Georgia Institute of Technology

Exploration of Commuting Patterns in San Francisco Bay Area

MS-GIST Capstone Project Report

Liu, Lier
7-1-2020

Abstract

Transportation has been an important aspect in life, both offering convenience and efficiency and potentially bringing inconvenience and delay to people. San Francisco Bay Area has always been famous for its comprehensive transportation networks and various transportation modes; meanwhile the issue of transportation congestion has caused severe consequences in people's lives such as wasted time and gasoline, polluted air, and traffic accidents. This project aims to explore the commuting patterns in the San Francisco Bay Area and analyze people's commuting behaviors and some demographic datasets. Correlations between demographic features and people's commuting choices and behaviors are explored, along with three representative new commuting methods, Bay Area Rapid Transit, Lyft bike sharing Bay Wheel, and Uber. Such thorough studies can potentially offer policy implications for transportation planners and policy makers, for the sake of establishing faster transportation networks and avoiding traffic delay to a larger extent.

Chapter 1: Introduction

Transportation has always played a vital role in human life as it has been rapidly developing in the past decades. People go out for multiple reasons, and their needs of exploring further places have been expanding, as the society has been growing. People from all professions have certain needs to travel, either to workplaces, or for long-distance travels and their preferred means could range from airplanes, trains, riding on personal vehicles or shared vehicles, or simply walking or biking. Apart from their professional needs, individuals transport for the sake of other purposes, including entertainment and recreation, travelling, grocery shopping, or family and friends visiting. Different means of transportation assist people in moving them a lot faster and saving a great deal of time.

Transportation has become evitable in human life and exist in every single aspect in life. Given the rapid development and expansion of various transportation means, a large number of choices can be made by people based on their needs and preferences. With efficient ways of transportation, people are allowed to arrive in time and achieve their tasks without wasting extra long span of time on the way. In the United States specifically, the transportation system is facilitated by air, rail, waterways and roads, while a large portion of people commute by railroads or planes for long-distance destinations, and by personal vehicles for short-distance destinations. Since 1957, the transit ridership on a national basis has developed by over 20% in the past decade, which has reached the highest level of the transit ridership since the middle of last century (United States Department of Transportation 2020).

The study area of this project is chosen to be San Francisco Bay Area, due to its fruitful available data for research purposes and its wide variety of available transportation means. San Francisco Bay Area has always been famous for its complex multimodal transportation infrastructure, including but not limited to highways, roads, bridges, rail, air, ferries, pedestrian and bike lanes. The maintenance, functioning, and development of highways, subway networks, rails, bus and trolley services, and airports are collectively managed by multiple agencies, California Department of Transportation, Metropolitan Transportation commission, and San Francisco Municipal Transportation Agency (Amin and Barz 2015). The public transportation in San Francisco Bay Area covers comprehensively and offers affordable commuting methods for residents and travelers. Along with the thorough transportation system in San Francisco Bay

Area, the issue of transportation congestion has been increasingly severe and negatively affecting people's quality of living. According to the Urban Mobility Report, commuters have wasted more than 107 hours on traffic congestion and delay, wasted 45 gallons of fuel on the roads, while in traffic delay in 2017 in San Francisco and Oakland Bay Area (Lasley 2019). The drivers in San Jose, south Bay Area, have wasted approximately 81 hours across the year, counted as the fifth worst traffic scenario over the nation.

For this reason, the commuting patterns and people's commuting behaviors are studied in this project, along with potential demographic features of the San Francisco Bay Area, potentially correlated with commuting patterns. First, the relationship between some selected demographic characteristics such as population and race are explored with commuting datasets. Second, commuting datasets are studied based on the geographic unit of census tracts and visualized for demonstration purposes. Third, three new popular transportation means in San Francisco are further explored to better comprehend people's choices of commuting. The three means of transportation are Bay Area Rapid Transit (BART), a heavy rail elevated and subway system, Lyft electric bike sharing services, Bay Wheel, and Uber, nation-wide ridesharing services.

This report is organized into five chapters. Chapter 1 summarizes the research background and need; Chapter 2 presents a comprehensive literature review. Chapter 3 introduces data acquisition and research methods and approaches. Chapter 4 offers demonstration of map outputs, screenshots from webpages, and results interpretation. Chapter 5 discusses potential limitations and summarizes the project.

Chapter 2: Litearature Review

Transportation weighs heavily in daily lives of human beings. People take various means of transportation for different reasons, such as school, work, grocery shopping, or traveling. Over the decades, transportation has been evolving rapidly, and it has been faster and more convenient than ever before. In San Francisco Bay Area specifically, congestion has always been a common significant social issue, and it has reached the highest level in the past few years. Roads and highways have particularly become more congested due to its rapid population growth, its fast-growing economy, and high demand for traveling. Congestion has negatively influenced life quality of local residents and experiences of travelers. The exploration of commuting patterns and popular means of transportation in San Francisco Bay Area can potentially inform policy makers, transportation departments, as well as local residents of the current situation and possible solutions and future plans.

San Francisco and San Francisco Bay Area have been considered a region with high traffic demand, and its accessibility has been featured as a key factor that influences the quality of living (Blanchard and Waddell 2017). Comprehensive analysis and understanding of accessibility of regional transit system can bring improvements on metropolitan planning organization (MPOs) for the sake of either equitable transportation resources allocation or better transportation service performance assessment. Grengs et al. (2010) mentioned the fundamental aspect of transportation systems evaluation is supposed to be the accessibility instead of mobility measured by delay per capita, time and money wasted while in a congested traffic, of LOS (level of service. Recent studies have found improvements can be made to promote the transit system services and accessibility in region job centers where people most frequently travel to (Blanchard and Waddell 2017), and possible tactics include establishing alternative rail systems, which both decreases greenhouse gas emissions and leads to positive impacts on the transit system overall. As a region with a series of places of interest, traffic demand from travelers could potentially affect the accessibility of many busy areas. Nevertheless, Lockwood et al. (2005)'s study showed that the total vehicle miles travelled (VMT) on a weekend day counts as much as 80% of the total VMT of a weekday in San Francisco Bay Area. This potentially suggests the traffic demand from local residents is more influential in determining traffic condition and accessibility in San Francisco Bay Area.

There are a set of factors that potentially shift the accessibility and commuting patterns, including but not limited to population growth and decentralization, job accessibility, public transit availability, housing density, geographic conditions, and urban expansion. In Ma et al. (2017)'s study, the patterns of urban transit commuters in Beijing, China were revealed, in which most people leave to work in the duration of morning peak hours, and head back in evening peak hours. Among all research samples, a substantial portion travel 21 days in a month, which can be interpreted as the working days in a month. Their average commuting distance is 10.99 km and public transit is not preferred when the distance between places of working and living exceeds 5 km. The imbalance between job accessibility and housing availability largely shapes such commuting pattern, in which the imbalance potentially wastes excessive time and results in traffic congestion. Another studied (Zhao et al. 2010) conducted in Beijing concentrated on the impact originated from urban expansion on urban commuting patterns. Zhao et al. presented similar suggestion that compact urban development can possibly save extra time wasted in commuting and address the job-house imbalance. This study also pointed out the impact from income, in which middle-income households are more likely to travel within suburbs, and high-income households have higher tendency to travel longer comparatively.

Demographic factors can also play a vital role in commuting behaviors and commuting patterns. A study conducted in Rochester Minnesota (Sang et al. 2011) has explored the correlation between gender, occupation and commuting patterns, particularly concentrated on journey-to-work. Sang et al. (2011) stated that women tend to spend less time on commuting before places of living and working. Such a gender rendered commuting patterns could be originated the social responsibility and professional preference. A researched based on datasets in Chicago (Wang 2000) also offered insights on people's socioeconomic factors. The author (2000) came with conclusions that, residents with lower income are expected to travel longer to their workplaces.

Apart from the public transportation and driving, some shared riding systems in not only San Francisco have been evolving rapidly, such as Uber and Lyft services and electric scooter and bike. The so-called ride-souring is characterized as the connection between drivers along with their vehicles with random passengers (Jin et al. 2019). The authors concluded that in New York City, Uber services both serve as compliment to public transit during midnight and in region where public transportation cannot fulfill people’s needs, and competition against public transportation means in a majority of daytime and places. Another study (Alley 2016) also mentioned the rapid growth in Uber services in the New York City, which offer affordable transportation to people in economically disadvantaged places and people who owned no personal vehicles. Schaller (2018) also commented on the impact from Lyft and Uber, also named Transportation Network Companies, and young, educated and relatively high-income people tend to ride on Lyft or Uber. Such expansion and development could bring benefits to areas where public transit is not fully covered or not as available and frequent as in centralized urban regions.

The rapid development of electric bikes and scooters in urbanized area has also contributed to more convenient, available and accessible transit (Hollingsworth 2019, McKenzie 2020, Puzio 2020). Such short-time and short-distance electric rental services have been accepted by an increase amount of people. The shared electric scooters as well as shared bikes provide optimal and affordable solutions when public transportation fail to directly connect to destinations, and extra walking is needed (Hollingsworth 2019). Other than the convenience brought by such shared services, traffic congestion can possibly be accelerated when shared scooters or bikes occupy road networks of other means of transportation (McKenzie 2020).

Chapter 3: Data and Methodology

This study analyzes four commuting related datasets acquired from U.S. Census Bureau, BART Ridership, Lyft System data, and Uber Movement respectively. To avoid potential bias in the study, datasets from the year of 2018 are obtained, and for datasets with certain time ranges, the study time is specified as August 2018. All primary datasets are further refined, cleaned and filtered for visualization and analysis purposes.

First, to acquire a more comprehensive view of the commuting patterns in San Francisco Bay Area, the demographic and geographic datasets are collected from U.S. Census Bureau and studied. The county boundary data of the study area is downloaded from U.S. Census Bureau in shapefile format, which can be used for clipping datasets in larger geographic scale and used as boundary file for visualization purpose. The geographic boundary file contains not only boundary information in polygon shapes, but also the county name, code and acres of land. The demographic datasets are obtained from U.S. Census Bureau as TIGER/Line files containing geographic features, census statistical boundaries, population information, race and income data covering the entire country.

Second, the commuting related datasets are downloaded from U.S. Census Bureau as TIGER/Line shapefiles, and the datasets are collectively called 2018 California 5-Year Tract. The tracts information is stored in tables in a geodatabase based on statistical areas. The downloaded geodatabase from Census Bureau gathers geography and demographics from the 2018 TIGER/Line shapefiles and data collected from the 2014 to 2018 American Community Survey, 5-year estimates. This set of data offers in-depth data and statistics associated with transportation and commuting in the study area of this research. Meaningful columns include the number of people choosing different means of transportation, such as carpool, bus, subway or driving, the amount of people living outside of the county they work in, and the amount of time people spend on their way to work. A series of useful commuting related data can be projected on maps and reflect people's transportation behaviors and commuting patterns San Francisco Bay Area. Nevertheless, this dataset, stored as tables in a geodatabase, contains no geographic coordinates information, which means it needs to be coupled with corresponding datasets containing latitudes and longitudes for the sake of visualization. The State-County-Tract FIPS Code CSV reference table is downloaded from The United States Department of Agriculture, Economic Research Service, in which the tracts codes, county names and geographic coordinates are used for map creation.

Third, three sets of data representing three different means of transportation frequently used by San Francisco Bay Area residents, being Bay Area Rapid Transit (BART), Lyft bike sharing, Uber Movement. The BART datasets include monthly ridership reports in 2018, from which daily average entries and exits in both weekdays and weekends can be extracted for visualization and analysis. The daily average entries are used because the number of entries show The Lyft bike sharing, the so-called Bay Wheel biking sharing system, datasets include CSV data tables containing origin and destination coordinates along with travel time. This set of data can potentially be visualized to exhibit traffic volume of Bay Wheel by census tracts. The Uber Movement dataset collects origin and destination coordinates of each Uber trips occurring in San Francisco Bay Area, together with its travel time in August 2018. This dataset will be visualized into an interactive online map showing trips' travel time from each origin place.

Due to software licensing issues, only ArcMap 10.7.1 is used for the first set of visualization tasks, while ArcGIS Pro and ArcGIS Enterprise were also considered as main software for visualization. ArcMap turns out to be as smooth and efficient, yet map publishing happens to be an issue and outputs are made into JPEG format for final results instead of story maps. The original datasets are tracts area codes, census tract shapefile, and commuting shapefile, which need further cleaning for maps creation. First, the tract area codes are studied, and area codes of nine counties are recorded for later use, being Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano, and Sonoma counties. The census tract dataset contains county names, tract codes and geographic coordinates of each piece of information, while the commuting dataset saved as tables in a file geodatabase has no geographic information and references. The second task is cleaning data, primarily the tract codes based on which the datasets can be joined. The GEOID of the demographic dataset contains irrelevant numbers and strings which need to be removed. For example, the code 14000US06001400100 contains irrelevant numbers and strings "14000US" and the right 11 numbers are extracted using field calculator to match the records in the commuting related table. With the common tract ID numbers, each row of data in the commuting related table is assigned geographic coordinates and county names, and the shapefile is ready to be visualized on base maps. The third step is to classify and calculate to avoid extra-long rows and categories in the process of visualization. For example, the leaving time to work parameter has been classified into 6:00 am to 6:29 am, 6:30 am to 6:59 am and so on, from which patterns can hardly be determined, because the intervals of those time periods are overly narrow. In this way, the leaving time to work parameter is reclassified into four groups being 6:00 am to 8:00 am, 8:00 am to 10:00 am, 10:00 am to 12:00 pm, and 12:00 pm to 4:00 pm, and each will be visualized on base maps with a unit of census tracts. The travel time to work is recategorized similarly into less than 30 minutes, 30 to 60 minutes, 60 to 90 minutes and more than 90 minutes, and the amount of people from original classes that belong to each new group will be added up, and then maps of four new categories can be made accordingly.

To better compare demographic variables and commuting behaviors and patterns of people in San Francisco Bay Area, some parameters are combined to show their relationships on the maps. For example, the variables people who choose to drive car, truck or van, and Asian/White population are collectively visualized on maps in two different symbology. In this way, the correlation between people's preference on means of transportation and their races can be more clearly visualized and analyzed. Moreover, five popular means of transportation are collectively visualized on maps, being "car, truck or van", "carpooled", "bus or trolley bus", showing the portion that each ways of commuting chosen by residents in San Francisco Bay Area. "Car, truck or van" and "carpooled" are subsequently removed from the maps once at a time, to show the percentages of the remaining means of commuting that people prefer.

The three sets of commuting data are visualized in three different ways, being Tableau Desktop and Tableau Public, Google Map API, and Python plus JavaScript. First, the table containing BART stations entry records in August 2018 is joined by station name abbreviations with the table that includes BART station longitudes and latitudes. Route identifier, Route Order and Route Location are subsequently created to define start BART stations and stop BART stations for each single trip. In the end, the datasets become one aggregated worksheet with entry counts, origin-destination records and station longitude and latitude for each BART trip. A customized base map is created and published through MapBox online, in which transportation elements are highlighted and the background is painted dark so that other characteristics could stand out. The customized map is inserted using its url and API access token. The purpose of using Tableau Desktop is to visualize distribution of stations' entry counts and learn about the busiest BART stations in San Francisco Bay Area. Also, an objective is to create an interactive map, in which trips from a BART station will be visualize based on the origin selection made by any user.

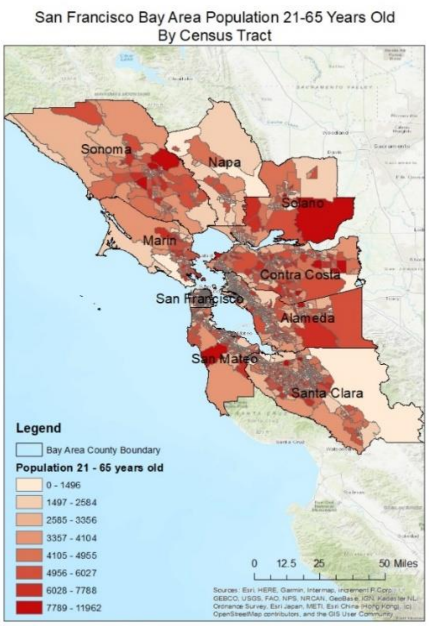
The second set of commuting data downloaded from Lyft bike sharing which contains the Bay Wheel ridership records. The dataset is obtained in a xlsx format and is therefore converted to json format for later use; the extra and unrelated data columns are eliminated from the dataset. With the json ridership data, a heat map can be created using Google Map API, and a common Google base map layer is added for better visualization. The heat map shows places where people frequently utilize Lyft Bay Wheel and places where it is less used, in which red suggests frequent uses and green reflects few uses. Some parameters are adjusted to better present the results. The radius and weights of the colors is made less than usual since there are fewer places with frequent uses of Bay Wheel compared to the situation as expected. By modifying the radius, the effect from each data point is changed in pixels. The data are mostly in the city of San Francisco and north Bay Area, while less or no data are available for Bay Wheel ridership in east and south Bay Area.

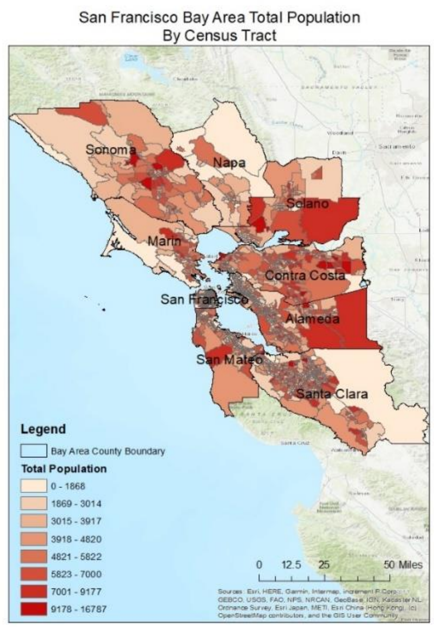
The third set of commuting data collected from Uber Movement in August 2018 and is proposed to generate an interactive web map presenting the travel time of each Uber trip. The dataset is downloaded in csv format and is cleaned because of its large volume and potential excessive computation time. Some of the columns in the csv data are therefore removed using Pandas in Python environment, with only relevant information remains. The next step is to assign 5 different classes for visualization purposes, in which the classes and color gradient are dependent of the travel time from each single origin place. Five classes are less than 8 minutes, 8 to 15 minutes, 16 to 25 minutes, 26 to 35 minutes and more than 35 minutes. Longer travel time trips are designed as blue, while shorter travel time are in light yellow. A json file containing census tracts ID and longitudes and latitudes of each origin and destination spots is subsequently visualized using JavaScript. As designed, each click on a random tract spot, a gradient map will show up, exhibiting the travel time of Uber trips from such tract spot to surrounding places. By clicking on another tract spot twice, a new gradient map will be rendered in a second showing similar travel time information of Uber. At this moment, the map is developed by Node.js and

Express at the backend, and visualized at the frontend using d3.js. In order to allow any user to access the map, the map application is deployed online via the open source server herokuapp.

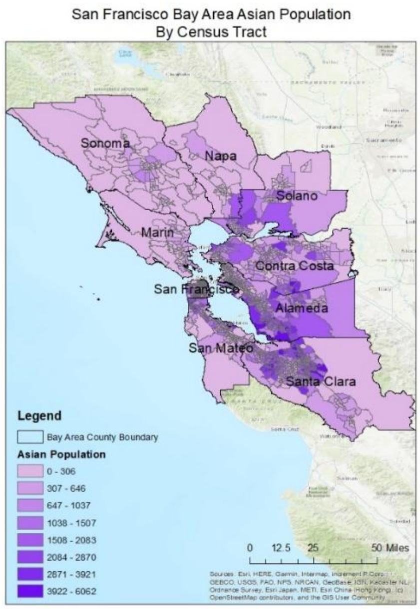
Chapter 4: Outputs and Results

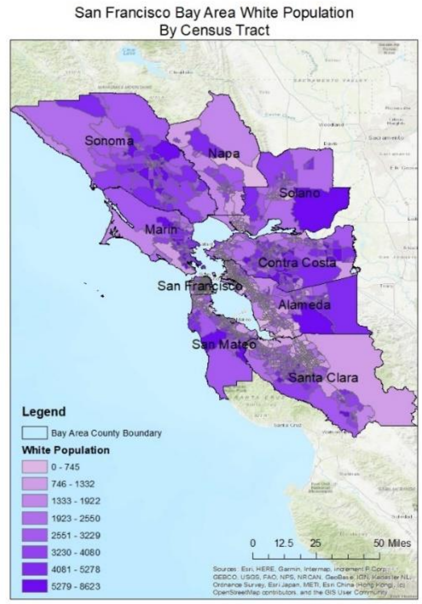
In The first three set of maps are created in ArcMap, demonstrating demographic information in San Francisco Bay Area associated with commuting behaviors and patterns. The first set of maps show population related information, including age, race, housing density, and income. The first two maps, San Francisco Bay Area Total Population by Census Tract in 2018 and 21 to 65 Years Old Population by Census Tract present similar spatial patterns, in which east bay tends to have high population and population between 21 to 65 density especially Alameda, Contra Costa and Solano Counties. The population distribution between 21 to 65 years old suggests people in desired ages of working, so that these counties can be assumed to have higher commuting demand than the rest of the counties.



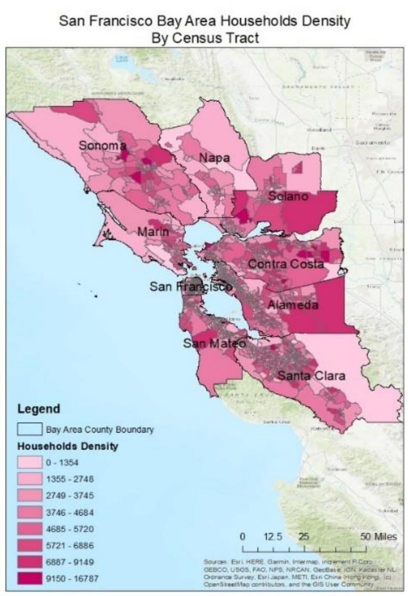


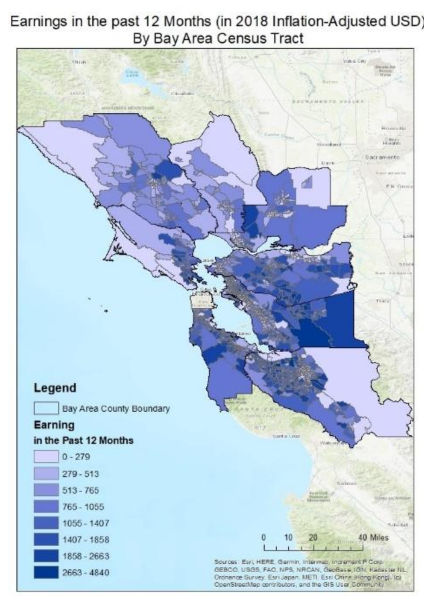
The second set of demographic maps contain race information, in which only Asian and White Population are used considering the facts that Asian population consists of a large portion of San Francisco Bay Area residents. The two maps show the Asian and White race distribution among all San Francisco Bay Area Counties. Based on the maps, more Asian people live in western Alameda, Contra Costa, and Solano Counties, along with north western Santa Clara County, while White population tend to appear in the rest of the areas such as eastern Sonoma, Contra Costa, and Alameda Counties, as well as other counties.



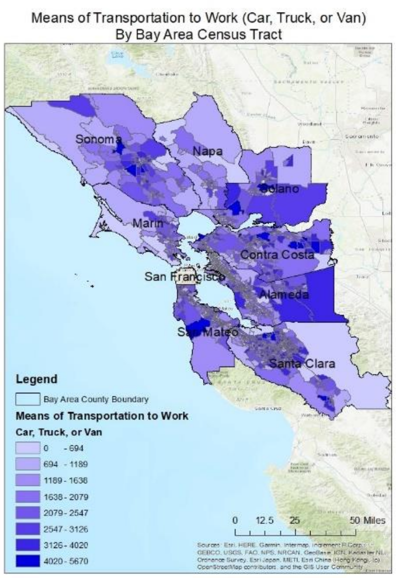


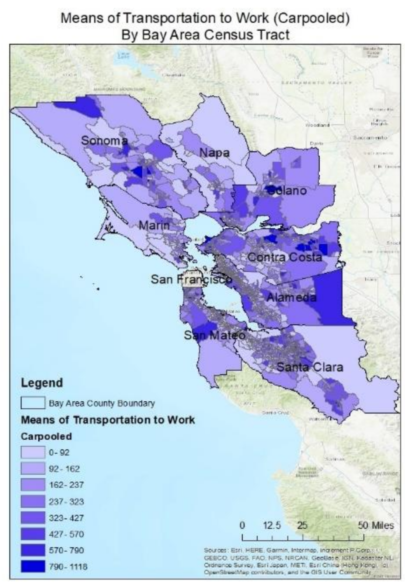
The third set of maps exhibit housing density and earnings in San Francisco Bay Area, in which darker color suggest high housing density or high earnings. Comparing the housing density and earning maps with race as well as population maps, both earnings and housing density maps have relatively closer relationship with White population compared to Asian population.

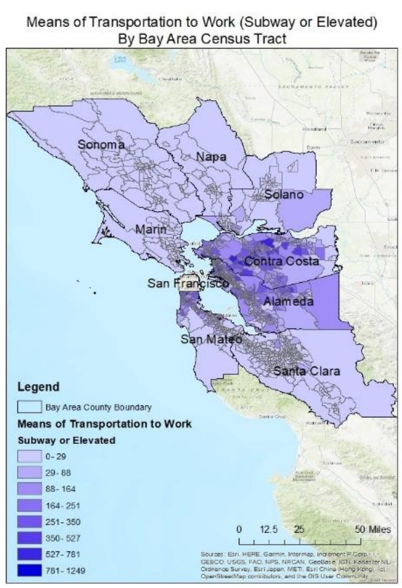


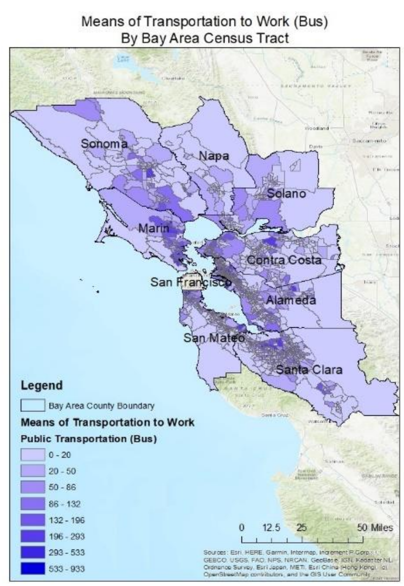


The next set of maps are also generated in ArcMap Desktop, which present the means of transportation preferred by people in San Francisco Bay Area. Major means of transportation include “Car, Truck, or Van” and “Carpooled”, and public transit means “Subway or Elevated”, and “Bus”. According to the set of four maps, people in Alameda, San Mateo, Solano, and Contra Costa tend to choose private vehicles or carpooled, while people living near San Francisco County tend to take public transportation such as subway or bus. In general, less people prefer public transit means comparatively.

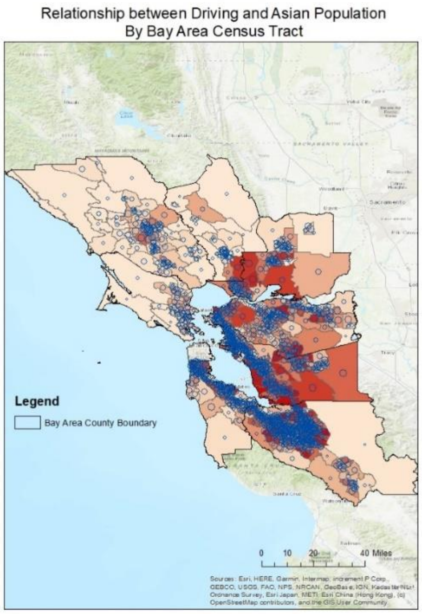


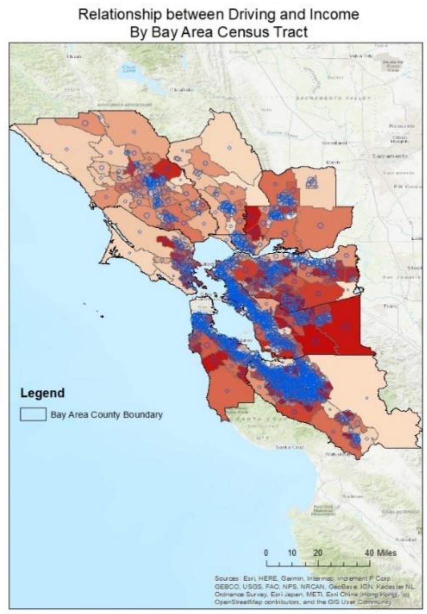




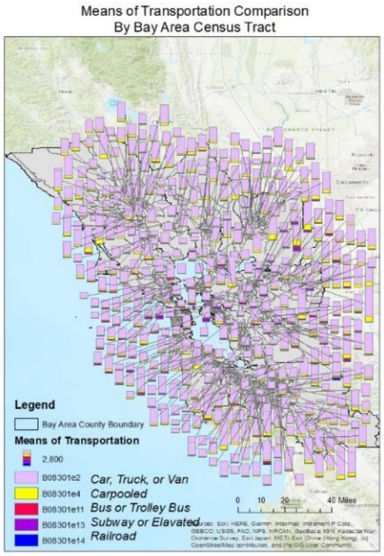


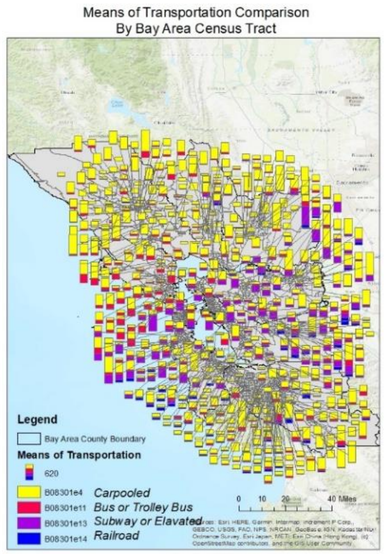
For comparison purposes, two sets of variables are combined for analysis. The legend of the map below is partially presented due to overly their large numbers of rows; the size of the circles shows the amount of people choosing to drive in car, truck or van in each census tract. From these two aggregated maps, Asian population have higher tendency to drive, and people with higher income are more likely to drive, yet the relationship between driving and income tends to be less solid than the correlation between Asian population and driving.

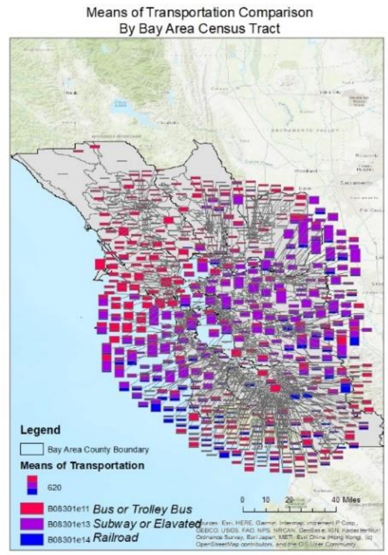




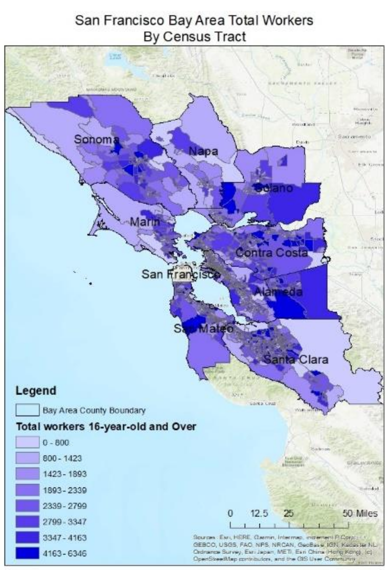
The following three maps collectively show the preference of transportation means of people in San Francisco Bay Area. In the first map, the relatively percentages of all popular means of transportation are presented, and driving personal vehicles takes up the largest portion. In the second map, with no “Car, Truck or Van” in the map, “carpooled” becomes the most popular among all remaining means of transit. The third map basically include three popular public transportation, “Bus or Trolley Bus”, “Subway or Elevated”, and “Railroad”. In east Bay Area and San Francisco County, a large portion of people choose to take subway, while in the north Bay Area, more people take buses, and in the south bay, people have nearly equal preference between railroad and bus.

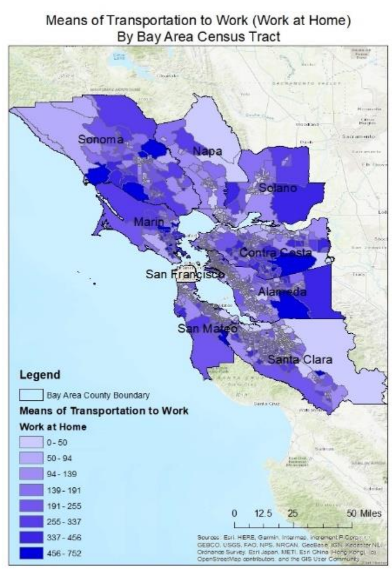




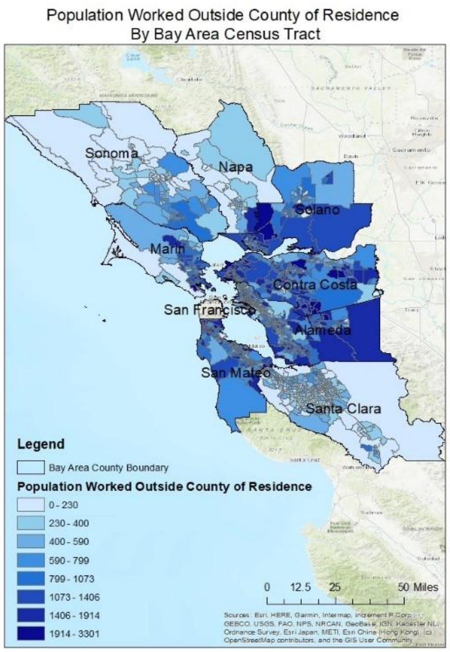


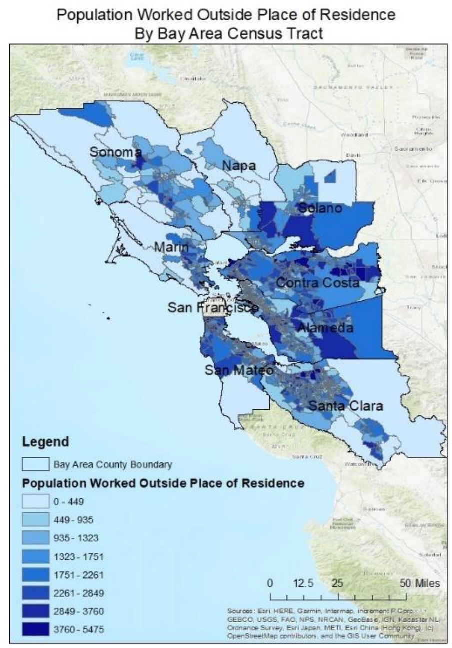
The following set of maps created in ArcMap Desktop further explore the working patterns of residents in San Francisco, in which the total number of workers and the amount of people who working at home. These two parameters are both related to traffic condition and commuting patterns, since places where people tend to work at home may have less traffic congestion. In the two maps below, people in Napa and Santa Clara Counties have apparent low tendency to work at home, while people in certain tracts in Sonoma, Contra Costa and Alameda County tend to work from home. In the total workers map, there are more workers in the east Bay Area.





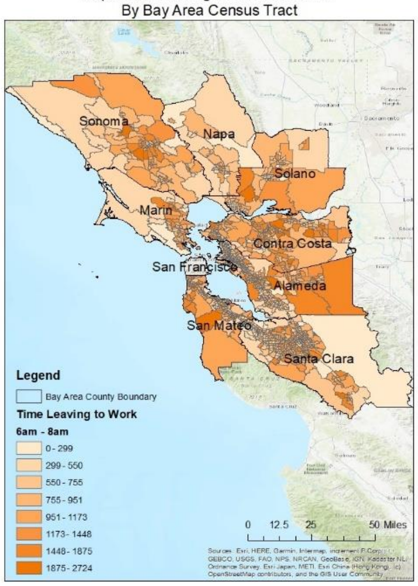
The two maps below show the people living in San Francisco Bay Area and work outsite of their living county or outside of the tract. These two maps may explain the large numbers of people choose to drive in personal vehicles due to the longer distance from home to work in Alameda, Contra Costa, San Mateo and Solano Counties.



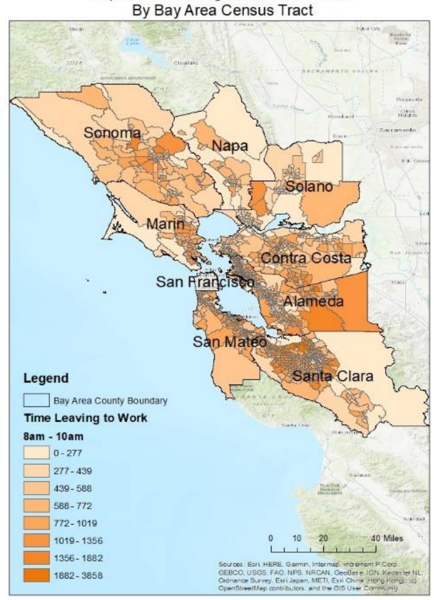


The next series of maps show the amount of people categorized by their time leaving to work. The a large portion of people in Alameda and San Mateo Counties, and a fair amount of people in Solano, Contra Costa and Sonoma Counties choose to or have to leave home during early morning to work.

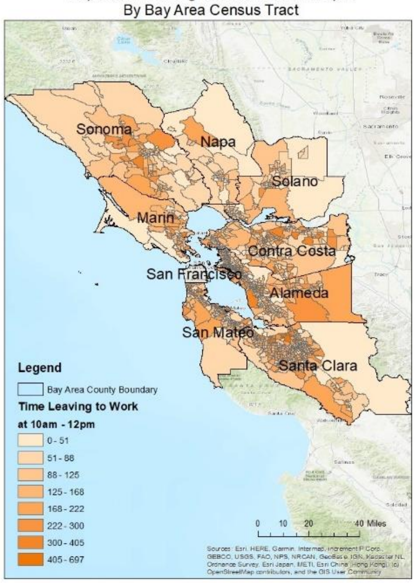
Population Leaving to Work at 6-8 am
By Bay Area Census Tract



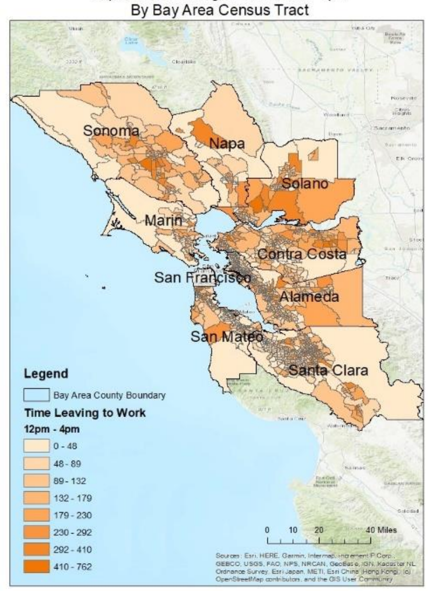
Population Leaving to Work at 8-10am
By Bay Area Census Tract



Population Leaving to Work at 10am-12pm
By Bay Area Census Tract

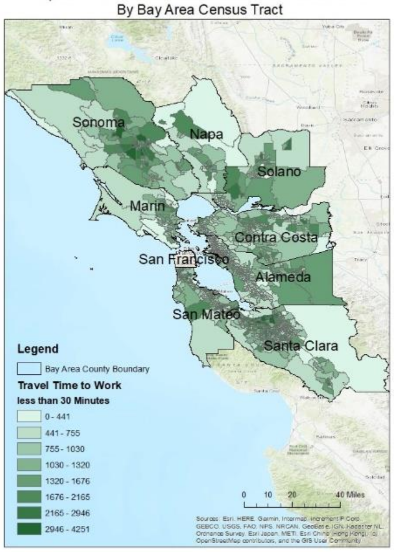


Population Leaving to Work at 12-4pm
By Bay Area Census Tract

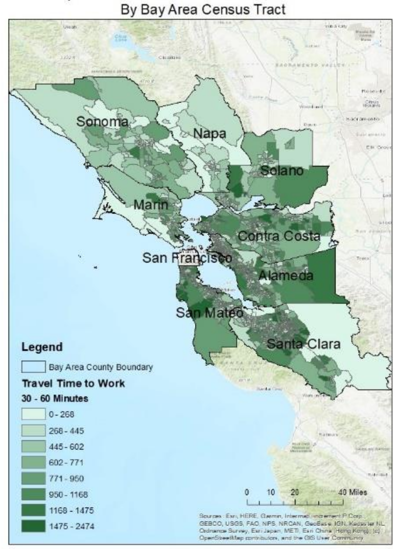


The maps below show the travel time to work of people in San Francisco Bay Area. A large protion of people in eastern Sonoma County and several tracts in Napa County are more likely to travel within 30 mintues to work while people in Alameda, San Mateo, and Contra Costa and eastern Santa Clara Counties are more likely to travel longer.

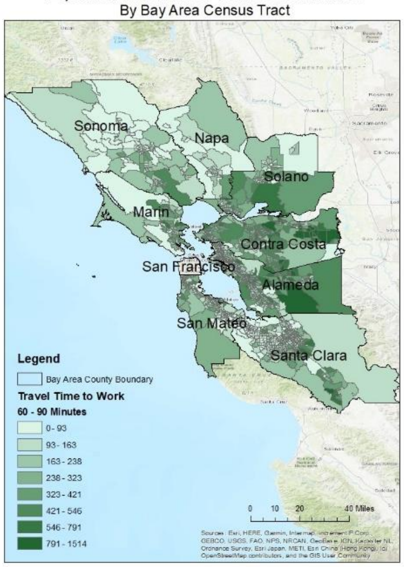
Population Travel Time to Work Less than 30 Minutes
By Bay Area Census Tract



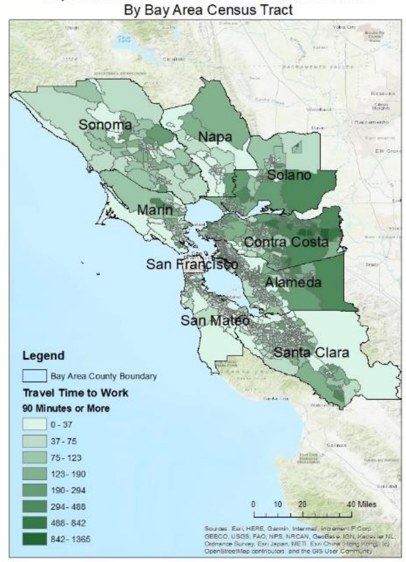
Population Travel Time to Work 30 to 60 Minutes
By Bay Area Census Tract



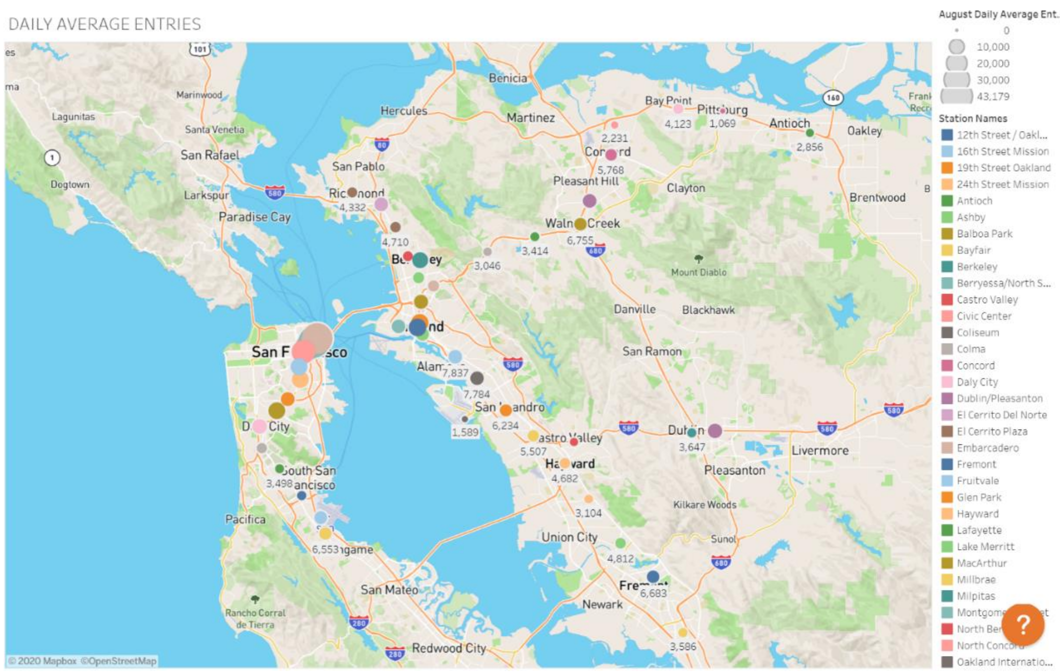
Population Travel Time to Work 60 to 90 Minutes
By Bay Area Census Tract

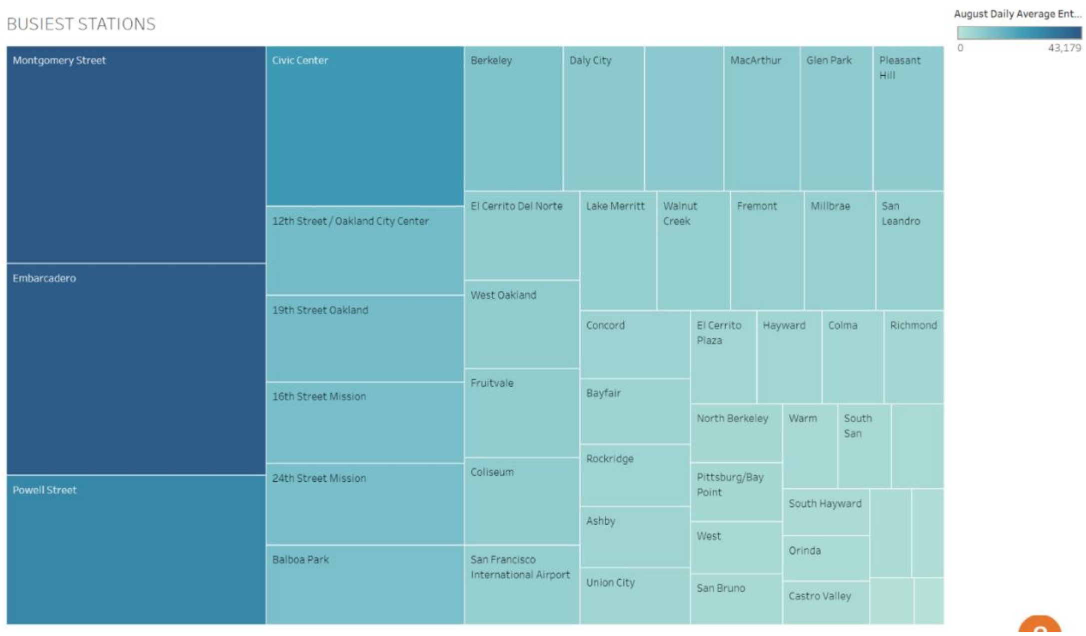


Population Travel Time to Work 90 Minutes or More
By Bay Area Census Tract



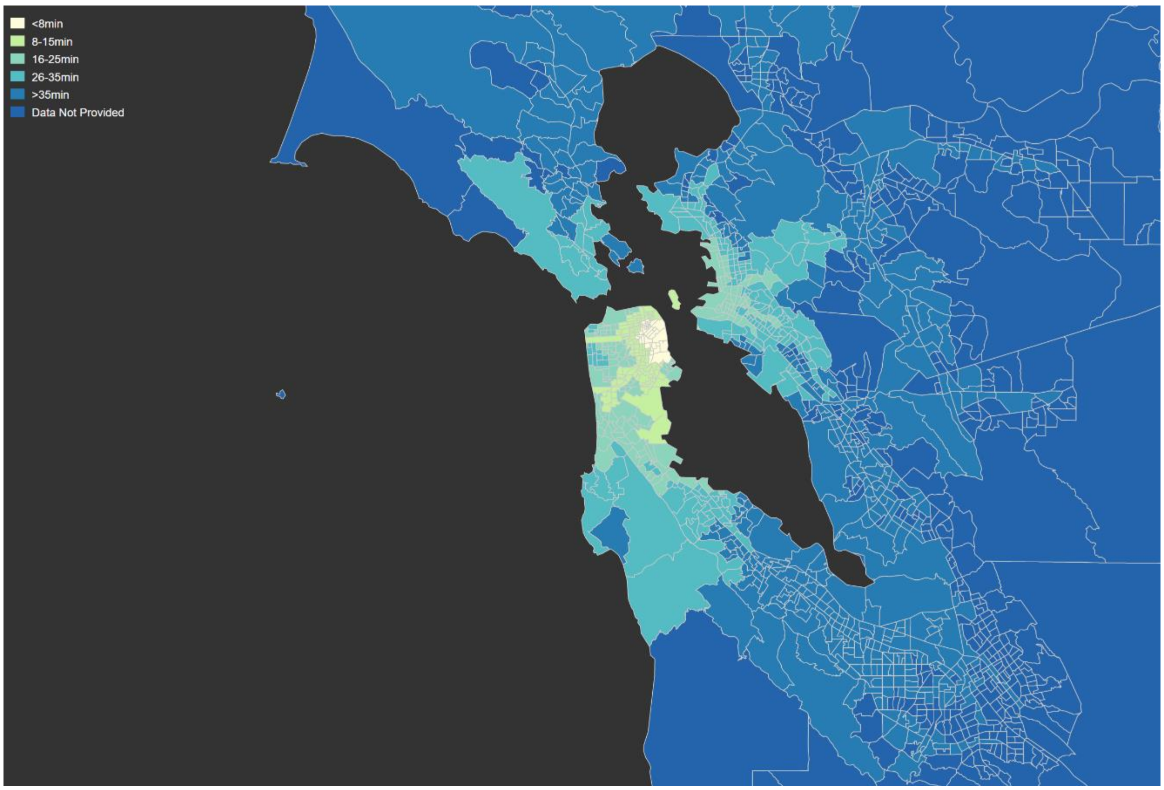
After exploring commuting data from U.S. Census Tracts, three new population transportaion means are studied separately, being Bay Area Rapid Transit (BART), Lyft electric bike, and Uber. The first set of graphs are created in Tableau Software and published through Tableau Public. The first two graphs are generated to obtain an overall view of the use of BART in San Francisco Bay Area, in which the first map gives daily average entries in August 2018 on a street map, and the second graph shows a sequnce of stations by their amount of entries. Among all BART stations, stations in or near the City of San Francisco and Concord have higher volume of passengers, and the busiest stations are Montgomery Streee, Embarcadero, Powell Street, and Civic Center.



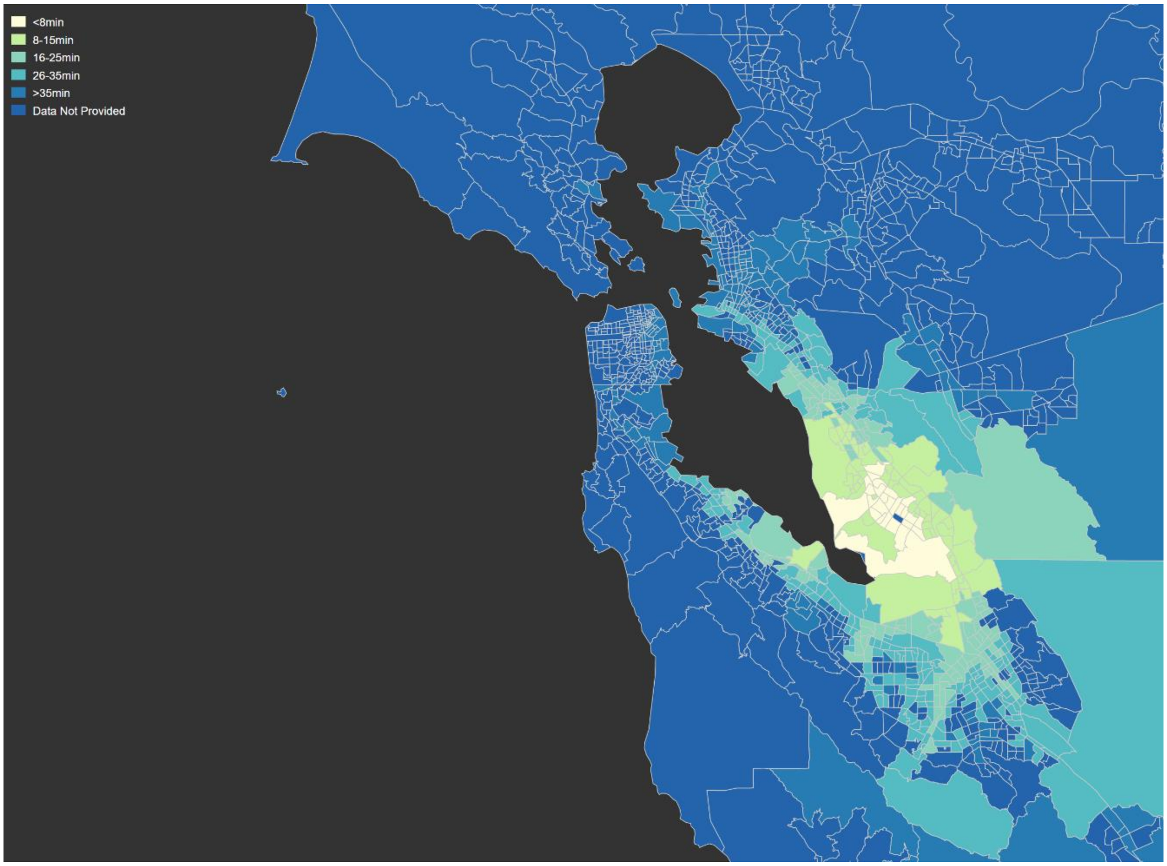

BUSIEST STATIONS
August Daily Average Ent...
0
43,179
Montgomery Street
Embarcadero
Powell Street
Civic Center
12th Street / Oakland City Center
19th Street Oakland
16th Street Mission
24th Street Mission
Balboa Park
Berkeley
Daly City
MacArthur
Glen Park
Pleasant Hill
El Cerrito Del Norte
Lake Merritt
Walnut Creek
Fremont
Millbrae
San Leandro
West Oakland
Concord
El Cerrito Plaza
Hayward
Colma
Richmond
Fruitvale
Bayfair
North Berkeley
Warm
South San
Coliseum
Rockridge
Pittsburg/Bay Point
South Hayward
Ashby
West
Orinda
San Francisco International Airport
Union City
San Bruno
Castro Valley

The next screenshot of the map is obtained from an interactive map. With a selection of Start ID on the top right, the map will show trip flows from the certain station to all destinations. The color of the circles shows the amount of trips, and circles of darker color, such as stations near the City of San Francisco suggest large numbers of BART trips in August 2018. The second map is set with a start station in Alameda county, and a large number of people take BART to travel to downtown San Francisco. Stations in San Francisco County are considered the busiest stations with both high volume of entries and exits.

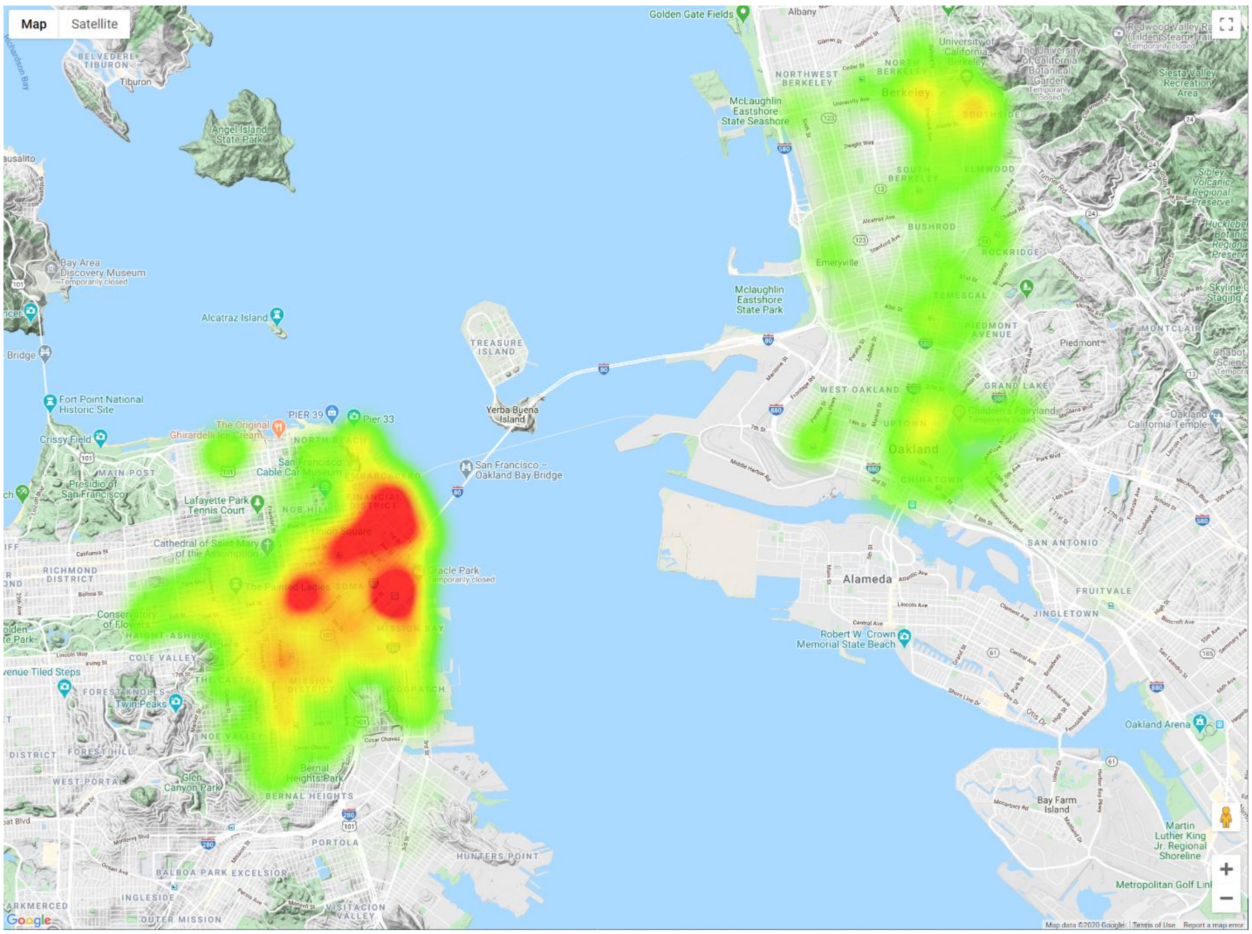
The two screenshots of Uber travel time are obtained from an interactive web, with datasets of Uber travel time of all Uber trips in the third quarter in 2018. With a single click on a random census tract, a gradient map will appear, in which lighter colors suggest shorter time spent on Uber trips, while darker colors mean longer time or no data available. Within the San Francisco County, when a start station is set in downtown San Francisco, some Uber trips take approximately equal time to travel to western San Francisco or to other counties in Bay Area. Such a pattern suggests higher probability of traffic congestion in San Francisco County.



When a start point is set in a tract in Alameda County, some trips within the county can probably take extra long time than other places inside the county, and less people choose to take Uber from Alameda County to San Francisco County compared to the means of transportation of BART, a relatively cheap and efficient alternative.



The screenshot of the map below is obtained from an interactive google map with a heat map layer presenting the usage of Lyft shared electric bikes trips, Bay Wheel. The map can be zoomed in and out depending on the users' objectives, and red represents high volumes and green represents relatively low usage of Lyft bikes. High usages of Bay Wheel are more to appear in downtown San Francisco, and comparatively Bay Wheel and BART are proved to be more popular than Uber in San Francisco where traffic congestion is a severe issue. Travelling in short distance with Bay Wheel and BART can be both efficient and cost-effective.



Chapter 5: Discussion and Conclusion

This project aims to explore the commuting patterns and potential demographic variables related to people's commuting behaviors. Through the study of demographic features datasets from U.S. Census Tracts in 2018, people's commuting behaviors related datasets in 2018, and three different popular means of transportation in San Francisco Bay Area, being Bay Area Rapid Transit (BART), a heavy rail elevated and subway system, Bay Wheel, Lyft electric bike sharing services and Uber, nation-wide ridesharing services in the month of August in 2018. Demographic datasets offer necessary socioeconomic background information, which could potentially explain the commuting patterns in study area. Through the exploration and visualization of the three commuting methods, in-depth findings of BART, Bay Wheel and Uber's usage can be learned.

First, among three selected demographic characteristics, population, race and income, white population, and population including total population and age 21 to 65 have moderate relationships with the commuting preference of driving car, truck or van. Asian population and income of households have relatively stronger correlation with the choice of driving personal vehicles. Second, through comparison between different modes of transportation, including driving, bus, carpooled, subway, people choose to drive personal vehicles take a large portion of the total population, and carpooled also accounts for a large portion. This variance in people's choices of commuting modes reflect their preferences as well as the more and less convenient modes overall. A comparison map between bus or trolley bus, subway or elevated, and railroad reflects an overall preference of riding on bus or trolley bus in north Bay Area, riding on subway or elevated in San Francisco County and Sonoma County, and no apparent preference other than driving and carpooled in east and south Bay Area. Third, large portions of people living in Alameda, Contra Costa, San Mateo and Solano Counties do not reside in the counties or places of their work places, which may explain the large portions of people choosing to drive in these counties for convenience and time-saving purposes. Fourth, relatively more people in Alameda and San Mateo Counties leave very early in the morning for work, and individuals in Alameda, San Mateo, and Contra Costa and eastern Santa Clara Counties tend to travel over an hour on their way to work. These demographic and commuting parameters collectively may explain the larger amount of people prefer driving, especially during morning and evening peaks in

Alameda, Contra Costa, San Mateo and Solano Counties overall. A limitation could be the less statistical evidence of such conclusions drawn from map visualization and interpretation, and if statistical analysis is combined to show the correlation, the conclusion will be more convincing.

The visualization of three popular commuting modes, BART, Bay Wheel and Uber are displayed in different formats. First, BART monthly ridership datasets in August 2018, daily average entries of each BART station are visualized through Tableau Desktop and Tableau Public. The static map and chart along with an interactive map, collectively show the high demand of BART in San Francisco downtown and nearby areas as both origin and destination. Second, the datasets from Uber Movement containing the Uber travel time data of the third quarter in 2018 are visualized based on census tracts using both Python and JavaScript. The interactive map shows extra-long delay occurred in San Francisco downtown area and south Bay Area compared with trips with similar travel distances. Third, the Bay Wheel ridership datasets in August 2018 are visualized using Google Map API, where a Bay Wheel usage data layer is created as a heat map placed above a normal street map layer, showing the high demand of Bay Wheel in San Francisco County especially downtown area compared with Sonoma County.

In conclusion, some certain demographic characteristics such as Asian population and income of households and distance from work places are proved to be closely correlated with the choices of driving personal vehicles, which leads to larger probability of traffic congestion and delay in Alameda, Contra Costa, San Mateo and Solano Counties. The usage of some popular ridesharing and public transit in the study area such as BART, Bay Wheel and Uber, collectively prove the high traffic demand in San Francisco County especially downtown area. Such findings can offer constructive advice for policy makers and researchers to come up with better transportation plans and policies to make commuting easier, faster and more accessible.

References

Alley, Jonathan K. "The impact of Uber technologies on the New York city transportation industry." (2016).

Amin, Ratna, and Sara K. Barz. "Seamless Transit: How to Make the Bay Area Public Transit Function Like One Rational, Easy-to-Use System." (2015).

Blanchard, Samuel D., and Paul Waddell. "Assessment of regional transit accessibility in the san francisco bay area of california with urbanaccess." Transportation Research Record 2654, no. 1 (2017): 45-54.

Grengs, Joe, Jonathan Levine, Qing Shen, and Qingyun Shen. "Intermetropolitan comparison of transportation accessibility: Sorting out mobility and proximity in San Francisco and Washington, DC." Journal of Planning Education and Research 29, no. 4 (2010): 427-443.

Hollingsworth, Joseph, Brenna Copeland, and Jeremiah X. Johnson. "Are e-scooters polluters? The environmental impacts of shared dockless electric scooters." Environmental Research Letters 14, no. 8 (2019): 084031.

Jin, Scarlett T., Hui Kong, and Daniel Z. Sui. "Uber, public transit, and urban transportation equity: A case study in new york city." The Professional Geographer 71, no. 2 (2019): 315-330.

Lasley, Phil. "2019 URBAN MOBILITY REPORT." (2019).

Lockwood, Allison M., Sivaramakrishnan Srinivasan, and Chandra R. Bhat. "Exploratory analysis of weekend activity patterns in the San Francisco Bay Area, California." Transportation Research Record 1926, no. 1 (2005): 70-78.

Ma, Xiaolei, Congcong Liu, Huimin Wen, Yunpeng Wang, and Yao-Jan Wu. "Understanding commuting patterns using transit smart card data." Journal of Transport Geography 58 (2017): 135-145.

McKenzie, Grant. "Urban mobility in the sharing economy: A spatiotemporal comparison of shared mobility services." Computers, Environment and Urban Systems 79 (2020): 101418.

Puzio, Thaddeus J., Patrick B. Murphy, Josh Gazzetta, Hannah A. Dineen, Stephanie A. Savage, Erik W. Streib, and Ben L. Zarzaur. "The electric scooter: A surging new mode of transportation that comes with risk to riders." Traffic injury prevention 21, no. 2 (2020): 175-178.

Sang, Sunhee, Morton O'Kelly, and Mei-Po Kwan. "Examining commuting patterns: results from a journey-to-work model disaggregated by gender and occupation." Urban studies 48, no. 5 (2011): 891-909.

Schaller, Bruce. "The new automobility: Lyft, Uber and the future of American cities." (2018).

United States Department of Transportation. "The National Transit Database (NTD)." Federal Transit Administration. (2020).

Wang, Fahui. "Modeling commuting patterns in Chicago in a GIS environment: A job accessibility perspective." The Professional Geographer 52, no. 1 (2000): 120-133.

Zhao, Pengjun, Bin Lü, and Gert de Roo. "Urban expansion and transportation: the impact of urban form on commuting patterns on the city fringe of Beijing." Environment and Planning A 42, no. 10 (2010): 2467-2486.