

In order to identify possible business opportunities, we thought it would be valuable to add demographic data. The original project goal aimed to predict the count of restaurants offering burritos or tacos using demographic characteristics as predictors using linear regression models. We considered variables such as total population, racial makeup of the zip code, median household income, median property values, average educational attainment, and age. Due to inexperience accessing Census data and the time constraint, the analysis was narrowed down to total population and median household income. One of the difficulties we came across is the fact that Census data is not arranged by ZIP codes but by Public Use Microdata Areas (PUMAs). This then required matching the PUMA data to ZIP codes using Geographic correspondence utility made available by the University of Michigan online and then merging it with the restaurant data set. Initial scatterplot viewing of the data indicated that a polynomial regression may better fit the data than a linear regression. The fitted model unfortunately doesn't seem to be a good fit, given the plot of the fitted line and the data.