STUDYING THE RELATIONSHIPS BETWEEN FOOD ACQUISITION AND HEALTH OUTCOMES DURING COVID-19

Chris Carey, Maia Guo, Nuoyi Wang

New York University, Center for Urban Science + Progress

ABSTRACT:

The COVID-19 pandemic had a complex impact on food acquisition behaviors in New York City (NYC). Pre-existing disparities in food access and food supply resiliency compounded negative health risks for certain populations by exacerbating food insecurity, increasing dependencies on unhealthy food, and creating disproportionate risks of COVID-19 infection during food acquisition. Mobility data of trips to food retail and service locations were used to supplement existing survey data to model and analyze changes in food acquisition in NYC during the pandemic. Demographic and socioeconomic clustering was employed in time series analyses of metrics modeling dietary changes and COVID-19 infection risk. These analyses identified neighborhoods and food location categories with the greatest need for more resilient food access and supply. While direct connections to health outcomes were not established in this paper, the produced measures identified the locations and relative degrees of disparities in dietary changes associated with negative health outcomes. Therefore, combined with community and expert knowledge, such a holistic model can inform the most effective means to direct equitable policy-driven improvements to the food system resilience of NYC, in preparation for future public health emergencies.

1 INTRODUCTION

1.1 Problem Definition

During the COVID-19 pandemic, vulnerable populations in New York City (NYC) had excessive difficulty in safely obtaining affordable and nutritious food as typical avenues for food access were massively disrupted for sustained periods of time. Access to healthier food options changed and food insecurity increased by 67% to two million NYC residents (NYC Food 20/20, 2020). Public health dangers and policy responses to COVID-19 exacerbated inequitable food access, and thus disproportionately increased food acquisition patterns associated with negative health outcomes. To correct this, lessons from the pandemic must inform policy responses to maintain food access during future public health emergencies.

1.2 Project Scope

Mobility data were used to supplement survey data to investigate changes in food acquisition and dietary behaviors in NYC during the pandemic. Direct links to health outcomes could not be established, due to the long time horizons for diet-related health outcomes to manifest. Therefore, the existing scientific basis of associations between food insecurity, nutrition, and diet-related diseases was relied upon.

1.3 Literature Review

1.3.1 Food Acquisition and COVID-19 The COVID-19 pandemic significantly increased food insecurity and changed food acquisition and consumption behaviors in NYC, including store trip frequency and fruit and vegetable intake (Litton and Beavers, 2021). The extent of these behavior changes varied by geographic region (Ellison and Ocepek, 2021), in regards to localized business operating decisions, employment disruption, and COVID-19 infection rates. NYC's food system was impacted in three critical ways: (1) school and food service closures reduced the availability of prepared food; (2) employment disruption reduced spending budgets and overwhelmed food benefit programs (NYC Food 20/20, 2020); (3) increased home meal preparation led to mixed consequences for diets, household expenditures, and the food retail sector (Ellison et al., 2021). Pre-existing social inequities disproportionately burdened those limited in food security and access (Cohen and Freudenberg, 2020).

1.3.2 Mobility Data During the COVID-19 pandemic, researchers investigated mobility patterns through mobility data provided by SafeGraph Inc. Initial studies focused on understanding localized COVID-19 spread by examining social distancing compliance by demographic group (Egorov et al., 2021); inferring the effects of "superspreader" events (Dave et al., 2020); examining influences of school reopenings (Courtemanche et al., 2021); and investigating social distancing behaviors in relation to xenophobia and political partisanship (Coston et al., 2021).

Research on food acquisition and health-related issues incorporating SafeGraph data has been growing. Ashby (2020) and Banerjee et al. (2021) found that unhealthy eating establishment visitations increased disproportionately in populations with high obesity rates and in rural counties during the pandemic, respectively. Furthermore, Kar et al. (2021) found that access to high-quality food by residents of low-income neighborhoods and food deserts in Columbus, Ohio became further constrained during the pandemic.

Recent studies also incorporated SafeGraph data to investigate crowding and COVID-19 spread. Hamidi and Hamidi (2021) and Verma et al. (2021) both found positive associations between crowding at points of interest (POIs) and localized COVID-19 case growth in NYC, with higher contact densities observed at food locations. The association between public transportation and COVID-19 spread was contested however, as Hamidi and Hamidi (2021) found no evidence of an association with subway ridership, whereas Chen et al. (2020) found a positive correlation through a geographically-weighted regression analysis.

2 DATA

2.1 Data Sources

Mobility and census datasets prepared by SafeGraph Inc. were used in this paper including: **Weekly Patterns** with attributes such as destination POIs, visit counts, and dwell times; **Core Places** with basic location, category, and brand information; and **2019 5-year American Community Survey (ACS) Data** with demographic and socioeconomic characteristics such as income and education level. The authors note that ACS data may not reflect recent conditions. Additionally, weekly trips were aggregated by destination POIs, which limited network and pattern analysis. However, data were analyzed with census block group (CBG) level granularity.

2.2 Data Curation

2.2.1 Data Pipeline SafeGraph data records were stored on the Hadoop Distributed File System of CUSP High-Performance Computing (Figure 1). These 470 GB of data were then filtered, re-coded, and aggregated with PySpark (1) into two tables (Appendix B, Table B.1, B.2). A POI table included attributes of 36,195 food retail and service POIs in NYC categorized by North American Industry Classification System (NAICS) code. A weekly trip table included 3,664,420 trip records to those POIs from December 2018 to February 2021. These tables were combined with census and geographic data for spatiotemporal analysis (A). A database and query processing server (2) enhanced the performance of a custom-built web application (3-4) which provided interactive analysis (B) to further direct analysis and modeling.

2.2.2 POI Categorization Only visitations to POIs under food retail or service NAICS code business categories were considered (Appendix B, Table B.3). Three POI categories were selected to assess nutritional changes in diets (Table 1). The authors considered these categorization criteria to be the best available approach to identify less healthy food locations at scale.

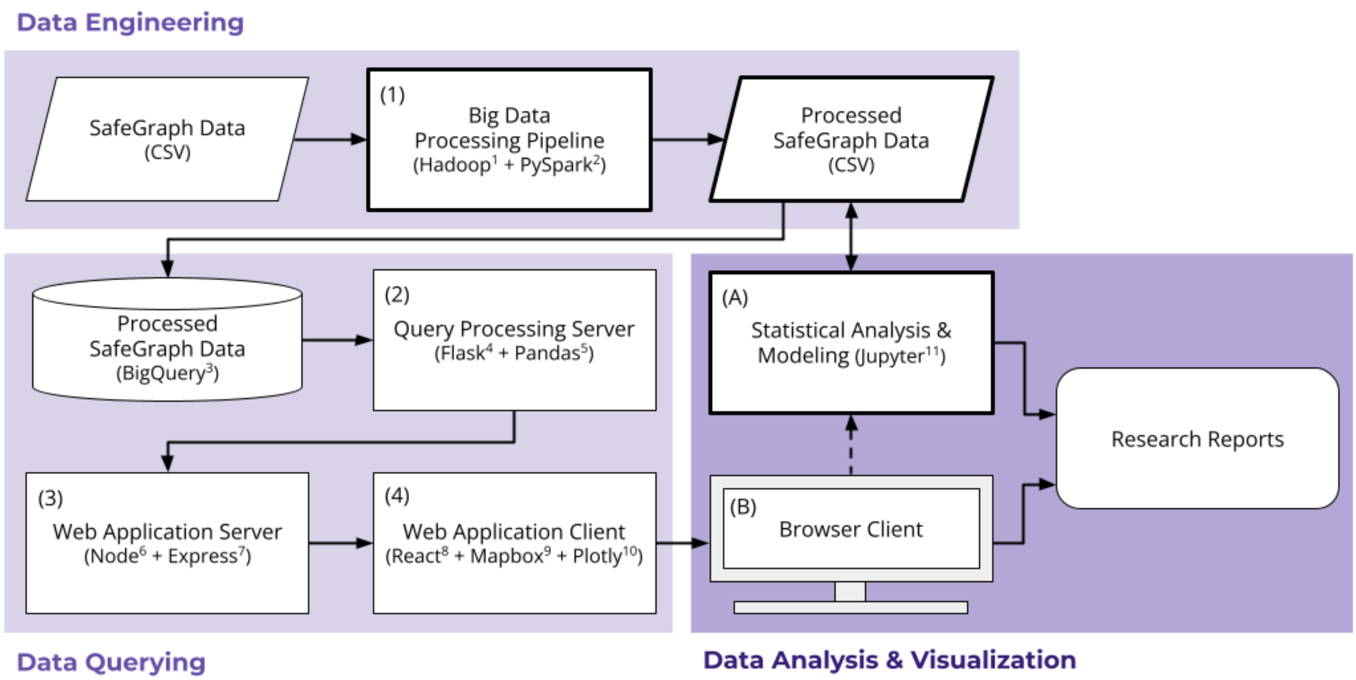

Figure 1. System design diagram. Technical footnotes in Appendix E.

Category	Nutritional Value	Rationale	Categorization Criteria
Supermarkets and Grocery Stores	Healthy	Offered greatest fresh produce (Morland et al., 2002)	NAICS Code
Delis and Convenience Stores	Less Healthy	Offered mostly packaged and processed food (NYC Food 20/20, 2020)	NAICS Code (Convenience Store) or "Deli" within POI name
Fast-Food Restaurants	Unhealthy	Widely considered unhealthy (Pereira et al., 2005; Stender et al., 2007)	POI name matched against set of fast-food chain restaurant names (Appendix B, Table B.4)

Table 1. Selected POI categories assessed for overall nutritional value.

2.3 Definitions

The following terms and definitions are used throughout paper for brevity:

- *Delis*: POIs categorized as Delis and Convenience Stores
- *Supermarkets*: POIs categorized as Supermarkets and Grocery Stores
- *Visitations*: A visitation to a food retail or service location destination as reported by SafeGraph
- *Home CBG*: The origin CBG of a visitation trip

3 METHODOLOGY

3.1 Analysis Metrics

The following metrics were developed to model the disparity in health impacts resulting from changes in food acquisition behaviors. These included both the long-term effects on dietary changes induced by food insecurity, and the short-term risks of COVID-19 infection during food acquisition trips.

3.1.1 Estimated Visitor Count Because SafeGraph device penetration rate may have varied across subpopulations, the weekly raw visitors count for each home CBG was adjusted by the ratio of device count to CBG population (Appendix D, Equation 1).

3.1.2 POI Category Visitation Proportion The proportion of visitations made to a particular POI category was the primary metric to model dietary dependency on a particular POI category and by extension, the overall nutritional value of visitors' diets (Appendix D, Equation 2). While the amount of food purchased per visitation varied, the authors considered visitation proportion the best metric available to model food acquisition and dietary changes at scale.

3.1.3 Contact Density Index (CDI) To evaluate disproportionate COVID-19 exposure risks encountered during food acquisition, an aggregated spatiotemporal contact density index (CDI) was developed (Appendix D, Equation 3). CDI estimated the amount of contact exposure to other individuals based on POI visitor count, floor area, and dwell time. This paper differed from Verma et al. (2021) by producing a CDI traceable to a home CBG, but with decreased accuracy in modeling simultaneous exposure.

3.2 Analytic Approaches

3.2.1 Population Clustering Food acquisition behaviors can vary by income, age, household size, cultural and ethnic backgrounds, and other demographic and socioeconomic attributes. Therefore, Principal Component Analysis (PCA) and *k*-means clustering were performed with 2019 trip and census data aggregated by CBG to produce subpopulations with similar characteristics (Appendix C, Table C.1). This enabled exploring spatially-related behaviors while avoiding privacy concerns surrounding location specificity. The number of clusters to form was determined through the Elbow method and result interpretability, leading to four groups of CBGs.

3.2.2 Time Series Analysis POI category visitation proportions and CDI were compared over time periods by cluster and home CBG income percentile. Rolling averages and seasonality alignments were applied.

3.2.3 Data Visualization An interactive web application was built to visualize metrics by CBG over selected time periods. Query results were displayed in a choropleth integrated with distribution and segmentation plots.

3.2.4 Limitations All analyses were limited by potential disparities in modeling real-world behavior through mobility data which were subject to various privacy-based aggregation techniques, device penetration rates, and potential model bias. Reported metrics in this paper were not adjusted for spatial or temporal autocorrelation.

While *k*-means clustering was a suitable method for population grouping, it approached model accuracy limitations with higher data dimensionality. PCA was employed to reduce dimensionality, but resulted in reduced interpretability. The final clustering method used balanced accuracy and result interpretability with the recognition of its limitations in modeling real-world behavior.

3.3 Data Ethics and Impact Considerations

While mobile location data have been utilized by scientists to study infectious disease spread, a key concern is the legitimization of surveillance tools to monitor mobility behaviors after pandemics have ended (Oliver et al., 2020). In the absence of a standardized privacy-conscientious framework for mobile phone data, the privacy protections provided by SafeGraph Inc. were relied upon.

SafeGraph data used in this paper were open and accessible to all public researchers. As opposed to cellular data records or GPS trajectories which can easily re-identify individuals (De Montjoye et al., 2013), SafeGraph data prohibited itinerary reconstruction by only exposing anonymized visitation counts. Additionally, spatial and temporal coarsening, minimum device count thresholding, differential privacy techniques, and Lapacian noise were added by SafeGraph (SafeGraph Inc., 2021).

Of additional importance was group privacy. Aggregated data might reveal sensitive mobility behaviors of groups and their members, including those characterized by protected classes. Therefore, groups were analyzed at coarse aggregation levels which combined multiple demographic and socioeconomic factors to describe impacts to minority groups, without analyzing such groups in isolation. Nevertheless, any resulting data-driven policy recommendations should be supplemented with community and expert knowledge to include human-driven perspectives, identify biases, and guard individual and group privacy.

4 RESULTS

4.1 Analysis Results

4.1.1 Visitation Totals Food location visitations in NYC sharply dropped after the initial stay-at-home order, before partially recovering during phased re-openings over the summer (Figure 2).

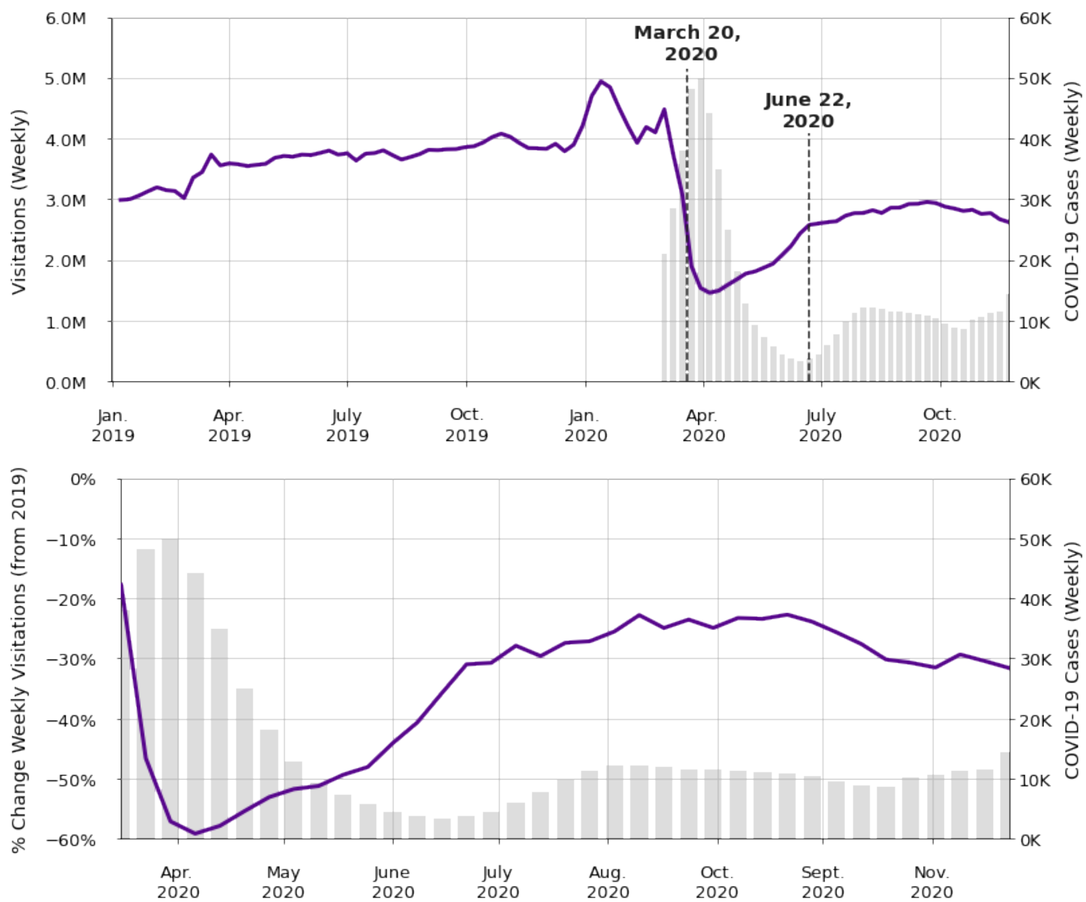

Figure 2. Three-week average of COVID-19 cases in NYC and weekly food location visitations (top) and percent change in weekly food location visitations from 52 weeks prior (bottom). Dates of initial stay-at-home order and Phase 2 re-opening marked (Details in Appendix A).

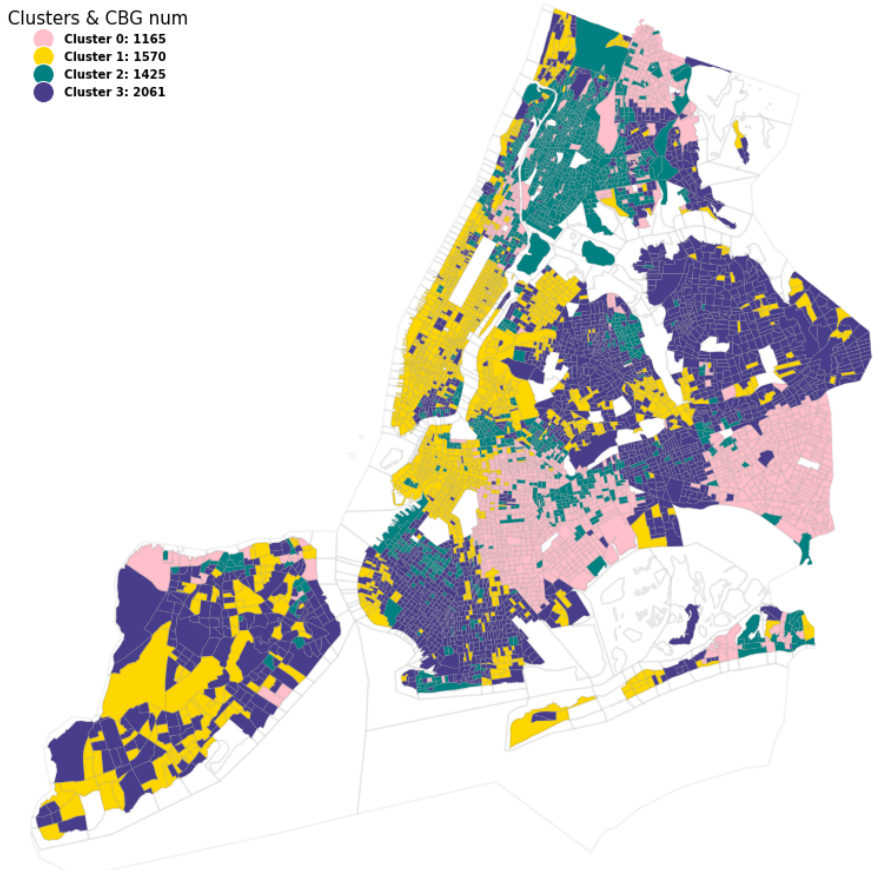

Figure 3. Clustered CBGs on the map (excluded CBGs with population under 10).

	Income Level	**Education Level**	**Families with Elderly & Children**	**Ethnic Composition**	**Families Receiving Food Assistance Benefits**	**POI Count**
Cluster #0	Moderate	Moderate	More elderly	More Black & African American	Moderate	More supermarkets, more fast-food
Cluster #1	Higher	Higher	Fewer	More White	Fewer	More food services, fewer supermarkets
Cluster #2	Lower	Lower	More children	More evenly-distributed	More	More supermarkets, more fast-food
Cluster #3	Moderate	Moderate	More elderly	More White & Asian	Moderate	Fewer supermarkets

Table 2. Demographic & socioeconomic characteristics and POI count of each cluster (detailed plots in Appendix C). POI count indicates the number of POIs of each category, determined by the unique placekey identifiers within the cluster area.

4.1.2 Clusters Clusters formed broad geographic regions (Figure 3) with distinct characteristics (Table 2).

4.1.3 POI Category Visitation Proportion In the immediate weeks surrounding the initial NYC stay-at-home order, food service visitation proportion dropped substantially before partially recovering over the summer (Figure 4). Changes in POI category visitation proportions varied by income percentile. Most notably, supermarket and deli visitation proportions generally increased by a greater amount for high-income CBGs (Figure 5). Additionally, fast-food visitation proportions increased for high-income CBGs and decreased for low-income CBGs. Clusters #0 and #2 maintained the overall highest supermarket and deli proportions whereas clusters #1 and #3 experienced the greatest initial increases. Clusters #0, #1, and #3 all temporarily increased fast-food proportions above pre-pandemic levels, with cluster #0 maintaining the highest overall value (Figure 6).

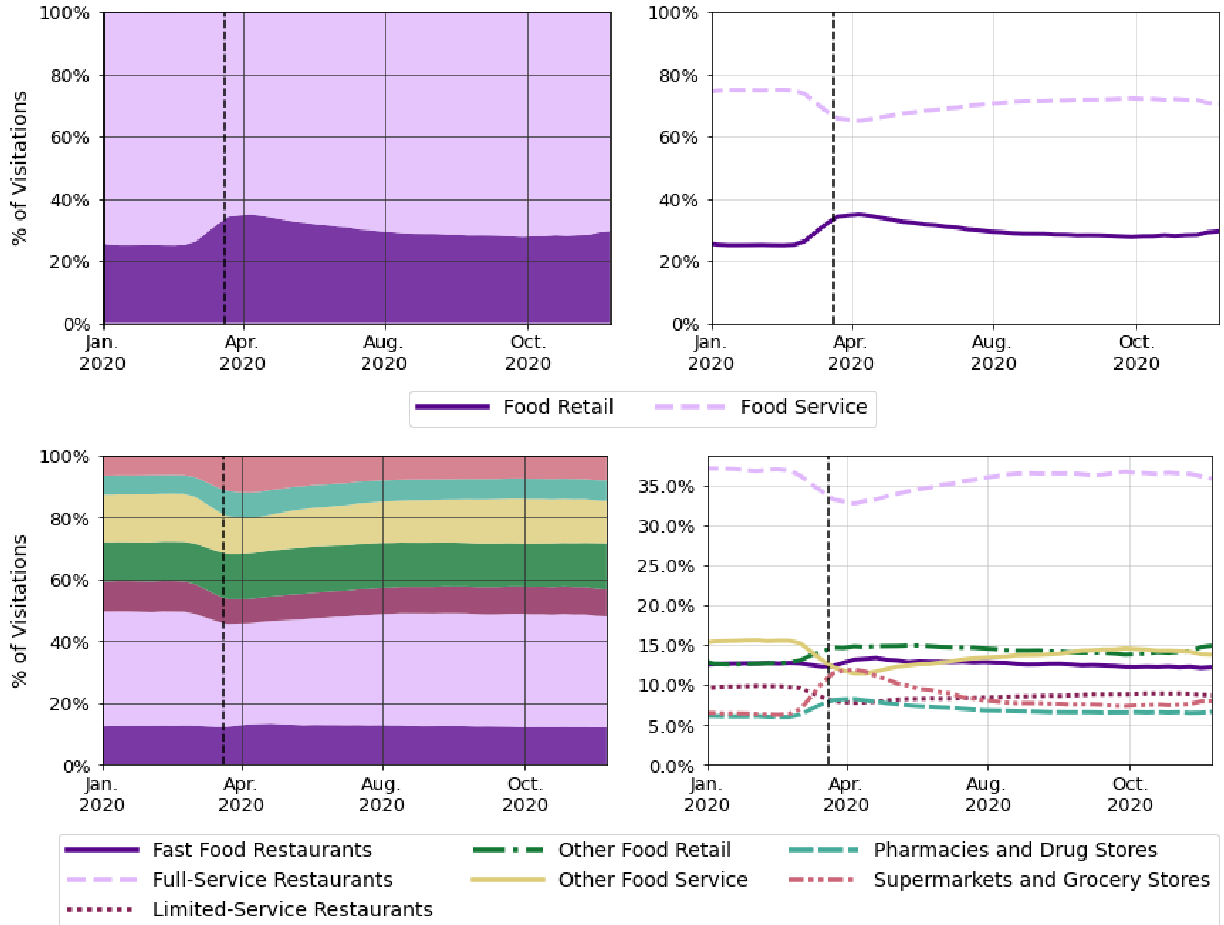

Figure 4. Weekly visitations to various POI categories over time (three-week average).

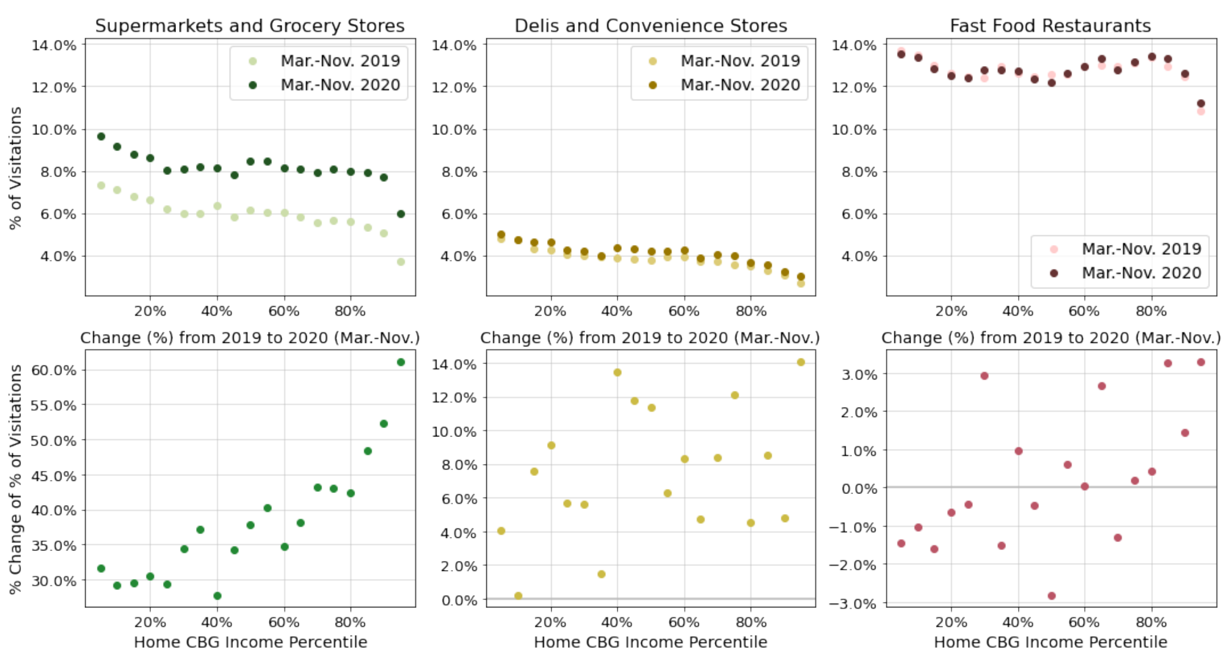

Figure 5. Selected POI category visitation distributions by home CBG income percentile: median percentage (top), change in median percentage (center), and percent change in median percentage (bottom).

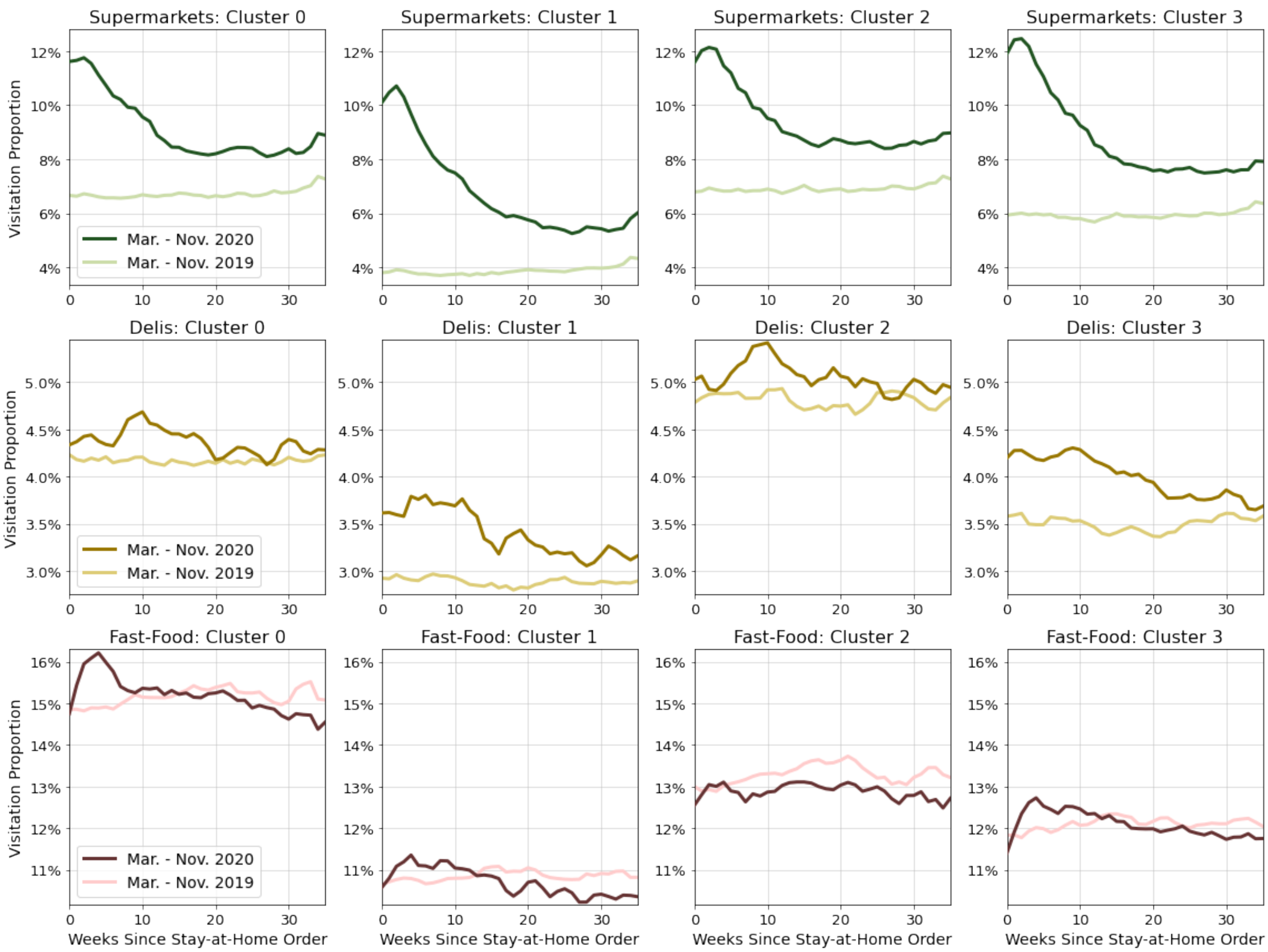

Figure 6. Selected POI category visitation distributions by demographic and socioeconomic cluster between March 16 - November 23, 2020 compared to 52 weeks prior: supermarkets (top), delis (center), and fast-food restaurants (bottom).

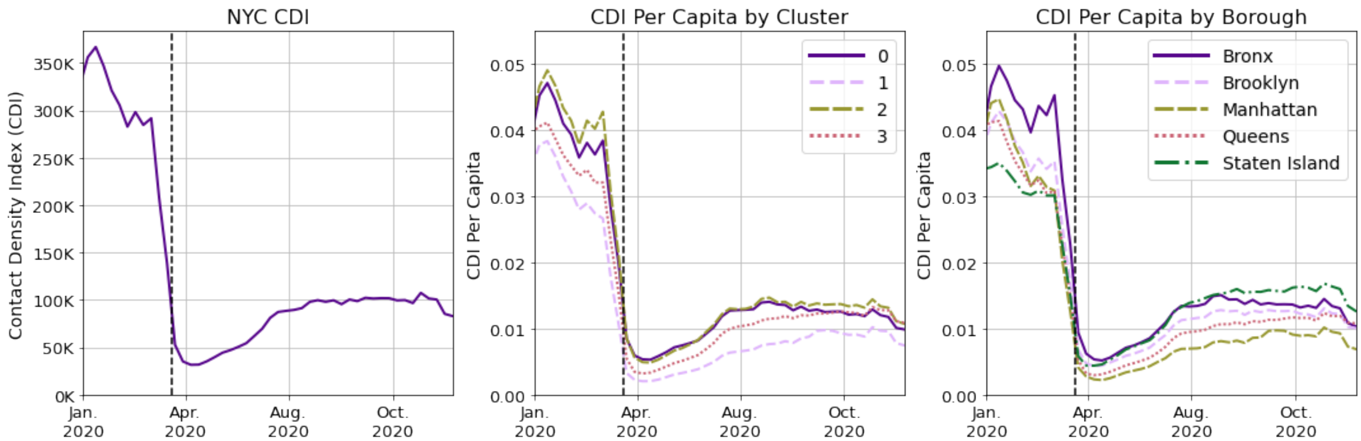

Figure 7. Three-week average CDI at food locations in NYC over time (left). Three-week average CDI per capita by cluster (center) and by borough (right) over time.

4.1.4 CDI After the initial stay-at-home order in NYC, CDI across all categories decreased (Figure 7) consistent with findings by Verma et al. (2021). Rankings of CDI per capita by home CBG cluster did not significantly change. However, Manhattan and Staten Island switched relative rankings and followed different distributions (99% CI).

4.2 Interpretation

Analyzed mobility data revealed trends consistent with survey data and offered additional insights into disparities in food retail dependency and resiliency, access to nutritious food, and contact density across geographic regions.

4.2.1 Consistency with Reported Food Acquisition Behaviors Visitation totals showed trends consistent with NYC survey data from May 2020 which reported that 64% of respondents shopped less frequently and 49% ate more packaged food compared to before the pandemic (CUNY School of Public Health, 2020); food retail visitation totals were 50% lower in May 2020 year-over-year, but formed a greater proportion.

4.2.2 Disproportionate Supermarket Resiliency Shoppers from low-income neighborhoods and clusters with higher rates of children and older residents had the greatest supermarket dependency. Low-income shoppers also reported the unavailability of essential and lower-priced food items at greater rates (CUNY School of Public Health, 2020) despite having the smallest increases in supermarket dependency. This suggested disparity in supermarket resiliency which exacerbated harm to high-poverty communities and individuals with diet-related diseases associated with food insecurity (Gundersen and Ziliak, 2015) who were already at greater risk from COVID-19 (Arasteh, 2021; Kimball et al., 2020).

4.2.3 Increased Dependencies on Unhealthy Eating Establishments Fast-food dependency temporarily increased citywide as non-fast food restaurant visitation proportions declined. CUNY School of Public Health (2020) surveys found that compared to pre-pandemic behavior: 54% of respondents reported having a less healthy diet, and lower-income, Black, and Latinx households consumed more packaged food at higher rates. These reports were consistent with observed fast-food and deli visitation proportions in total and by clusters.

4.2.4 CDI Disparity Followed COVID-19 Infection Disparity Clusters and boroughs with higher CDIs also included neighborhoods with the highest COVID-19 infections per capita (NYC Health, 2021). CDI model limitations prevented drawing policy recommendations specific to the food sector with confidence. However, CDI could potentially act as a broad diagnostic metric for outlier detection in future work.

4.3 Policy Implications

4.3.1 Preventing Food Supply Disruption Clusters #0 and #2 had the highest supermarket dependency, and thus were the most susceptible to food supply disruption when stay-at-home orders closed many food service options. This signaled that the greatest need to mitigate supermarket supply disruption was located in central and south Brooklyn, central and east Queens, and the Bronx.

A revised food resiliency strategy might seek to boost support for food banks and increase GrowNYC Greenmarket availability serving these neighborhoods; Greenmarkets located in Lower Manhattan and Downtown Brooklyn within cluster #1 were suspended due to low foot traffic (NYC Food 20/20, 2020). Expanding food supply in cluster #0 and #2 neighborhoods may be the most effective direction of food insecurity reduction efforts.

4.3.2 Supporting Non-Fast Food Restaurants After the initial stay-at-home order, most clusters and high-income neighborhoods increased their fast-food visitation proportions. These chain restaurants had the greatest financial resources to remain open. Meanwhile, over 1,000 NYC restaurants permanently closed with women and minority-owned businesses impacted disproportionately (NYC Food 20/20, 2020). Preserving greater nutritional choice to support stronger immune responses could be accomplished by ensuring immediate financial support to non-fast food services in future public health emergencies.

4.4 Visualization

Interactive choropleths enabled explorations of intra-cluster spatial distributions of visitations and CDI, such as investigating sharp increases in supermarket visitation proportions within cluster #0, or higher supermarket CDI in generally lower-income regions (Figure 8).

5 CONCLUSIONS

Survey data showed that shoppers from low-income and minority neighborhoods in NYC were more likely to experience supermarket food supply disruption, while mobility data showed that these same shoppers were also more dependent on supermarkets for a larger proportion of their diet. Therefore, a food resiliency strategy to reduce food insecurity during public health emergencies must be comprehensive and efficient. Such as strategy should expand food access in neighborhood clusters with the greatest supermarket dependency to prevent shortages, support non-fast food restaurants to ensure access to nutritious food, and boost financial support to emergency food programs serving low-income neighborhoods where supermarket resiliency is weaker.

The clustering techniques employed in this paper only provided guidance on broad geographic regions of NYC. Any data-driven policy strategy should be part of a holistic approach incorporating community outreach and expert knowledge. The rise in food insecurity in NYC and its associated negative health outcomes require persistent monitoring for adverse long-term health impacts, recognition of the disparity in the city's food resiliency system, and the collective support of New Yorkers to strengthen its weaknesses.

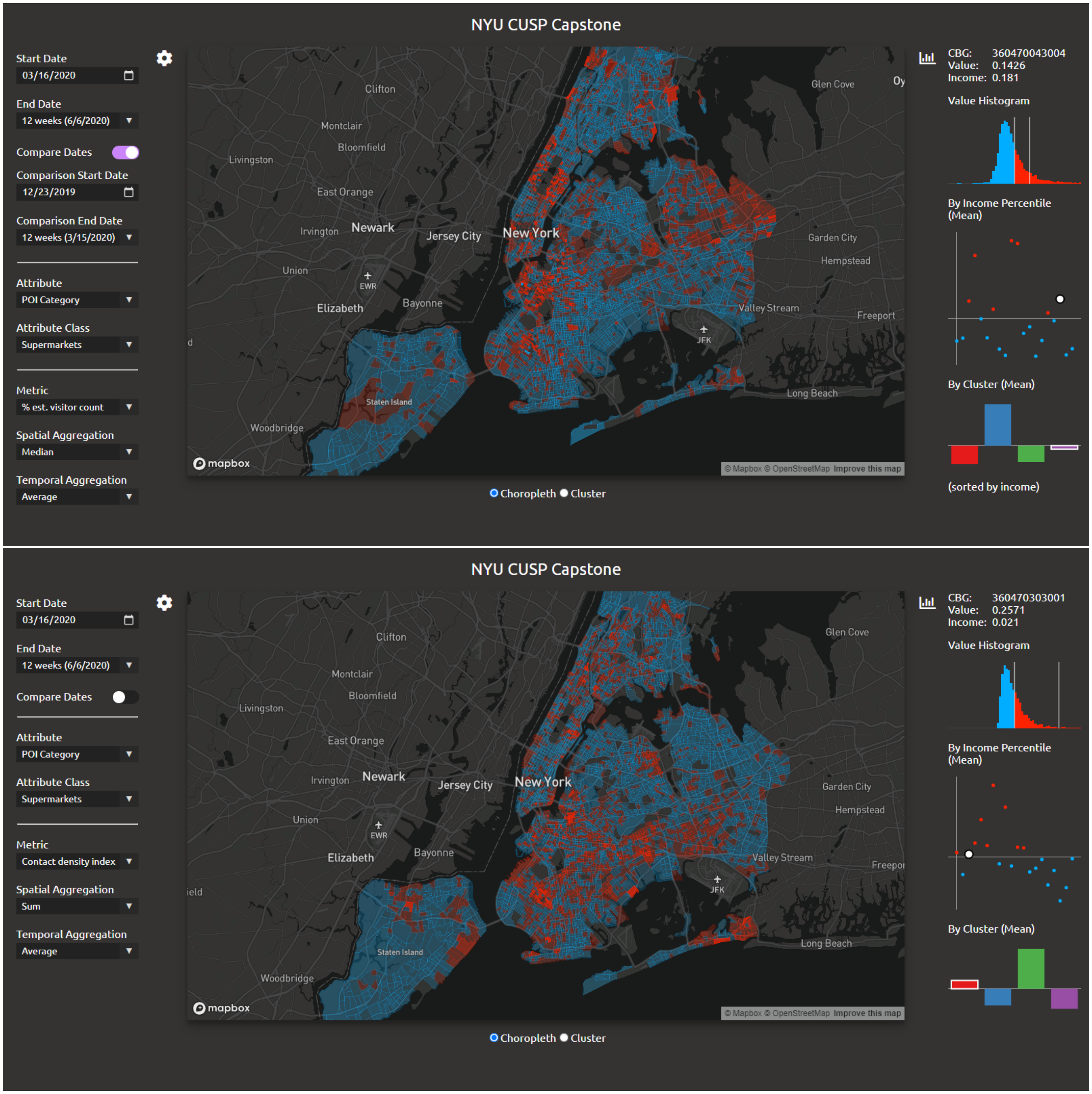

Figure 8. Change in mean weekly supermarket visitation proportion by home CBG between 12 weeks before and after March 16, 2020 (top); mean weekly supermarket CDI total by home CBG, March 16 - June 6, 2020 (bottom)

References

Arasteh, K., 2021. Prevalence of Comorbidities and Risks Associated With COVID-19 Among Black and Hispanic Populations in New York City: An Examination of the 2018 New York City Community Health Survey. *Journal of Racial and Ethnic Health Disparities*, 8(4), 863–869.

Ashby, N. J., 2020. Impact of the COVID-19 Pandemic on Unhealthy Eating in Populations With Obesity. *Obesity*, 28(10), 1802–1805.

Banerjee, T., Nayak, A., Zhao, H., 2021. A County-Level Study of the Effects of State-Mandated COVID-19 Lockdowns on Urban and Rural Restaurant Visits Using Consumers' Cell Phone Geo-Location Data. *Journal of Public Health*, 1–10.

Chen, Y., Jiao, J., Bai, S., Lindquist, J., 2020. Modeling the Spatial Factors of COVID-19 in New York City. *Available at SSRN 3606719*.

Cohen, N., Freudenberg, N., 2020. COVID-19's Effects on New York City's Food System: Lessons for Public Health Responses.

Coston, A., Guha, N., Ouyang, D., Lu, L., Chouldechova, A., Ho, D. E., 2021. Leveraging Administrative Data for Bias Audits: Assessing Disparate Coverage With Mobility Data for COVID-19 Policy. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 173–184.

Courtemanche, C. J., Le, A. H., Yelowitz, A., Zimmer, R., 2021. School Reopenings, Mobility, and COVID-19 Spread: Evidence From Texas.

CUNY School of Public Health, 2020. CUNY SPH COVID-19 Survey – Week 12 – Part 1: Food Insecurity.

Dave, D., Friedson, A., Matsuzawa, K., McNichols, D., Redpath, C., Sabia, J., 2020. Did President Trump's Tulsa Rally Reignite COVID-19? Indoor Events and Offsetting Community Effects. *NBER Working Paper*.

De Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M., Blondel, V. D., 2013. Unique in the Crowd: The Privacy Bounds of Human Mobility. *Scientific Reports*, 3(1), 1–5.

Egorov, G., Enikolopov, R., Makarin, A., Petrova, M., 2021. Divided We Stay Home: Social Distancing and Ethnic Diversity. *Journal of Public Economics*, 194, 104328.

Ellison, B., McFadden, B., Rickard, B. J., Wilson, N. L., 2021. Examining Food Purchase Behavior and Food Values During the COVID-19 Pandemic. *Applied Economic Perspectives and Policy*, 43(1), 58–72.

Ellison, B., Ocepek, M., 2021. Have Consumers' Food Values Changed During the COVID-19 Pandemic?

Gundersen, C., Ziliak, J. P., 2015. Food Insecurity and Health Outcomes.

Hamidi, S., Hamidi, I., 2021. Subway Ridership, Crowding, or Population Density: Determinants of COVID-19 Infection Rates in New York City. *American Journal of Preventive Medicine*, 60(5), 614–620.

Kar, A., Motoyama, Y., Carrel, A. L., Miller, H. J., Le, H. T., 2021. Impact of COVID-19 on Food Shopping: A Spatio-Temporal Analysis of Changes in Travel to Supermarket and Grocery Stores.

Kimball, A., Hatfield, K. M., Arons, M., James, A., Taylor, J., Spicer, K., Bardossy, A. C., Oakley, L. P., Tanwar, S., Chisty, Z. et al., 2020. Preliminary Estimates of the Prevalence of Selected Underlying Health Conditions Among Patients With Coronavirus Disease 2019—United States, February 12–March 28, 2020.

Litton, M. M., Beavers, A. W., 2021. The Relationship Between Food Security Status and Fruit and Vegetable Intake During the COVID-19 Pandemic. *Nutrients*, 13(3), 712.

Morland, K., Wing, S., Roux, A. D., 2002. The Contextual Effect of the Local Food Environment on Residents' Diets: The Atherosclerosis Risk in Communities Study. *American Journal of Public Health*, 92(11), 1761–1768.

NYC Health, 2021. NYC Coronavirus Disease 2019 (COVID-19) Data. github.com/nychealth/coronavirus-data. 24 Nov. 2021.

Oliver, N., Lepri, B., Sterly, H., Lambiotte, R., Deletaille, S., De Nadai, M., Letouzé, E., Salah, A. A., Benjamins, R., Cattuto, C. et al., 2020. Mobile Phone Data for Informing Public Health Actions Across the COVID-19 Pandemic Life Cycle.

Pereira, M. A., Kartashov, A. I., Ebbeling, C. B., Van Horn, L., Slattery, M. L., Jacobs Jr, D. R., Ludwig, D. S., 2005. Fast-Food Habits, Weight Gain, and Insulin Resistance (the CARDIA Study): 15-Year Prospective Analysis. *The Lancet*, 365(9453), 36–42.

SafeGraph Inc., 2021. Patterns: SafeGraph Docs. 27 Oct. 2021.

Stender, S., Dyerberg, J., Astrup, A., 2007. Fast Food: Unfriendly and Unhealthy. *International Journal of Obesity*, 31(6), 887–890.

The Hunter College NYC Food Policy Center, The Laurie M. Tisch Center for Food, Education Policy, and The CUNY Urban Food Policy Institute, 2020. New York Food 20/20: Vision, Research, and Recommendations During COVID-19 and Beyond.

Verma, R., Yabe, T., Ukkusuri, S. V., 2021. Spatiotemporal Contact Density Explains the Disparity of COVID-19 Spread in Urban Neighborhoods. *Scientific Reports*, 11(1), 1–11.

ACKNOWLEDGEMENTS

This work was directly supported by capstone sponsor Huy T. Vo [1,2]and the NYU Center for Urban Science + Progress (CUSP). Additional technical support and knowledge mentorship was provided by Claire Huang[2], Andrew Maroko[2], and Nevin Cohen[2].

[1] New York University (NYU)
[2] The City University of New York (CUNY)

APPENDIX A: DETAILED TIMELINE OF COVID-19 IN NYC

January 21, 2020	First confirmed COVID-19 case in the U.S.
January 30, 2020	COVID-19 declared a global health emergency by the World Health Organization.
February 29, 2020	First reported COVID-19 death in the U.S.
March 1, 2020	First COVID-19 case in New York State (NYS).
March 8, 2020	NYC issued guidelines to avoid crowded public transportation.
March 16, 2020	NYC public schools closed.
March 17, 2020	NYC bars and restaurants closed for in-person dining.
March 20, 2020	New York Governor Andrew Cuomo announced a statewide stay-at-home order effective on March 22; 100% of non-essential workforce mandated to work from home.
April 30, 2020	NYC subway began nightly closures from 1 a.m. to 5 a.m.
June 8, 2020	NYC began Phase 1 of reopening: a limited set of non-essential businesses were allowed to re-open with restrictions.
June 22, 2020	NYC began Phase 2 of reopening: a larger set of non-essential businesses and outdoor dining were allowed to re-open with restrictions.
July 6, 2020	NYC began Phase 3 of reopening, excluding indoor dining.
July 19, 2020	NYC began Phase 4 reopening, excluding indoor dining.
September 9, 2020	Malls in NYC reopened at 50% capacity with no indoor dining.
September 29, 2020	In-person learning resumed for NYC elementary students.
September 30, 2020	Indoor dining in NYC resumed with a 25% occupancy limit.
October 1, 2020	In-person learning resumed for NYC middle and high school students.
November 13, 2020	New COVID-19 restrictions went into effect: curfews on restaurants and gyms; private indoor gatherings limited to 10 people.
November 19, 2020	In-person learning halted for all NYC students; all learning went remote.
November 21, 2020	Indoor dining restrictions were renewed in NYC.
December 7, 2020	In-person learning resumed for NYC elementary students.
February 11, 2021	Indoor dining in NYC resumed with a 25% occupancy limit.
February 15, 2021	In-person learning resumed for NYC middle school students.
March 22, 2021	In-person learning resumed for NYC high school students.

APPENDIX B: DATA SCHEMAS AND POI CATEGORIES

Column Name	Description	Type	Example
placekey	Unique and persistent ID tied to this POI.	String	222-222@222-222-222
location_name	The name of the place of interest.	String	Salinas Valley Ford Lincoln
safegraph_brand_ids	Unique and consistent ID that represents this specific brand.	List	SG_BRAND_80ca06ab fa1a5104af9a770f485d ad07
brands	Name of larger brands that have been explicitly identified. A POI may have multiple brands.	List	Ford, Lincoln
top_category	Label associated with the first 4 digits of the POI's NAICS category.	String	Automobile Dealers
sub_category	Label associated with all 6 digits of the POI's NAICS category.	String	New Car Dealers
naics_code	4-digit or 6-digit NAICS code describing the business.	Integer	441110
latitude	Latitude coordinate of the place of interest.	Float	36.714767
longitude	Longitude coordinate of the place of interest.	Float	-121.662912
street_address	Street address of the place of interest.	String	1100 Auto Center Circle
city	The city in which this point of interest is located.	String	Irvine
postal_code	When iso_country_code == US, then this is the USA 5 digit zip code.	String	92602
open_hours	A JSON string with days as keys and opening & closing times (in the POI's local time) as values.	String	{ "Mon": [["8:00", "22:00"]]}
category_tags	An array of descriptive tags indicating higher-resolution category information.	List	[Mexican Food,Casual Dining,Lunch,Dinner]
opened_on	The outside year and month this POI opened in yyyy-mm format.	String	2019-10
closed_on	The outside year and month this POI closed in yyyy-mm format.	String	2020-03

Table B.1 - Key columns, description and data type of POI table (More details: https://docs.safegraph.com/docs/core-places)

Column Name	Description	Type	Example
placekey	Unique and persistent ID tied to this POI.	String	222-222@222-222-222
date_range_start	Start time for measurement period in ISO 8601 format of YYYY-MM-DDTHH:mm:SS±hh:mm.	String	2020-03-02T00:00:00-06:00
date_range_end	End time for measurement period in ISO 8601 format of YYYY-MM-DDTHH:mm:SS±hh:mm.	String	2020-03-09T00:00:00-06:00
raw_visit_counts	Number of visits in our panel to this POI during the date range.	Integer	1542
raw_visitor_counts	Number of unique visitors from our panel to this POI during the date range.	Integer	1221
visits_by_day	The number of visits to the POI each day (local time) over the covered time period.	JSON [Integer]	[33, 22, 33, 22, 33, 22, 22]
visits_by_each_hour	The number of visits to the POI for each of the 168 hours of the week, starting at midnight on date_range_start.	JSON [Integer]	[33, 22, 33, 22, 33, 22, 22, 21, 23, 33, 22, 11, 44, 22, 22, 44, 11, 33, 44, 44, 44, 33, 34, ...]
poi_cbg	The census block group the POI located.	String	560610112022
visitor_home_cbgs	A mapping of CBGs to the number of visitors to the POI whose home is in that CBG. Only cbgs with at least 2 devices are shown and cbgs with less than 5 devices are reported as 4.	JSON {String: Integer}	{"360610112021": 603, "460610112021": 243, "560610112021": 106, "660610112021": 87, "660610112021": 51}
visitor_home_aggregation	A mapping of census tracts to the number of visitors to the POI whose home is in that census tract.	JSON {String: Integer}	{"17031440300": 1005}
visitor_daytime_cbgs	A mapping of census block groups to the number of visitors to the POI whose primary daytime location between 9 am - 5 pm is in that census block group.	JSON {String: Integer}	{"360610112030": 9872, "880610112021": 8441, "569610112020": 5671, "160610112041": 2296}
distance_from_home	Median distance from home traveled by visitors (of visitors whose home we have identified) in meters. If we have fewer than 5 visitors to a POI, the value will be null.	Integer	1211
median_dwell	Median minimum dwell time in minutes.	Double	5

Table B.2 - Key columns, description and data type of weekly trip table
(More details: https://docs.safegraph.com/docs/weekly-patterns)

<table>
<tr><th>Category</th><th>NAICS Code</th><th>Top Category</th><th>Sub Category</th><th>Total Count</th></tr>
<tr><td rowspan="6">Specialty Food Stores</td><td>445210</td><td>Specialty Food Stores</td><td>Meat Markets</td><td rowspan="6">780</td></tr>
<tr><td>445220</td><td>Specialty Food Stores</td><td>Fish and Seafood Markets</td></tr>
<tr><td>445230</td><td>Specialty Food Stores</td><td>Fruit and Vegetable Markets</td></tr>
<tr><td>445291</td><td>Specialty Food Stores</td><td>Baked Goods Stores</td></tr>
<tr><td>445292</td><td>Specialty Food Stores</td><td>Confectionery and Nut Stores</td></tr>
<tr><td>445299</td><td>Specialty Food Stores</td><td>All Other Specialty Food Stores</td></tr>
<tr><td>Supermarkets</td><td>445110</td><td>Grocery Stores</td><td>Supermarkets and Other Grocery (except Convenience) Stores</td><td>2,830</td></tr>
<tr><td>Convenience Stores</td><td>445120</td><td>Grocery Stores</td><td>Convenience Stores</td><td>1,058</td></tr>
<tr><td rowspan="2">General Merchandise Stores</td><td>452319</td><td>General Merchandise Stores, including Warehouse Clubs and Supercenters</td><td>All Other General Merchandise Stores</td><td rowspan="2">680</td></tr>
<tr><td>453998</td><td>Other Miscellaneous Store Retailers</td><td>All Other Miscellaneous Store Retailers (except Tobacco Stores)</td></tr>
<tr><td>Big Box Grocers</td><td>452210</td><td>Department Stores</td><td>Department Stores</td><td>190</td></tr>
<tr><td>Full-Service Restaurants</td><td>722511</td><td>Restaurants and Other Eating Places</td><td>Full-Service Restaurants</td><td>13,363</td></tr>
<tr><td>Limited-Service Restaurants</td><td>722513</td><td>Restaurants and Other Eating Places</td><td>Limited-Service Restaurants</td><td>4,642</td></tr>
<tr><td rowspan="2">Snack and Bakeries</td><td>722515</td><td>Restaurants and Other Eating Places</td><td>Snack and Nonalcoholic Beverage Bars</td><td rowspan="2">5,316</td></tr>
<tr><td>311811</td><td>Bakeries and Tortilla Manufacturing</td><td>Retail Bakeries</td></tr>
<tr><td>Food Services</td><td>624210</td><td>Community Food and Housing, and Emergency and Other Relief Services</td><td>Community Food Services</td><td>2</td></tr>
<tr><td rowspan="2">Pharmacies and Drug Stores</td><td>446110</td><td>Health and Personal Care Stores</td><td>Pharmacies and Drug Stores</td><td rowspan="2">3,244</td></tr>
<tr><td>446191</td><td>Health and Personal Care Stores</td><td>Food (Health) Supplement Stores</td></tr>
<tr><td>Beer, Wine, and Liquor Stores</td><td>445310</td><td>Beer, Wine, and Liquor Stores</td><td>Beer, Wine, and Liquor Stores</td><td>1,012</td></tr>
<tr><td>Tobacco Stores</td><td>453991</td><td>Other Miscellaneous Store Retailers</td><td>Tobacco Stores</td><td>247</td></tr>
<tr><td>Drinking Places</td><td>722410</td><td>Drinking Places (Alcoholic Beverages)</td><td>Drinking Places (Alcoholic Beverages)</td><td>2,831</td></tr>
</table>

Table B.3 - Preliminary categorized POIs (Reference from https://docs.safegraph.com/docs/poi-types)

Fast-Food Chain Restaurant Name	Source
Arby's	Athens et al., 2016
Au Bon Pain	Bassett et al., 2011
Auntie Anne's	National Employment Law Project, 2015
Baskin Robbins	National Employment Law Project, 2015
Ben & Jerry's	New York State Department of Labor, 2021
Burger King	James et al., 2014
Carvel	National Employment Law Project, 2015
Chick-fil-A	James et al., 2014
Chipotle Mexican Grill	New York State Department of Labor, 2021
Domino's Pizza	National Employment Law Project, 2015
Dunkin'	James et al., 2014
Golden Krust Caribbean Bakery and Grill	New York State Department of Labor, 2021
Jamba	New York State Department of Labor, 2021
KFC	James et al., 2014
Little Caesars Pizza	National Employment Law Project, 2015
McDonald's	James et al., 2014
Nathan's Famous	New York State Department of Labor, 2021
Panera Bread	National Employment Law Project, 2015
Papa John's	Bassett et al., 2011
Pizza Hut	James et al., 2014
Popeyes Louisiana Kitchen	Bassett et al., 2011
Shake Shack	New York State Department of Labor, 2021
Starbucks	James et al., 2014
Subway	James et al., 2014
Taco Bell	James et al., 2014
Tim Hortons	New York State Department of Labor, 2021
Uno Chicago Grill	New York State Department of Labor, 2021
Wendy's	James et al., 2014
White Castle	New York State Department of Labor, 2021

Table B.4 - Set of POIs categorized as Fast-Food Chain Restaurants by scientific research studies, governmental organizations, and non-governmental organizations.

APPENDIX C: *k*-MEANS CLUSTERING

Type	Category	Feature Name
Mobility	Visitor Count	Median Estimated Visitor Count
Demographic	Population	Median Age
		Percentage of Male
	Race	Percentage of Two races excluding Some other race and three or more races
		Percentage of White alone
		Percentage of Black or African American alone
		Percentage of American Indian and Alaska Native alone
		Percentage of Native Hawaiian and Other Pacific Islander alone
		Percentage of Some other race alone
		Percentage of Two or more races
		Percentage of Asian alone & Two races including Some other race
	Household Type	Percentage of Households with one or more people under 18 years
		Percentage of Households with one or more people 60 years and over
	Educational Attainment	Percentage of Population over 25 without a high school diploma
		Percentage of Population over 25 with a bachelor's degree or higher
	Poverty Status	Percentage of Households Income in the past 12 months below poverty level
	Income	Median Household Income
	Food Stamps/Supplemental Nutrition Assistance Program (SNAP)	Percentage of Household received Food Stamps/SNAP
		Percentage of Households received Food Stamps/SNAP have disability
	Employment Status	Percentage of Civilian labor force
		Percentage of Employed

Table C.1 - Clustering Features (SafeGraph Open Census Data source, American Community Survey (ACS) table description)

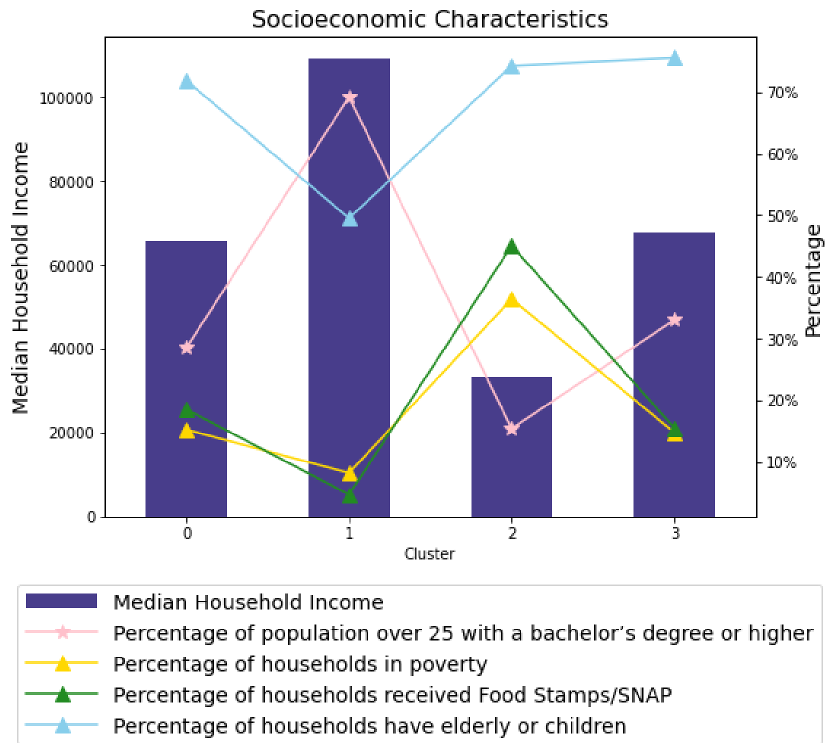

Figure C.1 - Income, education, and poverty etc. characteristics of each cluster

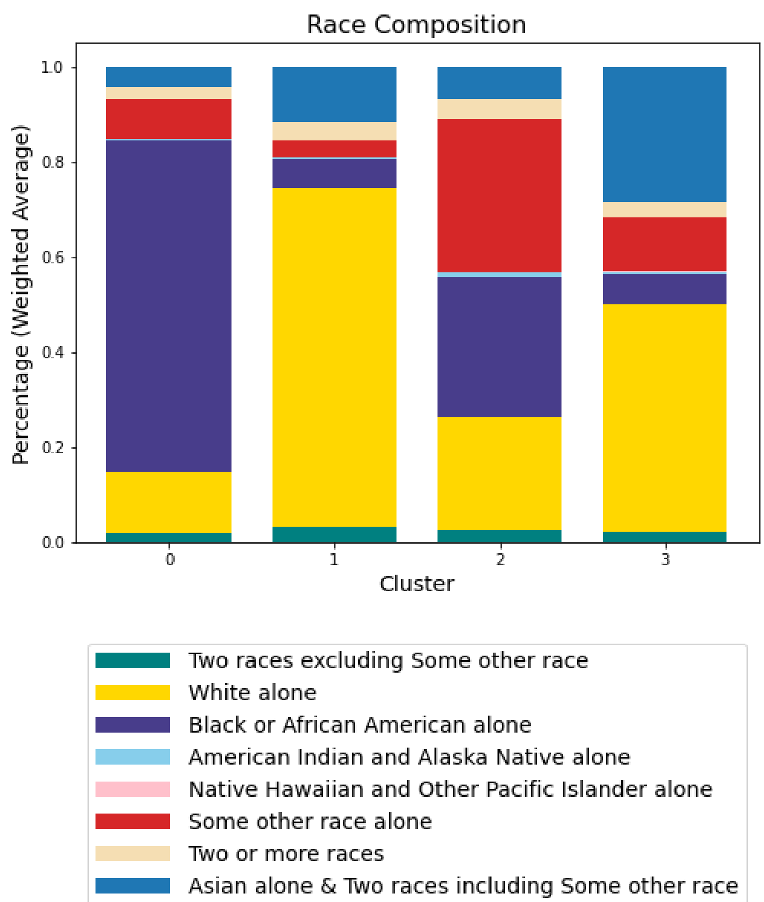

Figure C.2 - Ethic characteristics of each cluster

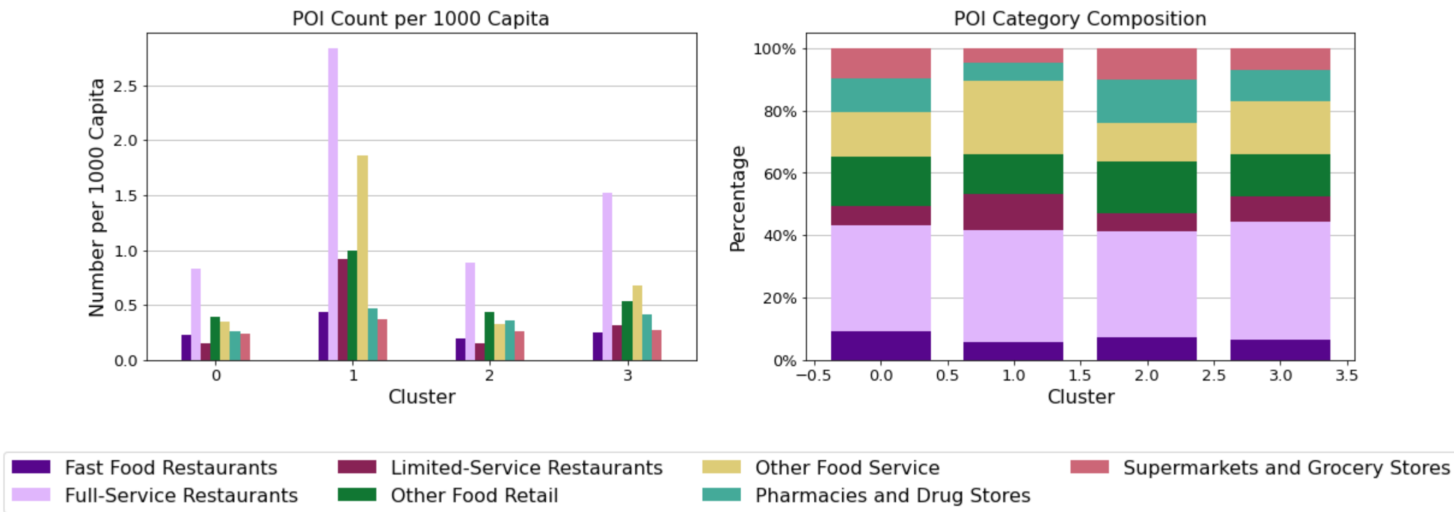

Figure C.3 - POI Count (the number of different types of POIs) by category in different clusters.

APPENDIX D: CALCULATION OF ANALYSIS METRICS

$$v_{ijk}^{\text{est}} = v_{ijk}^{\text{obs}} \times \frac{p_i}{d_{ik}} \tag{1}$$

Equation 1 - Estimated count of visitors from home CBG i at POI j during week k; given observed visitor count v_{ijk}^{obs}, population p_i, and device count d_{ik}.

$$w_{ick} = \frac{\sum_{j \in J_c} v_{ijk}^{\text{est}}}{\sum_{j \in J} v_{ijk}^{\text{est}}} \tag{2}$$

Equation 2 - Proportion of food location visitations made from CBG i to POIs of category c during week k; given the set of POIs J, and set of POIs J_c in category c.

$$c_{ijk} = v_{ijk}^{\text{est}} \times \frac{\sum_{m \in M} v_{mjk}^{\text{obs}}}{A_j} \tag{3}$$

Equation 3 - CDI for home CBG i at POI j during week k; given the set of CBGs M, observed visitor count v_{mjk}^{obs}, estimated visitor count v_{ijk}^{est}, and POI area A_j .

APPENDIX E: TECHNICAL REFERENCES

1. `https://hadoop.apache.org/`
2. `https://spark.apache.org/docs/latest/api/python/reference/index.html`
3. `https://cloud.google.com/bigquery`
4. `https://flask.palletsprojects.com/en/2.0.x/`
5. `https://pandas.pydata.org/`
6. `https://nodejs.org/en/`
7. `https://expressjs.com/`
8. `https://reactjs.org/`
9. `https://www.mapbox.com/`
10. `https://plotly.com/`
11. `https://jupyter.org/`
12. `https://scikit-learn.org/stable/tutorial/index.html`