

Getting To Know You: An Exploratory Data Analysis

Joseph T. Ornstein*

August 18, 2021

Abstract

At the beginning of most graduate-level courses, participants take a few moments to get to know one another before diving into the course content. This paper asks: why not do both? In a series of analyses, we simultaneously get to know a bit more about the participants in POLS 7012, while introducing some of the foundational statistical tools we'll be learning this semester.

1 Introduction

It is customary to begin journal articles with a section titled “Introduction,” to indicate that the reader has arrived at the beginning of the paper. As the very purpose of this paper is to serve as an introduction to the course and your classmates, I see no reason to dispense with custom here. In the following section, I will present some statistics and visualizations describing the personal data we collected at the beginning of class.

1.1 Where We're From

We know that political scientists come from a wide range of backgrounds.¹ But the *geographic* diversity of the discipline is, to my knowledge, understudied. Let's investigate where this particular cohort of graduate students is coming from.



Figure 1: A map of hometowns for students in POLS 7012

The median student in POLS 7012 is 135 km from their hometown, but the *average* student is 546 km from their hometown. A divergence between those two numbers suggests the presence of outliers and/or skew.

*University of Georgia

¹For some recent statistics on diversity in political science, see Reid and Curry (2019).

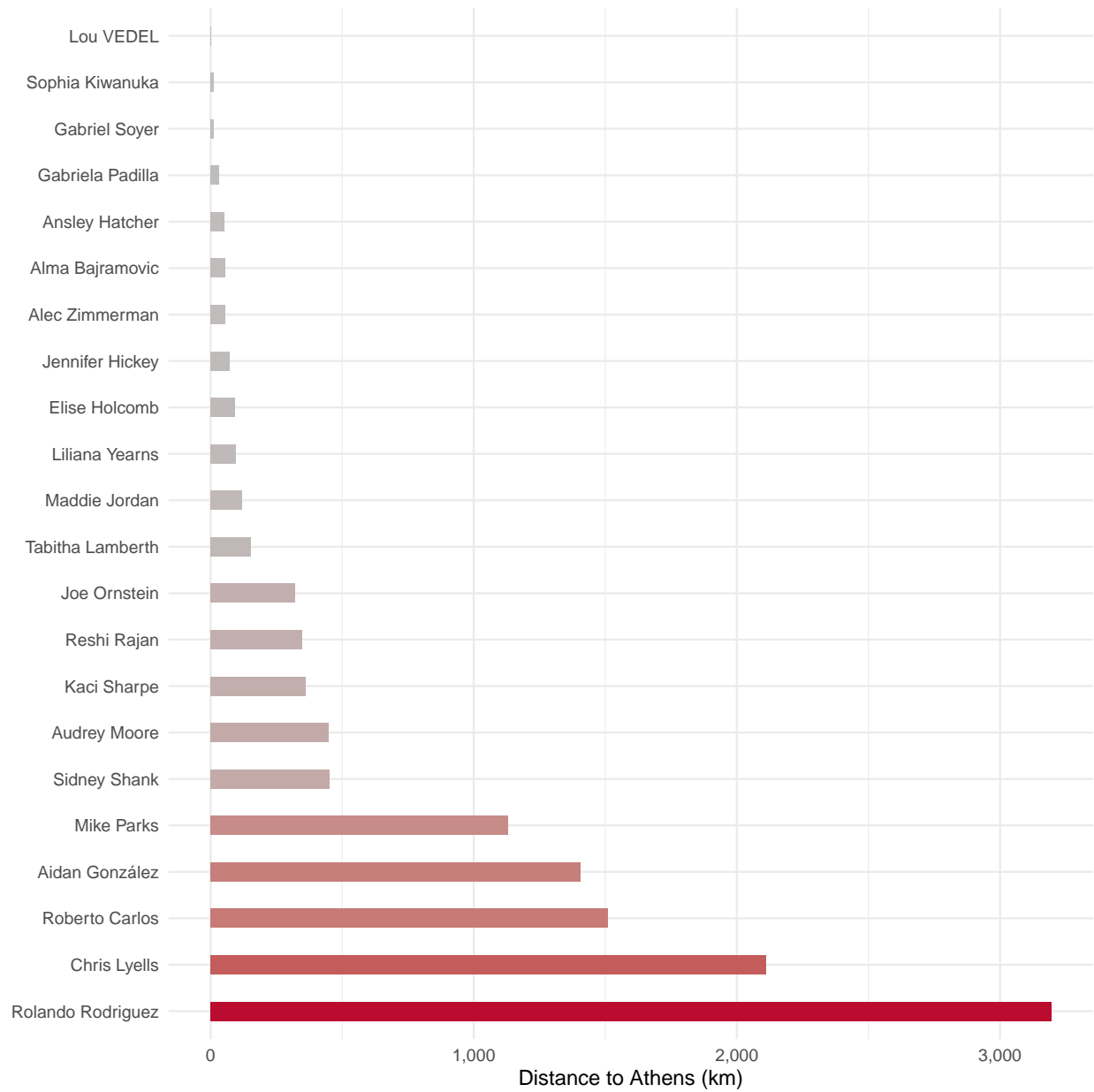


Figure 2: The distance of each student from their home town

1.2 Personal Characteristics

On average, we are NA feet, NA inches tall. 40.9090909% of us prefer dogs over cats. 68.1818182% of us prefer beaches over mountains. Here is one way to visualize all that information.

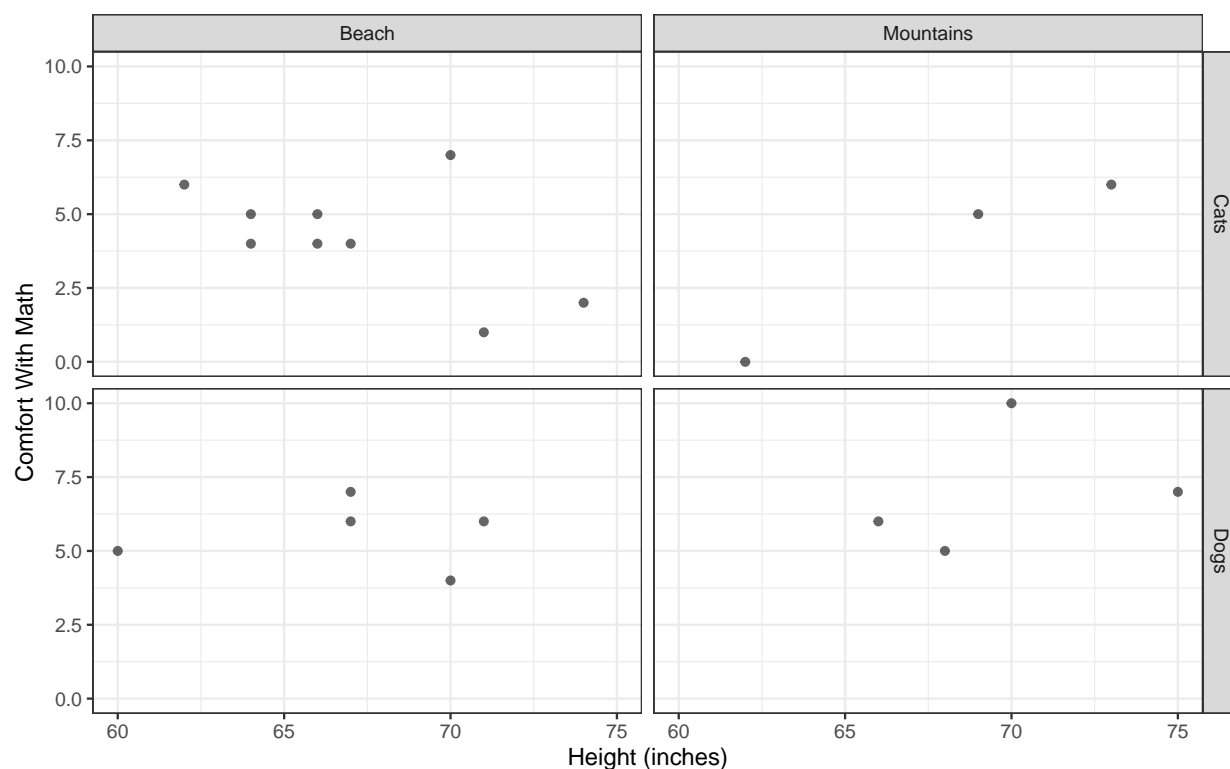


Figure 3: Distribution of personal characteristics in POLS 7012. I don't know if this is the *best* way to plot this data, but it is one way.

Here's an interesting correlation that shows up in the data: taller people seem to be more comfortable with math. On average, 1 additional foot of height is associated with 1.12 more points in math comfort. However, the correlation appears to be quite weak. The probability that a relationship between two variables this strong could arise purely from chance is roughly 46%. One of the central challenges in quantitative social science is to think carefully about what processes could generate the patterns you observe in the data. (And no amount of mathematical wizardry will ever relieve you of the need for careful thinking.)

2 Sentiment Analysis

Each participant was asked to write a short statement describing how they felt about taking a quantitative political methodology course. In total, 12 people expressed more positive than negative feelings, 7 people expressed more negative than positive feelings, and 0 people were somewhat neutral. It is not uncommon for students to feel anxious or uncertain – many people have unpleasant associations with math (Strogatz 2012), and for many of you, your last regular contact with the subject was in high school! During the semester, we'll discuss strategies to boost your confidence with math and coding, and I encourage you to lean on your cohort, senior grad students, and instructors for guidance and support.

A word cloud consisting of three words stacked vertically. The word 'excited' is at the top in a large, bold, black font. Below it is the word 'nervous' in a smaller, regular black font. At the bottom is the word 'anxious' in the same smaller, regular black font.

excited
nervous
anxious

Figure 4: Word cloud with words expressing positive or negative sentiments

References

- Reid, Rebecca A., and Todd A. Curry. 2019. “Are We There Yet? Addressing Diversity in Political Science Subfields.” *PS: Political Science & Politics* 52 (2): 281–86. <https://doi.org/10.1017/S1049096518002068>.
- Strogatz, Steven H. 2012. *The Joy of x : A Guided Tour of Math, from One to Infinity*. Boston: Houghton Mifflin Harcourt.