Final Project First Draft / Detailed Outline
PUBPOL 2130 / INFO 3130
May 1, 2025
Connie Chen, Amrith Samuel, Caroline Chan, Marianne Finn, Timothy Lin

Topic:

- **Explore features of public transit accessibility in relation to social demographics (race, gender, etc.) and community attributes (housing density, crime rate, etc.), and the correlation of public transit and the development of surrounding areas (economy, employment, etc.), as well as changes in these variables over time and in different regions.**

Literature Review Articles

1. https://pmc.ncbi.nlm.nih.gov/articles/PMC5476368/
2. https://pmc.ncbi.nlm.nih.gov/articles/PMC8111683/
3. https://www.sciencedirect.com/science/article/pii/S2214140517300853?via%3Dihub (Connie)
4. https://iopscience.iop.org/article/10.1088/1757-899X/1058/1/012049 (Connie)
5. https://www.sciencedirect.com/science/article/pii/S2214367X25000201 (Marianne)
6. https://wagner.nyu.edu/files/faculty/publications/State%20of%20Subway%20Ridership%20-%20Mar717.pdf (Marianne)
7. https://www.apta.com/wp-content/uploads/APTA-POLICY-BRIEF-Transit-Ridership-12.01.2023.pdf (Marianne)
8. https://repositories.lib.utexas.edu/server/api/core/bitstreams/32f52a27-5022-4b17-b5da-dd4b4be418fc/content (Caroline)
9. https://ascelibrary.org/doi/epdf/10.1061/9780784485521.052 (Caroline)
10. https://www.jstor.org/stable/26211733 (Timothy)
11. https://www.jstor.org/stable/48646229 (Timothy)
12. Fear of crime and transit use in Chicago neighborhoods (Amrith)
13. Transit-Oriented Development and Crime: Unpacking the evidence (Amrith)

Data Product Topics

1. ***Research Question*: How does urban design (combination of land use diversity/density and walkability & street design) correlate with the accessibility of public transportation?**

- Sub-Questions:
    1. Does higher transit service frequency and job accessibility via transit correlate with reduced car ownership and car usage?

2. Do residents in high income-entropy neighborhoods have greater access to jobs via transit (greater number of regional jobs accessible by transit) than those in low-entropy neighborhoods?
3. Does higher pedestrian-oriented urban design correlate with reduced greenhouse gas (GHG) emissions?

- Variables
    - D2B_E8MIX : 8-tier employment entropy (denominator set to observed employment types in the CBG)
    - D1A: Gross residential density (HU/acre) on unprotected land
    - D1B: Gross population density (people/acre) on unprotected land
    - D1C: Gross employment density (jobs/acre) on unprotected land
    - D2R_JOBPOP: Regional Diversity. Standard calculation based on population and total employment: Deviation of CBG ratio of jobs/pop from the regional average ratio of jobs/pop
    - D5DRI: Regional Centrality Index of Transit
    - NatWalkInd: Walkability Index
    - D3A: Road Network Density
    - D3B: Street intersection density (weighted, auto-oriented intersections eliminated)
    - D5BR: Jobs within 45-minute transit commute, distance decay (walk network travel time, GTFS schedules) weighted
    - D5DR: Proportional Accessibility of Regional Destinations by Transit (Employment accessibility expressed as a ratio of total MSA accessibility)
    - D4C: Aggregate frequency of transit service within 0.25 miles of CBG boundary per hour during evening peak period
    - D4E: Aggregate frequency of transit service per capita
    - Pct_AO0: Percent of zero-car households
    - Pct_AO1: Percent of one-car households
    - Pct_AO2p: Percent of two-plus-car households
    - VMT_tot_avg: Weighted average daily Vehicle Miles Traveled (VMT) per worker
    - W_P_Lowwage: Percent low wage workers
    - W_P_Medwage: Percent medium wage workers
    - W_P_Highwage: Percent of high wage workers
    - D3APO: Network density in terms of facility miles of pedestrian-oriented links per square mile
    - D3BPO3: Intersection density in terms of pedestrian-oriented intersections having three legs per square mile
    - D3BPO4: Intersection density in terms of pedestrian-oriented intersections having four or more legs per square mile
    - Workers_1: Number of workers

    - GHG_per_worker: Estimated average daily GHG generated by a worker in a workplace block group

- Data Product Type:
    - Research Question: Scatter Plot with Regression Line:
        - X-axis: Urban Design Composite Score (land use diversity + walkability)
        - Y-axis: Accessibility Index (transit accessibility)
        - Size: Population Density

    - Sub-Questions:
        1. Heat Map
            - Variables: Land use diversity/density variables, Walkability, Street design
        2. Scatter Plot with Regression Line:
            - X-axis: Transit service frequency
            - Y-axis: Car ownership
            - Color: Job accessibility via transit correlate
            - Size: Population Density
        3. Stacked Bar Chart + Line Graph
            - X-axis: Regions, Ordered by Income-Entropy (Labeled on the top of each bar)
            - Y-axis (left, bar): Percentage of workers that are low, medium, or high wage (by home locations)
            - Y-axis (right, line): Percentage of regional jobs accessible by transit

- Datasets
    - Smart Location Database: https://www.epa.gov/smartgrowth/smart-location-mapping
        - https://catalog.data.gov/dataset/smart-location-database8
    - US ACS Data: https://data.census.gov/

- Literature Summary
    - The article investigates how urban design influences transportation behaviors in Melbourne, Australia, using data from 16,890 participants in the 2009–2010 Victorian Integrated Survey of Travel Activity (VISTA09). The study assesses associations between urban design features—such as residential density, street connectivity, land use mix, and proximity to services—and four transport modes: walking, cycling, public transport, and private vehicle use. Key findings reveal that neighborhoods with residential densities exceeding 20 dwellings per hectare, well-connected street networks, access to ≥9 local destinations (e.g., supermarkets, schools), and proximity to public transport (≤400m for

bus stops, ≤800m for train stations) significantly increase the likelihood of walking, cycling, or using public transport while reducing car dependency.

- The article explores the critical role of public transportation in shaping sustainable urban environments. It identifies the interconnected challenges of urban sprawl, car dependency, and inefficient land use as key contributors to declining quality of life in cities. The authors argue that public transit systems, when integrated with strategic urban planning, can direct urban form toward sustainability through four primary factors: density of development, diversity of land uses, urban design, and proximity to urban centers. Higher residential and employment densities around transit routes increase ridership, enabling more frequent and efficient public transport services. Mixed land uses reduce reliance on private vehicles by creating self-sufficient neighborhoods where daily needs are met locally. Urban design elements, such as grid-based street patterns and pedestrian-friendly infrastructure, enhance walkability and connectivity. Proximity to city centers further diminishes car dependency, as residents in peripheral areas tend to use public transit more than those in central zones.
-

---

2. ***<u>Research Question</u>: How has overall public transportation ridership changed over time? Are there identifiable periods of significant increase or decline in ridership? What years do these occur?***

- <u>Sub-Questions</u>:
    - How have ridership trends varied by transit mode (e.g., bus, subway, commuter rail, light rail) over time? Which modes of transportation have seen the greatest increase or decrease in usage?
    - What are the monthly or seasonal patterns in ridership over the years? Have these patterns shifted? Has weekday vs. weekend ridership changed over time?
    - How have trends differed between transit-heavy cities (e.g., New York, Chicago, San Francisco) and car-dependent cities?
    - How did major events (e.g., COVID-19 pandemic, 2008 recession, gas price spikes) affect ridership trends?
    - How did fare changes (increases, reductions, free transit trials) align with ridership fluctuations?

- <u>Variables</u>
    - Method of transportation (bus, heavy rail, light rail)
    - Number of riders
    - Location
    - Fare Amount
    - Time (daily, monthly, yearly)

- Data Product Type:
    - Line Chart:
        - X-axis: Year (past 25 years)
        - Y-axis: Total annual ridership
        - Markers for major events
    - Multi-line chart:
        - X-axis: year
        - Y-axis: annual ridership
        - Line per mode of transportation

- Datasets
    - ACS data (2023, 2018, 2013, 2009)
    - NTD Monthly Ridership Data: https://www.transit.dot.gov/ntd/data-product/monthly-module-adjusted-data-release
    - MTA Daily Ridership Data: https://catalog.data.gov/dataset/mta-daily-ridership-data-beginning-2020
    - D.C. Daily Ridership Data: https://www.wmata.com/initiatives/ridership-portal/daily-summary.cfm
    - Chicago Daily Ridership Data: https://data.cityofchicago.org/Transportation/CTA-Ridership-Daily-Boarding-Totals/6iiy-9s97/about_data

- Literature Review Summaries (Rough Draft):
    - Report from NYU Rudin Center for Transportation analyzes subway ridership in Brooklyn, Bronx, Queens, and Manhattan from 1975-2015. (Moss et. al, 2017) The report uses data from the NYCT and historical MTA references to visualize how subway ridership has changed over time, and how that change mirrors changes in income and fares. The report divides their timeline into 10 year gaps and breaks down ridership by borough. While the report does not imply correlation nor causation between income, fares, major events, and ridership, the report demonstrates the fluctuation in ridership and what potential causes may be connected. A study conducted by Seyed Sajjad Abdollahpour, Huyen T.K. Le, and Steve Hankey investigated changes in transit ridership predictors pre- and post-pandemic across 35 United States metropolitan areas. The study defines three time periods: before the COVID-19 pandemic (1/1/2019 to 2/1/2020), during the pandemic (3/1/2020 to 2/1/2021), and the post-pandemic (recovery time) (3/1/2021 to 1/6/2023). In their research, they found that “socioeconomic factors, gasoline prices, telecommuting, and polycentric development collectively

contribute more to post-pandemic ridership than vehicle revenue mile (VRM), fare, transit coverage, and service areas, marking a reversal from the pre-pandemic era." Studies conducted previously focused on factors that influenced public transportation ridership pre-pandemic or during early phases of the pandemic. The study they conducted uses NTD data from 2019-2023. The research found that external factors, which includes gasoline prices, careless households, income, demographic, vaccination rate, and population density, contribute more to post-pandemic ridership. Here, we can further investigate some of these factors with data collected in 2024-2025 and see how ridership has changed. The APTA acknowledges that after falling to 20 percent of pre-pandemic levels in April 2020, ridership has recovered to more than 77 percent of pre-pandemic levels in 2023. With new data from New York City, Washington D.C., and Chicago we can visualize if the trend upward has continued.

---

3. ***<u>Research Question</u>: How does social vulnerability (characterized by factors such as poverty, disability, and minority status) correlate with public transit accessibility?***

- <u>Sub-Questions</u>:
    - Which specific aspects of social vulnerability (such as poverty, minority status, etc.) are most strongly associated with lower public transit accessibility?
    - Are neighborhoods with higher social vulnerability scores more likely to experience transit deserts (high need, low service)?
    - How do individual components of social vulnerability uniquely and collectively contribute to transit accessibility gaps?

- <u>Variables</u>
    - Social Vulnerability Index (SVI)
        - Employment %
        - Education %
        - Race
        - Age
        - Income
            - This data is readily available through ACS
    - Transit Accessibility
        - Transit service frequency
        - Jobs accessible within 45 minutes by transit
    - Transit Desert Index (TDI)
        - td_index > 1 → transit desert

        - td_index (-1,1) → transit adequate
        - td_index < -1 → transit oasis

- Data Product Type:
    - Research Question:
        - Scatterplot with Regression Line
            - Analyze whether there is an evident trend / correlation (e.g., downward trend of SVI as public transit accessibility goes up)
            - X-axis: SVI score
            - Y-axis: transit accessibility score
    - Sub-Questions:
        - Bar Chart of Correlation Coefficients
            - Show strength of association between SVI components (poverty, minority status, disability, etc.) and transit accessibility
        - Bivariate Choropleth Map
            - Overlay can potentially highlight the relationship between SVI and public transit accessibility
            - Can utilize TDI and SVI to show how they coincide spatially.

- Datasets
    - ACS Data
      https://data.census.gov/
    - SVI Data & Documentation
      https://www.atsdr.cdc.gov/place-health/php/svi/svi-data-documentation-download.html
    - Smart Location Database
      https://www.epa.gov/smartgrowth/smart-location-mapping
    - Transit Desert Dataset
      https://www.arcgis.com/sharing/rest/content/items/f1b6dd3c82d748ca9ac1e98972de5e5a/data

- Literature Review Summaries
    - The article examines "transit deserts" – areas in U.S. metropolitan regions where public transportation is lacking despite high demand – and finds that about 24.6 million people, or 11% of the population in these areas, live in such zones. Using a new measurement method and the CDC's Social Vulnerability Index, the research shows that residents of transit deserts are more likely to be socially vulnerable: nearly 19% live below the poverty line (compared to 12.8% citywide), 58.1% are minorities (versus 43.6% citywide), and 6.4% of households lack a vehicle (compared to 3.5% citywide). Transit desert residents also have a higher average vulnerability score (0.58 vs. 0.48 citywide) and are more likely to have limited English proficiency.

- The article, conducted in northeastern Illinois, investigates how public transit accessibility varies across different demographic groups, with a particular focus on racial and socioeconomic disparities. Using a needs gap index calculated from 2020 American Community Survey data at the census tract level, the study measures the mismatch between transit demand (driven by factors like poverty, minority status, and lack of vehicle ownership) and transit supply (proximity to transit stations). The analysis reveals that transit deserts are most pronounced in suburban counties (although not exclusively) and disproportionately affect Hispanic, African American, and Asian communities, with the Hispanic community experiencing the greatest underservice. Importantly, the spatial patterns of transit deserts for minority groups differ from those for the overall population, which highlights the need for targeted, group-specific interventions rather than one-size-fits-all solutions. The study's methodology and findings demonstrate that specific aspects of social vulnerability, such as poverty and minority status, are strongly associated with lower public transit accessibility and that these components contribute to gaps in service.

---

4. ***<u>Research Question</u>: How do transit accessibility and quality vary between US cities? More specifically, are some cities doing a better job at ensuring equitable access to reliable public transit options than others?***

- <u>Sub-Questions</u>:
    - How do US cities vary in terms of how close and frequent transit options are for various demographic groups?
    - How do US cities vary in their ability to provide minimally delayed transit services to all residents?

- <u>Variables</u> (* variable readily available at block group level in EPA Smart Location Variables Database or census/ACS)
    - Distance to nearest transit stop *
    - Transit service density *
    - Transit service frequency *
    - Jobs within a 45 minute transit commute *
    - Percent of trips that are on-time
    - Delay times by stop and line
    - Median household income *
    - Race *
    - Dissimilarity index (calculated between different demographic groups for blocks with satisfactory/unsatisfactory transit access)
- <u>Data Product Type</u>:
    - <u>Research Question</u>:
        - 2x2 grid of plots

        - Top row: bar chart of dissimilarity index values by city for race (left plot) and class (right plot)
        - Bottom row: bar chart of average delay times by city for race and class

    - Sub-Questions:
        - Choropleth of delay times for each block group in each city
        - Choropleth of demographic breakdown in each city
        - Bar chart of overall on-time performance figures for each city's associated transit authority
- Datasets
    - https://www.epa.gov/smartgrowth/smart-location-mapping
    - https://conservancy.umn.edu/items/85608086-03c2-4f93-a033-98125628b8f1
    - https://catalog.data.gov/dataset/?tags=on-time-performance
    - https://data.census.gov/
    - https://www.mta.info/developers
    - https://www.mbta.com/developers/gtfs
    - https://www.bart.gov/schedules/developers/gtfs-realtime
    - https://www.itsmarta.com/app-developer-resources.aspx
    - https://opendataphilly.org/datasets/septa-gtfs-alerts-updates/

- Literature Review Summaries
    - Existing literature finds multiple potential ways to gauge accessibility and disparities in accessibility, though few employ these metrics in a comparative context across US cities. Karlson and Owen examine accessibility across different demographic groups in the Minneapolis-St. Paul metropolitan area in their 2021 article "Accessibility: Distribution across diverse populations." Here, they use the "distribution of job accessibility" as a proxy through which broader accessibility patterns within cities may be investigated. Using annual transit accessibility data from the National Accessibility Evaluation and census employment data, they compute the average "cumulative accessibility" (number of jobs reachable within 30 minutes by transit) across census blocks. Importantly, since accessibility itself is a locational rather than individual metric, they weigh each block by its worker population. They find that younger and non-white workers tend to have more robust transit access across national, state, and local geographies. Additionally, they find that the most and least educated groups have greater accessibility than the middle group (high school diploma to an associate's degree).
    - Merlin takes a more comparative approach to accessibility in his 2017 article "A portrait of accessibility change for four US metropolitan areas," though his work centers around changes to accessibility over time rather than accessibility at large. In the article, he compares accessibility trends from 2000-2010 in Charlotte,

Chicago, St. Louis, and Seattle and uses an accessibility metric over traffic analysis zones (TAZs) that is a function of the number of jobs in other TAZs, an "impediment coefficient," and transit times. To compute these values, Merlin obtained data from each city's metropolitan planning organization and also used census labor data. The article finds that people are increasingly living in areas with low accessibility, though the benefits of transit accessibility vary by city, accruing to high density areas in Chicago and low density ones in St. Louis.

---

5. ***Research Question: How do local crime rates correlate with the use of public transportation across different urban neighborhoods?***

- Sub-Questions
    - How do specific types of crime (violent vs. property crimes) individually impact public transit usage rates?
    - Does a higher perceived safety score correlate with increased ridership during nighttime hours?
    - Are low-income communities more resilient to fluctuations in transit usage driven by changes in crime rates compared to high-income communities?

- Variables
    - Violent_crimes_x
    - Property_crimes
    - Nighttime_crime_pct
    - Transit_pct
    - Median_income
    - income_group

- Data Product Type:

- Research Question:
    - Transit Usage vs Violent Crime Count
    - Transit Usage vs Property Crime Count
    - Transit Usage vs Nighttime Crime Rate
    - Transit Usage Percentage by Income Group
    - Violent Crime Count by Income Group

- Sub-Questions
    1. How do specific types of crime (violent vs. property crimes) individually impact public transit usage rates?

        Graphs Used:

    - Transit Usage vs Violent Crime Count
    - Transit Usage vs Property Crime Count

2. Does a higher perceived safety score correlate with increased ridership during nighttime hours?
    Graph Used: Transit Usage vs Nighttime Crime Rate
3. Are low-income communities more resilient to fluctuations in transit usage driven by changes in crime rates compared to high-income communities?
    Graphs Used:
    - Transit Usage Percentage by Income Group
    - Violent Crime Count by Income Group

- Datasets
    - https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i
    - ACS data: https://data.census.gov/