

Transit Usage in Pittsburgh, Pennsylvania

Introduction:

The coronavirus pandemic had a profound impact on public transit systems across the country. During the height of the pandemic, lockdown requirements and work from home changes for non-essential workers caused reduced travel patterns. For the remaining commuter trips that were being taken, those who were able to often opted for private modes of travel in lieu of taking public transit. As a result, ridership rates across transit systems in every city plummeted.¹ Although most of the lockdown restrictions associated with the severest waves of the pandemic have disappeared, ridership levels in most cities never fully recovered.

Pittsburgh, Pennsylvania is no exception to this national trend. In 2019, ridership levels for the Pittsburgh bus system averaged 180K people every month. In 2023, those numbers are almost half of what they used to be at - a mere 110K people a month.² These persistently low ridership rates have led to financial difficulties and service reductions.³ Whether or not these ridership changes are a new permanent feature of Pittsburgh's transit system or simply a sluggish after effect of the pandemic remains to be seen. In the meantime, it is necessary for city transit planners to better understand ridership patterns - both their current state and how they have been changing since the pandemic - in order to best allocate the scarce public funding dedicated for transit systems.

A large body of research has shown that the effects of the pandemic were and continue to be greatly disparate across various demographic groups. Accordingly, changes in transit usage have also been markedly different depending on a wide variety of factors - some of which include people's income levels, educational backgrounds, employment status and racial/ethnic backgrounds.⁴ This project seeks to better understand the current and historical patterns of transit usage in Pittsburgh, Pennsylvania through the use of geospatial data analysis.

¹ Liu, Miller, and Scheff, "The Impacts of COVID-19 Pandemic on Public Transit Demand in the United States."

² "Performance Metrics and System Data."

³ Guza, "Port Authority."

⁴ Liu, Miller, and Scheff, "The Impacts of COVID-19 Pandemic on Public Transit Demand in the United States."

Methodology:

The primary tools used for this project were:

- Data visualization in ArcGIS
- Basic statistical analysis
- Heat map analysis using ArcGIS' Kernel Density Tool

Data:

This analysis was done on a collection of datasets related to transit usage in Pittsburgh as well as datasets related to demographics within the city. More specifically, the following data was collected and merged together:

- **Bus Stop Level Data**
 - Variables:
 - Total onboardings
 - Total offboardings
 - Average onboardings
 - Average offboardings
 - Data Level:
 - Bus Stop Level
 - Select Months (5 months represented across 3 years) aggregated by weekday, saturday and sunday
- **Bus Route Level Data**
 - Variables:
 - On Time Percentage
 - Average Riders
 - Trip Distance per Route
 - Trip Count
 - Average Riders per Trip (derived)
 - Data Level:
 - Bus Route Level
 - Every Month from 2016 to 2023 , aggregated by weekday, saturday and sunday

Demographic data was collected on the Block Group level as well as the more disaggregated Block Level from the Census and ACS 5 year estimate datasets:

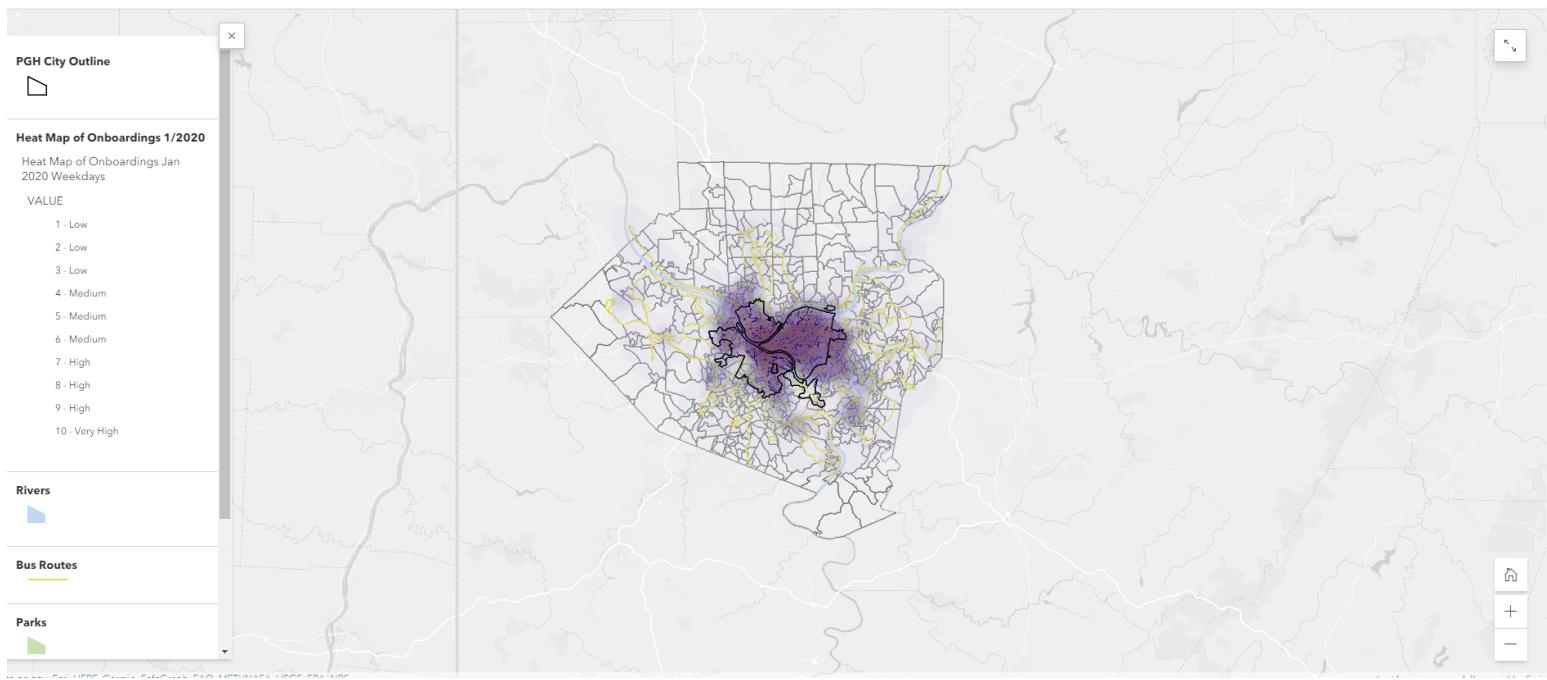
- **Block Level data:**
 - Variables:

- Population
- People under 18
- Housing Units
- Housing that is renter occupied
- Age
- Household Size
- Race
- Level:
 - Block Level
 - From 2010 Census
- **Block Group Level data:**
 - Variables
 - Households Below the Poverty Level
 - Total Population
 - Median Household Income
 - Median Home Value
 - Workers 16 or more who Took Public Transportation to work
 - Workers 16 or more who Worked at Home
 - Workers 16 or more who Drove Alone to Work
 - Population 35-64: Medicaid Coverage Only
 - Households Receiving Food Stamps/SNAP
 - Level:
 - Block Group Level
 - From 2020 ACS 5-yr Estimates

Results/findings:

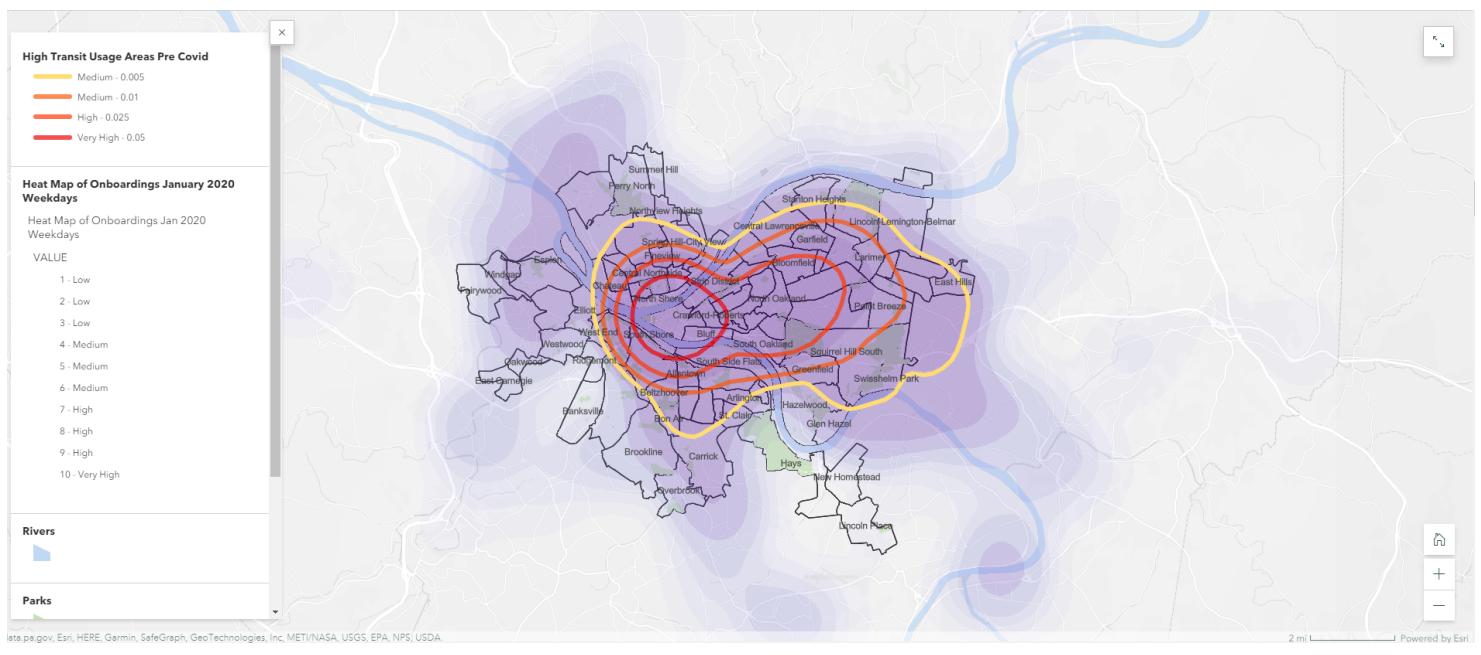
The Geospatial Distribution of Bus Ridership Before Covid

This heatmap was calculated using ArcGIS' Kernel Density tool for bus onboardings (by bus stop) in January 2020. Only weekday data was considered. As can be seen in the heatmap, although bus routes extend deep into Allegheny County, the vast majority of onboardings happen in the city limits of Pittsburgh.



Highest Use Areas of Bus Ridership Before Covid

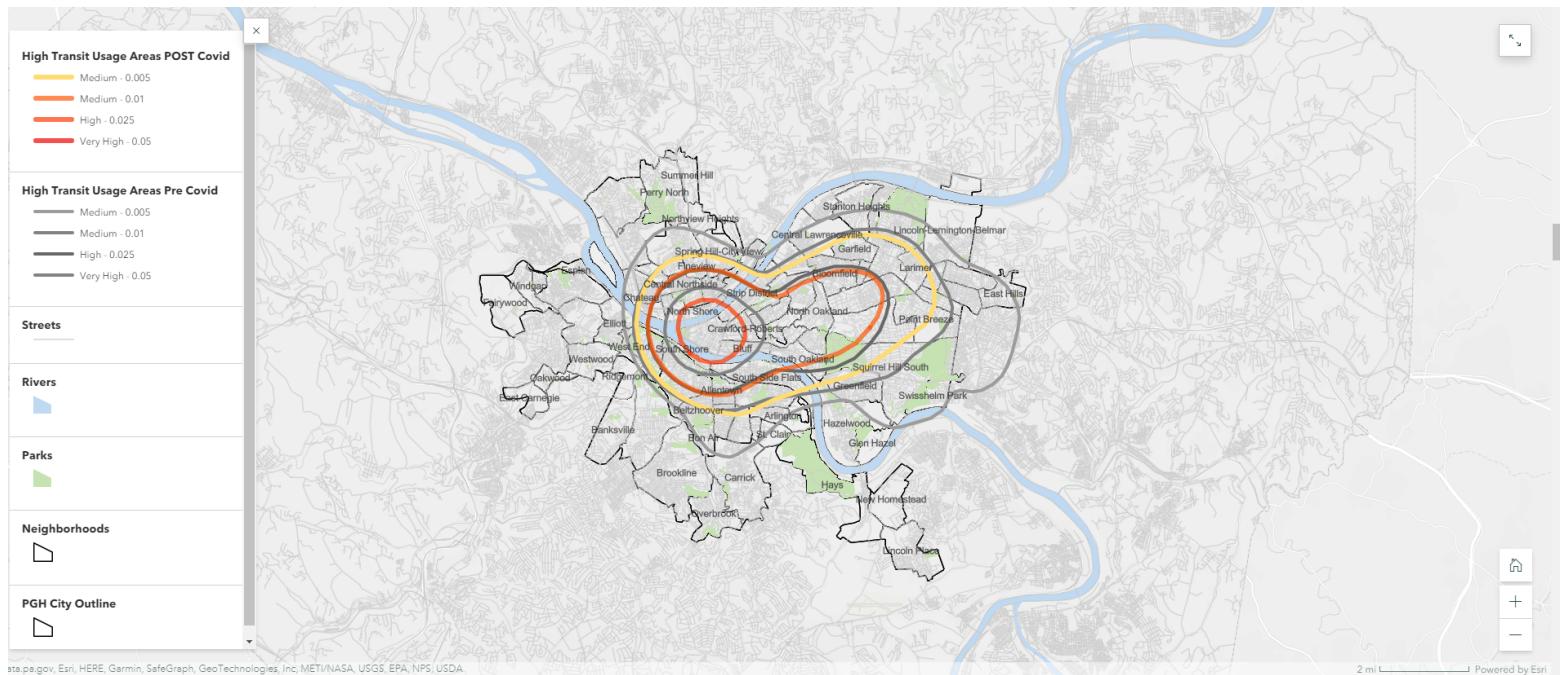
If we zoom into the previous heatmap, we can see there are several neighborhoods that had exceptionally high usage before the pandemic. Contours were drawn for different use levels, and it becomes clear that the Central Business District has the highest number of onboardings compared to any other neighborhood. Outside of the CBD, neighborhoods around the major universities such as Oakland, Shadyside, and Squirrel Hill all show high use levels. The high use contours spill out slightly onto the north side and south side of Pittsburgh but generally are more closely located to the center part of the city.



How did usage patterns change after Covid?

By comparing onboarding data from January 2020 to April 2021, we can see how ridership patterns have changed in both level and geospatial distribution. The map below has grayed out the high use contours from 2020 and redrawn them for data from April 2021. The first observation is that there are no neighborhoods in 2021 that reach the highest usage level (Red Contour) that was originally seen in 2020. This is reflective of the fact that overall ridership has declined across the city.

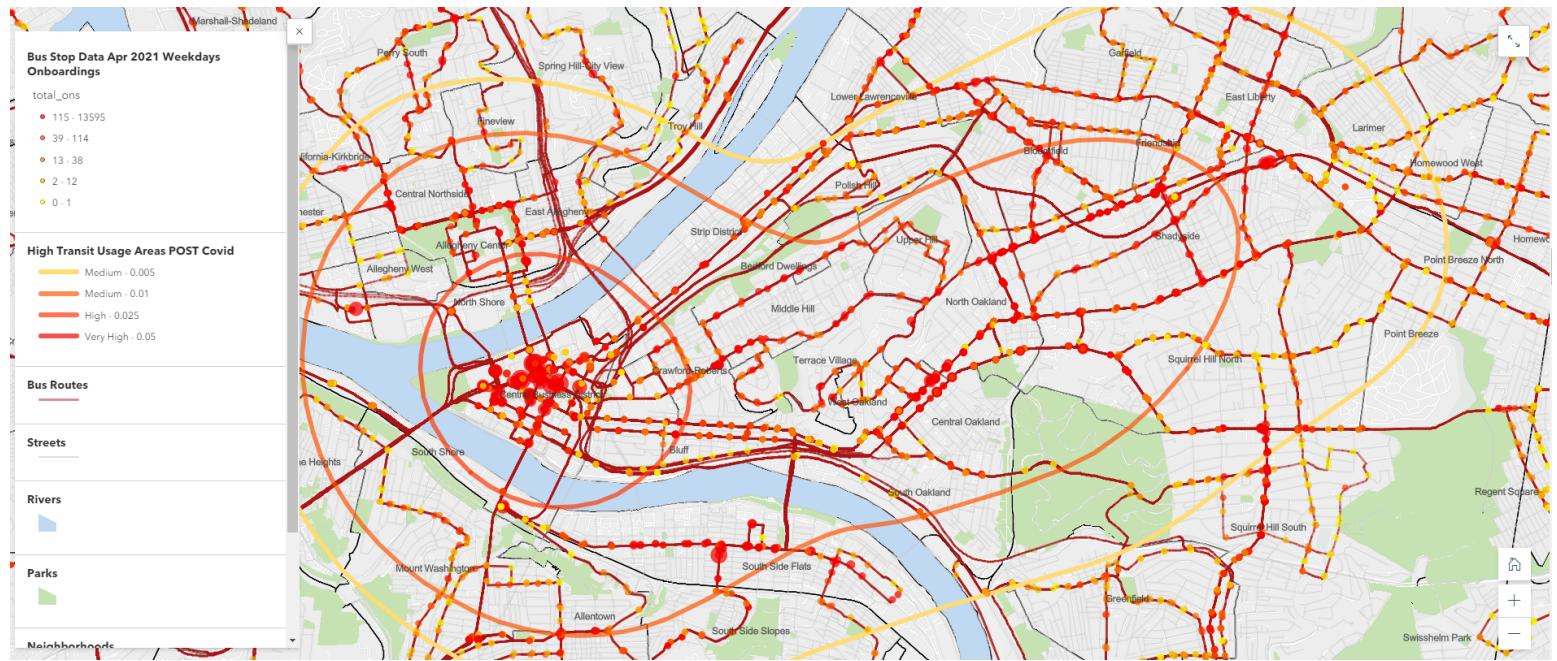
Apart from a general level decrease, these contrasted contour lines show that overall the geospatial patterns of ridership usage are broadly similar. What used to be medium use areas are now low use areas, and what used to be high use areas are now medium use areas. In other words, the geospatial distribution of ridership is broadly similar but simply has a smaller level of total ridership.



Interactive Map of Bus Stops and Bus Routes

Below is a map that shows the high use contours again with bus stop data and bus route data added. The color shows how many total onboardings there were at a stop. Some bus stops have exceptionally high onboarding levels, so those circles appear both red and enlarged. As can be seen, there is a high concentration of bus stops in the CBD that have exceptionally high onboardings. A couple other streets, such as Forbes Ave, Fifth St, and Center Ave also have

numerous bus stops with exceptionally high onboardings. If you zoom out, you can see that there are many routes on the outskirts of Allegheny County with low usage levels for the majority of the route.



Summary Statistics

Comparing High Transit Usage Areas to Allegheny County

Looking at the broadest high usage contour (yellow line), I calculated the Mean, Standard Deviation and Sum of a variety of ACS data on the Block Group Level.

These statistics show a variety of facts:

- Overall, there is more variation within the high transit use area. The standard deviation is generally higher for all metrics showing that this area is more heterogeneous than the broader Allegheny County
- Less than 10% of the total Allegheny County population lives within the high transit usage area
- Around 20% of all workers who take public transit to work live within the high transit usage area
- Around 6% of all workers who drive alone to work live in the high transit usage area
- Around 20% of all impoverished residents of Allegheny County live within the high transit usage area
- The mean poverty rate is higher in the high transit usage area, but there is much more variation in the high transit usage area than outside of it

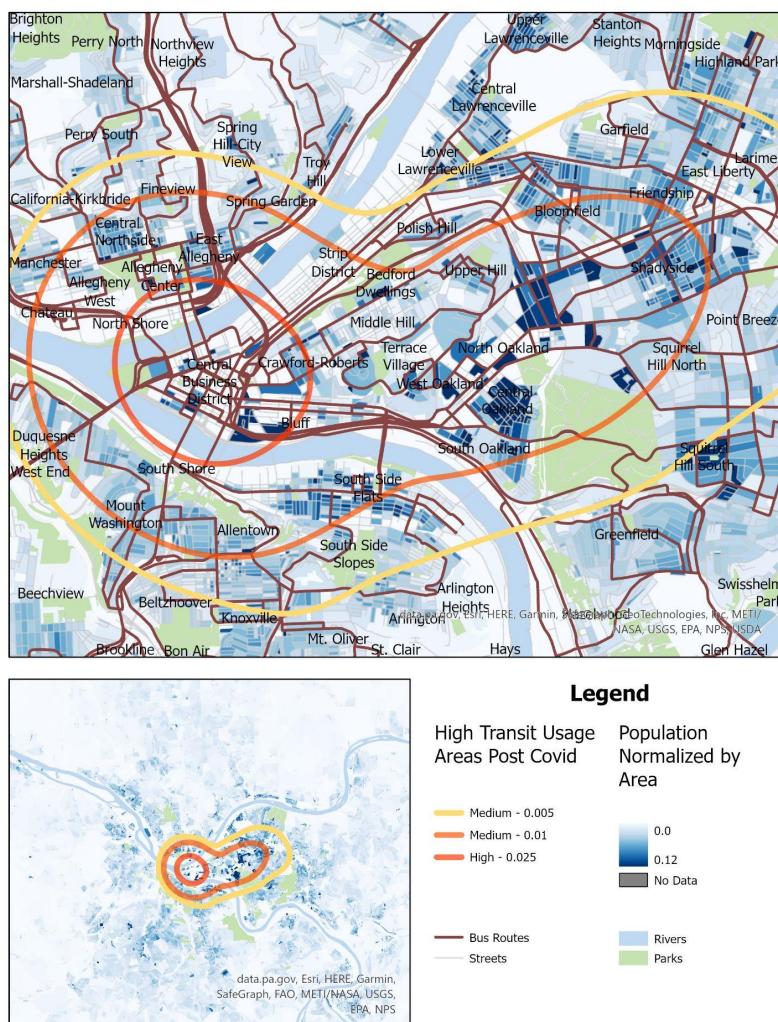
		Whole Dataset	Within 0.005 Medium Transit Usage Contour
Households Receiving Food Stamps	Mean	59.5	61.7
	Std Dev	59.5	81.6
	Sum	58K	6K
Households Below the Poverty Line	Mean	58.7	107.9
	Std Dev	68	109.5
	Sum	57K	11K
Total Population	Mean	1,127	970
	Std Dev	556	628
	Sum	1,102K	97K
Workers Who Took Public Transport to Work	Mean	50.6	103.5
	Std Dev	57.1	81.8
	Sum	49K	10K
Workers Who Drove Alone to Work	Mean	398.5	222.5
	Std Dev	249.3	150.8
	Sum	389K	22K
Poverty Percentage	Mean	5.70%	11.20%
	Std Dev	6.30%	9.30%
	Sum	NA	NA
Pop 35-64 Medicaid Coverage Only	Mean	34.6	25.5
	Std Dev	42.4	34.6
	Sum	33K	2K

Detailed Geographic Analysis of Transit Use Areas and Demographic Data

(below)

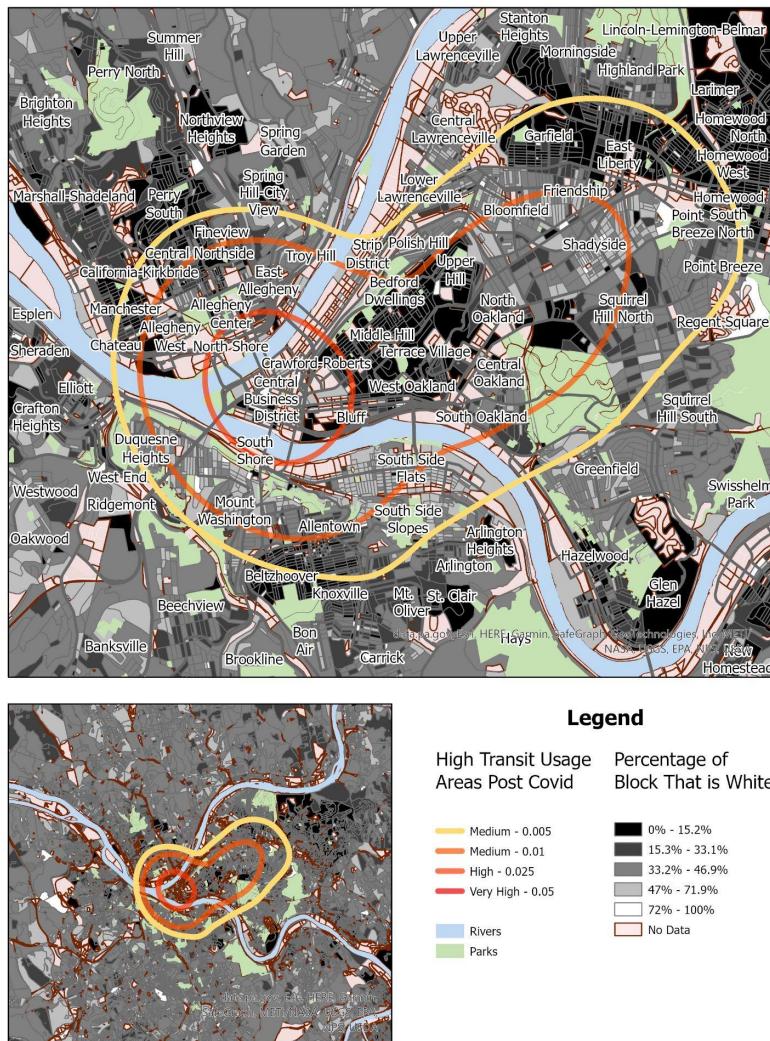
Population Density

To what extent can population density explain high transit usage areas? The map below shows the population density of each census block (Population divided by block area). As can be seen, several of the cities most dense areas, such as Oakland, are in the high transit usage areas. The Central Business District, however, is not particularly more dense than other blocks. This may reflect the fact that most transit usage in the CBD is related to employment, not residential purposes. Analyzing non weekday commuting patterns could reveal if this hypothesis is indeed true.



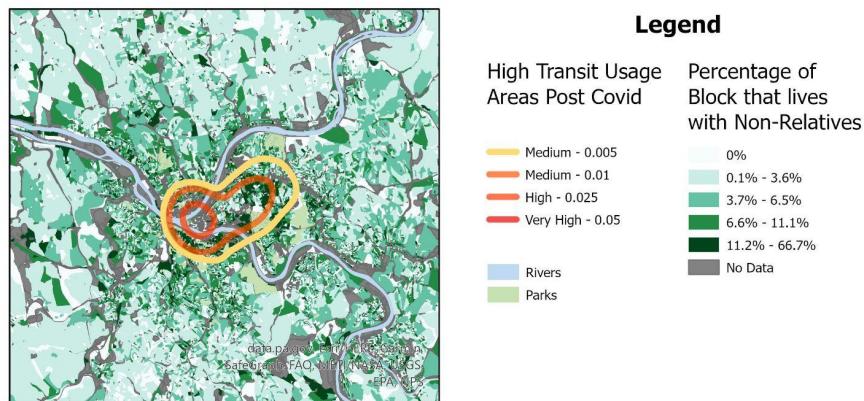
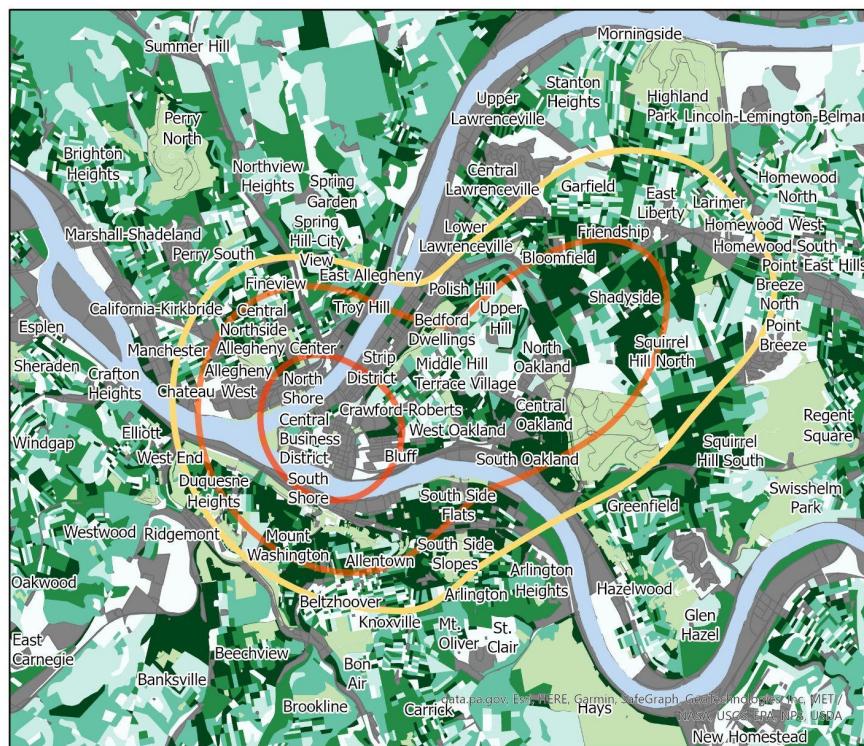
Race (% of Block that is white)

Can racial composition across the city explain transit usage patterns? The map below shows the percentage of people who identify as "white only" in the 2010 census. Although many blocks are missing data, we can see that there are both racially homogenous blocks and racially diverse blocks in the high transit use areas. Neighborhoods like Shadyside and Bloomfield are more white than neighborhoods like Middle Hill, Upper Hill and Terrace Village. Some of the most diverse blocks, however, lie well outside the high transit usage area as can be seen in the small map. The northeast section of Allegheny County is particularly diverse.



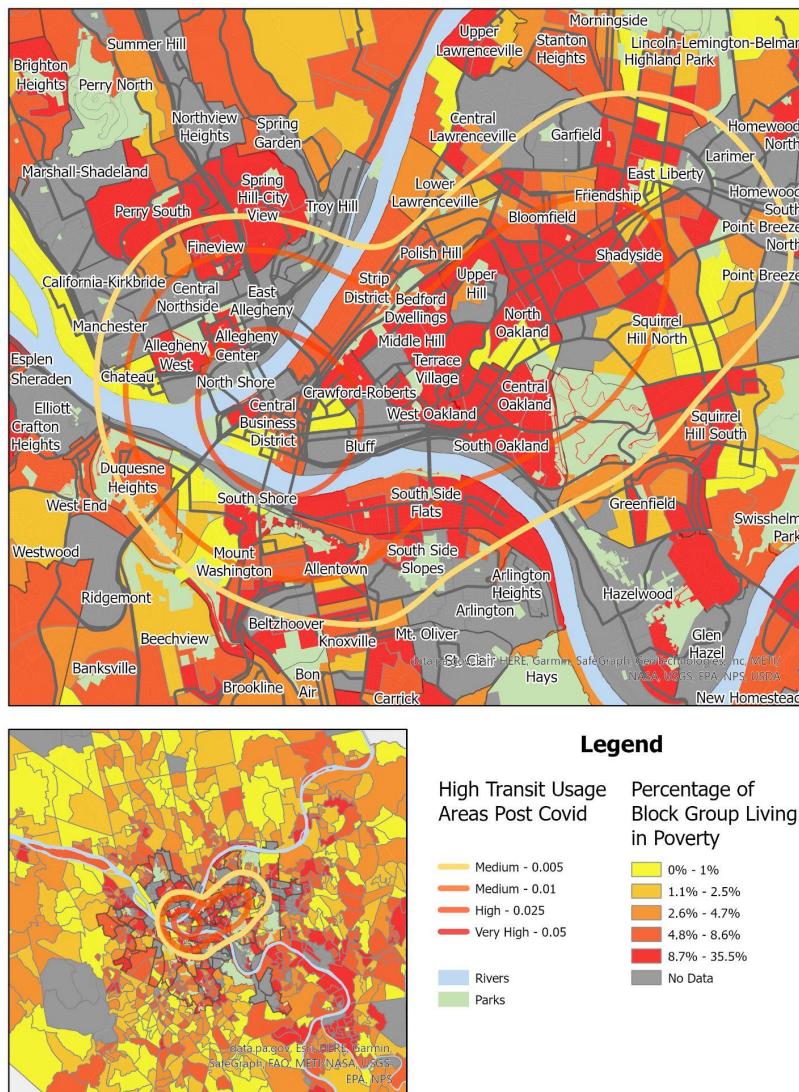
Percent of Residents who live with Non-Relatives

As a proxy for student populations, I analyzed the distribution of individuals who live with non family members. As can be seen, neighborhoods in Shadyside and Oakland have very high rates of residents living with non-relatives. Students may be less likely to have cars or be able to park at their universities, so this may help explain commuting patterns.



Poverty

This map shows the percentage of each block living in poverty. Compared to the broader Allegheny County patterns, it becomes clear that much of the high transit usage area experiences poverty at a greater rate than block groups outside of the transit usage area.



Conclusion and future work:**Conclusion**

From this initial analysis, we can better understand the geospatial distribution of bus ridership and how the pandemic has affected that. The heat map analysis shows that although the level of ridership has decreased, the geospatial distribution of the most high use areas is broadly similar. Investigating demographic details such as poverty, race, and non-relative living reveals that the high transit use area is more heterogeneous than the broader Allegheny community. Some areas are highly impoverished while others are more affluent. Some areas are highly diverse, but they are not necessarily the most diverse in the whole county. A large portion of the transit use seems to be connected to the central business district as well as the larger universities in the city.

Future Work

One potential area of future work would be to understand how to predict high levels of bus ridership based on demographic data. Intersecting census data with bus ridership data would give us a cross sectional dataset where every row was a census block/block group and an associated amount of bus ridership. We could then use regression analysis or some other predictive analytic technique to see which factors best predict high bus ridership in Pittsburgh. This could help inform policymakers where high use transit stops could be placed.

Data source(s):

- Bus Route Shapefiles From: [PRT Open Data](#)
 - Filename: *PRT routes current 2302*
 - URL:
https://open-data-pgh-transit.hub.arcgis.com/datasets/a26afd68bbc945898c8b51c1d1d5315e_0/explore?location=40.448770%2C-79.953480%2C11.80&showTable=true
- Pittsburgh Block Shapefiles: [Census.gov](#)
 - URL: <https://www2.census.gov/geo/tiger/TIGER2010BLKPOPHU/?C=N;O=D>
- Pittsburgh Block Group Shapefiles: [Allegheny County GIS Open Data](#)
 - URL:
<https://openac-alcogis.opendata.arcgis.com/maps/2c059d0501e5479ca90242a6d716d24c>
- Bus Ridership Data from: [Western Pennsylvania Regional Data Center](#)
 - URL: <https://data.wprdc.org/dataset?q=prt>
 - People getting on/off the bus by stop
 - Weekday/Saturday/Sunday for select year/month combos
 - Specifically 2019/09, 2020/01, 2020/09, 2021/04
 - Average riders by route
 - Weekday/Saturday/Sunday from 2017-2023
 - Number of trips for every route completed and distance of trip
 - Weekday/Saturday/Sunday from 2017-2023
 - On time percent by route (1 min early or 5 min late)
 - Weekday/Saturday/Sunday from 2017-2023
- Pittsburgh ShapeFiles From: [Allegheny County GIS Open Data](#)
 - URLs:
 - [City outline](#)
 - [Neighborhoods](#)
 - [Parks](#)
 - [Rivers](#)
 - [Block Groups](#)
- Allegheny County Street CenterLines From: [Pennsylvania Spatial Data Access \(PASDA\)](#)
 - URL: <https://www.pasda.psu.edu/uci/DataSummary.aspx?dataset=1224>
- Census Block Group Level Data From: [ACS 5 Year Survey, downloaded via Census API/Python](#)
 - [API Documentation](#)
- Census Block Level Data From: <https://data.census.gov/> (downloaded manually)

References:

- Guza, Megan. "Port Authority: Service Cuts Coming next Month Will Lead to More Reliable Service." TribLIVE.com, May 20, 2022.
<https://triblive.com/local/port-authority-service-cuts-coming-next-month-will-lead-to-more-reliable-service/>.
- Liu, Luyu, Harvey J. Miller, and Jonathan Scheff. "The Impacts of COVID-19 Pandemic on Public Transit Demand in the United States." *PLoS ONE* 15, no. 11 (November 18, 2020): e0242476. <https://doi.org/10.1371/journal.pone.0242476>.
- "Performance Metrics and System Data." Accessed April 24, 2023.
<https://www.rideprt.org/inside-Pittsburgh-Regional-Transit/Transparency/performance-metrics-and-system-data/>.