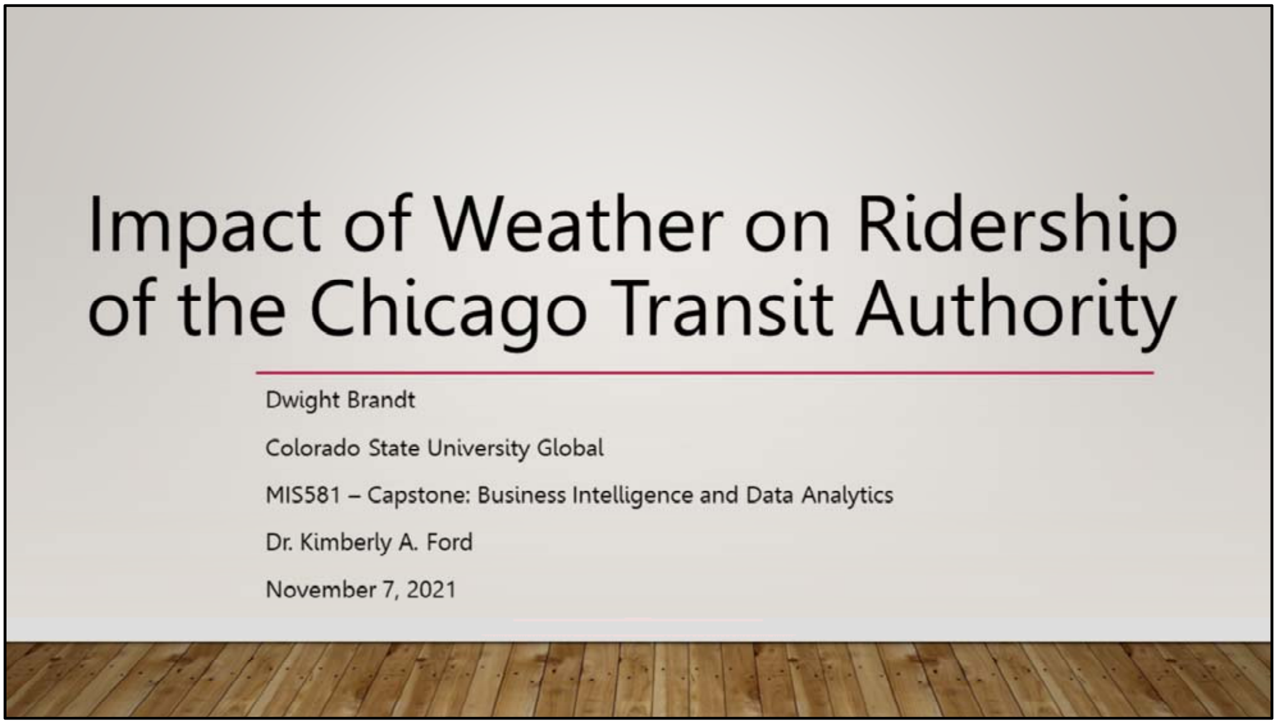
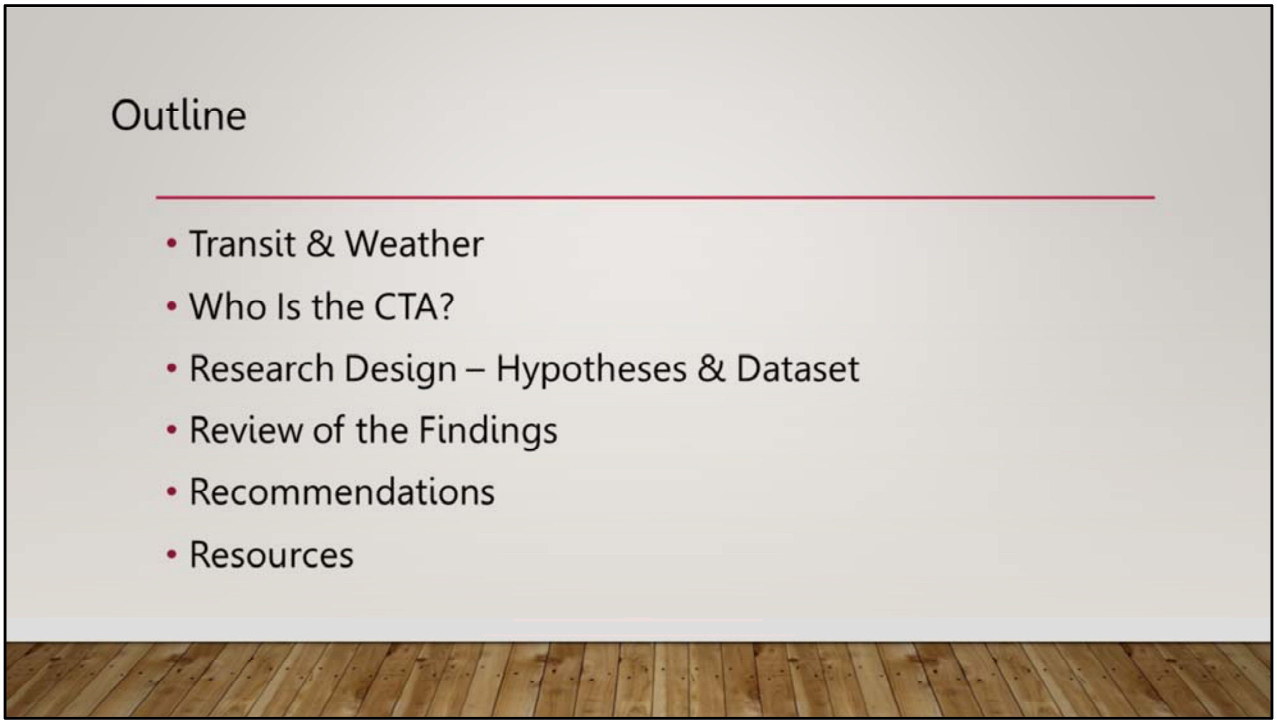


Impact of Weather on Ridership of the Chicago Transit Authority



In this presentation, several topics will be covered. To begin, I will provide a brief introduction on how weather and transit are linked.

Next, some points on the Chicago Transit Authority's, also referred to as "the CTA" background will be noted. Then, I will dig deeper into the study itself. The first element will involve the research design. This will include the hypotheses developed and the dataset used in the analysis.

After that, I will outline some of the key findings of the research. Lastly, recommendations will be suggested.

At the end, I will provide you with where to find information on the project itself, including the dataset and the research paper.

Let us get started.

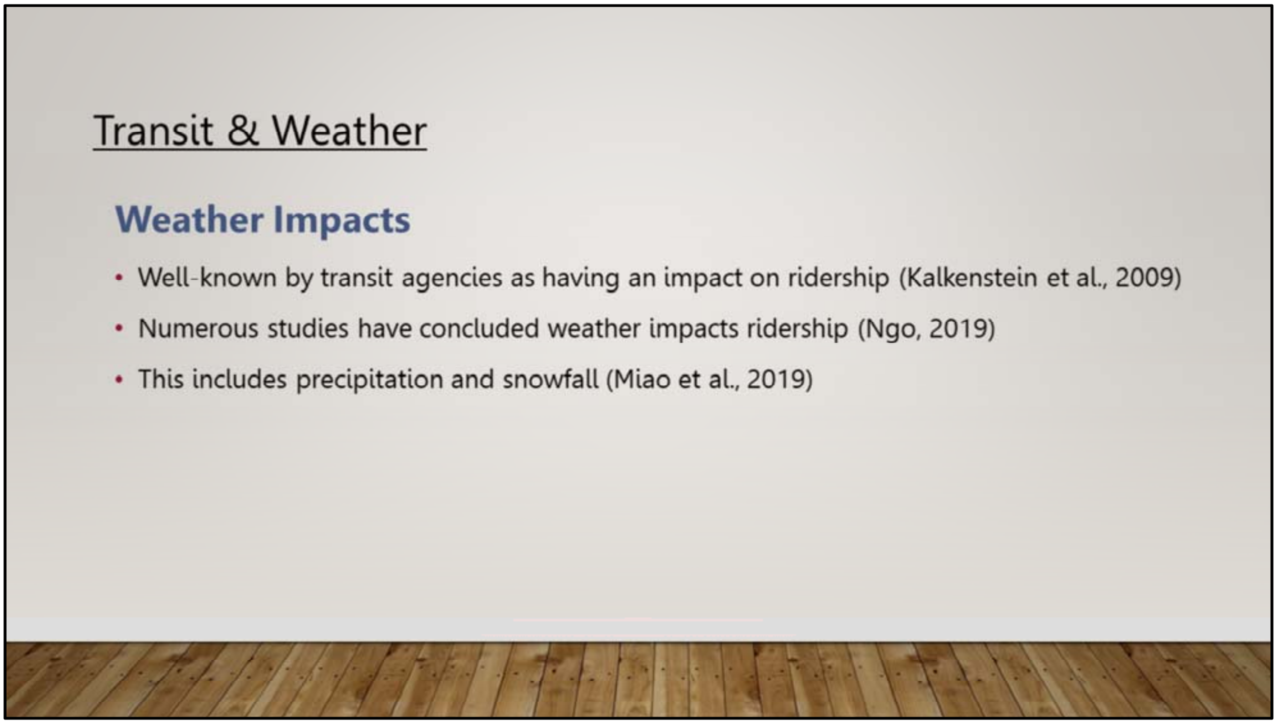
(next slide)

Humanity has looked towards transportation as a means to get from one place to another throughout much of history. Over time, the modes have expanded to include such high-volume methods as bus and rail, capable of moving large numbers of people. Referred to as mass transit systems, they play an important role especially in urbanized settings (Farmer & Noonan, 2014) and to general economic growth of a region (Vajjarapu et al., 2020).

One component that is directly related to the success of any mass transit system is ridership, since often a large share of a budget that funds a mass transit system comes directly from those paying for the services in which they ride (Merlin et al., 2021). There are a multitude of factors that can impact ridership including the price of fuel, vehicle ownership, and accessibility to transit centers (Kalkstein et al., 2009).

One major disruptor that has had a large negative impact on travel in general has been the COVID-19 pandemic. Parker et al. (2021) note that results of a survey indicated that 75% of participants stated a reduction in transit usage in the United States since the start of the pandemic. A reduction of this magnitude would likely have a substantial impact on any transit agency reliant on revenue from its riders.

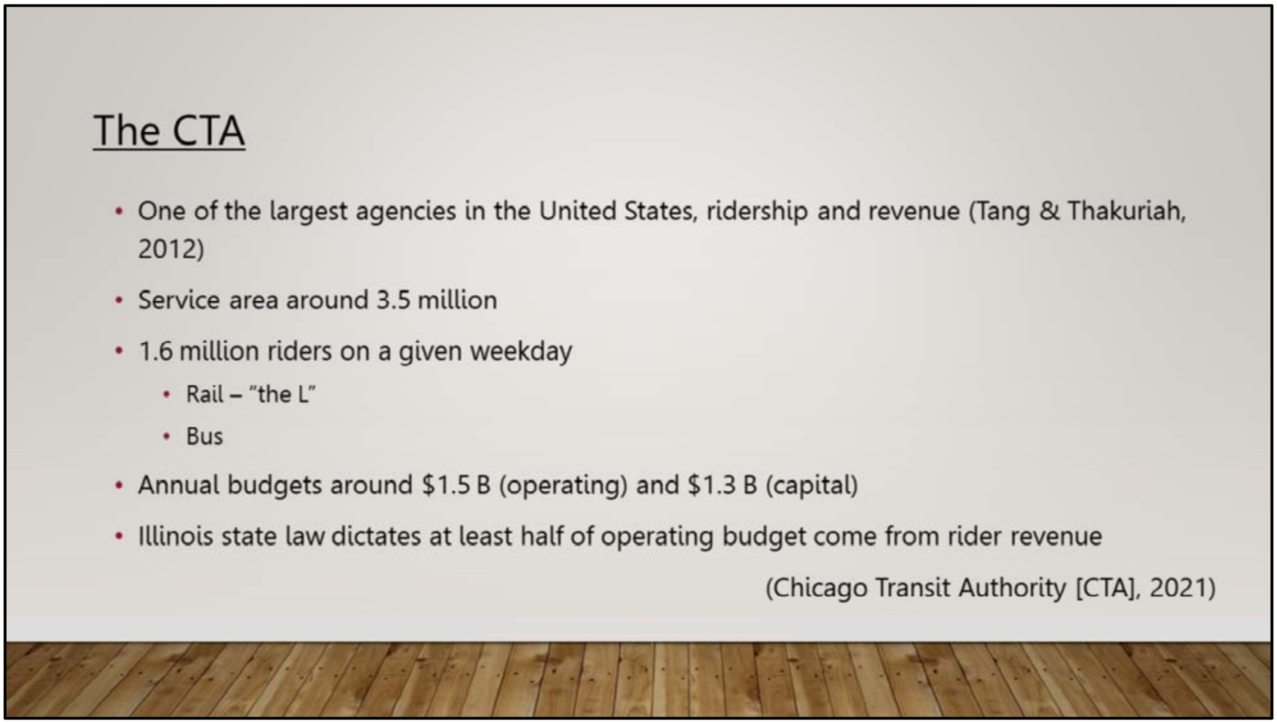
(next slide)



Weather has been long noted as a factor that impacts transit ridership by those overseeing transit agencies (Kalkstein et al., 2009). There have been numerous studies that have looked at the relationship between transit ridership and weather (Ngo, 2019). For example, riders are not wanting to walk or wait in snow or heavy rain for public transit, and as a result, this can reduce ridership (Miao et al., 2019). The general consensus of these studies concluded that weather can negatively impact transit ridership.

Now, let us go into the topic of this research, the CTA.

(next slide)



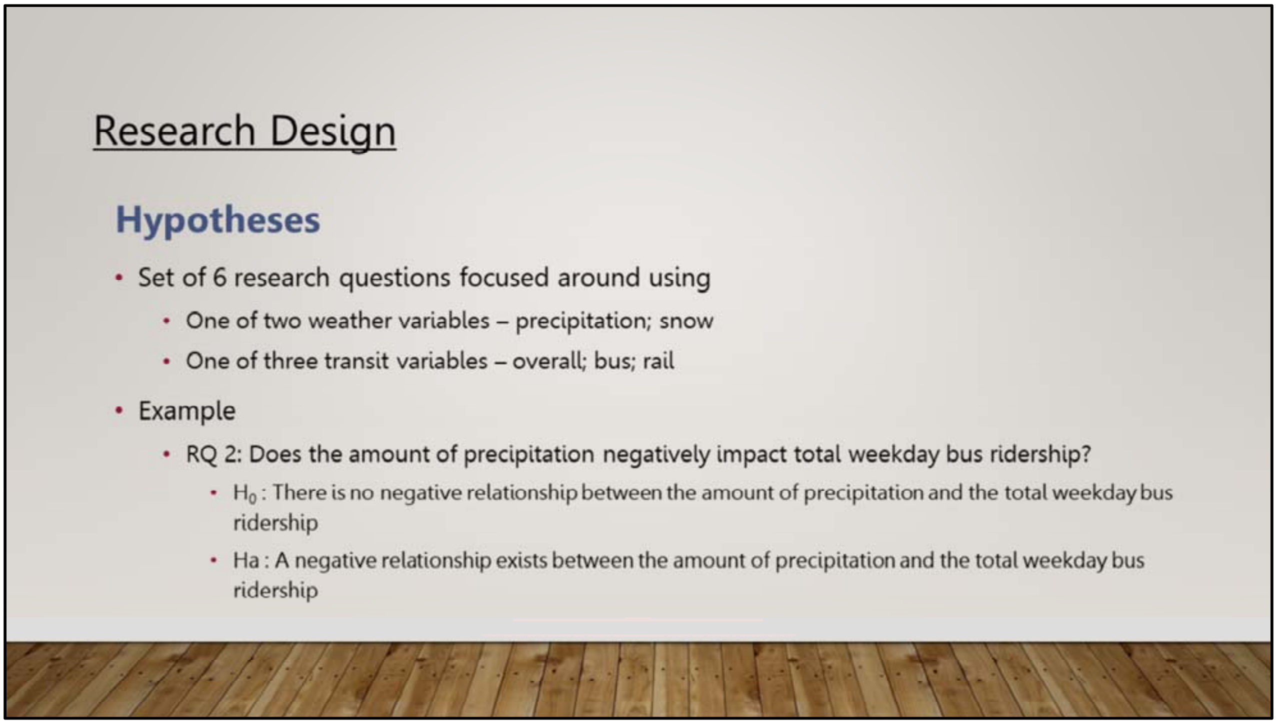
Formed in the 1940s, the CTA today is among the largest mass transit agencies in the United States both in revenue and ridership, and it has grown over its existence to include not only the city of Chicago, but also the around three dozen bordering municipalities of the Greater Chicago Area (Tang & Thakuriah, 2012), amounting to a total service population of 3.5 million (Chicago Transit Authority [CTA], 2021). On any given weekday, around 1.6 million riders take the L or a bus on its system, with ridership about evenly split between the two modes.

As an organization, there are close to 10,000 employees on its payroll, positions that are funded through and included in its annual operating budget of over $1.5 billion. In addition, the CTA has an annual capital budget of close to $1.3 billion (CTA, 2021). An Illinois state law requires the CTA to recover a minimum of one half of its operating budget from ridership revenue (CTA, 2021). Such an organization of this size, with its large budgets and number of employees, would likely be impacted by ridership swings, not only from substantial impacts of a pandemic, but perhaps weather-related impacts as well, adding to the constraints of the system.

Now thinking of the CTA, how does weather impact ridership on its system? This is the premise of the research project.

We will now transition in this presentation, to get into the details.

(next slide)

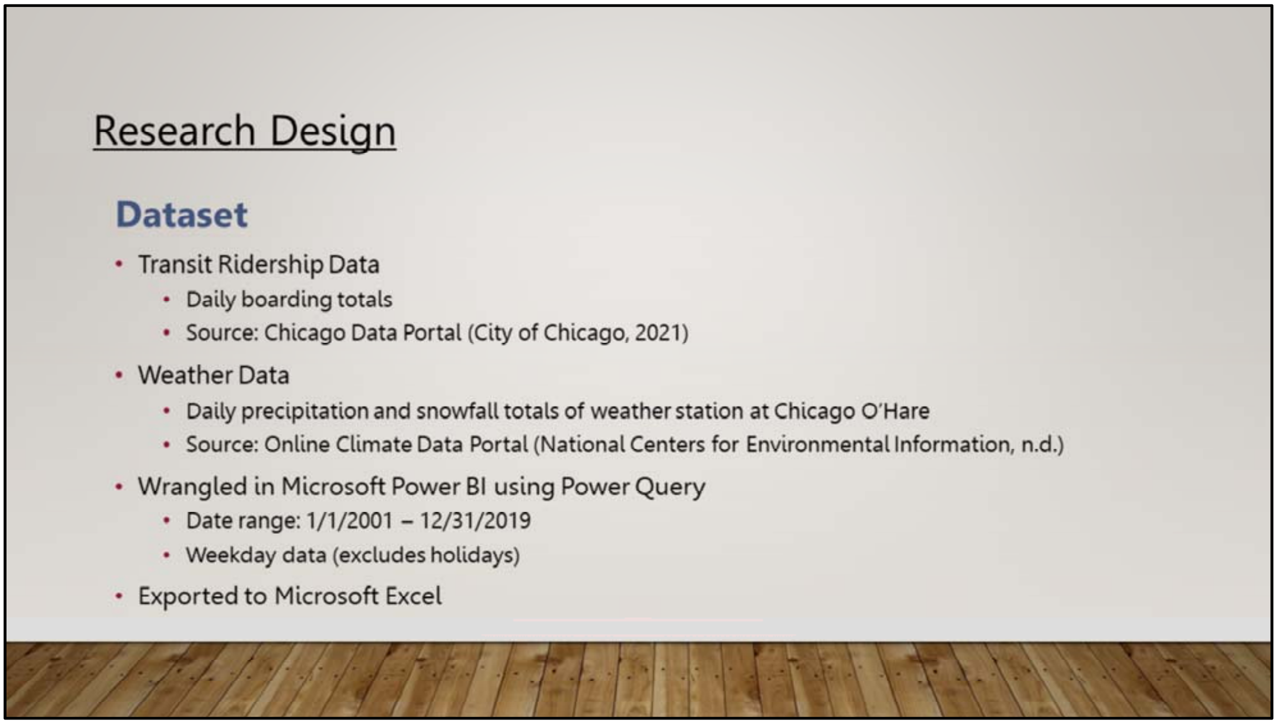


There were a total of six research questions (RQs) developed as part of this project, each with a null and alternative hypothesis. Each RQ was focused on one of the two weather variables, precipitation or snowfall, and one of the three transit variables, overall ridership or bus or rail. The analysis focused in on weekday transit ridership. This is because Li et al. (2018) indicate that past research points to weekday transit ridership as being more concrete and solid during times of increased weather like snow or rain, as opposed to the weekends where behaviors in riders are more determined by leisure activities that include shopping or attending sporting events.

Most of the RQs noted a negative relationship in the hypotheses, save for RQ1 being neutral and RQ3 being positive, the latter of which looked at the relationship between precipitation and rail ridership. The thought here is that with more precipitation, roads could be more congested delaying buses, but not rail service, so there could potentially be more riders using that mode.

Now that we have the sets of hypotheses defined, let’s move on to see how the dataset was developed.

(next slide)



Two data sources were used to create the single dataset used for this project. The transit ridership data was sourced from the Chicago Data Portal (City of Chicago, 2021). This data noted daily boarding totals, reflected as an aggregation of these transactions, going back to 2001. This was broken down by mode in addition to the overall total. The weather data was sourced from the Online Climate Data Portal of the National Centers for Environmental Information (n.d.). A dataset search for daily precipitation and snowfall totals at the weather station for Chicago O'Hare International Airport (ID USW00094846) was conducted for the time period of January 1, 2001 through December 31, 2019. This particular weather station was selected due to its central geographic location in the greater Chicago area.

These separate datasets were then pulled into Microsoft Power BI, and using the Power Query capabilities, were combined into a single dataset. The final dataset reflected a date range to match that of the weather data and excluded weekend and holidays. It was then exported to Microsoft Excel.

(next slide)

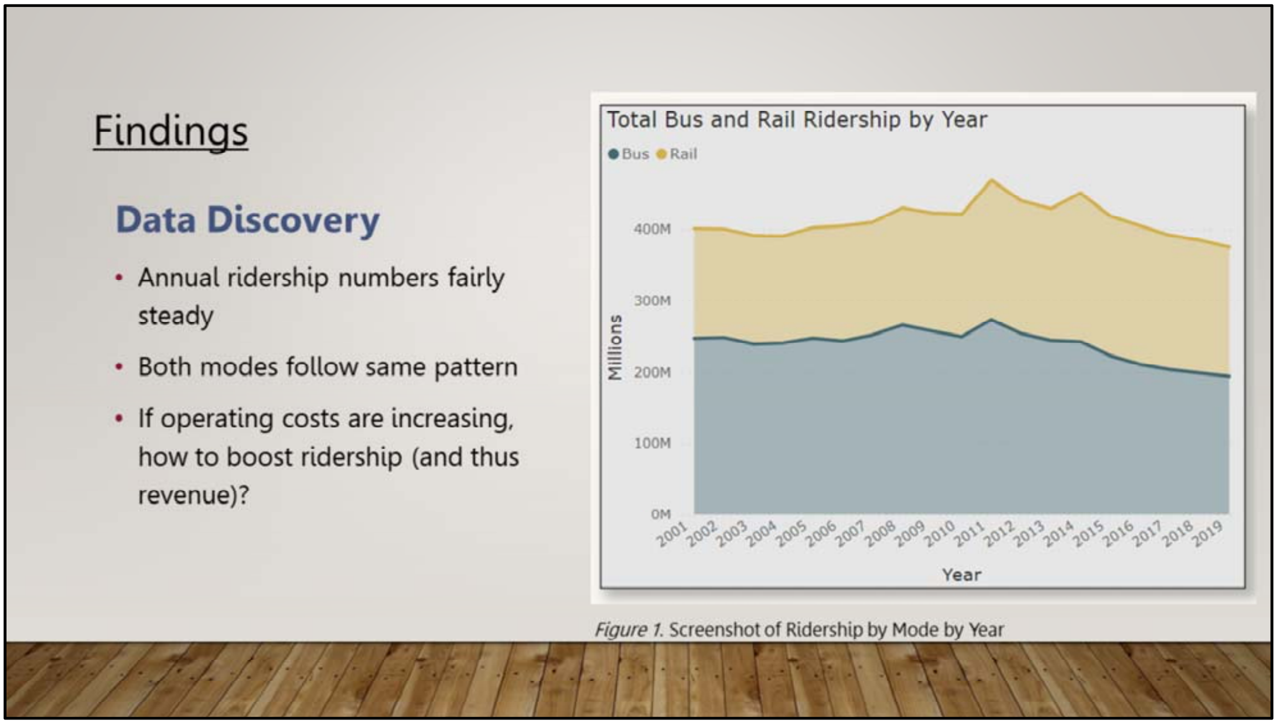


Figure 1. Screenshot of Ridership by Mode by Year

Before we jump into the findings themselves, I would like to call out that I did some initial data discovery to see if there might be anything of interest in the data. On this slide we can see how ridership numbers of each mode are performing annually.

As the graph (Figure 1) shows, the numbers for the most part have held steady since 2001, with a slight downward trend since 2011. Both modes appear to mirror the ups and downs. One thought on this graph is that it is likely operating costs for the CTA have been going up year after year. If ridership numbers are not increasing, what mechanisms could there be to boost ridership, and thus, boost revenue?

(next slide)

Findings

Data Discovery

- Breaking data into days of the week
- Fairly consistent
- Monday could be lower due to number of holidays falling on that day

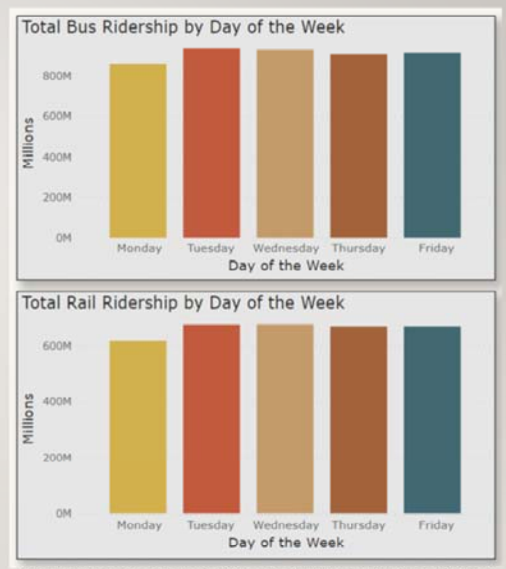


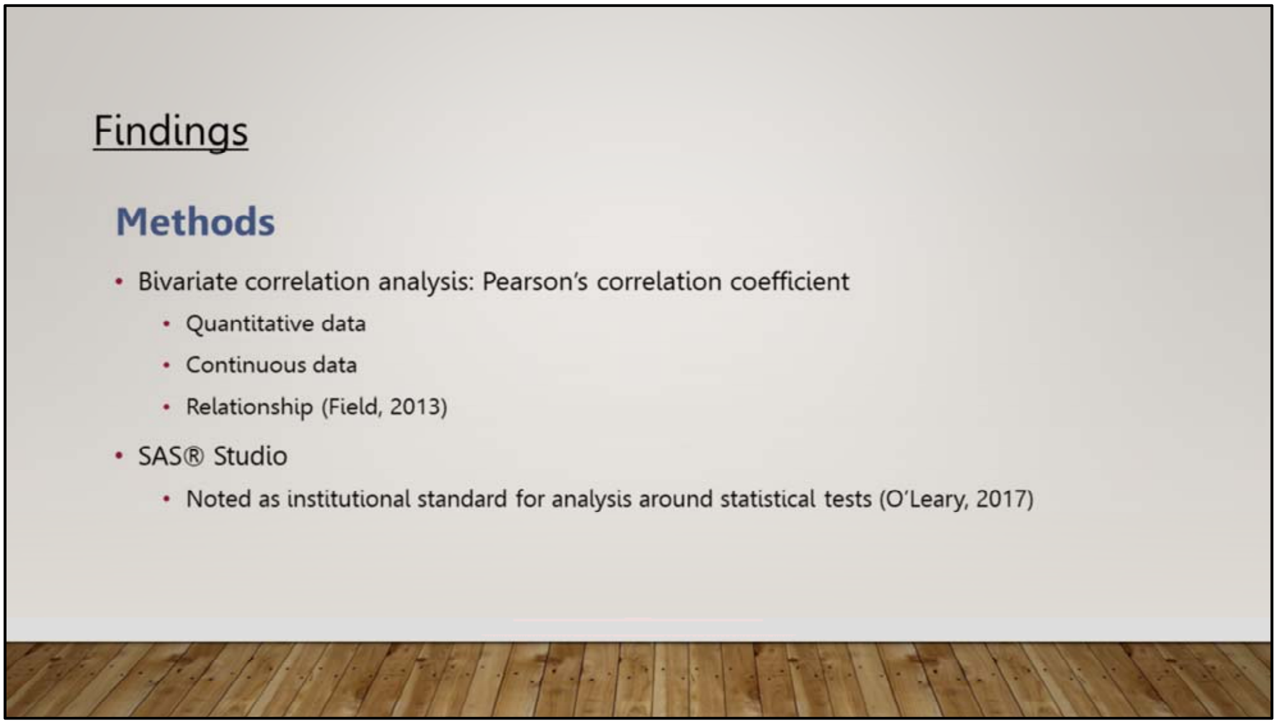
Figure 2. Screenshot of Ridership by Mode by Day of Week

In addition to looking at the data year over year, another thought that came to mind was whether one day of the week in particular had any more ridership than another. Recall earlier that Li et al. (2018) indicated that past research points to weekday transit ridership as being more concrete and solid.

As we can see, for the most part, ridership numbers are fairly consistent among the days of the week and between the modes of travel taken. Keep in mind that Monday could likely be lower than the rest of the days since many holidays fall on a Monday.

We will now move into the findings on the next slide.

(next slide)



Before we jump into the results of the hypothesis tests, I would like to cover briefly the methods used. The dataset variables are both quantitative and continuous in nature. As the hypotheses focus on a relationship between two variables, Pearson's correlation coefficient ("Pearson's r") was used, as Field (2013) suggests. To conduct these statistical tests, SAS Studio, through the SAS OnDemand for Academics platform, was used. O'Leary (2017) notes SAS as being the institutional standard for analysis around statistical tests.

(next slide)

Findings

Overall Ridership

- Results of each correlation test show significance
- Variables are weakly and negatively correlated
- As amount of precipitation/snowfall increases, ridership numbers decrease

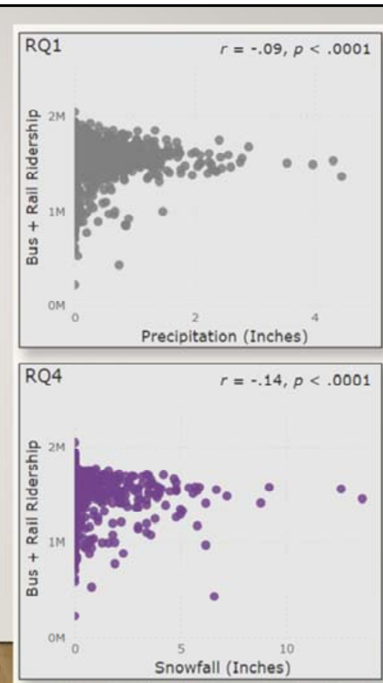


Figure 3. Screenshot of Scatterplots for RQ1 and RQ4

In looking at the results, we will first look at RQ1 and RQ4, which analyzed the combined bus and rail ridership and each respective weather variable. As the graph (Figure 3) shows, the results indicate that each of the respective weather variables is significantly related to the overall total weekday ridership, though weakly correlated. As the amount of precipitation or snowfall increases, the ridership numbers decrease.

(next slide)

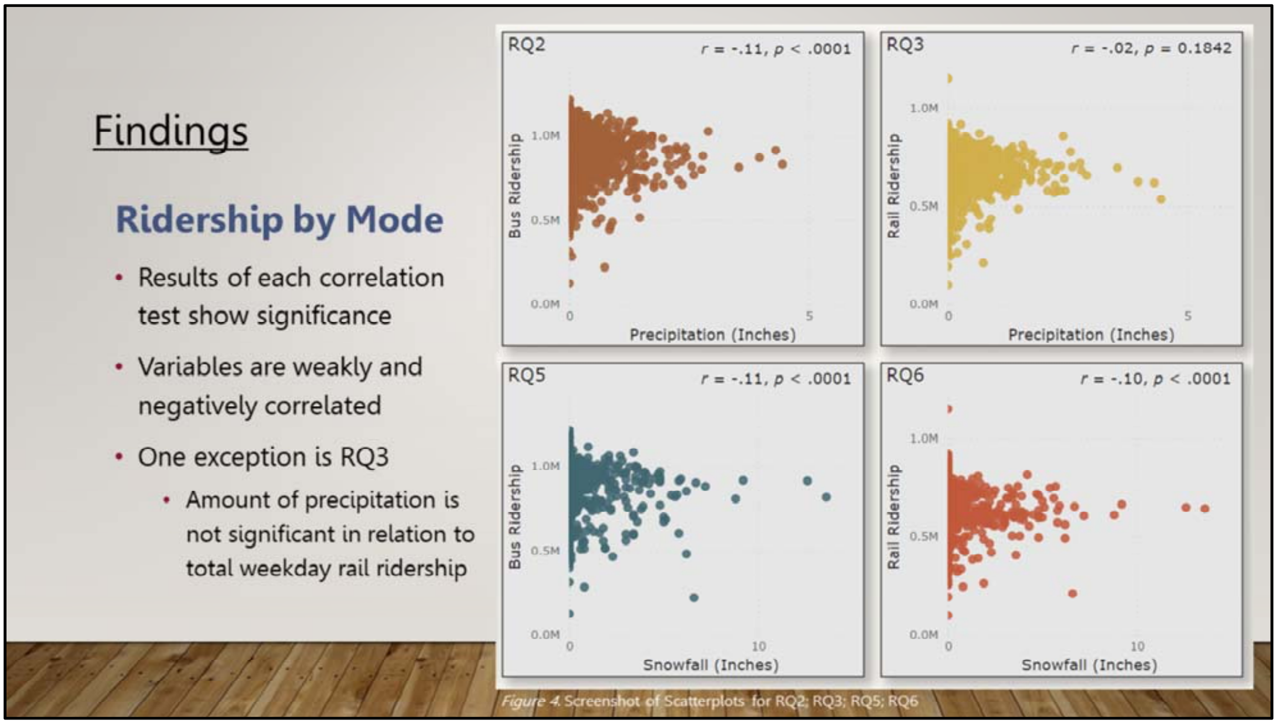


Figure 4. Screenshot of Scatterplots for RQ2; RQ3; RQ5; RQ6

This slide looks at each mode of transit broken out. The graph on this slide (Figure 4) shows the results of each correlation test for the RQ noted. The set here looks at each mode of transit with each type of weather. Like the results of RQ1 and RQ4 that looked at overall ridership regardless of mode, the results shown here also indicate that each of the respective weather variables is significantly related to the weekday ridership numbers of each respective mode, though weakly correlated. And like the results of RQ1 and RQ4, this indicates that as the amount of precipitation or snowfall increases, the ridership numbers decrease.

Now, there is one except to this. If you can cue in on RQ3, which is the top right graph. This correlation test looked precipitation with rail ridership. Here, the amount of precipitation is not significant in relation to the total weekday rail ridership, r = -.02, p = .1842. Because the *p-value* is not less than the significance level of 0.05, the null hypothesis is not rejected and the alternative hypothesis is not accepted.

(next slide)

Findings

Overall Results

- Results of each correlation test show significance
- Variables are weakly and negatively correlated
- As amount of precipitation/snowfall increases, ridership numbers decrease
- One exception is RQ3
  - Amount of precipitation is not significant in relation to total weekday rail ridership

RQ	Variables	Relationship Noted in Ha	p-Value	Correlation	H_0 Result
1	Ridership_Total_All Precipitation_Inches	Neutral	<.0001	-.09	Rejected
2	Ridership_Total_Bus Precipitation_Inches	Negative	<.0001	-.11	Rejected
3	Ridership_Total_Rail Precipitation_Inches	Positive	0.1842	-.02	Not Rejected
4	Ridership_Total_All Snowfall_Inches	Negative	<.0001	-.14	Rejected
5	Ridership_Total_Bus Snowfall_Inches	Negative	<.0001	-.11	Rejected
6	Ridership_Total_Rail Snowfall_Inches	Negative	<.0001	-.10	Rejected

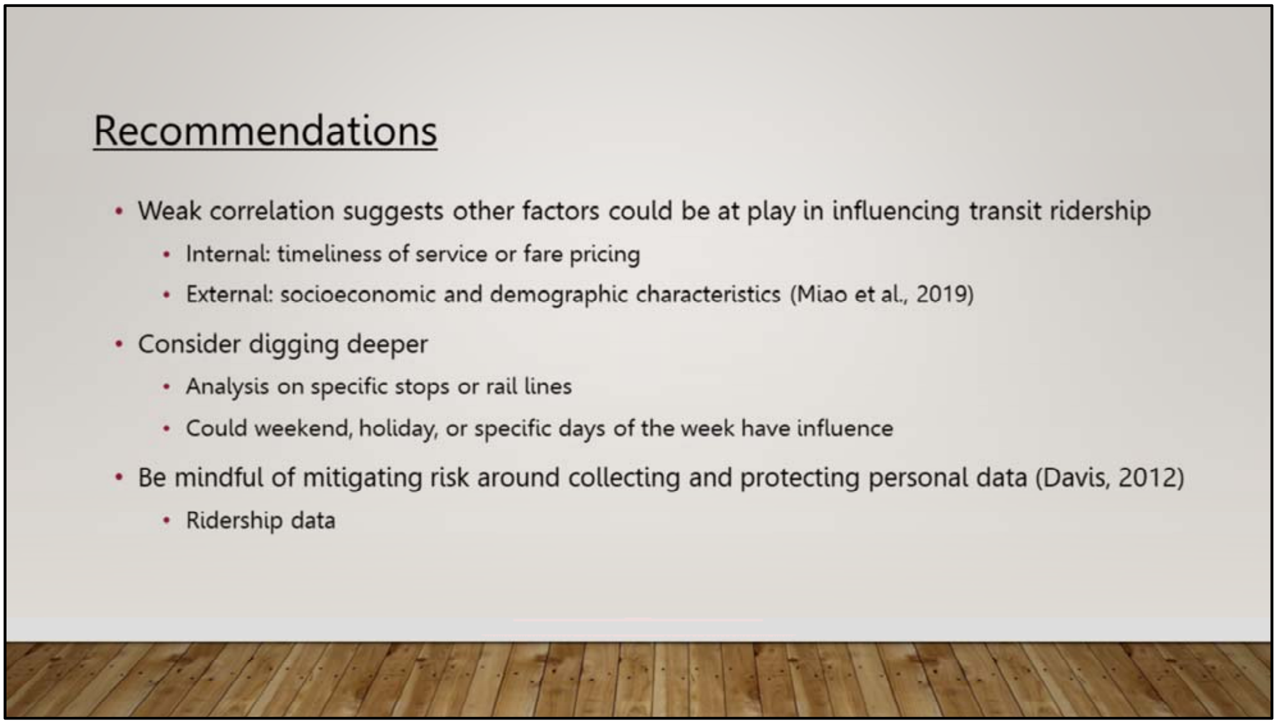
Table 1. RQ Hypothesis Comparison

In looking at all the results compiled into a single table (Table 1), we can see that as mentioned on the previous slides that overall, the results of each correlation test show significance and that the variables analyzed are weakly and negatively correlated. This means, that as the amount of precipitation or snowfall increases, ridership numbers decrease. Again, the one exception was RQ3 that tested the relationship between precipitation and rail ridership.

This overarching assessment aligns with past studies that have researched the impact of weather on other mass transit systems across the globe and have generally come to the same conclusion.

So, how can the CTA use this knowledge?

(next slide)

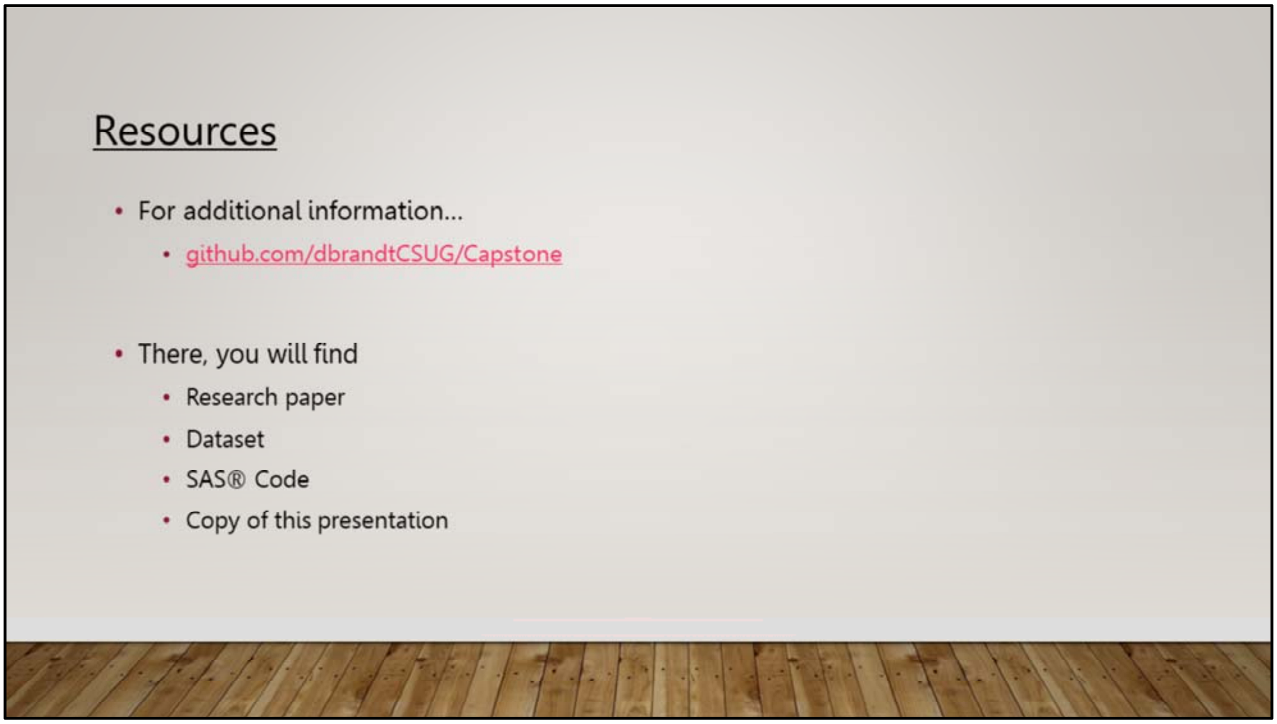


The weak correlation between weather and transit ridership suggests that other factors could be at play that may be influencing ridership numbers. As Miao et al. (2019) points out, there are internal factors such as the timeliness of service and fare pricing, as well other external factors in addition to weather like socioeconomic and demographic characteristics that can influence public transit ridership.

Another consideration may be to dig deeper by conducting analysis at specific stops or on specific rail lines, to better understand whether weather and these additional factors could have a more granular impact. While this study only analyzed weekday ridership data, there may be an opportunity to analyze the entire dataset to understand how weekends, holidays, or specific days of the week have any insight that can be provided.

While there can be absolute value in data to an organization for opportunities like innovation as well as social benefit from the analysis and subsequent reporting of insight, it remains nonetheless important though to value the privacy of information like that of a rider. Thus, the organization should allow for the mitigation of risk around collecting and protecting data (Davis, 2012), including personally identifiable information like that of a rider’s name.

(next slide)

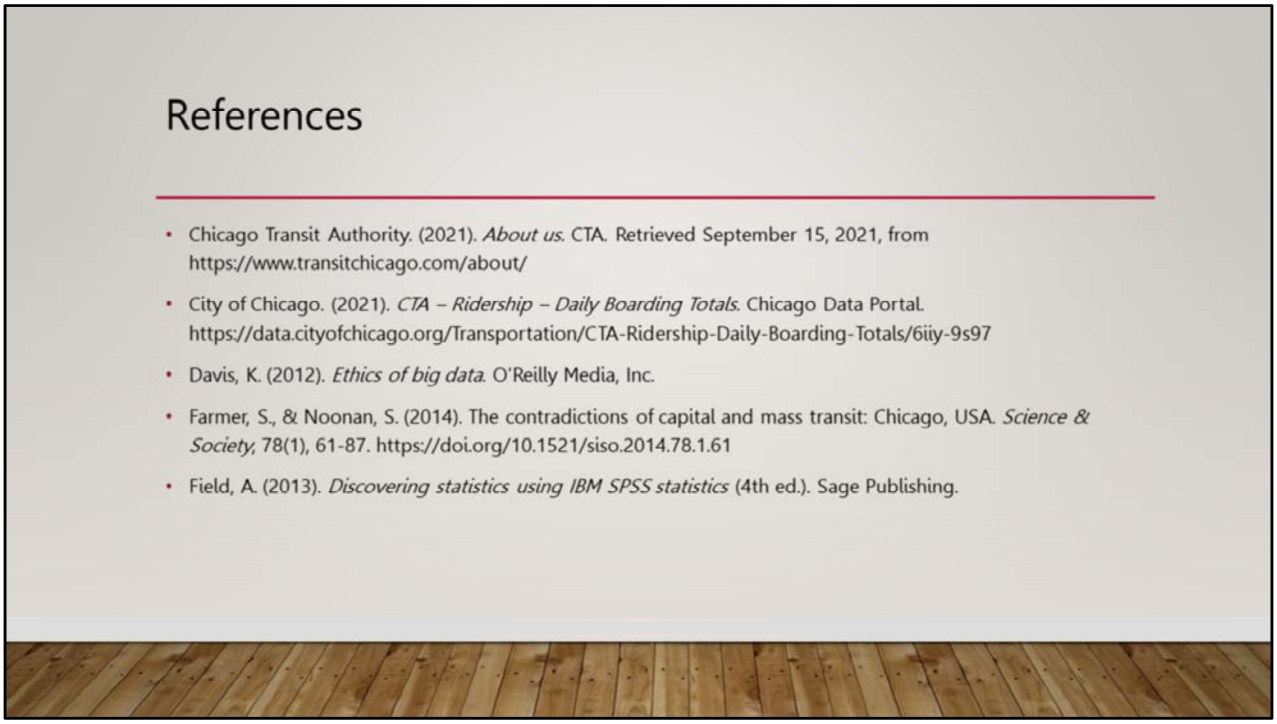


Thank you for watching my presentation. I hope you enjoyed it.

For additional information, please visit my GitHub repository, which can be found at the link on the screen (www.github.com/dbrandtCSUG/Capstone.

There you will find all the materials for my Capstone project, including my research paper, the dataset, SAS® code, and a copy of this presentation.

Have a wonderful day!


References
• Chicago Transit Authority. (2021). *About us*. CTA. Retrieved September 15, 2021, from https://www.transitchicago.com/about/
• City of Chicago. (2021). *CTA – Ridership – Daily Boarding Totals*. Chicago Data Portal. https://data.cityofchicago.org/Transportation/CTA-Ridership-Daily-Boarding-Totals/6iiy-9s97
• Davis, K. (2012). *Ethics of big data*. O'Reilly Media, Inc.
• Farmer, S., & Noonan, S. (2014). The contradictions of capital and mass transit: Chicago, USA. *Science & Society*, 78(1), 61-87. https://doi.org/10.1521/siso.2014.78.1.61
• Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage Publishing.

References

References

- Kalkstein, A. J., Kuby, M., Gerrity, D., & Clancy, J. J. (2009). An analysis of air mass effects on rail ridership in three US cities. *Journal of Transport Geography*, 17(3), 198–207. https://doi.org/10.1016/j.jtrangeo.2008.07.003
- Li, J., Li, X., Chen, D., & Godding, L. (2018). Assessment of metro ridership fluctuation caused by weather conditions in Asian context: Using archived weather and ridership data in Nanjing. *Journal of Transport Geography*. 66. 356–368. https://doi.org/10.1016/j.jtrangeo.2017.10.023
- Merlin, L. A., Singer, M., & Levine, J. (2021). Influences on transit ridership and transit accessibility in US urban areas. *Transportation Research Part A: Policy & Practice*, 150, 63–73. https://doi.org/10.1016/j.tra.2021.04.014
- Miao, Q., Welch, E. W., & Sriraj, P. S. (2019). Extreme weather, public transport ridership and moderating effect of bus stop shelters. *Journal of Transport Geography*, 74, 125–133. https://doi.org/10.1016/j.jtrangeo.2018.11.007
- National Centers for Environmental Information. (n.d.). *Climate Data Online Search*. NCEI. Retrieved September 16, 2021, from https://www.ncdc.noaa.gov/cdo-web/search

References

References

- Ngo, N. S. (2019). Urban bus ridership, income, and extreme weather events. *Transportation Research Part D*, 77, 464–475. https://doi.org/10.1016/j.trd.2019.03.009
- O'Leary, Z. (2017). *The essential guide to doing your research project.* (3rd ed.). Sage Publishing.
- Parker, M. E. G., Li, M., Bouzaghrane, M. A., Obeid, H., Hayes, D., Frick, K. T., Rodríguez, D. A., Sengupta, R., Walker, J., & Chatman, D. G. (2021). Public transit use in the United States in the era of COVID-19: Transit riders' travel behavior in the COVID-19 impact and recovery period. *Transport Policy*, 111, 53–62. https://doi.org/10.1016/j.tranpol.2021.07.005
- Tang, L., & Thakuriah, P. V. (2012). Ridership effects of real-time bus information system: A case study in the City of Chicago. *Transportation Research Part C*, 22, 146–161. https://doi.org/10.1016/j.trc.2012.01.001
- Vajjarapu, H., Verma, A., & Allirani, H. (2020). Evaluating climate change adaptation policies for urban transportation in India. *International Journal of Disaster Risk Reduction*, 47. https://doi.org/10.1016/j.ijdrr.2020.101528

References