Request #: 580 - ITLS - Publication/Article

Comparing the characteristics and composition of three online scientific communities

Lisa Lundgren [A02339821] - ITLS Faculty

October 12, 2021

Background

We collected data from Twitter in three different instances: one set of data associated with four paleontology-specific hashtags; and two other sets associated with specific usernames. From these datasets, we used a framework to categorize users who created tweets. The categories included Public, Scientist, and Education and Outreach entities. We are now looking to make comparisons about these categorizations and the different instances

Sample

Data were usernames and user biographies from three Twitter "habitats" that were categorized. Habitat 1 was called Four Hashtags and included Public (n = 2199) Scientists (n = 872) Education and Outreach (n = 440) Habitat 2 was called project X and included Public (n = 509) Scientists (n = 526) Education and Outreach (n = 311) Habitat 3 was called projectY and included Public (n = 238) Scientists (n = 632) Education and Outreach (n = 243)

Hypothesis

Our hypothesis is that there will be differences at the categorical level, e.g. that The Four Hashtags habitat will be different than the projectX habitat in the number of Scientists.

Progress

Exploratory summaries and descriptive statistics have been run, I just need help deciding which statistical tests we can use to test our hypothesis, and if our data is normal and if we should even do statistical testing!

Request

I just need help deciding which statistical tests we can use to test our hypothesis, and if our data is normal and if we should even do statistical testing!

Timeline

We hope to submit this for a conference by Nov 5, 2021.