Determining Covid Similarities in Orange County Neighborhoods

Ronnie Howard

***Abstract*—**TODO: For now see the Introduction and Data sections below

1 Introduction

In this project I have decided to look at the problem of COVID-19 outbreaks in the different neighborhoods of Orange County. I will define a **neighborhood** as each area in Orange County, Florida that has a unique zip code. Interestingly, a quick glance on zip-codes.com [1] shows that there are over 50 zip codes in the Orange County area. The plan is to cluster these neighborhoods into like neighborhoods. I will then use a Chloropleth map to show how severe the outbreaks are in each neighborhod. My hope is that similar neighborhoods will show similar amounts of COVID-19 cases so that more research can be done to determine what in particular makes a certain neighborhood have more cases than others. The **null hypothesis** is that there is no relationship between similarly clustered neighborhoods and the number of outbreaks that occur. The **alternative hypothesis** is that similarly clustered neighborhoods will have similar outbreaks of **COVID-19.**

1.1 Audiences

This is important to a variety of different audiences and I will discuss a few now. The first audience that this problem affects is the Department of Health, along with other government officials such as the mayor. This stakeholder is responsible for tracking outbreaks and quarantining different areas based on how bad the pandemic is. This stakeholder will be interested in my analysis because if my hypothesis is correct, it will allow them to zone in on neighborhood clusters that have the most outbreaks. They could use my analysis to begin further research into determining what makes some neighborhood clusters for susceptible than others.

Another audience is the general population. Most people want to avoid areas that have high concentrations of outbreaks. My analysis would tell them specific types of neighborhoods that they should avoid.

Finally, I want to mention that the leading theory on outbreaks is that higher populated areas will have higher outbreaks. I agree that this is most likely true. However, my

analysis goes a bit deeper because Orange County is already a largely populated county in Florida. Orange County includes Orlando which implies that Orlando's cases will most likely be higher than the rest of the county but further analysis of the types of neighborhoods may still be important.

2 DATA

There are two main sources where I plan to gather my data but time will tell if I need more as I continue on the project.

1. The Florida Department of Health maintains an open database of known cases of COVID-19 based on zipcode. This dataset can be accessed through a RESTful API. I will use this database to gather the zipcodes of Orance County and also to gather the number of reported outbreaks based on that zipcode. The COVID-19 data will be used to create a Chloropleth map based on outbreaks. An example of their dataset can be seen below.

OBJECTID	condigo postal	OBJECTID_1	DEPCODE	COUNTYNAME	FieldMatch	PONname	Places	OB
402	32801	760	48	Orange	Orange-32801	Orlando	Orlando	714
403	32803	761	48	Orange	Orange-32803	Orlando	Orlando, Winter Park, Fairview Sh...	715
404	32804	762	48	Orange	Orange-32804	Orlando	Orlando, Fairview Shores	716
405	32805	763	48	Orange	Orange-32805	Orlando	Orlando, Holden Heights	717
406	32806	764	48	Orange	Orange-32806	Orlando	Belle Isle, Edgewood, Orlando, C...	718
407	32807	765	48	Orange	Orange-32807	Orlando	Orlando, Winter Park, Azalea Park	719
408	32808	766	48	Orange	Orange-32808	Orlando	Orlando, Fairview Shores, Lockha...	720

Figure 1—FDOH data on COVID-19 cases based on zipcodes in Orange County, FL [2]

2. I will use the **Foursquare API** to gather venue information within each of the aforementioned zipcodes of Orange County, FL. This data will allow me to cluster the neighborhoods in Orange County based on each neighborhood's venues. Below is an example of the type of data Foursquare provides. *Note that this data comes from Ontario Canada and is shown as an example and does not represent Orange County Florida. I will update this section with Orange County Florida Data when I complete my project.*

	Neighborhood	Accessories Store	Airport	American Restaurant	Art Gallery	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Bakery	Bank	...	Supermarket	Supplement Shop	Sushi Restaurant	Tea Room
0	Parkwoods	0	0	0	0	0	0	0	0	0	...	0	0	0	0
1	Parkwoods	0	0	0	0	0	0	0	0	0	...	0	0	0	0
2	Parkwoods	0	0	0	0	0	0	0	0	0	...	0	0	0	0
3	Victoria Village	0	0	0	0	0	0	0	0	0	...	0	0	0	0
4	Victoria Village	0	0	0	0	0	0	0	0	0	...	0	0	0	0

5 rows x 103 columns

Figure 2—An example of the data Foursquare will return [3]

3.

3 REFERENCES

1. Datasheer, LLC. (n.d.). Zip-Codes.com. Retrieved June 15, 2020, from https://www.zip-codes.com/county/fl-orange.asp
2. FDOH. (2020). Florida COVID19 Cases by County. Retrieved June 15, 2020, from https://open-fdoh.hub.arcgis.com/datasets/florida-covid19-cases-by-county/geoservice
3. Foursquare. (2019). Foursquare Developer. Retrieved June 15, 2020, from https://developer.foursquare.com/
4.