The Team: Charlotteans

A list of the members of your team and their expected roles.

Matt Clewis

Teya Robinson

Zacharias Girmay

Christopher Sloat

Clint Jansson

Problem Statement a.k.a Decompose The Ask

Has the migration to Charlotte NC resulted in an increase of housing prices, and as a driving an increase homelessness?

Identify Data Sources

Describe the data source(s) you will use. Make sure you have access to the data you want to use \*in the quantity and quality you need\*. Ideally you should get all the data in your hands before starting your project.

Describe how you plan to obtain the data, or how you got it if you already have it.

Data source: Homelessness from Kaggle

<https://www.kaggle.com/adamschroeder/homelessness/home>

State by state Migration from Census.

<https://www.census.gov/data/tables/time-series/demo/geographic-mobility/state-to-state-migration.html>

Population estimates from UNC

<https://demography.cpc.unc.edu/resources/data-tables/>

Cost of living

<https://www.opendatanetwork.com/entity/0400000US37/North_Carolina/economy.cost_of_living.index?component=all&year=2016>

Define Strategy and Metrics

List some goals of your analysis, ideally in the form of testable hypothesis, or via well-defined success metrics. These can be tentative, and you don’t need to stick to them throughout your project. Again, since you haven’t done any exploratory analysis yet, you might assume that the data has structure that it doesn’t, and you might not have seen other interesting patterns in the data. But you should always approach the data with some expectations so that your efforts are focused.

We believe that cost of living, population and homelessness have a positive correlation to each other.

1. Looking for an increase in Charlotte population
2. Looking for an increase in cost of living
3. Looking for an increase in homeless population

Description of Data Analysis Tools You Plan to Use

Describe the tools you plan to use throughout the project. As you might expect, there will be several stages in the project. Make sure to use **at minimum** the required tools (see Project1 power point).

Python and Jupyter Notebook for the analysis of CSV data,

Pandas to aggregate and manipulate data frames

API connection for data on the cost of living

Matplotlib for generating plots and visualizations

PowerPoint to present the data and talking points

Describe the Data Products Your Project Will Produce

Data products include results of statistical tests, visualizations of the data, websites/apps, or model parameters. What will be the tangible evidence of your discovered insights?

1. Line graphs to track changes year over year change in cost of living, homelessness, and overall population.
2. Using Pandas, we will use the correlation coefficient to the rates of change.
3. This will establish the relationship and provide justification to the hypothesis.