### **Write-up (30 points)**

* Introduction: The introduction provides a clear explanation of the question and the dataset used to answer the question, including a description of all relevant variables in the dataset. (3 points)
* Justification of approach: The chosen approach and visualizations are clearly explained and justified. (3 points)
* Code: Code is correct, easy to read, properly formatted, and properly documented. (10 points)
* Visualization: The visualizations are appropriate, easy to read, and properly labeled. (10 points)
* Discussion: Discussion of results is clear and correct, and it has some depth without begin excessively long. (4 points)

Introduction

In this project, we utilize interactive and animated spatio-temporal visualization to illustrate how the composition of non-emergency requests differ across communities with varying sociodemographic characteristics in the Chicago 311 Service Requests data.

The 311 system is a non-emergency response system where people can make a request to find information about services, make complaints, or report non-emergency problems, such as potholes and trash collection. While the system was initially designed to reduce call volume on the overloaded 911 system, 311 request systems have become an integral part of the e-government movement in which technological innovations are deployed to help local governments deliver more efficient and effective services to residents. Thus, we employ the 311 data in Chicago to provide insights on the variation of communities’ needs and assist the city to better allocate resources accordingly. We incorporate the demographic distribution or socioeconomic measures in order to determine if areas with unusually high or low requests for different services may correlate to a certain distribution of the area.

We focus on two measures of interest regarding the 311 data: number of requests and the amount of time it takes to complete a request. While the first measure informs the demand of non-emergency services, the second measures reflects on the quality of responses from the city to its residents. We use zip codes to identify our geographical areas and merge in socio-demographic information from the American Community Survey (ACS) 2019.

For our analysis, we use the data on 311 Service Requests received by the City of Chicago that are publicly available on the Chicago Data Portal. The dataset includes requests created after the launch of the new 311 system on 12/18/2018 and was last updated on May 11, 2022. Currently, it has 6 million rows and 37 columns, where each row is a request. Useful features from the data include request type, owner department, create date, closed date, and zip code. Since we are interested in the response time, we restrict observations to requests that have been completed. It is noted that the address for requests of the type “311 INFORMATION ONLY CALL” is often the address of the City’s 311 Center. See the codebook in data folder for all variables of interest.

For demographic characteristics, we use data from The American Community Survey (ACS) that is a questionnaire conducted by the United States Census Bureau yearly to collect information about American citizens. Relevant sociodemographic elements were selected from this survey in the year 2019, and converted into a workable dataset using tidyCensus. Our subset of data includes 24 columns and 296 rows correlating to Chicago and other outlying areas included in the Chicago Metropolitan Statistical Area (MSA). We limit the data to observations with zip codes that match with the zip codes included in Chicago 311 dataset. Among the selected variables include factors relevant to race, education, age, gender, and socioeconomic status. While the already selected variables are robust, the many questions included in the original survey allow for us to add or reduce the number of variables included as necessary. See the codebook in the data folder for all variables of interest.

Approach:

We believe that spatial visualizations are best suited to examine volume of different 311 service request types, response time to address those requests and varying sociodemographic characteristics across different areas of Chicago. To aid in illustrating the characteristics of a selected service domain, we use additional basic ggplot2 visualizations i.e., bar charts to show the distribution of different request types within a domain/department for the top and bottom quartile of the selected demographic group. Switching to a different demographic group updates these bar charts along with the choropleth for the share of the selected sociodemographic group in different areas.

This helps to show the relation between the department with initial responsibility for the service request or service type and socioeconomic indicators. For example, there is a positive correlation between influx of a DOB – Building service request and the share of native population. The bar charts help look at this relation in a more clear way as we are able to see the volume and request time for each request type for both the top and bottom quartile of selected characteristic. The comparison becomes much easier and detailed.

We add an element of interactivity to this analysis using Shiny app. The app provides the feature to toogle certain categories of service requests, as well as select a community area in order to see the distribution of service requests and the demographic patterns associated with the frequency of requests in the area.

Discussion: