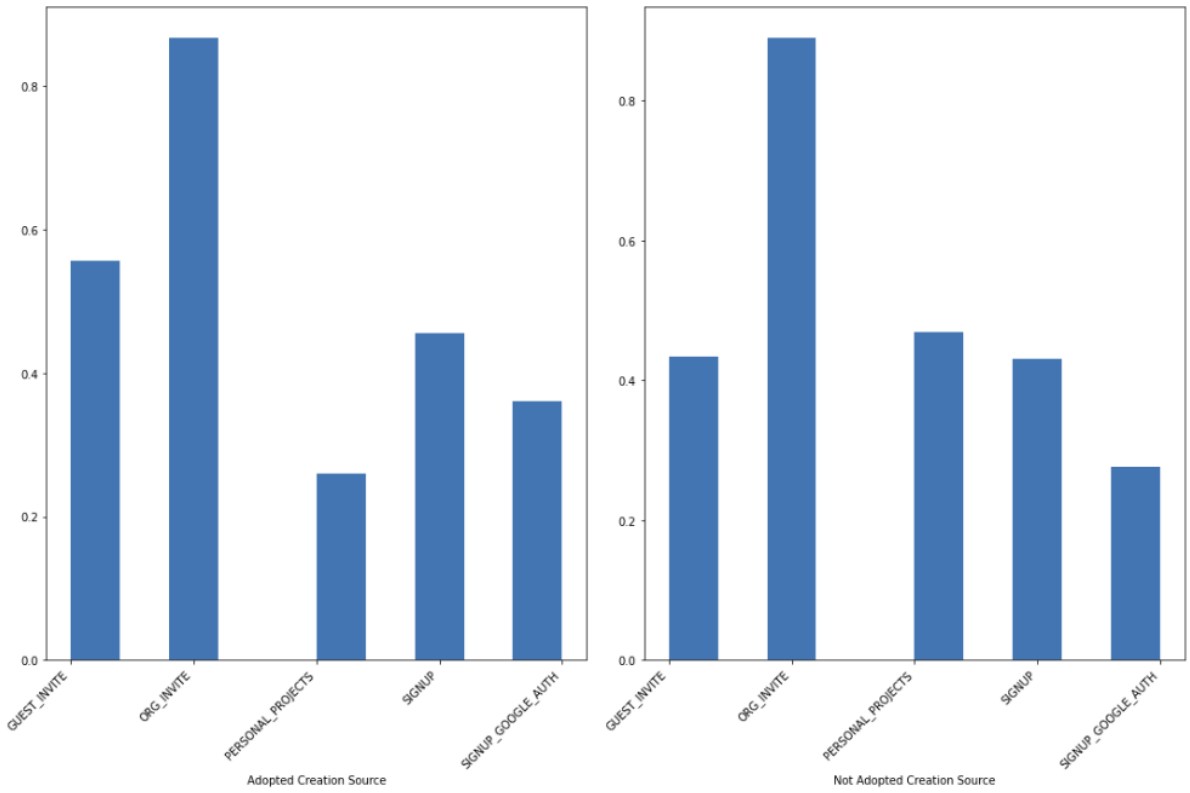
Rest Inc. Take-Home Challenge Report

Based on my analysis of the data given by Rest Inc., I am finding it challenging to discern any demographic information that would allow us to predict whether or not a user will become an “adopted” user. I looked at grouping the users first by whether or not they would be considered adopted. I then visually assessed the demographic data based on this information. For each histogram I explored, the distributions between the two groups were nearly identical. The only difference was their magnitude. The only difference I found was that between the not adopted and adopted groups, there was a higher ratio of accounts created through a “guest invite” that became adopted accounts and a higher proportion of accounts created for “personal projects” were not adopted (see Fig. 1) Perhaps those who create guest accounts are more likely to become adopted users.



I attempted to create a clustering model using K-Modes which looks at categorical data to create groups. I tried it with just 2 clusters as I was hoping to differentiate between those who would become adopted users and those who would not. Ultimately what happened was that there was almost a perfect 50/50 split between the two groups that were created meaning the model was unable to truly capture any difference. I think that with the data provided it would be very challenging to create a predictive model which could anticipate whether a user will become adopted.

Perhaps looking at some data regarding the kinds of projects that people are working on when joining the website or their industry we can parse out some differences. For future research, I definitely recommend gathering a wider breadth of demographic data that can help better identify users. We could look at age and whether age of a user impacts their potential for adopting the website. There definitely is room for future improvement and analysis.