Transit Analysis Proposal

My proposed project would investigate the effectiveness and efficiency of public transit in the City of Charlotte, NC. Charlotte is known for its lackluster public transit and hyper-sprawling development, forcing low income individuals and those lacking a vehicle to deal with increased difficulties getting around the city and paying for transportation. However, I believe there are ample opportunities to expand transit service in certain areas of the city, regardless of the notorious car-dependent development patterns exhibited around Charlotte. To attempt to identify prime locations for expanded transit service, I will be analyzing transit frequency per square mile in the city in relation to a number of different variables at the census block group level, as well as producing spatial representations that I hope will be informative for future policy decisions. I will conclude by making recommendations for transit corridors that are well suited for increased public transit service.

The two online sources I am utilizing for data are the Environmental Protection Agency's (EPA) Smart Location Database, known as SLD, and the Simply Analytics web app, which provides access to historical census data and various consumer and business survey providers. From Simply Analytics I have selected metrics from the 2022 Consumer Expenditure Estimates, which primarily supplies data from the 2022 Consumer Expenditure Survey (CE), and uses data from the 2022 American Community Survey (ACS) to estimate expenditures at the census block group level. Additionally sourced through Simply Analytics is 2022 median household income data from the ACS. The variables I have selected from Simply Analytics are at the census block group level, and in dollar amounts: Household average expenses for transportation, and Median household income. I have downloaded the entire SLD and filtered for the variables I plan to use,

which are as follows: Number of low wage workers (home location), Number of low wage workers (work location), Gross activity density (jobs + housing units per acre), Jobs per acre, and Aggregate frequency of transit service per square mile. I plan to join these two datasets using the FIPS code at the block group level. Both datasets are filtered to Mecklenburg county, with the first dataset including only block groups within the official Charlotte city limits.

First, I will investigate a relationship between transit frequency and density. I will create a bivariate choropleth map of the transit frequency and gross activity density variables. If the results show that there are high density areas with low transit service frequency, this can be informative for policy discussions about increasing transit service in the city. Second, I will utilize the average household transportation costs variable to analyze whether individuals in areas with lower transit service frequency spend more or less on transportation. I will create a proportion of transportation costs to income in order to perform this analysis based on the percentage of income spent on transportation, rather than gross dollars spent on transportation. Then, I will use linear regression, first to analyze a relationship between transit frequency and income to identify if transit serves high income areas more adequately than low income areas, and second, to analyze a relationship between transit frequency and percentage of income spent on transportation to see if transit service is better or worse in areas where people might need it the most. I will visualize the trend line of these linear models on a scatter plot, with transit service as the dependent variable on the Y axis, and income/percentage of income spent on transportation as the independent variable on the X axis. Investigating this relationship will reveal, to some extent, how successful the city has been at providing public transit to lower income individuals.

Finally, I will attempt to identify transportation corridors well-suited for expanded public transit service and frequency. To do this, I will utilize the jobs per acre and number of low wage worker (work location) variables in the SLD to create a bivariate choropleth map aiming to find areas with a high volume of job opportunities and a high number of low wage workers. Then, I will create a similar map with the transit frequency and number of low wage worker (home location) variables to find areas of low income workers poorly served by public transit. With the help of these two visualizations, I will identify and recommend transportation corridors that are the most well suited for increased public transit service, with the goal of promoting transit connectivity between areas with large numbers of low income workers, and areas where those people frequently commute to work. Finding routes in the city to make public transit more accessible and more effective will ease the burden of transportation costs to citizens who commute by car but could potentially take public transit if it were efficient and connected to their place of work. Increased public transit ridership over car travel also goes a long way towards reducing local air pollution and carbon emissions. The results of the first two parts of this project will be informative on a broader scale for how the city has approached transit service to date and how it wants to move forward. The results of the final part could have more immediate benefits in the form of efficient and effective transit service for those who serve to benefit the most.