Cape Fear Collective
EQUITY FOCUSED. DATA DRIVEN. HUMAN CENTERED.

DS Technical Screen

Context:

Food insecurity is a major issue in the Cape Fear region, and throughout the country. Listen to the podcast at https://capefearcollective.org/collective-voice-old/food-deserts/ for some context on food deserts in our region.

Assignment:

Explore the provided data to understand the problem of food deserts in North Carolina and New Hanover County.

1. Join data sets to create a clean df with one row per tract.
2. Prepare visualizations and insights (at least 3) to illustrate interesting findings from the data.
3. Build a predictive model with a binary output (1 if food desert, else 0).
4. Organize findings into an hour-long presentation (saving 15 minutes for questions and discussion, so 45 minutes of content). Your audience will come from a spectrum of data literacy and data comfort.

Things we will be looking for:

- All analysis should be done in R and should be reproducible, meaning someone else can run the code and get the same results (or very similar…if sampling or some other random process is involved).
- Visualizations and insights that are clear and demonstrate candidate's grasp on data/statistics techniques.
- An assessment of the performance, accuracy, and pros/cons of the predictive model.

Data:

You have been provided with five North Carolina datasets. No additional context about these data will be provided other than the provided data dictionary. However, all data were sourced from USDA and Census Bureau (ACS) so they should be widely documented in publicly available locations. Additional data and sources may be used, if necessary.

- food.desert – binary classification of food deserts, poverty rates, median household income.
- population – county population over time
- ruca.usda – ruca codes, tract population, land_area
- stores – number of grocery stores, supercenters, convenience stores, and specialty food stores
- vehicles – number of cars per household

Submission:

Email your presentation and code (or link to Github) to data@capefearcollective.org by 8am on the day of your presentation.