The Ease and Equity of Point of Interest Accessibility via Public Transit in the US - Team 12

Alexander Li
Georgia Institute of Technology

Mengyang Liu
Georgia Institute of Technology

Aurimas Racas
Georgia Institute of Technology

Tejas Santanam
Georgia Institute of Technology

Junaid Syed
Georgia Institute of Technology

Przemek Zientala
Georgia Institute of Technology

1 INTRODUCTION AND PROBLEM

The tool developed as a result of this project analyzes the ease and equity of access to major POI categories (e.g. vaccination centers, grocery stores, hospitals) using public transit in major US cities. We built an interactive website that enables easy exploration of current access equity and allows performing scenario analysis by introducing/removing POIs.

Many populations rely on public transit for access to work, food, family, and more. By giving city planners a tool to see the gaps in public transit access, we hope to improve consideration of transportation accessibility in future infrastructure projects. Currently, decision-making around site placement in urban environments often fails to consider accessibility and equity as major criteria. An even larger issue lies with the fact that even transit planning, at least in Atlanta, is largely accessibility-blind and is focused more on ridership counts at different points in the city. However, what that approach fails to consider is places that have lower ridership due to accessibility and connectivity issues with POIs that people need to access. This can lead to a cycle where infrastructure development and site placement is not focused on the opportunity areas where ridership can develop if access to resources like work, food, vaccination, and more is prioritized.

Our work aims to provide a holistic overview of public transit system effectiveness and equity in major US cities, which may reveal previously unknown gaps. We will provide interactive interfaces that will allow decision-makers and the general public to explore and understand what drives public transit network effectiveness and how difficult would it be to make improvements to accessibility and equity. We base our work on data sources that are also available outside the U.S. whenever possible, and thus our work may serve as a framework for future work that extends to non-U.S. geographies.

2 SURVEY

Access to facilities by public transit is key to reducing the vulnerability of population segments that do not have access to a private car, with researchers calling for more attention to public transit in accessibility modelling since late 2000's [11].

COVID-19 generated a flurry of research showing the importance of public transit. Multiple studies looked at vaccination rates, showing that they are lower in rural areas [13], minority communities [17], and persist despite efforts to address them [10]. Public transit was shown to be an important factor. Research on informal population in Rome found access to public transit especially important to elderly people [2]; a review of US state-level COVID-19 vaccination strategies found that public transit was a critical component in reducing inequities, with strategies used including ride subsidies, financial support for the public transit authorities, and deployment of ride-sharing services [19].

Importance of public transit access is not limited to vaccination centers. Spatial access is critical to ensuring quality health services [21], especially in socially vulnerable regions [20]. Planners and policy makers are interested in identifying areas with low grocery store coverage [9], understanding how public transit impacts equity of job location access [8] and measuring access to educational facilities [22], among others.

With public transit providers suffering from lower ridership and service cuts due to COVID-19 [4], an updated and comprehensive research into public transit accessibility and its equity is critical, with such information useful to policy makers, city planning officials and for everyday citizens looking to be better informed on the social problems in their neighborhoods.

Most recent research focuses on vaccination center (VC) ease of access and typically relies on catchment area approaches. In the US, "vaccine deserts" were identified using enhanced two-step catchment area approaches (E2SFCA), and concluded that VCs were not effectively targeting the most vulnerable groups [16]. A study in Florida utilized two step floating catchment (2SFCA) and E2SFCA methods to show hospital access differences across the state [7]. In England, research has concluded that taking travel time into account reveals significant VC coverage disparities not visible in the official statistics [6]. In Germany, it was found that the VCs are distributed equitably in urban/rural areas, and the key determinant to access is the transport mode used instead [14].

There is existing research focused on other types of POIs, too. In Brazil, research looked at healthcare system capacity under COVID-19 incorporating data about income and racial inequities, they utilized the B2SFCA method and found that intensive care unit equipment availability was substantially lower among Black and poor communities [15]. A study in Tamil Nadu utilized the 3SFCA method to calculate the accessibility index of healthcare facilities and find optimal locations to setup new ones [18]. Research in Iran has shown that spatial analysis techniques (E2SFCA) can be used to select public hospitals and public healthcare centers for optimal coverage of the population [12]. Research in Hong Kong utilized gravity models to measure the spatial accessibility of urban parks and found that public transport significantly increased the spatial inequality of urban park accessibility between public and private housing residents [3]. A study in US proposed the KD2SFCA method by incorporating the kernel density function into the 2SFCA method to highlight the geographic disparities in accessibility to food stores [5].

3 METHODOLOGY

3.1 Data

3.1.1 Data sources. Our data consists of three primary categories. The first one is the US Census (American Community Survey - 2020) which contains information about demographics and census block group polygons. This data is collected by the US government on a census block group level and obtained via SafeGraph. We extracted population amounts, race breakdowns, no-car percentages, income levels by bracket, age breakdown, and sex breakdown for each census block group. The next category is Point of Interest (POI) data from OpenStreetMap (OSM) (collected using Overpass API). We extracted the latitude/longitude and names of schools, theaters, restaurants, hospitals, and grocery stores. This data is joined with vaccination site data obtained from the Center for Disease Control (CDC). Finally, the last category of data consists of public transit scheduling and stop information obtained via the General Transit Feed Specification (GTFS) data feeds that are provided by public transit operators (we used transitfeeds.com website to identify where the datafeeds are available).

The aforementioned datasets were collected for five of the largest metro areas in the US (New York City, Los Angeles, Chicago, Dallas, and Atlanta). The total data size in our database is 11 gigabytes. Our dataset consisted of 60,000 vaccination centers, 14,000 restaurants, 4,500 grocery stores/supermarkets, 4,200 (pre-)schools, 700 clinics/hospitals, and 500 cinemas/theatres.

3.2 Data preparation

3.2.1 Catchment area estimation. For each POI, we estimated their coverage areas (referred to as catchment areas thereon) by generating isochrones - areas delineating a boundary reachable within a given travel time. We generated 30-minute isochrones based on morning, afternoon and evening public transit schedules. We used a Dockerized OpenTripPlanner server for this purpose; it relies on aforementioned GTFS transit feeds and OSM street-network information to generate accurate isochrones.

3.2.2 Hierarchical spatial indexing system. To ensure easier and consistent calculations and more readable visualizations, we mapped all our data points (demographics data, POI coordinates and catchment areas) to the H3 hierarchical spatial indexing system (developed by Uber). We used H3 hexagons with an area size of $\approx 0.1 km^2$, corresponding to H3 resolution level 9. This resulted in M:N relationships between census polygons and H3 hexagons.

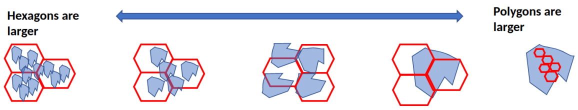

Figure 1: Overlap types between H3 hexagons and census polygons. Note that, with decreasing hexagon area, the relationship of polygon to hexagon IDs becomes closer to pure 1:N

Using smaller hexagons resulted in smaller margins of error in our analysis. In the even that information (e.g. demographic data of a census block group) spanned multiple H3 hexagons, the allocation was made in proportion to the percentage area covered in each hexagon. In total, we calculated information for 137,000 unique H3 hexagons across the United States.

3.3 Accessibility index calculation

The accessibility index was calculated using a 2SFCA (2-step floating catchment area) approach [7]. Out of all the possible methods, this method was chosen as it is the simplest method for policymakers to understand. Since we wanted our tool to have implementable value, we felt that the 2SFCA approach would not confuse policymakers while still conveying significant insights into accessibility. While the other methodologies may be more complex, the accessibility in understanding for a non-technical person seemed to be a larger benefit. The calculation for each census block (group) was done as follows:

1. For each POI ID, get all catchment areas. For each catchment area, calculate POI-to-population ratio as follows:

$$R_j = \frac{|supplyunits|}{\sum_j catchmentpop_j}$$

2. For each H3 hexagon, sum R_j of all catchment areas which overlap with the hexagon. Weigh that sum by proportion of population for a given group as a fraction of entire population in that hexagon. The sum is the final accessibility score:

$$A_h = \sum_{j \in j \cap h} R_j$$

3.4 User interface

We built an interactive website where users can explore the accessibility statistics and engage in scenario-like analysis. The main technologies used include Vue.js and Deck.gl Javascript libraries, supported with a FastAPI Python backend. A screenshot of the user interface is shown in the figure below.

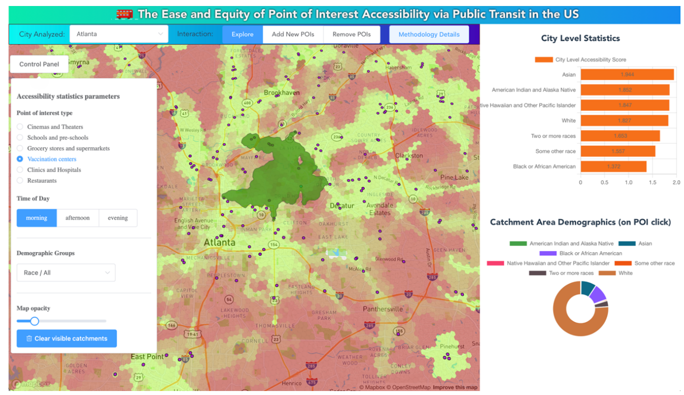

Figure 2: User Interface

Upon entering the site, users are shown the accessibility of vaccination sites in Atlanta. In the control panel dropdown on the left, the users can adjust different factors and select different cities, POIs (Cinemas and Theaters, Schools and Pre-schools, Grocery Stores, Vaccination Centers, Hospitals and Clinics, Restaurants), times of day (morning, afternoon, evening), and demographic factors (Race, Age & Sex, Income, Vehicle Availability). The users can then go through different demographic options to see how transit accessibility to the chosen POI type differs for different demographic groups. Visually, the H3 hexagons on the map will display in different colors to indicate accessibility scores to that POI type, and upon hovering, a tooltip will show for each hexagon containing the selected demographic information. The tool will also allow users to place new potential POIs and remove existing POIs and see the recalculated impact on the demographic accessibility metrics displayed on the charts on the right hand side of the page. This kind of situational analysis is a unique feature of our tool. New POIs will be added in a different color from existing POIs so that the user does not conflate the two types. Users can also click on each POI to see the 30-minute isochrone on the map. This isochrone can also be generated for new hypothetical POIs added by the user. Additionally, the analysis displayed can be changed based on time of day, POI type, and desired demographic analysis. For example, the placement of some POIs might use age as a motivating factor (hospitals might want to locate closer to older age groups). Or perhaps, the placement process might require time of day analysis (theaters are mostly attended in the evening). Thus, the tool allows decision makers to view and get a deeper understanding on the factors that matter to them. The tool also contains analysis for Atlanta, New York City, Los Angeles, Chicago, and Dallas. The user interface is simple to use and allows decision makers to clearly see where the gaps in access are, as well as how significant those gaps are, for both transit system planning and POI location planning purposes.

3.5 Innovations

The primary innovation is the ability to quantify and compare accessibility metrics across cities, POI categories, and populations segments beyond solely data reporting. Additionally, the focus on creating an interactive visual tool using H3 polygons for population discretization for decision makers is something that hasn't been done before to our knowledge. Since census blocks are of differing sizes and can have changing boundaries every census year, H3 hexagons allow for a standardized way to compare different parts of the city in equally sized geographies. Furthermore, the scenario analysis that recalculates accessibility statistics based on adding and removing POIs is a unique feature. Lastly, the granularity down to the census block group level in analysis allows for accuracy in equity analysis that commonly has only been done to the metropolitan statistical area (MSA) level.

4 EXPERIMENTS AND EVALUATION

4.1 Comparisons and Atlanta Takeaways

In the figure below, we can see the comparison of weighted accessibility across cities broken down by POI type.

Weighted accessibility is the accessibility for each H3 hexagon normalized by population. Atlanta actually has the highest accessibility for theaters, hospitals, and schools out of the five cities analyzed. In addition, Atlanta does not have the lowest accessibility for any

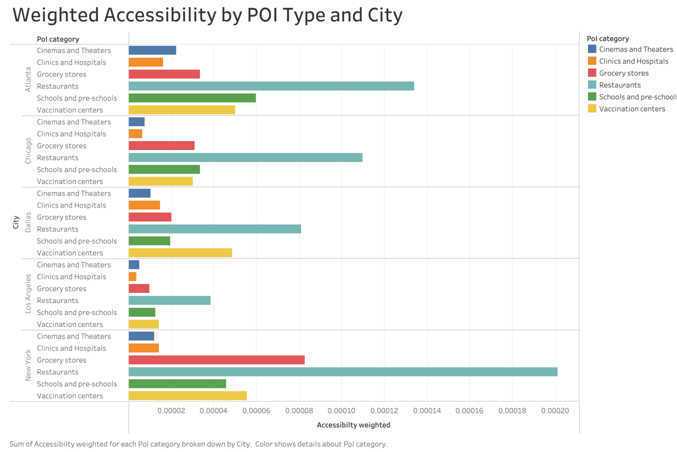

Figure 3: Weighted Accessibility by POI Type and City

of the POI types. New York City has the highest accessibility for grocery stores, restaurants, and vaccination centers. Los Angeles has the lowest accessibility for all POI types, a testament to the dysfunction of its transit system. However, even though a city has a higher accessibility score, that does not mean it works for everyone. For example, even though Atlanta has the highest accessibility for clinics and hospitals, the accessibility of those hospitals is twice as high for White populations in Atlanta as compared to Black populations, as seen in the figure below. The accessibility scores can be validated by, for example, the fact that New York City has the highest restaurant accessibility. That intuitively makes sense due to the population and restaurant density New York is known for, thus serving as a sanity check for our numbers.

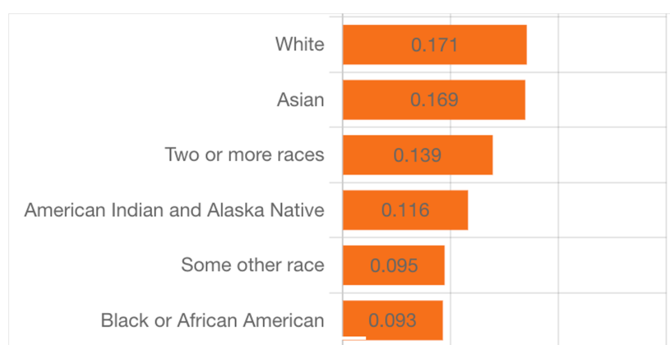

Figure 4: Hospital Accessibility Index by Race in Atlanta

Similarly, access to grocery stores in Atlanta is 70% higher for White populations than Black populations. For example, in Chicago, the gap in access for grocery stores is much smaller between White and Black populations, while for hospitals in Chicago, Black populations have a higher accessibility than White populations. Even Los Angeles, which has the worst overall accessibility across POIs, has similar accessibility for White and Black populations to all POIs. This highlights a suboptimal transit system in Los Angeles, but at least that system serves people equitably. In Atlanta, the major issue is the lack of transit infrastructure and reachability in the South and Southwest portions of the city. MARTA often lacks reach in those neighborhoods that are home to majority Black populations. This creates a transit divide where entire groups of people are left disconnected from the city. The fact that the ML models run were able to predict majority race easily solely on accessibiltiy score for a hexagon underscores this fact. Just like racial divides, there are also divides in income in Atlanta. For grocery stores and hospitals, those who make over $200,000 a year have the highest accessibility via public transit. However, in Los Angeles, those who make under $10,000 a year have the highest accessibility to grocery stores and hospitals, which are often the people who use public transit the most.

4.2 Statistical analysis

Based on the accessibility metrics, we wanted to see if we could predict the majority race based on accessibility index for each H3 hexagon in Atlanta. As an example, we present results for vaccination centers, with time of day being the afternoon. We developed a Random Forest model for this classification task. We found that the model was able to predict majority race in case of Black and While populations with high certainty (F1 scores 0.85) indicating that there is an association between race and accessibility and thus pointing to the inequity of accessibility between them. The model did not perform well in case of minority races (e.g. Asian majority was predicted with a F1 score of 0.11, other races had an F1 score of 0) which is not unexpected given that there are very few locations where these races are in the majority. A confusion matrix for test set can be seen below in Figure 3 illustrating this finding.

We also analyzed the inequity dynamics by using demographic characteristics of H3 hexagons to predict accessibility scores to answer the opposite question. We used an Extra Tree Regressor model for the prediction. The most important demographic characteristics for accessibility prediction, in order, were race, age/sex, vehicle availability, and income. There were also explorations made using Linear Regression models, but the Extra Tree Regressor performed the best (RMSE of 0.2).

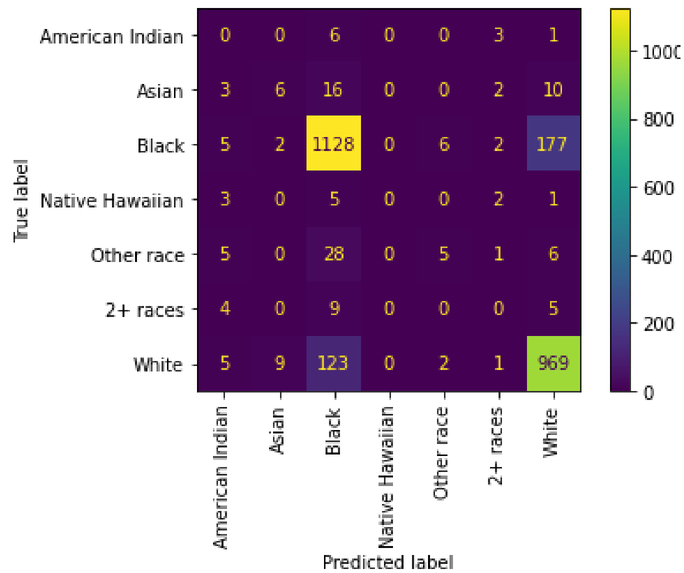

Figure 5: Test set confusion matrix. Note the high predictive power for Black and White populations.

We plan to release the application for use upon completion of the semester, and our success will be measured by the interest of policy makers, city planners, and corporations in the tool, as well as the value they find in the tool's analysis.

4.3 External validation with policy makers

Last week, we took this tool to the City of Atlanta to get feedback. Our team met with Doug Nagy, the Deputy Commissioner of the Atlanta Department of Transportation (ATLDOT) and Jordan Dowdy, the Head of Data and Analytics. They loved the features the tool had, and mentioned that it was in line with factors they have been thinking about. Evidence of Doug's praise can be seen in the LinkedIn post below.

They validated the analysis of the tool by mentioning that the areas highlighted in the tool as lower/higher accessibility areas were in line with the general idea

Figure 6: Doug Nagy praised the tool on LinkedIn and encouraged us to go to market

they had, but that this tool allowed them to quantify they extent of that accessibility. Doug and Jordan found the tool extremely useful overall and also encouraged us to explore possible market opportunities by selling the tool to companies so they can see the accessibility of their offices, warehouses, HQs, etc. to their workforce and labor market. Doug Nagy mentioned that he would be willing to hand-deliver a proposal of our tool to the Atlanta Committee for Progress next week, the members of which comprise the CEOs and leaders of the top corporations in Atlanta. He believed there could even be interest in our tool to the tune of $50,000 per company. As far as site performance, the load times across cities, times of day, POI types, and demographic groups were consistent which is a good sign for usability. Doug and Jordan did not see any further usability issues with the site either. From a use case perspective, an alternative idea they suggested was to also look at zoning, land use, and land value relative to the transit infrastructure in Atlanta.

5 CONCLUSION AND DISCUSSION

Throughout this project, we have been able to create a tool that empowers city and transit planners to make better decisions with regards to accessibility and equity. By providing granular analysis and an easy-to-understand visual tool, we were able to impress officials in the Atlanta government. There is also an ability to help companies in Atlanta better access their labor markets, as outlined by leadership at ATLDOT. Depending on POI types, the analysis may literally lead to saving lives [1].

In future city and transit planning operations by the city, extra care needs to be taken to ensure that these currently underserved populations are receiving the vital transportation connections needed for their livelihood. Despite relatively higher overall accessibility scores seen in the previous section, Atlanta is actually one of the worst in this regard as compared to other cities. We hope these eye-opening stats will drive change in the way Atlanta develops.

The main limitation of our tool relates to insufficient data, especially as it pertains to the completeness of POI locations. Another limitation of our work is that only consider public transit and not other modes of transport, which may prevent our work from being complete for certain types of decision-making situations. However, at a minimum, everyone has the ability to use public transit, so it provides a good check for whether or not it is even possible for someone to be served by a POI.

Overall, Atlanta's transit system and city planning vision has a lot of potential to grow in the coming years. We can hope that the city can seize the opportunity to better serve racial minority and lower-income populations and increase their access to vital POIs.

6 WORK BREAKDOWN

We did not incur costs beyond server setup to host our application (for which we leveraged free Azure credits for students). Effort-wise, the team split the workload roughly equally, with an average workload of ≈ 5-10 hours per week per person.

All team members have contributed a similar amount of effort. As far as specific work items go, all members contributed to literature review. The reports were primarily written by Tejas, Przemek, Aurimas, Junaid. Data collection and cleaning was carried out by everyone. The frontend was developed primarily by Mengyang, Tejas, Aurimas, and Przemek. The backend was developed by Aurimas, Junaid, Alex, and Przemek - this included storing the data in the PostgreSQL DB, computing catchment areas, as well as writing SQL functions to easily expose that data to the API. Insights from data, as well as Machine Learning models were created by Alex and Junaid. The poster and final report were created by Tejas and Przemek. Finally, each member individually recorded their own poster presentation.

REFERENCES

[1] 2022. Vaccine equity. https://www.who.int/campaigns/vaccine-equity

[2] Enrico Bentivegna, Silvia Di Meo, Anita Carriero, Nadia Capriotti, Alberto Barbieri, and Paolo Martelletti. 2022. Access to COVID-19 Vaccination during the Pandemic in the Informal Settlements of Rome. *International Journal of Environmental Research and Public Health* 19, 2 (2022), 719.

[3] Zheng Chang, Jiayu Chen, Weifeng Li, and Xin Li. 2019. Public transportation and the spatial inequality of urban park accessibility: New evidence from Hong Kong. *Transportation Research Part D: Transport and Environment* 76 (2019), 111–122. https://doi.org/10.1016/j.trd.2019.09.012

[4] Katherine L Chen, Madeline Brozen, Jeffrey E Rollman, Tayler Ward, Keith Norris, Kimberly D Gregory, and Frederick J Zimmerman. 2020. Transportation Access to Health Care During the COVID-19 Pandemic: Trends and Implications for Significant Patient Populations and Health Care Needs. (2020).

[5] Dajun Dai and Fahui Wang. 2011. Geographic Disparities in Accessibility to Food Stores in Southwest Mississippi. *Environment and Planning B: Planning and Design* 38, 4 (2011), 659–677. https://doi.org/10.1068/b36149 arXiv:https://doi.org/10.1068/b36149

[6] Catherine Duffy, Andy Newing, and Joanna Górska. 2022. Evaluating the Geographical Accessibility and Equity of COVID-19 Vaccination Sites in England. *Vaccines* 10, 1 (2022), 50.

[7] Mahyar Ghorbanzadeh, Kyusik Kim, Eren Erman Ozguven, and Mark W Horner. 2021. Spatial accessibility assessment of COVID-19 patients to healthcare facilities: A case study of Florida. *Travel Behaviour and Society* 24 (2021), 95–101.

[8] Mark W Horner and Jessica N Mefford. 2005. Examining the spatial and social variation in employment accessibility: A case study of bus transit in Austin, Texas. In *Access to destinations*. Emerald Group Publishing Limited.

[9] Junfeng Jiao and Amin Azimian. 2021. Measuring accessibility to grocery stores using radiation model and survival analysis. *Journal of Transport Geography* 94 (2021), 103107.

[10] KFF. 2022. Latest data on covid-19 vaccinations by race/ethnicity. https://www.kff.org/coronavirus-covid-19/issue-brief/latest-data-on-covid-19-vaccinations-by-race-ethnicity/

[11] David Martin, Hannah Jordan, and Paul Roderick. 2008. Taking the bus: incorporating public transport timetable data into health care accessibility modelling. *Environment and Planning A* 40, 10 (2008), 2510–2525.

[12] Alireza Mohammadi, Abolfazl Mollalo, Robert Bergquist, and Behzad Kiani. 2021. Measuring COVID-19 vaccination coverage: an enhanced age-adjusted two-step floating catchment area model. *Infectious diseases of poverty* 10, 1 (2021), 1–13.

[13] Bhavini Patel Murthy, Natalie Sterrett, Daniel Weller, Elizabeth Zell, Laura Reynolds, Robin L Toblin, Neil Murthy, Jennifer Kriss, Charles Rose, Betsy Cadwell, et al. 2021. Disparities in COVID-19 vaccination coverage between urban and rural counties—United States, December 14, 2020–April 10, 2021. *Morbidity and Mortality Weekly Report* 70, 20 (2021), 759.

[14] Stefan Neumeier. 2022. Accessibility of COVID-19 Vaccination Centers in Germany via Different Means of Transport. *KN-Journal of Cartography and Geographic Information* (2022), 1–18.

[15] Rafael HM Pereira, Carlos Kauê Vieira Braga, Luciana Mendes Servo, Bernardo Serra, Pedro Amaral, Nelson Gouveia, and Antonio Paez. 2021. Geographic access to COVID-19 healthcare in Brazil using a balanced float catchment area approach. *Social Science & Medicine* 273 (2021), 113773.

[16] Benjamin Rader, Christina M Astley, Kara Sewalk, Paul L Delamater, Kathryn Cordiano, Laura Wronski, Jessica Malaty Rivera, Kai Hallberg, Megan F Pera, Jonathan Cantor, et al. 2021. Spatial Accessibility Modeling of Vaccine Deserts as Barriers to Controlling SARS-CoV-2. *medRxiv* (2021).

[17] Marissa B Reitsma, Jeremy D Goldhaber-Fiebert, and Joshua A Salomon. 2021. Quantifying and benchmarking disparities in COVID-19 vaccination rates by race and ethnicity. *JAMA network open* 4, 10 (2021), e2130343–e2130343.

[18] R Shanmathi Rekha, Shayesta Wajid, Nisha Radhakrishnan, and Samson Mathew. 2017. Accessibility analysis of health care facility using geospatial techniques. *Transportation Research Procedia* 27 (2017), 1163–1170.

[19] A Thoumi, H Tewarson, and K Johnson. 2021. Prioritizing Equity for COVID-19 Vaccinations: Promising Practices from States to Reduce Racial and Ethnic Disparities. *Washington, DC: Duke-Margolis Center for Health Policy and National Governors Association Center for Best Practices* (2021).

[20] Au Vo, Rahul Bhaskar, Tai-Yin Chi, and Guillaume Faddoul. 2020. Identifying socially vulnerable regions with persistent low accessibility to emergency care through a spatial decision framework. *Journal of Decision Systems* 29, 4 (2020), 201–222.

[21] Neng Wan, Bin Zou, and Troy Sternberg. 2012. A three-step floating catchment area method for analyzing spatial access to health services. *International Journal of Geographical Information Science* 26, 6 (2012), 1073–1089.

[22] Pavan Teja Yenisetty and Pankaj Bahadure. 2020. Spatial accessibility measures to educational facilities from public transit: a case of Indian cities. *Smart and Sustainable Built Environment* (2020).