

A Social Network Analysis on the Impact of Zip Code on COVID-19 Transmission



Jay Mantuhac
SOCIOLOGY 280
November 30, 2021



1

Introduction

Why zip codes?

HEALTH • PUBLIC HEALTH

Your ZIP Code Might Determine How Long You Live—and the Difference Could Be Decades

> [Breastfeed Med.](#) 2016 Oct;11:396–7. doi: 10.1089/bfm.2016.0113. Epub 2016 Aug 11.

Why Your ZIP Code Matters More Than Your Genetic Code: Promoting Healthy Outcomes from Mother to Child

Story from **Blue Cross Blue Shield Association** ⓘ

Up to 60% of our health is determined by zip code

Why zip codes?

TABLE 1 | Comparison of COVID-19 infection rate per 10,000 persons by sociodemographic characteristics in 108 zip codes^a in Indiana, United States.

	Median infection rate in zip codes above median demographic	Median infection rate in zip codes below median demographic	Median infection rate of zip codes with available demographic data (Reference)	<i>P</i>-value	Pearson correlation coefficient	<i>P</i>-value
African American residents	17.4	6.7	9.9	<0.0001	0.67	<0.0001
Hispanic residents	15.9	7.0	9.6	<0.0001	0.59	<0.0001
Foreign-born residents	13.5	6.4	9.6	<0.0001	0.52	<0.0001
Geographic density ^b	15.0	4.5	9.4	<0.0001	0.68	<0.0001
Living in poverty	13.5	8.5	9.6	<0.0001	0.49	<0.0001

(Hansen et al., 2020)



2

The Data and the Network



A Brief Timeline of the Start of COVID-19 Pandemic (AJMC, 2021)

- Jan 21, 2020: First confirmed case of COVID-19 in the US
- March 11: WHO Declares COVID-19 a Global Pandemic
- March 19: CA declares its first Stay at Home order (i.e. “lockdown”)



Data Source

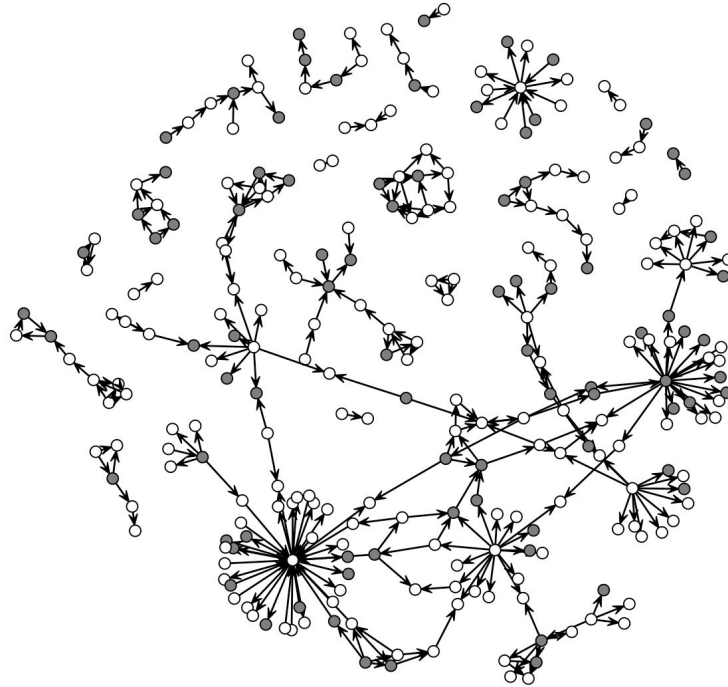
- MicrobeTrace, Centers for Disease Control
 - Software used to visualize network data and calculate network-level statistics
- Represents individual level COVID-19 transmission data from an outbreak across 11 zip codes in Georgia in March 2020
 - Unfortunately, the COVID-19 dataset is **dummy data** and does not represent an actual individual-level COVID-19 outbreak



Basic Network Statistics

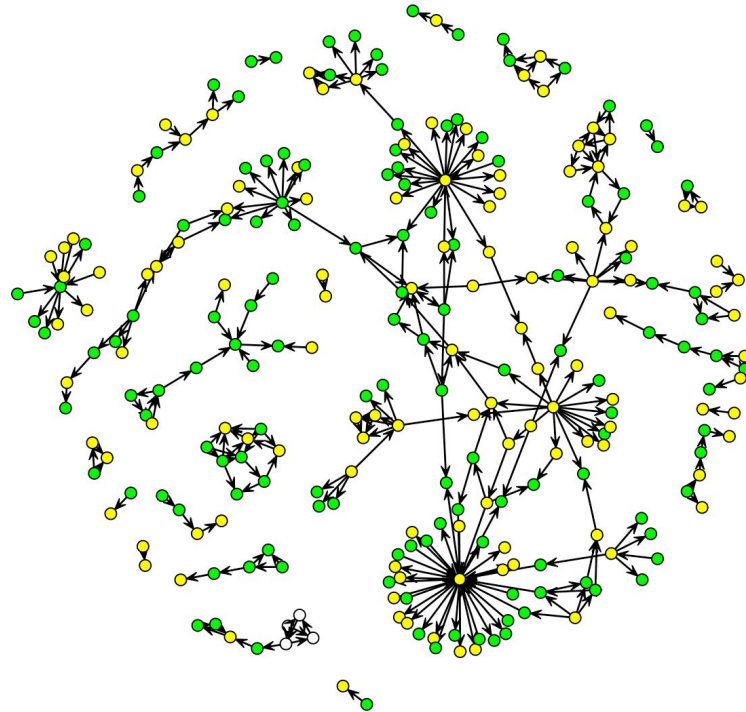
- Directed Network
- 263 Nodes, 297 Directed Ties
- Demographics
 - 85 Black, 172 White
 - 120 Female, 139 Male
- Some missing data points

Sociogram by Race/Ethnicity



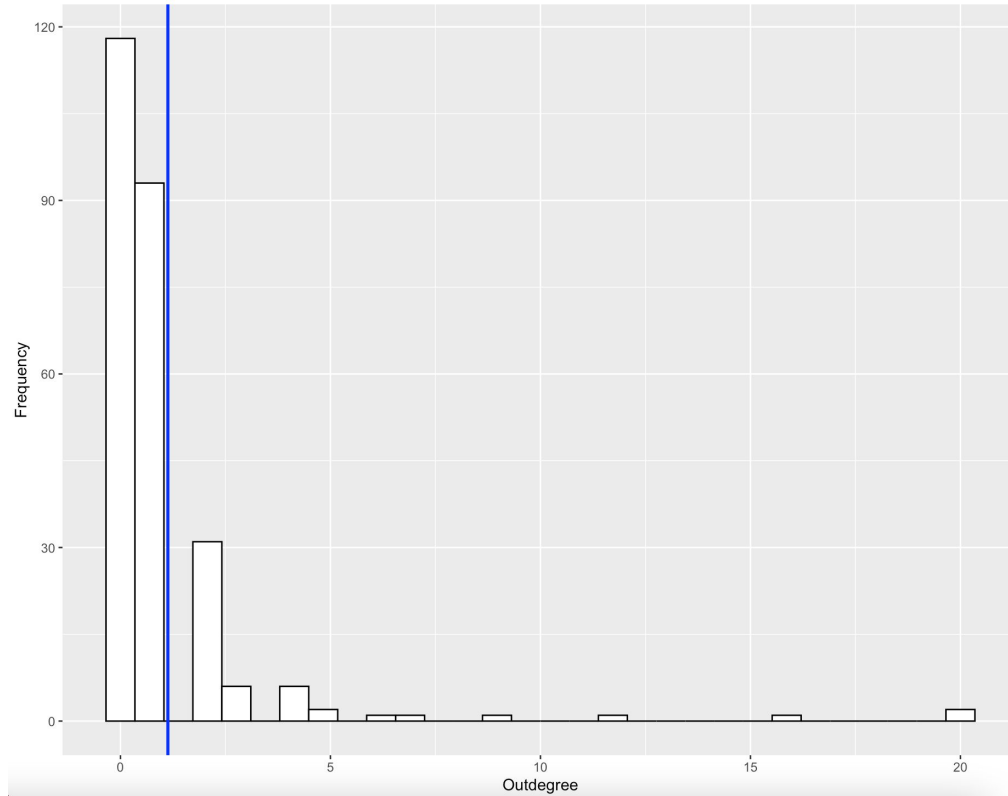


Sociogram by Gender





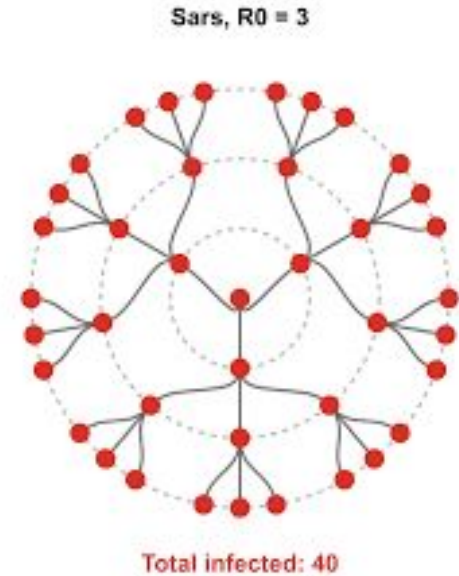
Outdegree Distribution





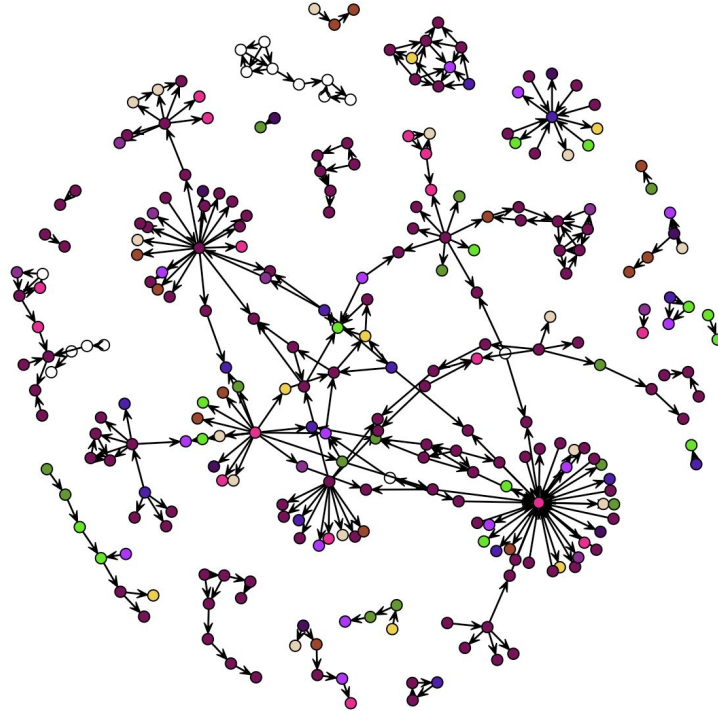
Side Note: R_0 Value

- If you get infected with a disease, how many people are you expected to infect?
 - Dependent on external factors
- In the context of SNA, $R_0 \sim$ a node's outdegree





Sociogram by Zip Code





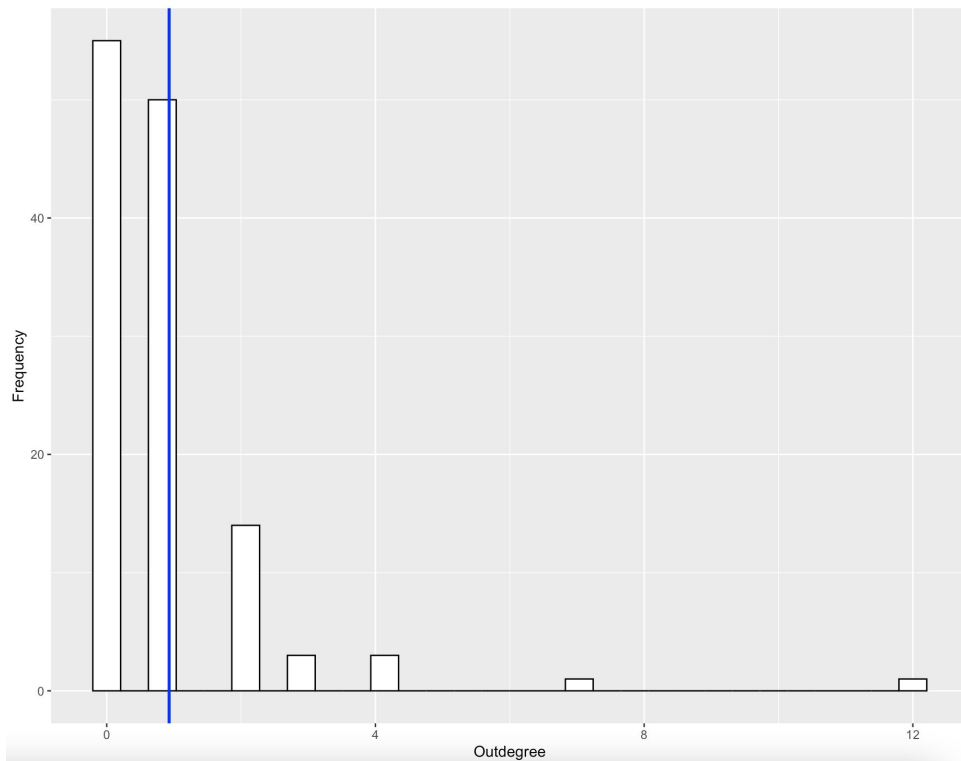
3

The Scientific Question of Interest

“

*Do we have evidence that zip
code demonstrates a
homophilous relationship in a
simulated COVID-19
outbreak?*

Outdegree Distribution of Zip Code 30338





The EI Index and Odds Ratio

- Baseline $E-I^*$ of 0.26
 - Implies there's some degree of homophily in our network at baseline
- Observed OR for a Homophilous Tie: 1.71



ERGM Development (Still in Development)

```
Call:
ergm(formula = net2 ~ edges + nodefactor("zip"))

Maximum Likelihood Results:

              Estimate Std. Error MCMC % z value Pr(>|z|)
edges          -5.43579    0.33952     0 -16.010 < 1e-04 ***
nodefactor.zip.30306 -0.55376    0.29798     0  -1.858 0.063119 .
nodefactor.zip.30324  0.74948    0.20758     0   3.611 0.000306 ***
nodefactor.zip.30329 -0.12620    0.25624     0  -0.492 0.622377
nodefactor.zip.30338  0.02127    0.18194     0   0.117 0.906936
nodefactor.zip.30340 -0.36902    0.27459     0  -1.344 0.178994
nodefactor.zip.30341 -0.57079    0.29241     0  -1.952 0.050936 .
nodefactor.zip.30602 -0.28145    0.33698     0  -0.835 0.403596
nodefactor.zip.30605 -0.46488    0.28275     0  -1.644 0.100150
nodefactor.zip.32606 -0.68912    0.39399     0  -1.749 0.080272 .
nodefactor.zip.35064 -0.12620    0.30439     0  -0.415 0.678446
nodefactor.zip.37116  0.43924    0.22441     0   1.957 0.050312 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null Deviance: 95524 on 68906 degrees of freedom
Residual Deviance: 3758 on 68894 degrees of freedom

AIC: 3782 BIC: 3892 (Smaller is better. MC Std. Err. = 0)
```



4

Final Thoughts



Discussion + Future Research Steps

- ◆ Zip code appears to demonstrate a homophilous relationship
 - ◇ Infections are highly dependent on where individuals are
- ◆ How do other factors play a role in transmission?
 - ◇ e.g. Race/Ethnicity, Gender



Limitations

- Use of dummy data!
 - The data are based out of infections that would have occurred very early on in the pandemic
- Dataset only captures a limited number of variables



Next Steps (for this project)

- Continue to develop the ERGMs + fit a model that best fits the ability to predict an infection
 - Could other factors that underlie the zip code influence transmission dynamics?
- Find out how the dummy dataset was generated



References

- Ducharme, J., & Wolfson, E. (2019, June 17). *How Your Zip Code Could Affect Your Lifespan—And the Difference Could be Decades*. Time.Com. <https://time.com/5608268/zip-code-health/>
- Graham, G. N. (2016). Why Your ZIP Code Matters More Than Your Genetic Code: Promoting Healthy Outcomes from Mother to Child. *Breastfeeding Medicine*. <https://doi.org/10.1089/bfm.2016.0113>
- Hanson, A. E., Hains, D. S., Schwaderer, A. L., & Starr, M. C. (2020). Variation in COVID-19 Diagnosis by Zip Code and Race and Ethnicity in Indiana. *Frontiers in Public Health*, 8, 593861. <https://doi.org/10.3389/fpubh.2020.593861>
- Morgan, K. (2019, April 22). *Story from Blue Cross Blue Shield Association: Up to 60% of our health is determined by zip code*. USA TODAY. <https://www.usatoday.com/story/sponsor-story/blue-cross-blue-shield-association/2019/04/22/up-60-our-health-determined-by-zip-code/3542001002/>