Data Analysis for United Way

Our objective with this project was to find out more information about “food deserts” in Topeka, KS or areas where there is less access to food. We wanted to determine if the following hypothesis was accurate: “In areas with higher median household income, there will be more food resources available in that area. In areas with lower median income, there will be fewer resources available in that area”. We concluded that our hypothesis should be rejected due as we didn’t not see a strong positive correlation between median household income and access to food resources. Our data indicated that there is actually a slight negative correlation between median household income and food resources meaning, if you live in a census tract with a lower household median income, you are slightly more likely to have access to more food resources than a census tract with a higher household median income. We have a few explanations for why that might be the case.

In order to dig into this issue, we used Census data from census.gov and Topeka income data from datausa.io to gather median household income data and census tract data in Topeka, KS. We also made a call to the Google Places API using a hand drawn area around Topeka to pull all potential food resources within that area. Our Tableau dashboard visualizes all those resources and is filterable for whatever resources the user is interested in looking at. Each point on the visualization also shows what that resource is if you hover over it.

From that point, we used machine learning to complete a linear regression analysis to check the value of the correlation between median household income and the number of food resources in a given census tract. This showed a slight negative correlation.

To test a few other potential factors, we also pulled data for the Topeka population by census tract as well as housing units by census tract. We ran regression analyses on these and they seemed to show a positive correlation with access to food resources with housing units having a stronger relationship than population.

On review of the Tableau visualization, it appears most of the Topeka, KS area census tracts fall into the lower end of the household median income scale, which is one reason why the food resources might skew towards the lower end. If we were to measure access to food resources against median household income across the entire US, it is possible on average we might find a different result. However, our results do show that population or the density of housing in a given area does seem to have a stronger correlation to access to food resources.

Zach Wilson, Mark Messick, Annie Di Persio, Ben Anderson