Title
Using Data Science for Equity at SFMTA

Author
Caro, Ryan

Publication Date
2022-06-01

CAPSTONE PROJECT REPORT

UCLA Institute of Transportation Studies

Using Data Science for Equity at SFMTA

Project Lead: Ryan Caro
Faculty Advisor: Dr. Adam Millard-Ball
Client: San Francisco Municipal Transportation Agency

May 2022

Rebuilding California SB1

PSR Pacific Southwest Region UTC University Transportation Center

Technical Report Documentation Page

1. Report No.	2. Government Accession No. N/A	3. Recipient's Catalog No. N/A
4. Title and Subtitle Using Data Science for Equity at SFMTA		**5. Report Date** 2022-06-01 **6. Performing Organization Code** UCLA-ITS
7. Author(s) Ryan Caro		**8. Performing Organization Report No.** LA2132
9. Performing Organization Name and Address Institute of Transportation Studies, UCLA 3320 Public Affairs Building Los Angeles, CA 90095-1656		**10. Work Unit No.** N/A **11. Contract or Grant No.**
12. Sponsoring Agency Name and Address UCLA Institute of Transportation Studies www.its.ucla.edu		**13. Type of Report and Period Covered** Final **14. Sponsoring Agency Code** UC ITS

15. Supplementary Notes

DOI: doi:10.17610/T6602P

16. Abstract

The San Francisco Municipal Transportation Agency (SFMTA) cut back service at the onset of the COVID-19 pandemic in 2020 and has been slowly rebuilding its service since. The SFMTA Board asks staff to conduct an equity analysis after each service change. To measure equity, SFMTA calculates the number of jobs that can be reached in a 30-, 45-, and 60-minute commute from specified "equity neighborhoods," selected based on the percentage of households with low incomes, low rates of private vehicle ownership, and race and ethnicity
demographics.

My research is the first step in automating these analyses in order to build a tool to optimize SFMTA service for job access from equity neighborhoods. I use open-source tools to build a model of transit across the Bay Area and calculate the number of jobs available within the specified commute times from each equity neighborhood.

I find that although this tool is planned to help SFMTA improve its service, proximity to BART is the greatest predictor of job access within 45 and 60 minutes. I find that among open-source tools, OpenTripPlanner's point-to-point methods are significantly more accurate than its one-to-many methods, despite being several orders of magnitude slower. I also make recommendations for how SFMTA can continue to build this tool. I find that equity neighborhoods are a very rough proxy for disadvantaged groups, and I suggest that SFMTA create an equity index to weight tracts across the entire city rather than focusing exclusively on equity neighborhoods.

17. Key Words	18. Distribution Statement
transit Bay Area equity analysis transportation planning	No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	41	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

UNIVERSITY OF CALIFORNIA
Los Angeles

Using Data Science for Equity at SFMTA

A comprehensive project submitted in partial satisfaction of the requirements for the degree Master of Urban & Regional Planning

By Ryan Caro

Client: San Francisco Municipal Transportation Agency
Faculty Chair of Committee: Dr. Adam Millard-Ball

2022

Disclosure

This report was prepared in partial fulfillment of the requirements for the Master in Urban and Regional Planning degree in the Department of Urban Planning at the University of California, Los Angeles. It was prepared at the direction of the Department and of the San Francisco Municipal Transportation Agency as a planning client. The views expressed herein are those of the authors and not necessarily those of the Department, the UCLA Luskin School of Public Affairs, UCLA as a whole, or the client.

Executive Summary

The San Francisco Municipal Transportation Agency (SFMTA) cut back service at the onset of the COVID-19 pandemic in 2020 and has been slowly rebuilding its service since. The SFMTA Board asks staff to conduct an equity analysis after each service change. To measure equity, SFMTA calculates the number of jobs that can be reached in a 30-, 45-, and 60-minute commute from specified "equity neighborhoods," selected based on the percentage of households with low incomes, low rates of private vehicle ownership, and race and ethnicity demographics.

My research is the first step in automating these analyses in order to build a tool to optimize SFMTA service for job access from equity neighborhoods. I use open-source tools to build a model of transit across the Bay Area and calculate the number of jobs available within the specified commute times from each equity neighborhood. I find that although this tool is planned to help SFMTA improve its service, proximity to BART is the greatest predictor of job access within 45 and 60 minutes.

I find that among open-source tools, OpenTripPlanner's point-to-point methods are significantly more accurate than its one-to-many methods, despite being several orders of magnitude slower. I also make recommendations for how SFMTA can continue to build this tool. I find that equity neighborhoods are a very rough proxy for disadvantaged groups, and I suggest that SFMTA create an equity index to weight tracts across the entire city rather than focusing exclusively on equity neighborhoods.

Acknowledgements

I would like to thank my faculty advisor, Dr. Adam Millard-Ball, for his unwavering patience and extensive technical support of this project, and my client advisor, Steph Nelson, for their time and flexibility. I am greatly indebted to Dr. Willem Klumpenhouwer and Dr. Diego Silva at the University of Toronto, who built the TransitCenter dashboard on which much of my research was based, and who generously offered to provide me more information about their code. I am also thankful to Dr. Yoh Kawano, who first introduced me to data science, Python, and Geographic Information Systems at UCLA, and without whom I could not have started on this path.

In addition, I would like to acknowledge the following people who made my time at UCLA possible: Linda Munitz & Rob Caro; Vivienne Lipsitz; Chris, William, & Erik Levenberg; Luis Mancheno & Eric Felleman; Ellen & Nathan Caro; Jimmy Milby; Sandra Sohn; Wilson Velasco & Dante Ramos; Ari Peña & Luka Uršič; Ja-Rei Wang & Jose Martinez Saiz; Bryan St. Amand & Dr. Christian Gomez; Dr. Paul Melton & Miguel Otero-Adran; Dr. Xin He & Dr. Peter Zhao; Alice Lyons & Kyle Bland; Kevin Jobe; Ramiro Maravilla Flores; Braeden Mansouri & Mason Gossard; and all the others who sheltered, fed, comforted, and otherwise helped me along the way.

I would also like to thank the UCLA Institute of Transportation Studies for their generous support of this project and of my time at UCLA.

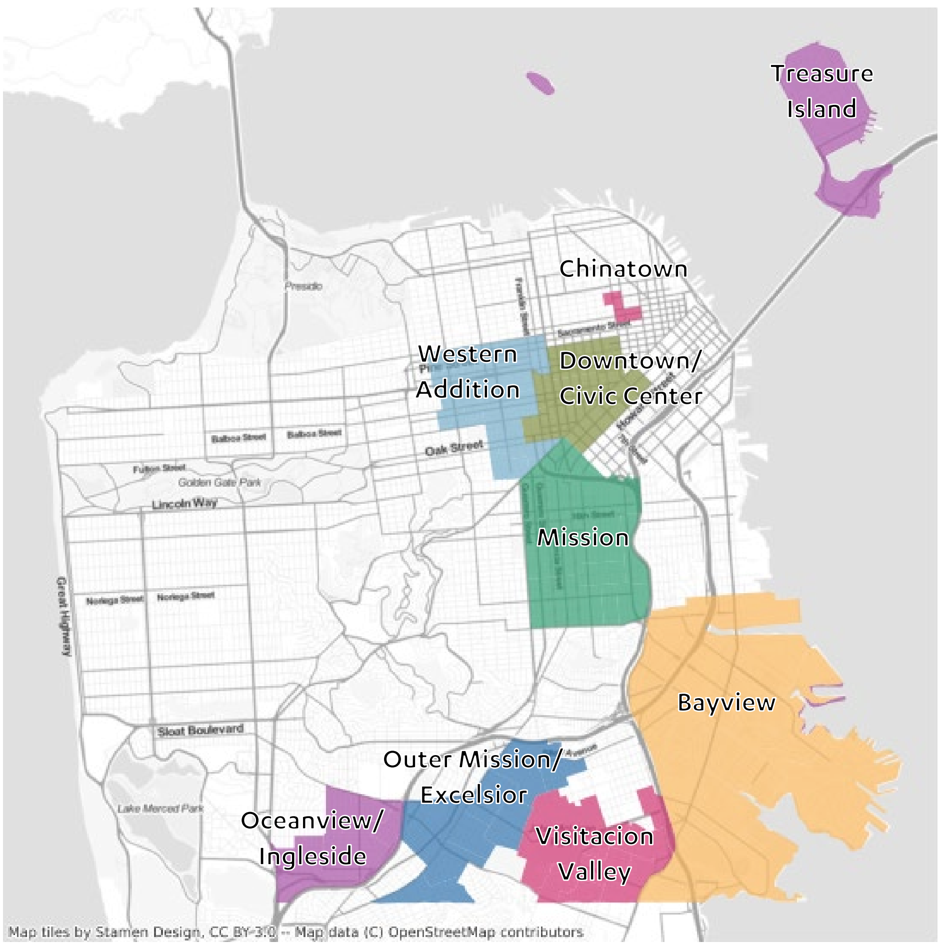
The Institute of Transportation Studies at UCLA acknowledges the Gabrielino/Tongva peoples as the traditional land caretakers of Tovaangar (the Los Angeles basin and So. Channel Islands). As a land grant institution, we pay our respects to the Honuukvetam (Ancestors), ‘Ahiihirom (Elders) and ‘Eyoohiinkem (our relatives/relations) past, present and emerging.

Contents

Introduction

The San Francisco Municipal Transportation Agency (SFMTA) cut back service at the onset of the COVID-19 pandemic in 2020 and has been slowly rebuilding its service since. The SFMTA Board asks staff to conduct an equity analysis after each service change. To measure equity, SFMTA calculates the number of jobs that can be reached in a 30-, 45-, and 60-minute commute from specified "equity neighborhoods," a metric which all Bay Area transit agencies use. SFMTA selected the neighborhoods based on the percentage of households with low incomes, low rates of private vehicle ownership, and race and ethnicity demographics. There are nine equity neighborhoods in San Francisco: Treasure Island, Chinatown, Tenderloin/SOMA, Western Addition, the Mission, Bayview, Visitacion Valley, Outer Mission/Excelsior, and Oceanview/Ingleside, shown in Figure 1 below.

Figure 1. Equity Neighborhoods defined by SFMTA



My research is the first step in answering the following question: "How can current SFMTA bus and rail service be redistributed to maximize the number of jobs available to residents of the nine equity neighborhoods within a 30-, 45-, or 60-minute commute?" While the literature review below points out some possible flaws with this question, I seek to produce recommendations that are practical within the limited scope of SFMTA's power. This research is important because hundreds of thousands of San Franciscans ride SFMTA services every day. However, not everyone has equal access to service. Overall, white San Franciscans can reach more jobs within a given time period, while Latinx and Black San Franciscans can reach fewer (Klumpenhouwer et al., 2021). Focusing on service to equity neighborhoods may help to eliminate some of these disparities.

First, I examine the demographics of the nine equity neighborhoods. While equity neighborhoods contain disproportionate numbers of Black and Latinx people, low-income households, and households without access to a vehicle, they only contain about half of these populations. I suggest that future work should use a total figure for underserved populations from all parts of the city, rather than relying on the equity neighborhoods as a proxy.

Currently SFMTA analyzes the number of jobs available to residents after making service changes, but the Board of Directors wants to be more proactive about measuring the effects of possible service changes. My project serves to document some possible data science methods that future consultants could use to optimize service. In particular, I use open-source, Python-based methods from OpenTripPlanner (OTP) to measure the number of jobs available to residents of equity neighborhoods within the specified commute times. I compare the results of two different OTP calculations for transit times: point-to-point, which calculates the shortest path from each origin point to each destination point, and one-to-many, which calculates the transit time from one origin to every destination at the same time. I find that calculating point-to-point times, while about 100 times slower, gives more accurate results than the one-to-many method.

Finally, I make recommendations for how future consultants could improve this analysis, by using more granular geographic data, realtime trends, and more information on travel patterns over the course of the day and week, in addition to more computational power. These changes would ensure that SFMTA can propose changes quickly and produce suggested service modifications with the most benefit to disadvantaged people in San Francisco.

Literature Review

This literature review covers metrics of accessibility, measures of transit equity, and finally how equity in transit relates to improved economic outcomes for marginalized people. Measuring accessibility is difficult and always imprecise, as there is no way to understand the precise desires, abilities, and priorities of every individual or household. When measuring accessibility via transit, most researchers focus on access to jobs, either by comparing transit access to vehicular access, comparing different cities to one another, or comparing the access of different demographic groups to each other. This last topic is most important for my research. When looking at urban areas as a whole, researchers often find that low-income people of color have better access to jobs via transit than high-income or white residents, but this ignores the fact that those who can afford to own a car have the option to live far from transit. More useful is a comparison of access among different groups of carless people, or at least of those who live in the urban core. Finally, there are differing results when looking at the economic outcomes of living in transit-rich neighborhoods. Though, as the literature reviewed below show, transit is not the most important factor in economic outcomes, it is the one that transit agencies have the power to change, and they should strive to do so.

Measures of accessibility

Yeganeh et al. (2018) cite Hansen's (1959, p. 73) definition of "the intensity of the potential for interaction" as the first to introduce the concept of accessibility to urban planning. Wachs and Kumagai (1973, p. 437) provide a more useful definition: "the ease with which citizens may reach a variety of opportunities for employment and services." Though the concept of accessibility dates back half a century or more, only recently have transportation agencies started using it in their work. Traditionally transportation agencies have been more concerned with mobility, a measure of how fast people (and usually people in cars) can move around a city (Levine et al., 2019). Mobility only calculates speed, however, and not the ease with which people can access destinations. For example, adding a grocery store in a neighborhood that does not have one does not improve mobility, but it does improve accessibility. Vehicular congestion, a sign of low mobility, is often also a sign of high accessibility, as a concentration of desirable destinations draws many people. Though many traditional transportation planners think primarily about mobility, they must also consider land use to measure accessibility (Sundquist et al., 2021).

While Wachs and Kumagai's definition of accessibility is simple, measuring it is not (Duranton & Guerra, 2016). What, after all, is a unit of ease? How does one measure an opportunity? Recent advances in data science have allowed transportation agencies to be more proactive in measuring accessibility (Governor's Institute on Community Design, 2017). Even with these advances, however, measures of accessibility are different from experiences of accessibility. Both Curl et al. (2015) and Lättman et al. (2018) find that perceived or reported accessibility is consistently different from actual accessibility as measured with GIS or data science.

The most common measure of accessibility is access to work. Though the opportunity to earn income is fundamental, less than a quarter of trips in large U.S. metropolitan areas are for commutes to work (Duranton & Guerra, 2016). More complex measures for an individual household involve adding the time and cost to the nearest destinations: school, grocery store, restaurant, house of worship, etc. This method breaks down quite quickly, however, as people rarely go to the closest of these destinations (Duranton & Guerra, 2016). Even more complex measures involve summing the number of destinations within a specific time and cost threshold (Miller, 2018). These measures can weight all destinations equally, or give more weight to larger and closer ones in a gravity model (Handy & Niemeier, 1997). Neither of these models, however, takes into account individual preferences; a butcher shop has very little utility to a vegan, for example (Duranton & Guerra, 2016). To account for this, a final method involves measuring utility of various destinations for individuals, though this requires researchers to collect large amounts of data on preferences (Handy & Niemeier, 1997). A creative shortcut to the large amount of data or computational power required for the methods presented above is to measure the time or cost to the *n*th destination of a certain type, with the value of *n* depending on the destination's category. For example, Klumpenhouwer et al. (2021) measure accessibility in time to the closest urgent care facility but time to the *third* closest grocery store, on the grounds that individuals are more likely to have a preference for where they buy food compared to where they receive emergency medical care.

Measures of transit accessibility and equity

Research on accessibility via public transportation has three main purposes: to compare accessibility via transit to accessibility via personal vehicles, to compare accessibility via transit between different regions, and to compare accessibility via transit between different demographic groups in the same region.[1] Of course, these purposes are not mutually exclusive. Kawabata and Shen (2006) compare job accessibility via transit with job accessibility via car both within and between Tokyo, Los Angeles, and Boston. These metrics also lend themselves to rankings. The Accessibility Observatory at the University of Minnesota Center for Transportation Studies regularly publishes rankings of the largest U.S. metropolitan areas by access to jobs, comparing lists from year to year (Owen & Murphy, 2020).

Levine et al. (2019) summarize the difficulties inherent in measuring accessibility via transit. Unlike walking, biking, or driving a car, taking transit involves multiple stages, including getting to the origin stop, waiting for the vehicle, possible transfers to other vehicles, and moving from the final stop to the destination. Because service patterns change over the course of the day and week, no single number can capture the sum of destinations available via public transportation. These challenges are compounded by the fact that while road networks change slowly and are well-documented by state and local transportation authorities, transit patterns can change from month to month and are often managed by multiple poorly funded agencies. While the General Transit Feed Specification (GTFS), discussed later in this report, has standardized much of this information, it does not help in determining pedestrian conditions in the vicinity of transit stops, which can cause great variability in walking, and thus transit, time calculations (Levine et al., 2019).

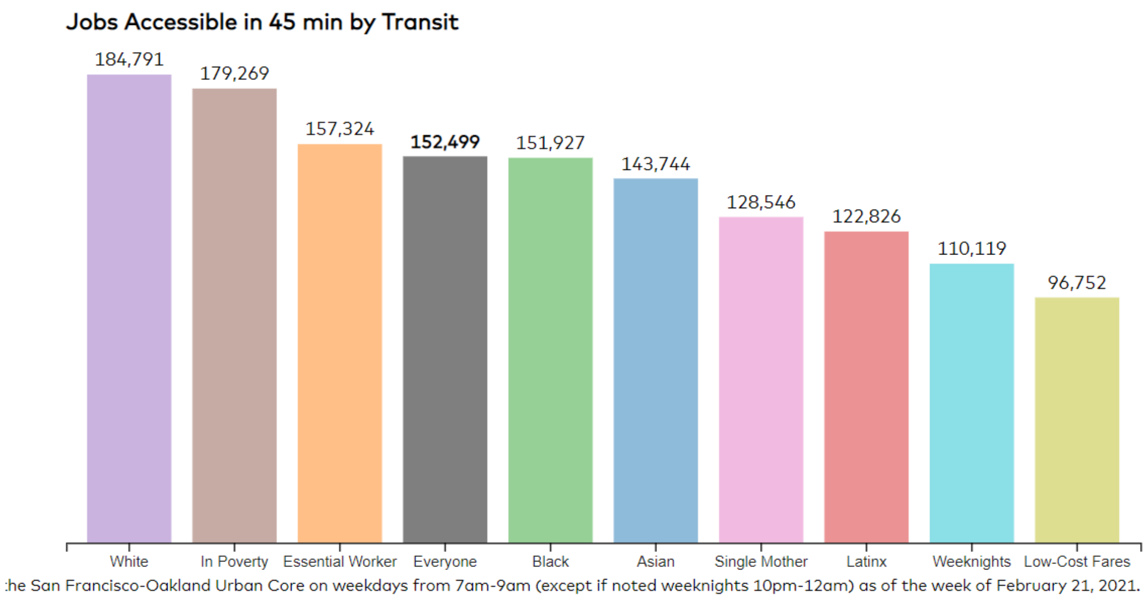
Despite these challenges, many researchers have performed the technically difficult tasks of comparing transit accessibility between regions (Levine et al., 2019; Owen & Murphy, 2020) and comparing different demographic groups within the same region. Studies which have measured this second form transit equity have found conflicting results, often based on whether they are looking at an entire metropolitan region or just the urban core. For

[1] Here "transit" and "public transportation" refer to fixed-route modes such as buses, subways or heavy rail, regional rail, light rail, or ferries.

example, Yeganeh et al. (2018) find that in the majority of the 45 largest U.S. metropolitan areas, low-income non-whites have a higher number of jobs accessible via transit than high-income whites. While this finding seems to imply that transit systems are equitable, or somehow favor marginalized people, it does not take into account the fact that low-income people of color are generally concentrated in the transit-rich urban core, while high-income whites in most regions are more likely to live in far-flung suburbs without transit access (Glaeser et al., 2008; Nelson, 2020). In addition, as Carleton and Porter (2018) point out, aggregate metrics for marginalized groups can obscure the disparities experienced by people within those groups.

Klumpenhouwer et al. (2021) provide more nuanced and more useful results by focusing on the urban cores of the seven metropolitan areas with the greatest number of transit users and disaggregating their metrics. As Figure 1 shows, the average white resident of the San Francisco-Oakland urban core has access to 22% more jobs via a 45-minute transit commute than residents overall. This disaggregated data shows that workers in poverty have access to nearly as many jobs by the same metric as white workers, Black workers have slightly less job accessibility than average, and Latinx workers have the lowest accessibility of listed ethnic or racial groups.

Figure 2. Jobs accessible in 45 minutes by transit in the San Francisco-Oakland urban core for the average person in each demographic category (Klumpenhouwer et al., 2021)



Even disaggregated data does not necessarily show the whole picture. As Figure 1 shows, weeknight accessibility is significantly lower than weekday morning accessibility. Other researchers (Al Mamun & Lownes, 2011; El-Geneidy et al., 2016; Jomehpour Chahar Aman & Smith-Colin, 2020) have added time-of-day measures to determine if accessibility via transit is equitably distributed at different hours, but their research is hampered by a lack of data on the number of jobs starting and ending at each hour. Furthermore, the analyses listed in this section all draw on census data to estimate ridership, but Karner and Golub (2015) have shown that census data does not align with ridership data collected from on-board surveys.

Though there are many ways to measure access to opportunity, current research accepts it as the most important benefit of a transit system (Karner et al., 2020), despite the lack of precision. Other researchers, however, present more justice-oriented methods for analyzing equity impacts. Pereira et al. (2017) compare five different philosophies of justice: utilitarianism, libertarianism, intuitionism, Rawls' egalitarianism, and Capability Approaches. They argue for a combination of Rawls' egalitarianism and Capability Approaches, which results in a distribution of "transport investments and services in ways that reduce inequality of opportunity" (Pereira et al., 2017, p. 184). Their theory of justice, however, relies on changing current systems and future investments, but only briefly touches on redistribution of burdens such as pollution and noise. Karner et al. (2020) argues for transportation justice that centers society rather than state actors. Agency-led analysis finds justice in a "quantitative result showing that disadvantaged populations benefit at least as much as a reference population and are not burdened more than they are" (p. 445). Community-led mobilization, on the other hand, does not have a single just outcome but instead relies on whether the community supports the outcome.

Transit access and outcomes

Much of the research on transit accessibility and equity assumes that transit accessibility is *ipso facto* desirable, but findings from outcomes-based studies are mixed. Most of these studies focus on employment or income, which are both relatively straightforward to measure and vital for wellbeing. For example, in a longitudinal study of poor families, immigrants, and people of color, Smart and Klein (2015) find that living in a transit-rich area does not correlate with future earnings but instead with a slightly higher chance of being unemployed. The authors point out that there are likely confounding variables not included in their model. In a study of the 113 largest U.S. urbanized areas, Lyons and Ewing (2021) find that transit access does indeed have a small effect on lowering poverty and unemployment, though it is dwarfed by job/population balance, education, and urban compactness. The authors provide some context for this finding by pointing out that because most Americans commute by car, increased transit service does not improve access to most commuters. In this vein, Ong and Houston (2002, p. 344) also find that "the level of transit service near a recipient's home makes a moderate, yet statistically significant, contribution to increasing the probability of employment."

Comparing economic outcomes from transit accessibility for people without access to a vehicle produces more promising results. In a study of low-skilled workers in Los Angeles, San Francisco, and Boston, Kawabata (2003) finds that job accessibility via transit significantly correlates with overall employment and full-time employment for those without a car in the two California cities. Raphael (1998) performed a similar study in the Bay Area and found that, controlling for race, job accessibility by transit explained about 20% of the difference in employment between young Black and white men. Millard Ball et al. (2022), however, found that while transit access did not have a significant effect on employment for those living in affordable housing in San Francisco, neither did parking requirements. Though transit access is not the silver bullet to solve the United States' long history of segregation and racism, it remains an important part of improving opportunities for marginalized people, particularly those without access to a car.

Conclusion

Though measuring accessibility, especially by transit, is often difficult and always imprecise, it is important for understanding the possibilities for those without a car. Improving accessibility for these people shows promising correlations with better economic outcomes, though it is only one part of much larger movements towards transportation justice. Though the literature overall has often contradictory findings, transit agencies have an obligation to serve those who need them most. Despite the importance of equity in transit, only 14% of agencies surveyed listed providing service for disadvantaged or transit-dependent riders as a goal, and only 8% did so explicitly (Taylor & Morris, 2015). SFMTA is taking a laudable approach to modifying service to improve the job accessibility of the most vulnerable San Franciscans.

Methodology

Data

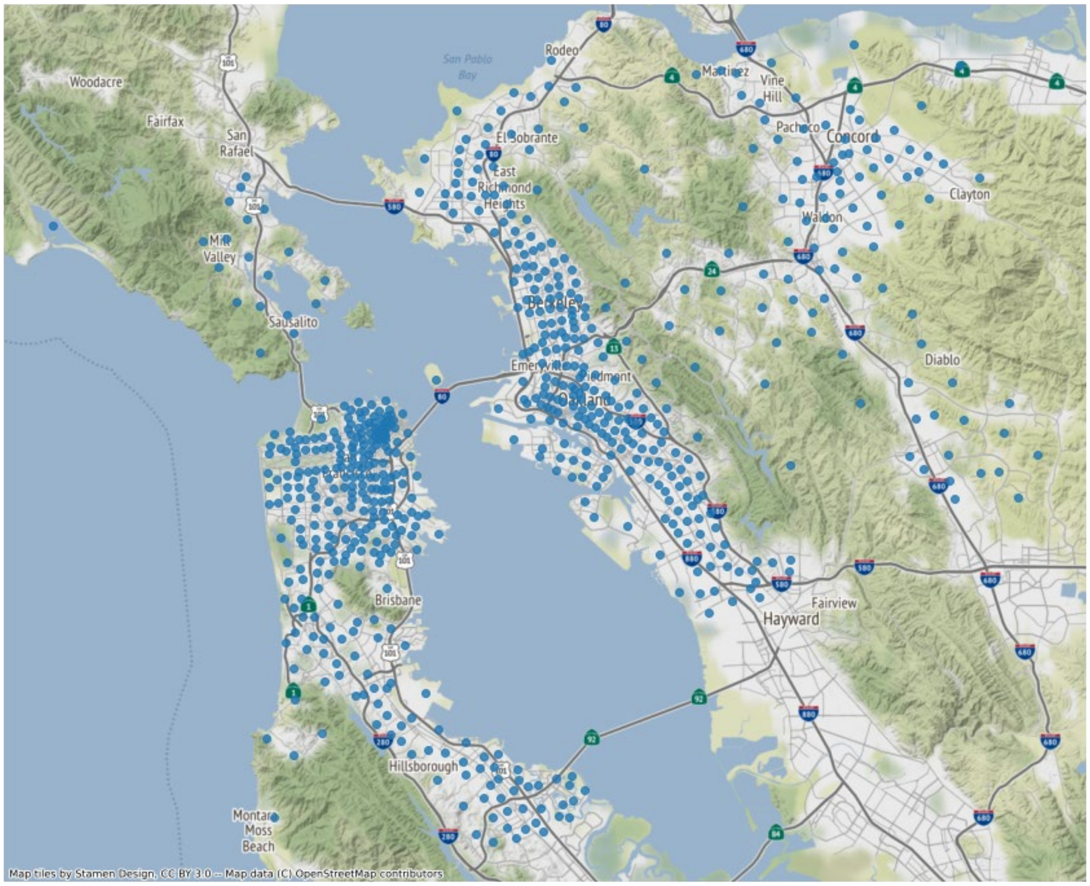
The nine equity neighborhoods are Treasure Island, Chinatown, Tenderloin/SOMA, Western Addition, Mission, Bayview, Visitacion Valley, Outer Mission/Excelsior, and Oceanview/Ingleside, shown in Figure 1 above. SFMTA has provided me with a shapefile of census tracts in the equity neighborhoods. Each census tract is represented by its centroid, rather than a polygon of its borders. To complete this data, I use tract geometries and demographic data from the 2015-2019 American Community Survey from the U.S. Census Bureau, imported using Census Reporter (Census Reporter, n.d.).

SFMTA chose these neighborhoods based on the following characteristics, reported in their Muni Service Equity Strategy (San Francisco Municipal Transportation Agency, 2016):

- Higher than average proportion of households below 200% of the federal poverty rate, consistent with the agency's cutoff for its low-income pass, compared to San Francisco as a whole. The agency uses 200% of the federal poverty rate because of the high cost of living in San Francisco.
- Higher than average proportion of people of color, compared to San Francisco as a whole.
- Lower than average rates of vehicle ownership, compared to San Francisco as a whole.
- Concentration of affordable and public housing developments, provided by the Mayor's Office of Housing.
- Neighborhoods with Muni routes "heavily used by persons of color and low-income transit riders."

SFMTA has also provided me with job counts, also by census tract centroid, for the San Francisco Bay Area. The data extend as far north as San Rafael in Marin County, as far south as San Mateo in San Mateo County, and across the Bay as far east as Clayton and San Ramon in Contra Costa County and San Lorenzo in Alameda County. Figure 3 shows the extent of jobs data provided.

Figure 3. Census tracts with job data for analysis



To calculate travel times by transit to these jobs, I use the General Transit Feed Specification (GTFS) feed for Bay Area transit agencies, downloaded using Transitland's Application Programming Interface (API) (Interline Technologies, 2021). GTFS provides a standardized format for transit agencies to publish their data so that both consumer-facing applications and researchers can use it (McHugh, 2013). Each agency's GTFS feed shows scheduled departure times for every vehicle at each stop on each route. Some agencies also include realtime information, reporting expected departures rather than simply scheduled times, but my analysis, like SFMTA's in-house analysis, only uses schedule data. The list of agencies is presented in table 1 below:

Table 1. Bay Area transit agencies

AC Transit	Mountain View Community Shuttle
Altamont Corridor Express (ACE)	SamTrans
Angel Island Tiburon Ferry	San Benito County Express
BART	San Joaquin Regional Transit District
Blue & Gold Fleet	San Francisco Bay Ferry
Caltrain	Santa Cruz Metropolitan Transit District
Capitol Corridor	SFMTA
City of Palo Alto Shuttle	SMART
Commute (San Mateo County)	SolTrans
County Connection	Sonoma County Airport Express
Dumbarton Express	Stanford Marguerite Shuttle
Emerygoround	Tideline Water Taxi
Fairfield and Suisun Transit	Tri Delta Transit
Golden Gate Ferry	Union City Transit
Golden Gate Transit	VTA
Marin Transit	WestCAT
Mission Bay TMA	Wheels Bus
Monterey-Salinas Transit	

Analysis

I first measured what proportion of targeted demographic groups live in the equity neighborhoods. Since the neighborhoods are defined by census tracts, I was able to use data from the American Community Survey 2015-2019 five-year estimates by census tract. I calculated the number of low-income households, households without car access, and different racial and ethnic groups.

I then measured the number of jobs currently available within a 30-, 45-, and 60-minute commute from the census tracts in the equity neighborhoods at each hour of the day. I am basing my calculations off of code from the TransitCenter Equity Dashboard (Klumpenhouwer et al., 2021). TransitCenter's code uses OpenTripPlanner (OTP), an open-source multi-modal trip planner, which in turn uses GTFS feeds and OpenStreetMap, a collaborative user-edited map of the world, for first/last mile journeys (OpenTripPlanner Contributors, 2016).

After downloading the GTFS data from Transitland (Interline Technologies, 2021), I created a graph object which models the composite transit system in the Bay Area. Querying this graph object produces travel times for any combination of origin and destination points in the service area. For each origin and destination pair, OTP finds the fastest route at a given time, producing the walk time to the first stop, transit route information including any transfers, and walk time from the final stop. My analyses, however, only includes the total travel time.

By looping through the origin tracts in equity neighborhoods and the destination tracts across the Bay Area, I was able to find the transit travel time for every pair of points. I then looped through the equity neighborhood tracts again, summing the number of jobs in tracts that were accessible in under 30, 45, and 60 minutes. I also ran the same calculations with OpenTripPlanner's one-to-many function (OpenTripPlanner Contributors, 2020) rather than calculating each set of origin and destination points individually and compared the results with the initial job access numbers. The code used for this project is linked in Appendix 1. After performing this analysis, I used Matplotlib (The Matplotlib development team, 2022) with Contextily basemaps (Arribas-Bel & Contexily Contributors, 2020) to create maps of the level of job access by origin tract within equity neighborhoods.

I had planned to use an optimization algorithm to redistribute SFMTA service to increase job access for residents of equity neighborhoods, but have limited the scope to this initial exploration. I present my process, findings, and recommendations here so that SFMTA can hire consultants to continue this project.

Findings

Equity Neighborhood Demographics

I first looked at the demographics of the equity neighborhoods compared to San Francisco as a whole, using 2015-2019 ACS data by census tract for race/ethnicity, poverty status, and households without vehicle access. Table 2 shows the demographics of the Equity Neighborhoods compared to San Francisco as a whole.

Table 2. Selected demographic characteristics for Equity Neighborhoods and San Francisco as a whole

	San Francisco	Equity Neighborhoods
Latinx	15%	24%
Black	5%	10%
White	41%	27%
Asian	34%	35%
Other	5%	5%
Below 200% of poverty line	21%	30%
Households without vehicle	31%	44%

As Table 2 shows, compared to San Francisco as a whole, equity neighborhoods have higher proportions of Latinx and Black residents, residents living below 200% of the federal poverty line, and households without vehicles. They have lower proportions of white residents and approximately the same proportion of Asian and "other race" residents.

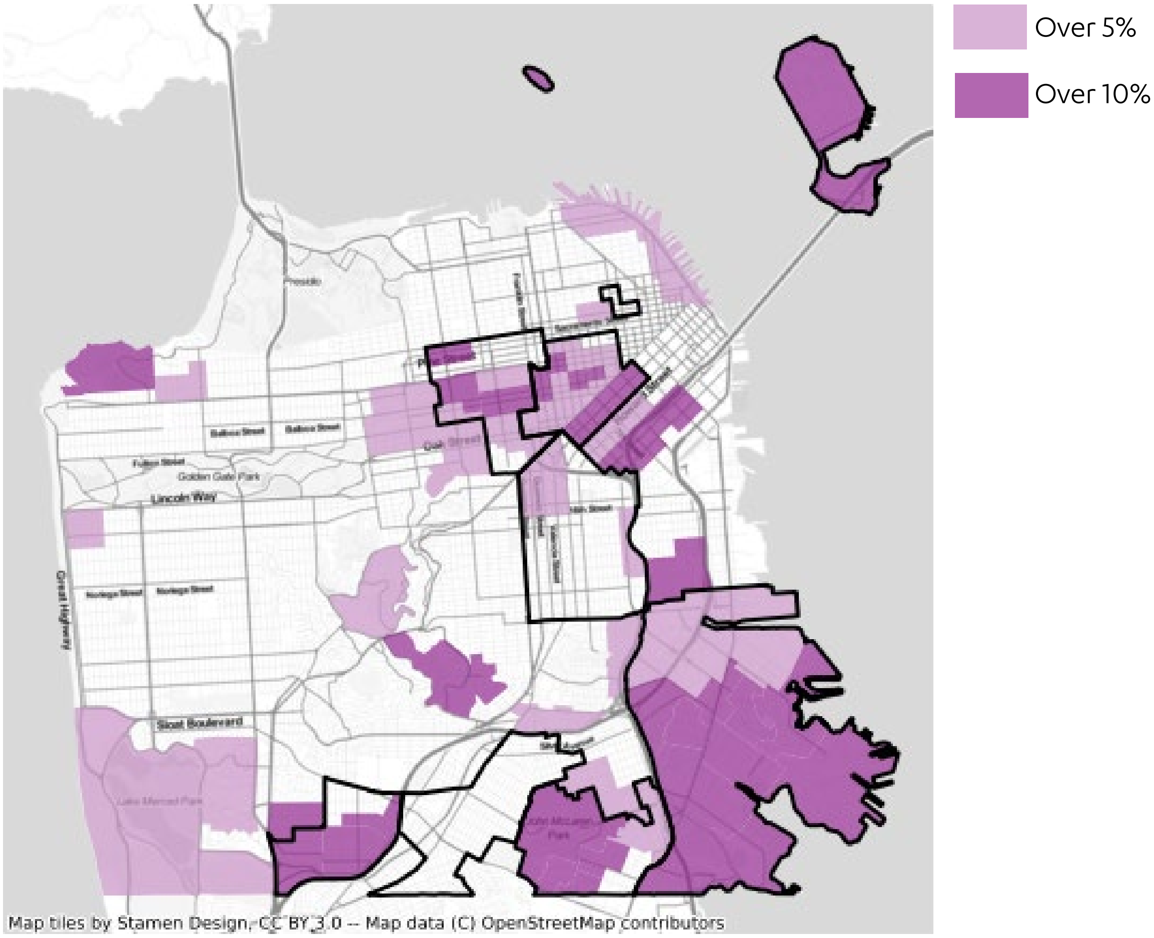
Table 3 shows the proportion of each group living in equity neighborhoods. Equity neighborhoods make about one third of San Francisco's total population and number of households but hold close to two thirds of its Black residents and around half of its Latinx residents and households with low income or no car access.

Table 3. Proportion of each group living in Equity Neighborhoods

Population	33%
Latinx	52%
Black	63%
White	22%
Asian	34%
Other	32%
Below 200% of poverty line	46%
Households	32%
Households without vehicle	45%

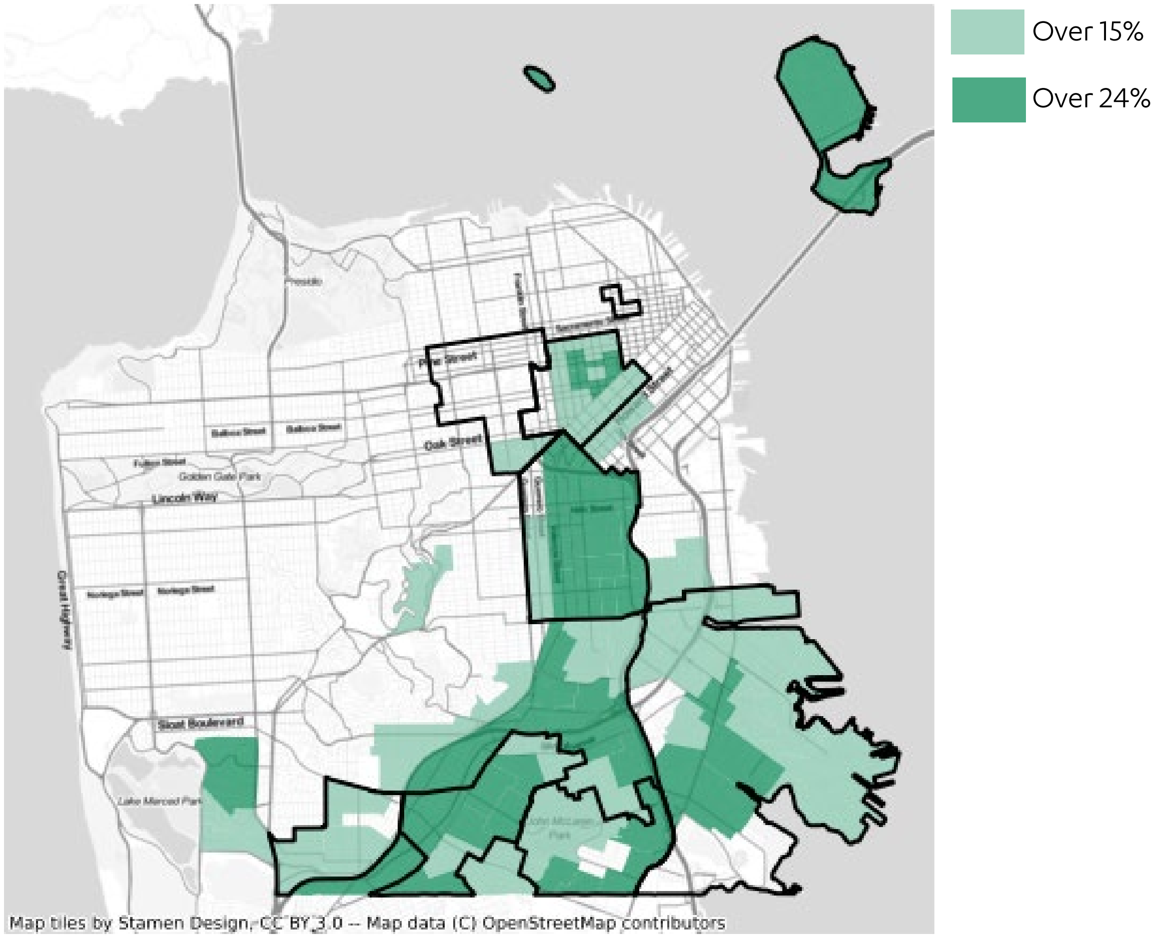
Figures 4, 5, 6, and 7 on the following pages illustrate all parts of San Francisco that have higher than average proportions of Black residents, Latinx residents, low-income households, and households without vehicle access, respectively. Each map shows areas with higher than the citywide average in a lighter color and areas with higher than the equity neighborhood average in a darker color.

Figure 4. SFMTA Equity Neighborhoods and proportion of Black residents



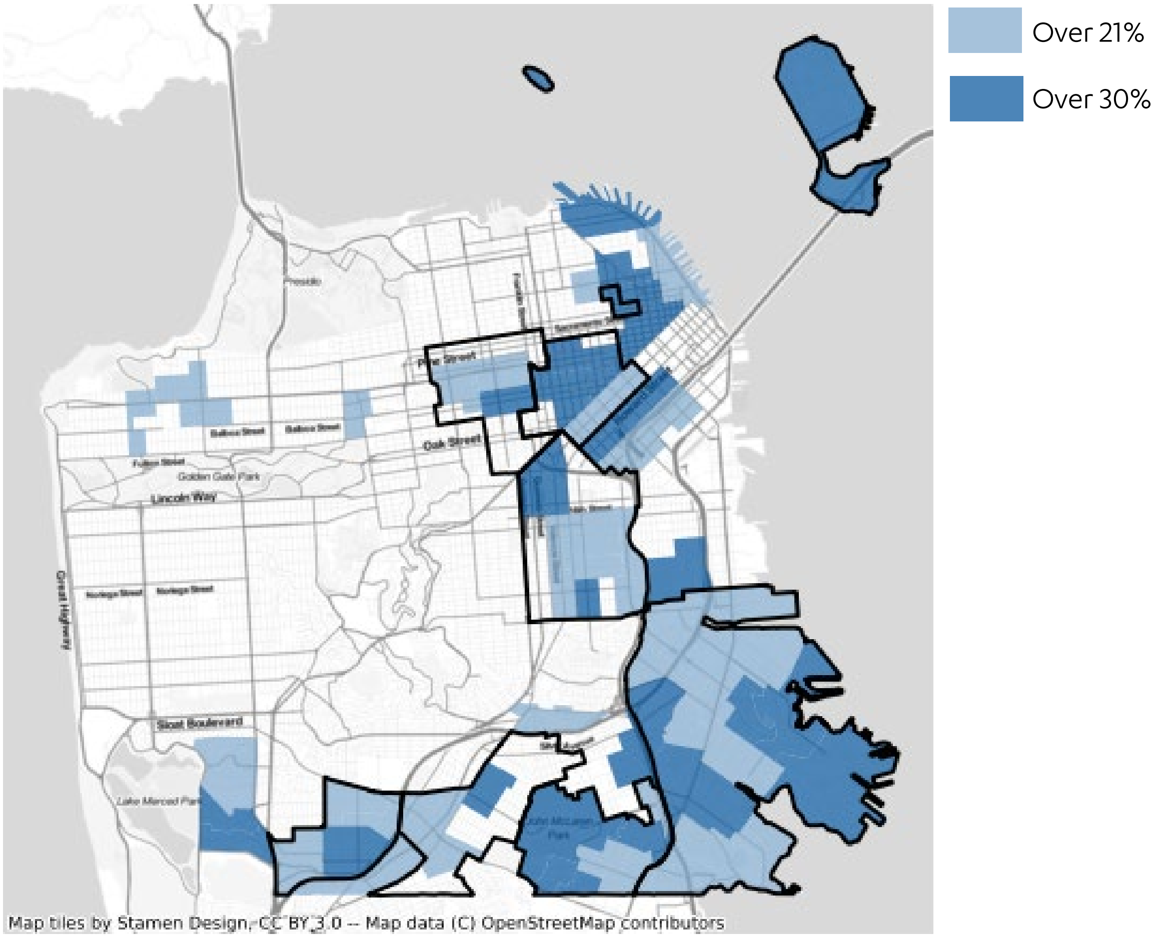
Of all demographic groups measured, Black residents are most clustered in equity neighborhoods, particularly in Bayview, Western Addition, Visitacion Valley, and Treasure Island. However, there are still several areas outside of the equity neighborhoods with a higher proportion of Black residents than the equity neighborhood average, in Lands End, Soma, Potrero Hill, Diamond Heights, and Merced Heights.

Figure 5. SFMTA Equity Neighborhoods and proportion of Latinx residents



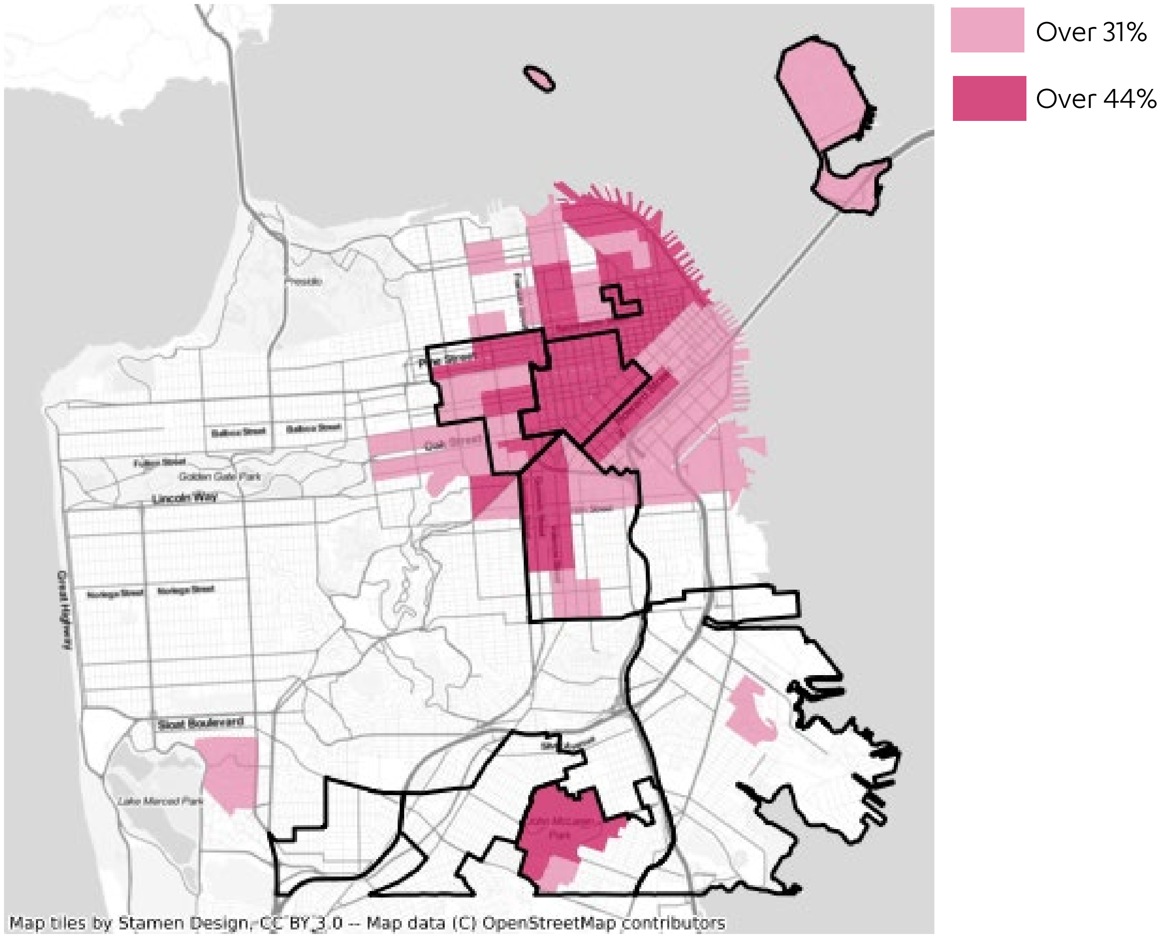
Latinx residents make up 15% of the population of San Francisco but 24% of the population of equity neighborhoods. They are particularly concentrated in the Mission, Bayview, and Outer Mission/Excelsior. However, there are also significant concentrations in Stonestown, Mission Terrace, and Bernal Heights.

Figure 6. SFMTA Equity Neighborhoods and proportion of population with low income



Despite the fact that equity neighborhoods house close to half of households with incomes under 200% of the federal poverty limit, several tracts with high proportions of these households are outside of these neighborhoods. These include parts of Chinatown outside the neighborhood boundary, the north waterfront, Parkmerced, and parts of Soma and southern Potrero Hill, which has affordable housing developments.

Figure 7. SFMTA Equity Neighborhoods and proportion of households without car access



Many neighborhoods adjacent to the Financial District, such as Telegraph Hill, North Beach, and Nob Hill, as well as Duboce Triangle, have rates of carlessness higher than the equity neighborhood average.

As these figures show, for no metric do the equity neighborhood boundaries line up with demographic reported by the census.

Jobs Access

Figure 8 shows the number of essential jobs available to residents of Equity Neighborhoods via transit in 30 minutes or less at 8:15am on a weekday morning. Residents of Downtown/Civic Center can access the greatest number of jobs. In particular, the area immediately north of Market Street between Powell and Leavenworth Streets has access to close to nearly 600,000 jobs in 30 minutes or less. This area is directly between the Powell and Civic Center combined BART and SF Muni stations.

Figure 8. Essential jobs available by transit in 30 minutes to residents of equity neighborhoods at 8:15am on a weekday

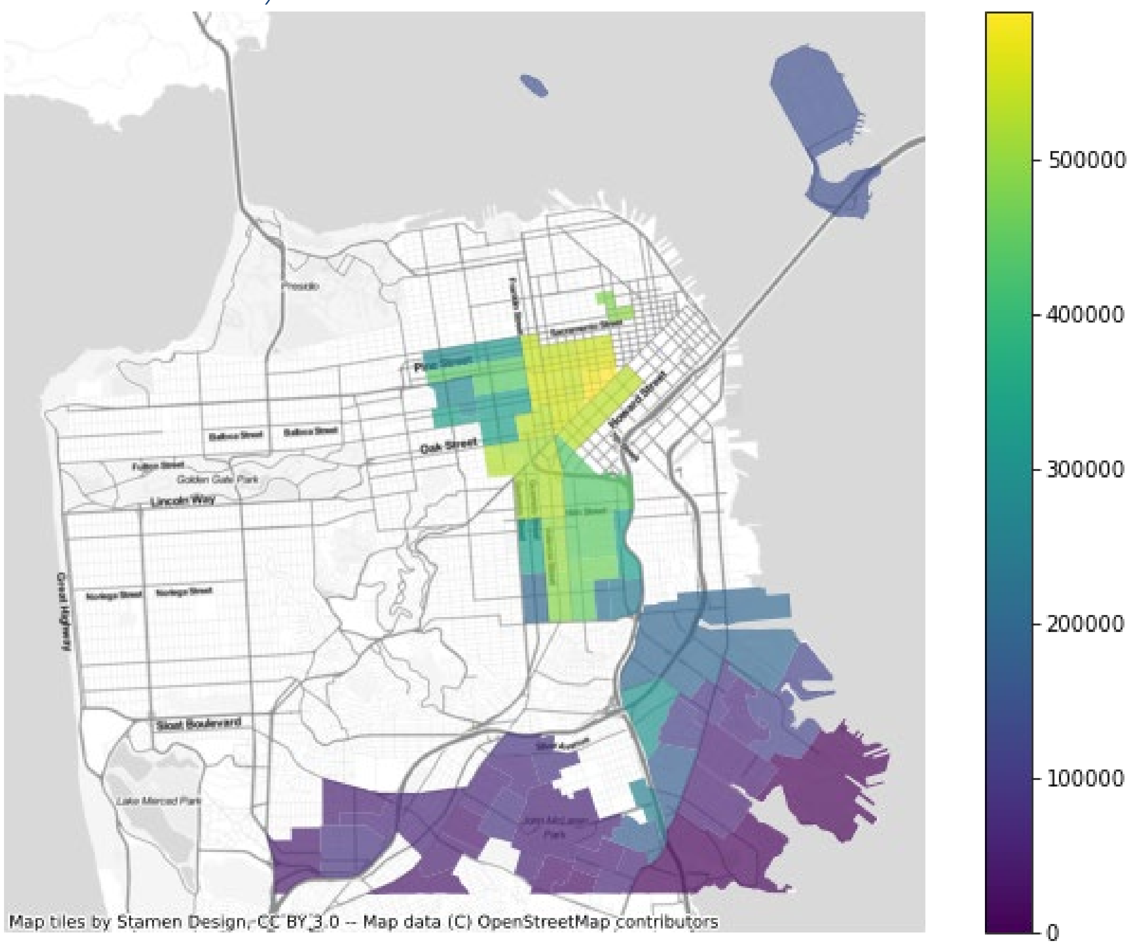


Figure 9 shows the number of essential jobs available to residents of Equity Neighborhoods via transit in 45 minutes or less at 8:15am on a weekday morning. The Downtown/Civic Center neighborhood still has relatively high job access, but the highest job access is in the area of the Mission immediately surrounding the 24th Street BART station.

Figure 9. Essential jobs available by transit in 45 minutes to residents of equity neighborhoods at 8:15am on a weekday

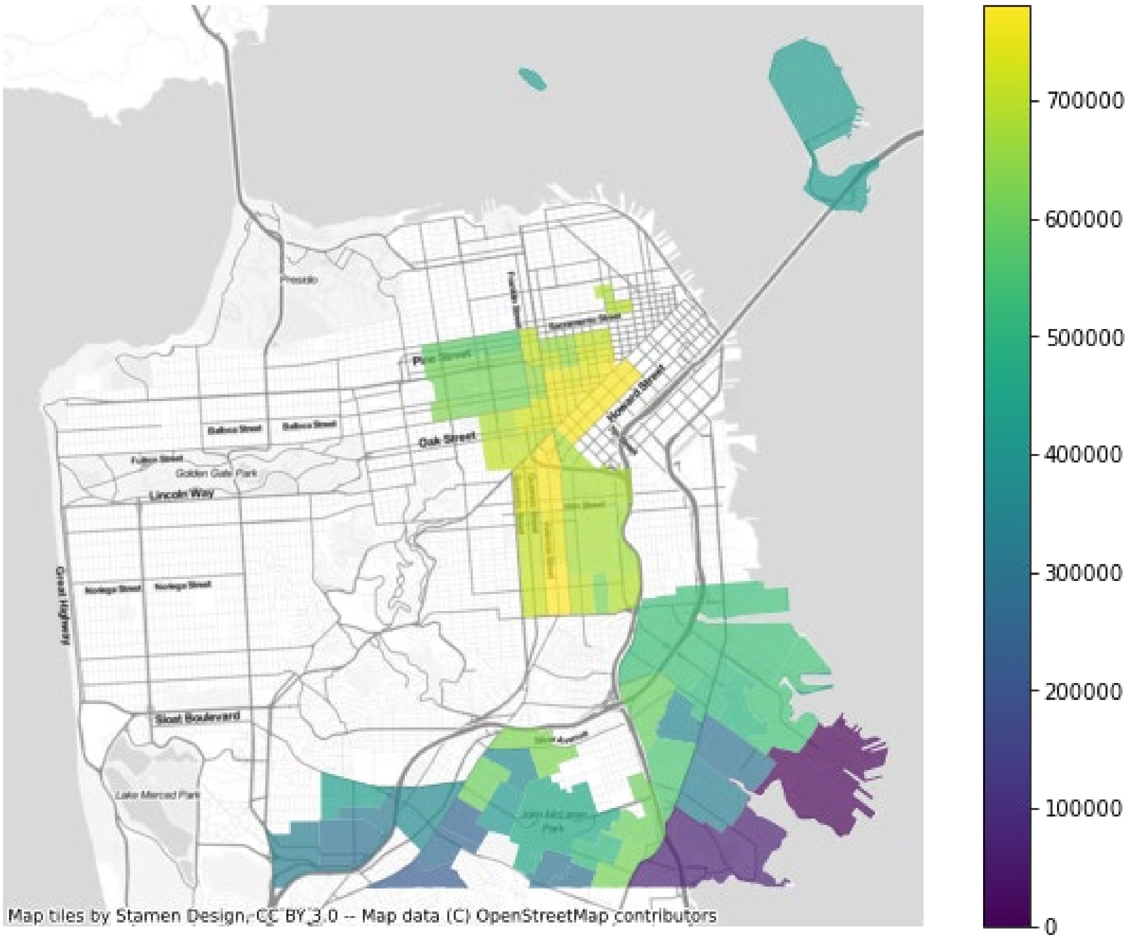
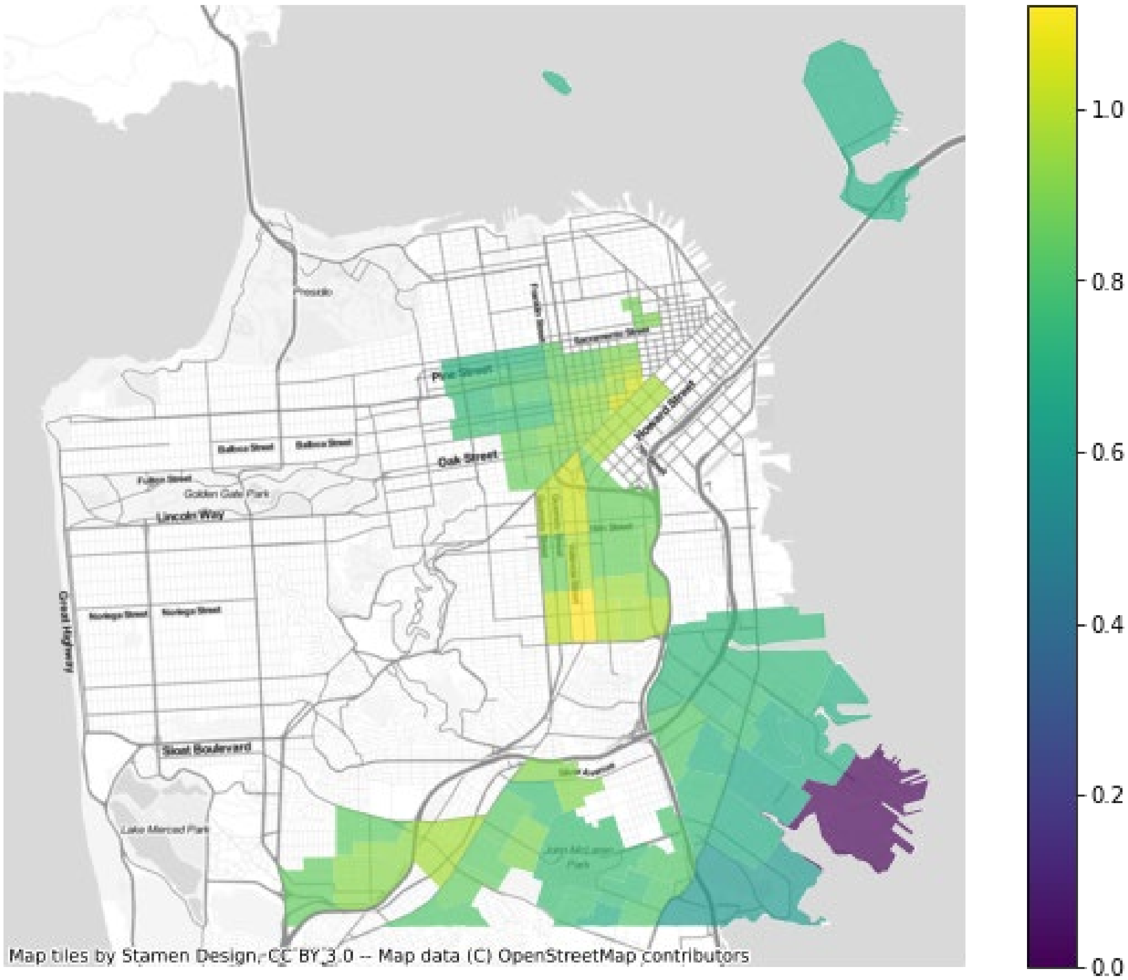


Figure 10 shows the number of essential jobs available to residents of Equity Neighborhoods via transit in 60 minutes or less at 8:15am on a weekday morning. The part of the Mission around the 24th Street BART station has the highest job access.

Figure 10. Essential jobs available by transit in 60 minutes to residents of equity neighborhoods at 8:15am on a weekday (in millions)



Perhaps surprisingly, at 45 and 60 minutes job access via transit in San Francisco is driven not by SFMTA but by BART. However, this analysis includes on job access, and does not measure where residents of equity neighborhoods actually work. Since BART extends much further across the Bay Area than SFMTA service, proximity to BART stations opens up many job opportunities.

One-to-Many Results

TransitCenter's code was written using point-to-point transit time calculations, so I attempted to improve upon it by using OpenTripPlanner's one-to-many function. Because one-to-many runs much more quickly (in about 3 minutes, compared to 5 hours for the point-to-point analysis), it would be much more useful in performing multiple measurements. This is particularly important for the sake of optimization, since most optimization algorithms involve rerunning the analysis dozens or hundreds of times with slightly modified input data in order to find the best output.

Unfortunately, though one-to-many was significantly faster, it was also less accurate. One-to-many often produced identical times for multiple origins, even though I configured it to take walking times to the rail station or bus stop into account. Comparison with Google Maps' transit directions showed that the point-to-point calculations were correct. Though I attempted to fix the issues with the one-to-many function with help from Dr. Adam Millard-Ball, my faculty advisor, I was unable to get closer to the point-to-point calculations. He pointed out that the one-to-many function has been deprecated in the most recent version of OpenTripPlanner, so consultants working on this project would not be able to use it. Figure 11 shows the difference in job accessibility by tract for the three different time periods. In many cases there is a difference of over 200,000 jobs, though the difference is not consistently in one direction or the other.

Figure 11. Difference in jobs accessibility calculations using point-to-point (blue) and one-to-many (red) analyses

Recommendations

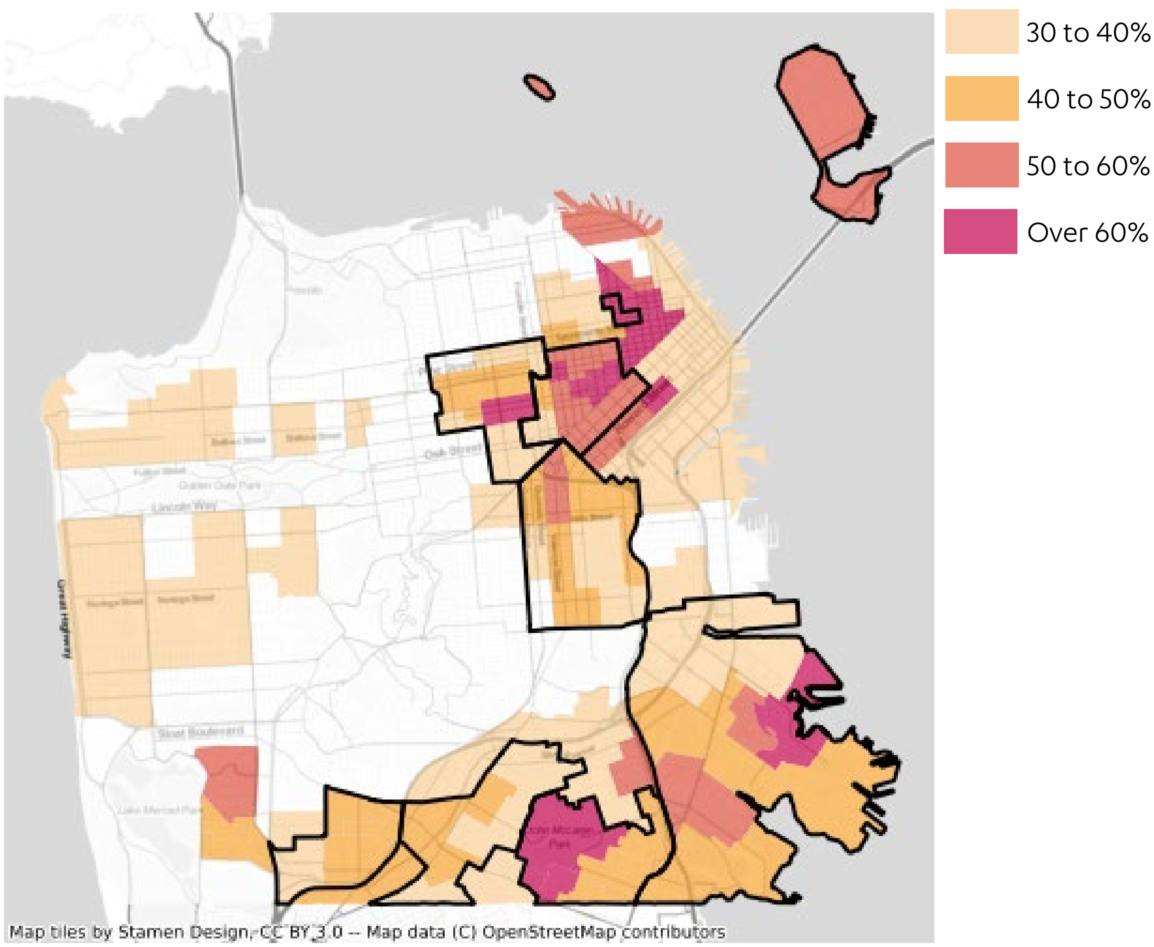
This project is the first step in using data science to optimize SFMTA's service. If SFMTA plans to continue this work using open-source tools like OpenTripPlanner, their team should invest enough computational power to be able to run point-to-point calculations much more quickly. The Accessibility Observatory at the University of Minnesota Center for Transportation Studies uses OTP to measure job access via transit, calculating job access for each origin point in 50 American cities at every minute between 7 and 9 am to calculate an average accessibility metric (Owen & Murphy, 2020). This is similar to SFMTA's current equity analysis, which uses Esri's ArcGIS products to create a transit isochrone from each origin point for each minute in a half-hour period, and then average the total number of jobs accessible at each minute. Beyond speeding up the current calculations in order to perform the analysis, this section makes recommendations to improve the final tool.

A step to improve this tool would be to use block group rather than census tract data, particularly for the job data across the Bay Area. Though calculating the precise job access for every address in the city would be computationally impossible, using block groups would add a layer of granularity that would improve the accuracy of the results. Within the City of San Francisco itself, census tracts are quite small, often no more than a few blocks wide, because of the city's density. However, in other, less dense, parts of the region, tracts can be much more spread out, as shown in Figure 3 above. Using smaller geographies more accurately estimates the walk time at the beginning and end of the transit trip, which, while geographically a small portion of the trip, is a significant part of the time spent in transit (Levine et al., 2019).

Although San Francisco policymakers are eager to use the Equity Neighborhoods as a metric, Table 3 above shows that they hold less than half of all households with incomes under 200% of the federal poverty level or households without a car. Additionally, while their residents are disproportionately Black and Latinx, they hold less than two thirds of the Black residents of San Francisco, and only about half of the Latinx residents. Figures 4 to 7 above shows that some areas of the city outside the Equity Neighborhoods have equal or greater proportions of the selected demographic characteristics.

To take into account all marginalized people in San Francisco, rather than only those who live in Equity Neighborhoods, this tool could take as its input the total number of low-income people in all block groups, rather than just total population of just those census tracts in the equity neighborhoods. If stakeholders want a more holistic picture, SFMTA staff could calculate the "Equity Population" of each block group, using a combination of race/ethnicity, income, and vehicle access similar to that used to determine the equity neighborhoods now. The rest of the tool would run as it does now, since I have designed it to take in any file with population data and associated geometries. Figure 12 shows one version of such an equity index with an even average of populations of color, households below 200% of the federal poverty line, and households without vehicle access. SFMTA could, of course, choose to incorporate other variables as well, such as people with disabilities.

Figure 12. Equity index and equity neighborhoods



In fact, SFMTA could take these analyses further by considering all disadvantaged people across the Bay Area, optimizing service so that residents of other parts of the Bay can access jobs in San Francisco, rather than only optimizing for San Francisco residents' accessibility. This is a particularly important as housing prices rise and more low-income people are pushed further from San Francisco.

As currently planned, this tool does not use any data on when jobs start and end. Especially as more white-collar workers continue to work from home after the COVID-19 pandemic, SFMTA may need to move service to early morning or late evening to accommodate essential workers who work different shifts. Ideally future iterations of this tool would include data on work schedules.

Though as planned tool would not generate new routes, SFMTA staff could test out possible new routes by adding them to the GTFS feed with an initial frequency of zero, running the tool, and seeing what the output frequency is.

Bibliography

Al Mamun, Md., & Lownes, N. (2011). A Composite Index of Public Transit Accessibility. *Journal of Public Transportation*, *14*(2). https://doi.org/10.5038/2375-0901.14.2.4

Arribas-Bel, D. & Contexily Contributors. (2020). *Contextily: Context geo tiles in Python—Contextily 1.1.0 documentation*. https://contextily.readthedocs.io/en/latest/

Carleton, P. R., & Porter, J. D. (2018). A comparative analysis of the challenges in measuring transit equity: Definitions, interpretations, and limitations. *Journal of Transport Geography*, *72*, 64–75. https://doi.org/10.1016/j.jtrangeo.2018.08.012

Census Reporter. (n.d.). *Census Reporter: Making Census Data Easy to Use*. Retrieved November 21, 2021, from https://censusreporter.org/

Curl, A., Nelson, J. D., & Anable, J. (2015). Same question, different answer: A comparison of GIS-based journey time accessibility with self-reported measures from the National Travel Survey in England. *Computers, Environment and Urban Systems*, *49*, 86–97. https://doi.org/10.1016/j.compenvurbsys.2013.10.006

Duranton, G., & Guerra, E. (2016). *Developing a Common Narrative on Urban Accessibility: An Urban Planning Perspective* (Moving to Access, p. 43). Brookings.

El-Geneidy, A., Buliung, R., Diab, E., van Lierop, D., Langlois, M., & Legrain, A. (2016). Non-stop equity: Assessing daily intersections between transit accessibility and social disparity across the Greater Toronto and Hamilton Area (GTHA). *Environment and Planning B: Planning and Design*, *43*(3), 540–560. https://doi.org/10.1177/0265813515617659

Glaeser, E. L., Kahn, M. E., & Rappaport, J. (2008). Why do the poor live in cities The role of public transportation. *Journal of Urban Economics*, *63*(1), 1–24.

Governor's Institute on Community Design. (2017). *The How and Why of Measuring Access to Opportunity: A Guide to Performance Management*.

Handy, S. L., & Niemeier, D. A. (1997). Measuring Accessibility: An Exploration of Issues and Alternatives. *Environment and Planning A*, *29*, 1175–1194.

Hansen, W. G. (1959). How Accessibility Shapes Land Use. *Journal of the American Institute of Planners*, *25*(2), 73–76. https://doi.org/10.1080/01944365908978307

Interline Technologies. (2021). *Documentation • Transitland*. https://www.transit.land/documentation/

Jomehpour Chahar Aman, J., & Smith-Colin, J. (2020). Transit Deserts: Equity analysis of public transit accessibility. *Journal of Transport Geography*, *89*, 102869. https://doi.org/10.1016/j.jtrangeo.2020.102869

Karner, A., & Golub, A. (2015). Comparison of Two Common Approaches to Public Transit Service Equity Evaluation. *Transportation Research Record*, *2531*(1), 170–179. https://doi.org/10.3141/2531-20

Karner, A., London, J., Rowangould, D., & Manaugh, K. (2020). From Transportation Equity to Transportation Justice: Within, Through, and Beyond the State. *Journal of Planning Literature*, *35*(4), 440–459. https://doi.org/10.1177/0885412220927691

Kawabata, M. (2003). Job Access and Employment among Low-Skilled Autoless Workers in US Metropolitan Areas. *Environment and Planning A: Economy and Space*, *35*(9), 1651–1668. https://doi.org/10.1068/a35209

Kawabata, M., & Shen, Q. (2006). Job Accessibility as an Indicator of Auto-Oriented Urban Structure: A Comparison of Boston and Los Angeles with Tokyo. *Environment and Planning B: Planning and Design*, *33*(1), 115–130. https://doi.org/10.1068/b31144

Klumpenhouwer, W., Allen, J., Li, L., Liu, R., Robinson, M., Silva, D. D., Farber, S., Karner, A., Rowangould, D., Shalaby, A., Buchanan, M., & Higashide, S. (2021). A Comprehensive

Transit Accessibility and Equity Dashboard. *Findings*, 25224. https://doi.org/10.32866/001c.25224

Lättman, K., Olsson, L. E., & Friman, M. (2018). A new approach to accessibility – Examining perceived accessibility in contrast to objectively measured accessibility in daily travel. *Research in Transportation Economics*, *69*, 501–511. https://doi.org/10.1016/j.retrec.2018.06.002

Levine, J., Grengs, J., & Merlin, L. A. (2019). *From mobility to accessibility: Transforming urban transportation and land-use planning*. Cornell University Press.

Lyons, T., & Ewing, R. (2021). Does transit moderate spatial mismatch? The effects of transit and compactness on regional economic outcomes. *Cities*, *113*, 103160. https://doi.org/10.1016/j.cities.2021.103160

McHugh, B. (2013). Pioneering Open Data Standards: The GTFS Story. In B. Goldstein & L. Dyson (Eds.), *Beyond Transparency: Open Data and the Future of Civic Innovation*. Code for America Press. https://gtfs.org/

Millard-Ball, A., West, J., Rezaei, N., & Desai, G. (2022). What do residential lotteries show us about transportation choices? *Urban Studies*, *59*(2), 434–452. https://doi.org/10.1177/0042098021995139

Miller, E. J. (2018). Accessibility: Measurement and application in transportation planning. *Transport Reviews*, *38*(5), 551–555. https://doi.org/10.1080/01441647.2018.1492778

Nelson, E. (2020). *Where Do the Poor Live in Cities? Revisiting the Role of Public Transportation on Income Sorting in US Urban Areas* (SSRN Scholarly Paper ID 3628712). Social Science Research Network. https://doi.org/10.2139/ssrn.3628712

Ong, P. M., & Houston, D. (2002). Transit, Employment and Women on Welfare1. *Urban Geography*, *23*(4), 344–364. https://doi.org/10.2747/0272-3638.23.4.344

OpenTripPlanner Contributors. (2016). *OpenTripPlanner 2*. https://docs.opentripplanner.org/en/latest/

OpenTripPlanner Contributors. (2020). *One-to-many Travel Times*. http://docs.opentripplanner.org/en/dev-1.x/Surface/

Owen, A., & Murphy, B. (2020). *Access Across America: Transit 2019* (p. 172). University of Minnesota Accessibility Observatory.

Pereira, R. H. M., Schwanen, T., & Banister, D. (2017). Distributive justice and equity in transportation. *Transport Reviews*, *37*(2), 170–191. https://doi.org/10.1080/01441647.2016.1257660

Raphael, S. (1998). The Spatial Mismatch Hypothesis and Black Youth Joblessness: Evidence from the San Francisco Bay Area. *Journal of Urban Economics*, *43*(1), 79–111. https://doi.org/10.1006/juec.1997.2039

San Francisco Municipal Transportation Agency. (2016). *Muni Service Equity Strategy Report: Fiscal Year 2016-17 and 2017-18*. https://www.sfmta.com/sites/default/files/agendaitems/2016/4-5-16%20Item%2013%20Muni%20Equity%20Strategy%20Report_1.pdf

Smart, M. J., & Klein, N. J. (2015). *A Longitudinal Analysis of Cars, Transit, and Employment Outcomes* (CA-MNTRC-15-1244). https://rosap.ntl.bts.gov/view/dot/29394

Sundquist, E., McCahill, C., & Brenneis, M. (2021). *Measuring Accessibility: A Guide for Transportation and Land Use Practitioners*. https://trid.trb.org/view/1859213

Taylor, B., & Morris, E. (2015). Public transportation objectives and rider demographics: Are transit's priorities poor public policy? *Transportation*, *42*(2), 347–367.

The Matplotlib development team. (2022). *Matplotlib—Visualization with Python*. https://matplotlib.org/

Wachs, M., & Kumagai, T. G. (1973). Physical accessibility as a social indicator. *Socio-Economic Planning Sciences*, *7*(5), 437–456. https://doi.org/10.1016/0038-0121(73)90041-4

Yeganeh, A. J., Hall, R. P., Pearce, A. R., & Hankey, S. (2018). A social equity analysis of the U.S. public transportation system based on job accessibility. *Journal of Transport and Land Use*, *11*(1), 1039–1056.

Appendix 1: Github Code

For the code behind this analysis, please visit https://github.com/rctect101/sfmta_equity