

Can We Predict Big 5 Personality Traits from Demographic Characteristics?

Mr. Ethan Woods and Dr. David Han

Department of Management Science and Statistics
University of Texas at San Antonio, TX

Abstract

Here we aim to predict the Big Five personality traits based on the demographic information using a generalized linear model. Data was obtained from openpsychometrics.org, pre-processed in MS Excel, and imported to R for statistical analysis. First, it was attempted to predict each individual response item using an ordinal regression model. It was however found to be not viable, even after various weightings were applied to the demographic data. The response variables were then aggregated to form five categories, one for each personality trait: conscientiousness, agreeableness, neuroticism, openness to experience, and extraversion. We then applied a dimension reduction technique to the country variable as well as the race variable in order to achieve an adequate model fit. It was determined that although the demographic information could be useful, precise prediction of the Big Five traits require other information that was not captured in the dataset.

Keywords: Big 5 personality traits, demographic characteristics, dimension reduction, generalized linear model