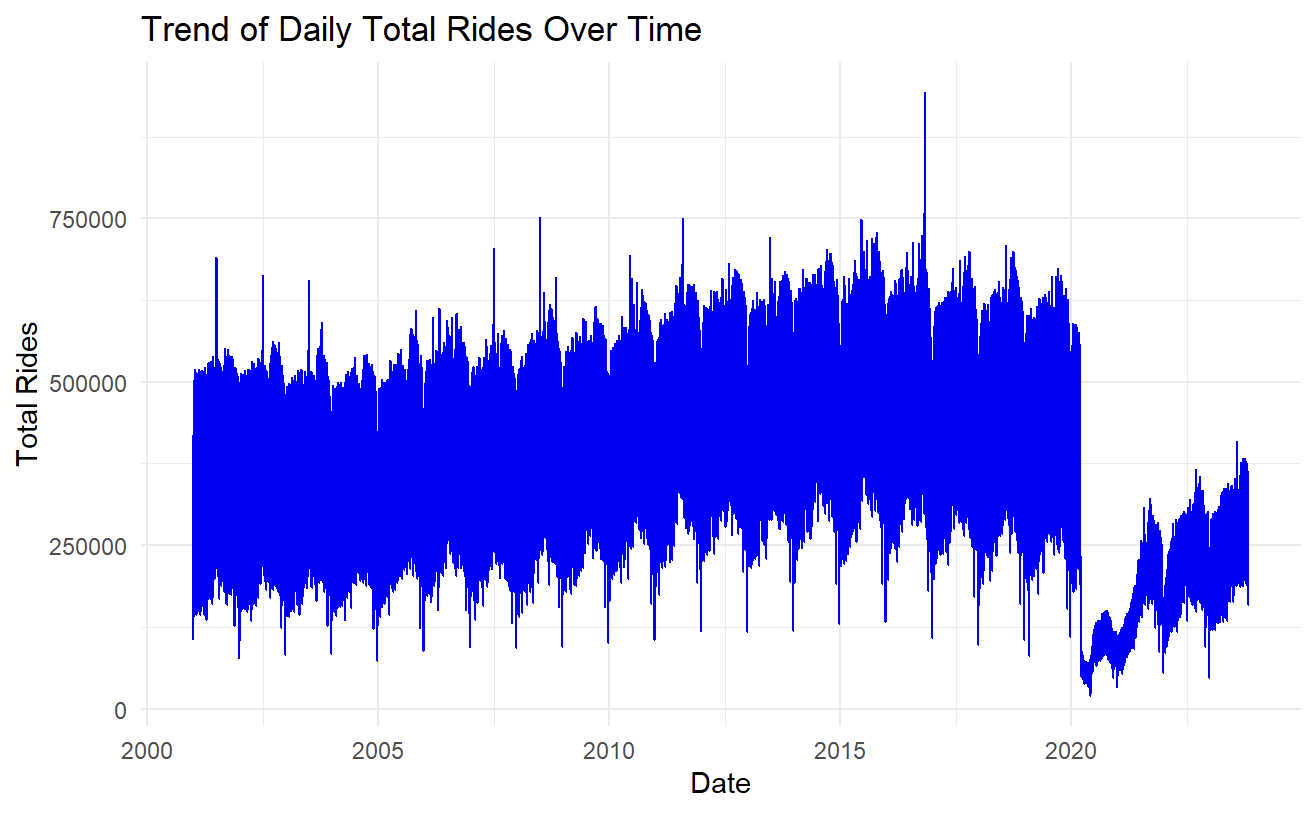
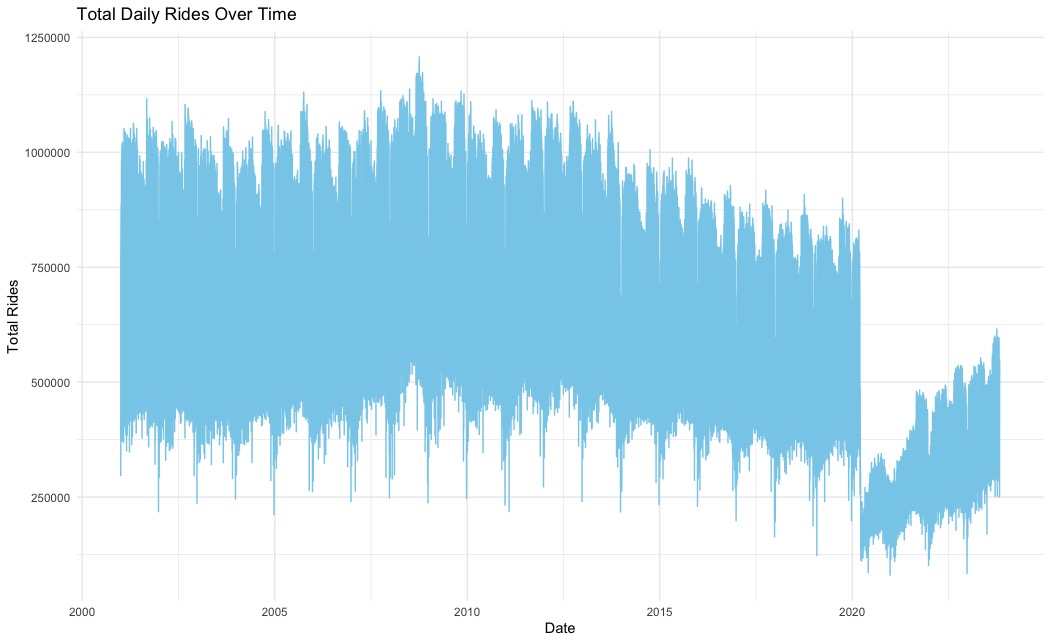
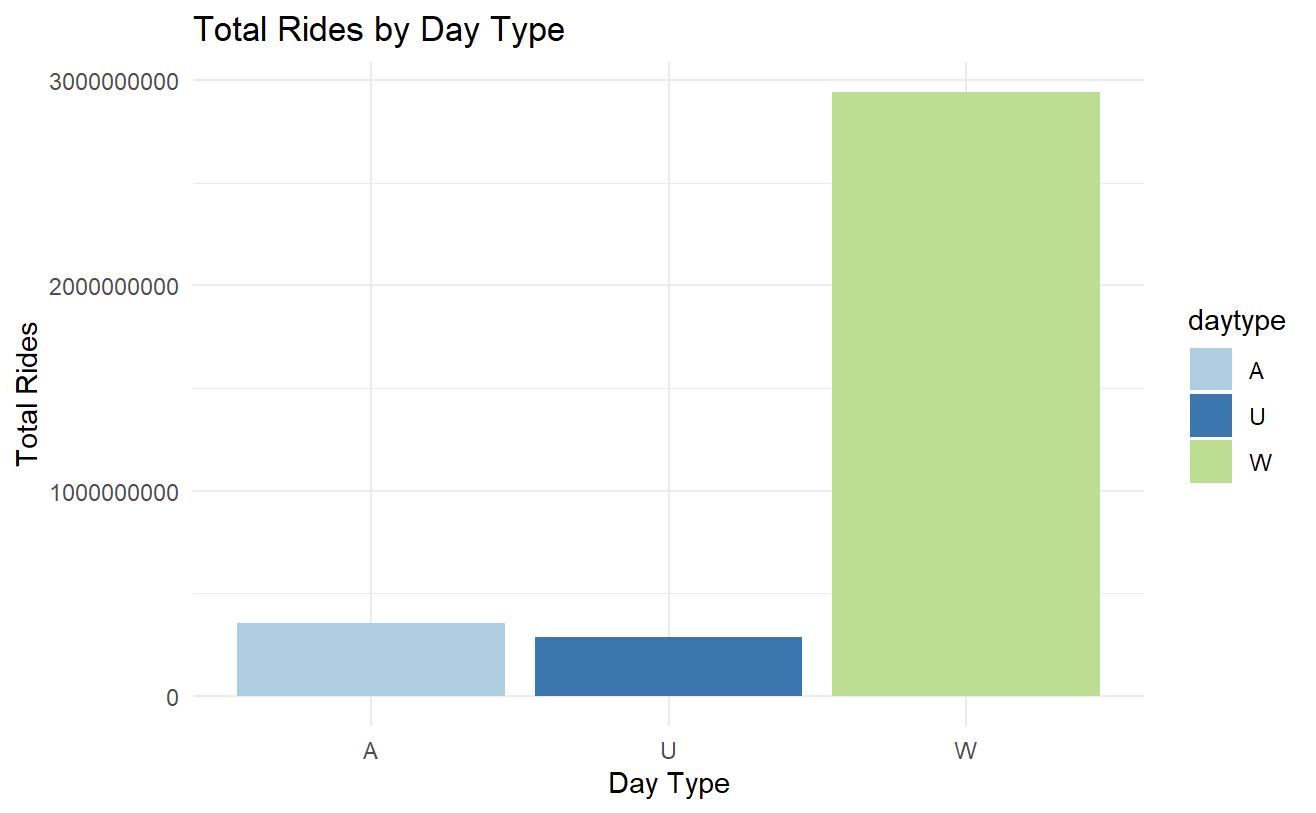
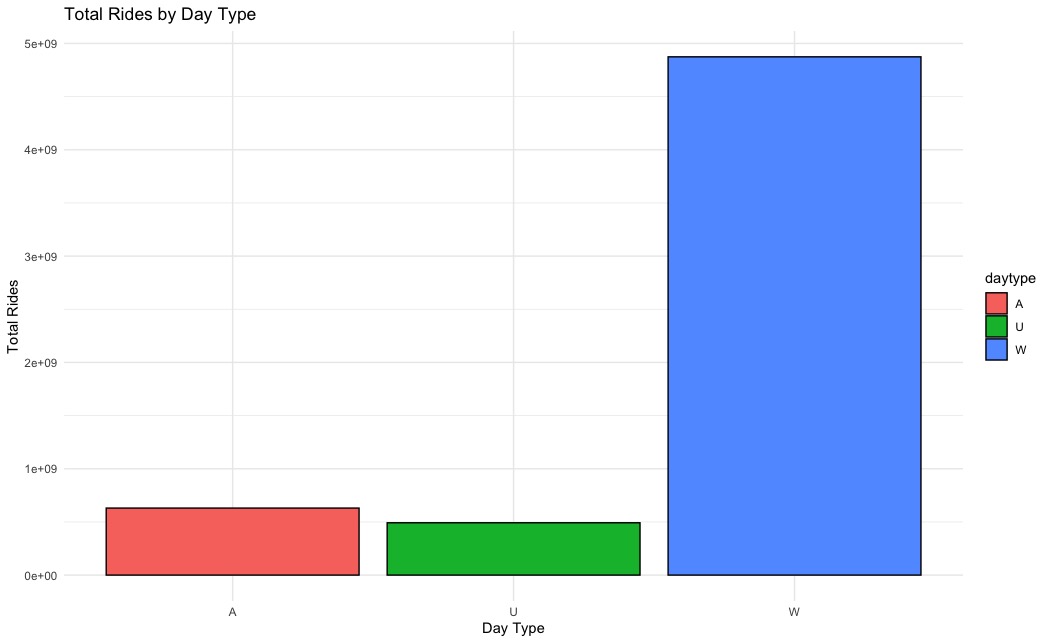
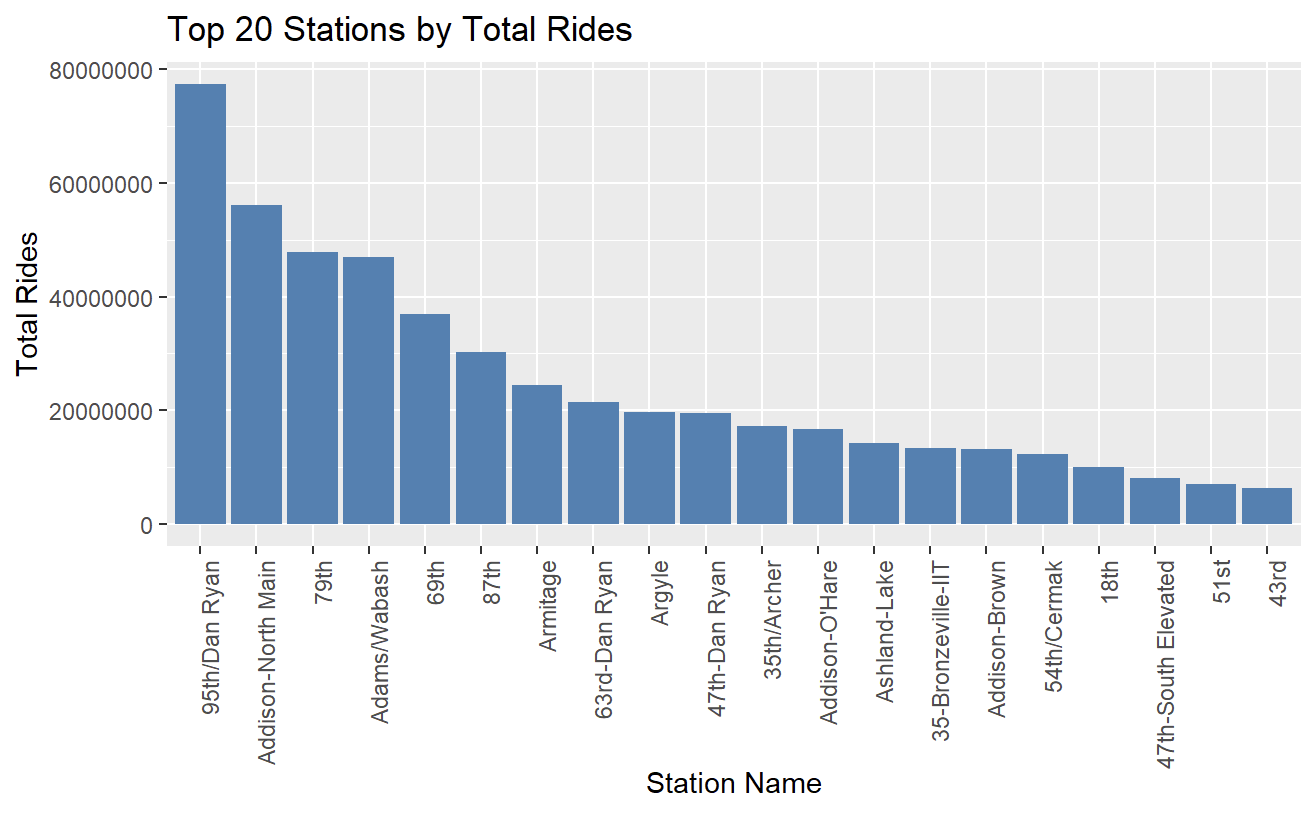
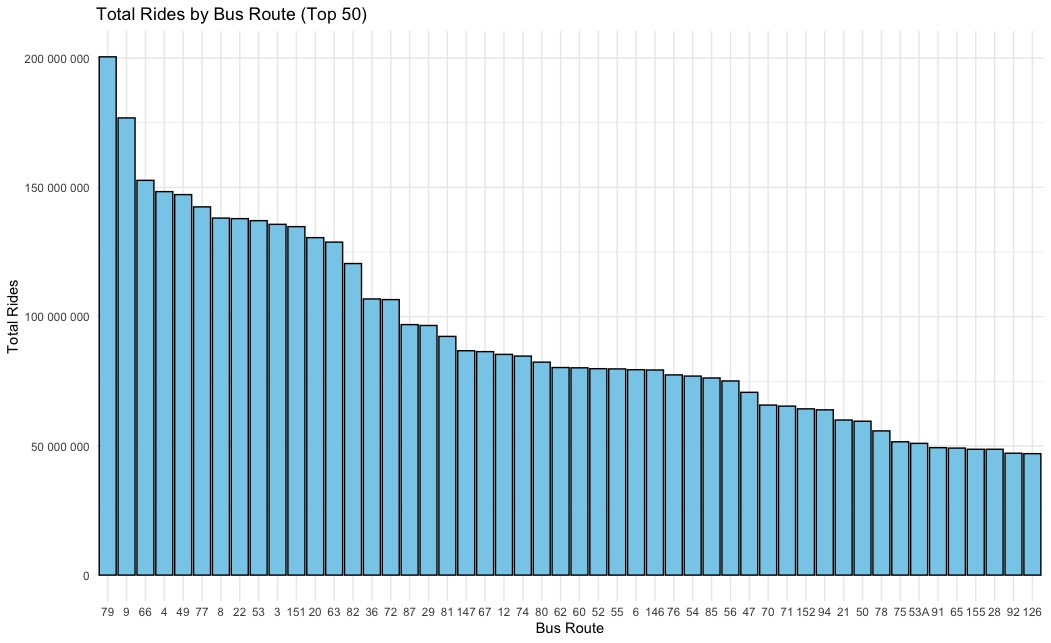
**Data Introduction**  
*CTA Ridership Data*  
This data set has the daily ridership entries for CTA stations in Chicago broken down by station and weekday, Saturday, or Sunday/Holiday. This will allow us to see which stations and lines are most used and what days they are most used on. This should allow us to see the differences between weekday and weekend transit patterns as well as see how special days (for example the Cubs World Series) impact the data. This data, combined with the bus system data and demographic data, should allow us to see which communities in the city are underserved and which stations or areas require additional capacity.   
  
Total Daily Rides Over Time (2001-2022)

   
*CTA ’L’ (right), Bus (left)*

Total Rides by Day Type (W: Weekday, A: Saturday, U: Sunday/Holiday) 

*CTA ’L’ (right), Bus (left)*

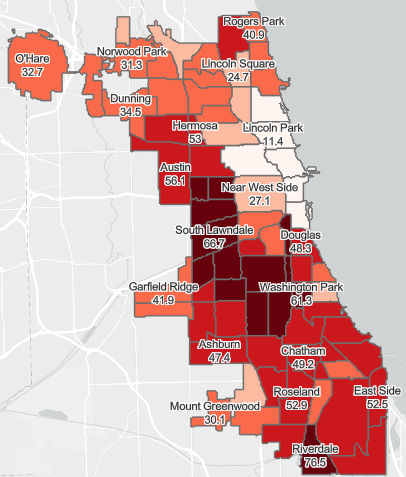
Top ‘L’ Stations and Bus Routes

*CTA ’L’ (right), Bus (left)*

The CTA ‘L’ graph highlights the top 20 most used stations. We can see that 95th Dan Ryan is the most used. Since that is the last stop on the South side for the red line it likely is pulling in people from surrounding communities that do not have any other access to public transport. 79th St. The station is the next on the red line, so it is likely having a similar effect. The Addison station is the main station to get to Wrigley Field which is why it would be the second most used.

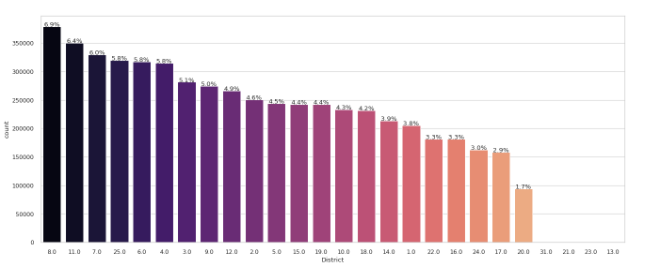
*Chicago Community Hardship Index*

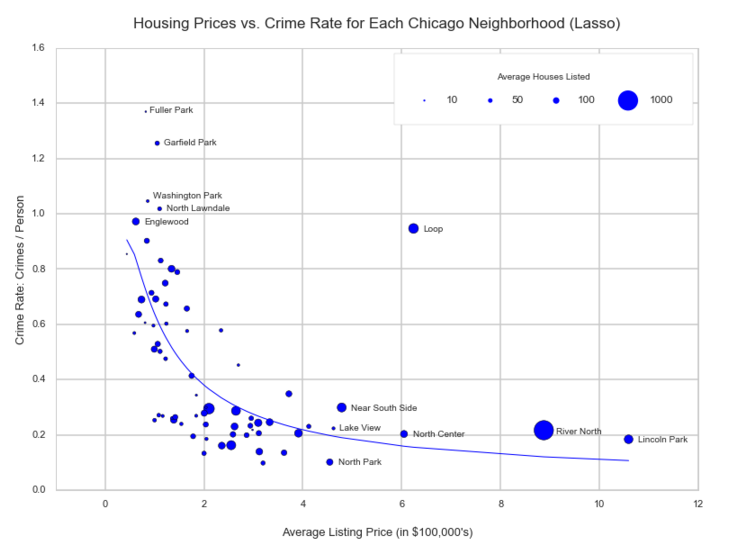
UIC Great Cities Institute’s dataset uses data from the U.S. Census Bureau to examine the conditions of economic and social hardship within Chicago community areas. The dataset includes Community Area (the names of the community areas), Community Area Number, Hardship Index Score, and the variables that make up the hardship score: Percent of Crowded Housing (housing units with more than one person per room), Percent of Households with Income Below Poverty Level (below the federal poverty level), Unemployment Rate for Population Age 16 and Over, Percent Aged 25 and Over with No High School Diploma, Percent of Population Under Age 18 and Over Age 64 (dependency), and Per Capita Income.  

*Chicago Population Counts*

The dataset contains population totals by age group and by race for ZIP codes within Chicago. The dataset has estimates for 2018 to 2021 and is from the U.S. Census Bureau.

*Housing and Crime Data*

Chicago crime data informs CTA analysis for safer routes and reliable public transportation services. Analyzing crime rates across different districts in Chicago, District 8 emerges as a focal point for attention when considering safety and security measures within the Chicago Transit Authority (CTA) system. Understanding the concentration of crime in this district provides valuable insights for enhancing safety protocols and addressing potential concerns for CTA passengers and staff.



Chicago housing data offers key insights into neighborhood dynamics like housing prices, crime rates, and demographics, informing CTA analysis for route planning, safety measures, and equitable transit access. Identifying neighborhoods with favorable combinations of housing prices and crime rates can help build strategies to promote mixed-use development and transit-oriented growth. By targeting areas with high housing demand and crime rates that are relatively low, the CTA can support development projects that enhance walkability, connectivity, and livability around transit hubs.

*Multiple Dataset Acknowledgements*  
Going forward, the largest issue we will likely face is combining the various datasets in order to gain a more holistic understanding of the transit patterns in the city. However, Chicago also provides geographic data for all transit stops, so that will also be used as an aid. We should be able to use geographic visualization techniques as well as predictive techniques like regression analysis to predict what a new station's usage could be based on the surrounding demographic data.  
  
**Datasets**  
<https://data.cityofchicago.org/Transportation/CTA-Ridership-L-Station-Entries-Daily-Totals/5neh-572f/about_data><https://data.cityofchicago.org/Transportation/CTA-Ridership-Bus-Routes-Daily-Totals-by-Route/jyb9-n7fm/about_data><https://greatcities.uic.edu/2022/05/16/great-cities-releases-updated-hardship-index-for-chicago-community-areas/><https://catalog.data.gov/dataset/chicago-population-counts><https://www.kaggle.com/code/paultimothymooney/explore-housing-prices-in-chicago-il/input><https://data.cityofchicago.org/Public-Safety/Crimes-2022/9hwr-2zxp/data>**Problem Definition/Goals**  
Is there a way that we can optimize urban development in Chicago by using more up-to-date CTA, Pace, demographic, and socioeconomic information to identify patterns to optimize transportation and forecast future developments?

*Current Concerns/Identified Issues*  
Due to the city’s demographic and socioeconomic changes over time, the efficiency and accessibility of public transportation may not be as reliable and optimal as it once was and has the potential to be.  
  
**Shared Interest in Data and Problem Exploration**  
Analyzing CTA data allows us to contribute to improving public transportation services in Chicago. By identifying patterns, optimizing routes, and understanding rider behavior, we can help make the transit system more efficient and user-friendly.

**Research Questions/Queries**

1. Can we identify opportunities for public transportation optimization based on our findings?
   1. What do those opportunities look like?
2. How do we locate and identify areas that would benefit from development?
3. How does the implementation of new CTA stations affect property values in surrounding neighborhoods, and what factors contribute to variations in these impacts?
4. What conclusions/ideas can we draw from other examples that would supplement/advocate for our findings?
   1. Precedents: The 78, Lathrop Homes Redevelopment