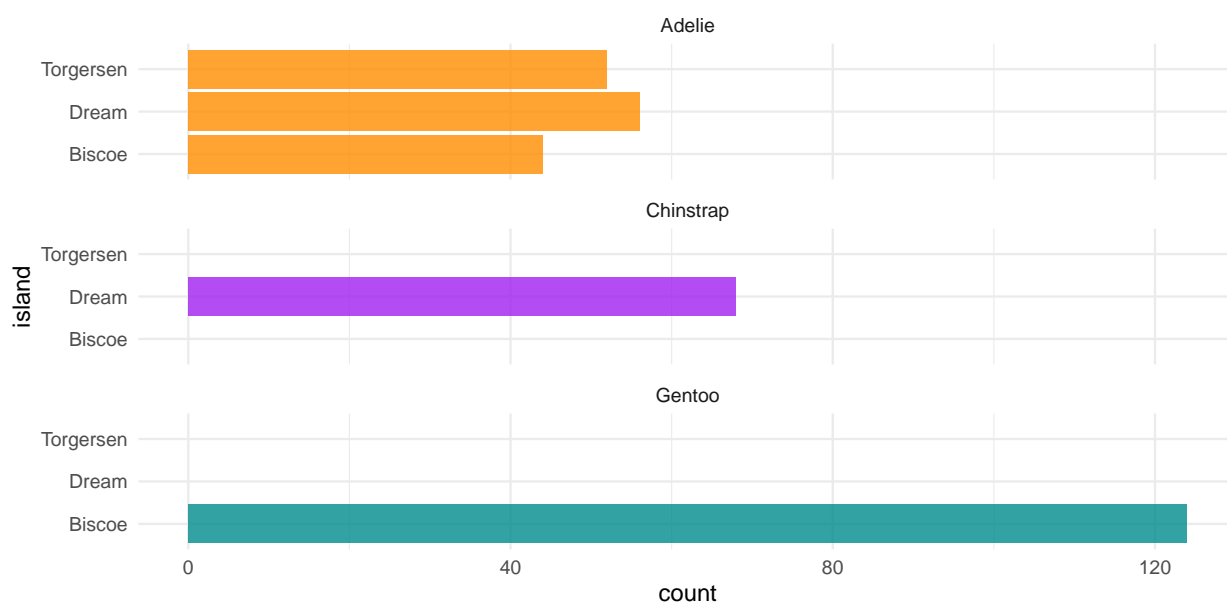


Figure 2: More bills of penguins

Appendix

6 References

- R Survey and ACS dataset https://www.voterstudygroup.org/uploads/reports/Data/Nationscape-User-Guide_2020sep10.pdf <https://www.voterstudygroup.org/uploads/reports/Data/NS-Methodology-Representativeness-Assessment.pdf>
- Tausanovitch, Chris and Lynn Vavreck. 2020. Democracy Fund + UCLA Nationscape, October 10-17, 2019 (version 20200814). Retrieved from [URL].
- Friendly, Michael, Chris Dalzell, Martin Monkman, and Dennis Murphy. 2020. *Lahman: Sean “Lahman” Baseball Database*. <https://CRAN.R-project.org/package=Lahman>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Tausanovitch, Chris, and Lynn Vavreck. 2019. “Democracy Fund + Ucla Nationscape.” <https://www.voterstudygroup.org/publication/nationscape-data-set>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemond, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.