Technical Appendix
Public Transit Fare Reinstatement Assessment for King County

1. Data Overview

This report uses data provided by King County Metro Transit and the American Community Survey (Census Tract Data). From King County Metro, the report takes a sample of overall system ridership (5%) with counts for passengers boarding and departing (alighting) at metro stops across the county between September 1, 2020 and October 31, 2020.

Part 1 - Time Frame: Due to substantial changes in the metro system that may have affected ridership prior to September 19, this analysis only uses data starting on September 19, 2020. In addition, the analysis uses one month worth of data in the post period instead of matching the number of days in the pre period because it was determined that more data overall would improve the analysis. To support this decision, and as described in the body of the memo, a delay in rider response occurred two weeks after the fare reinstatement. Furthermore, no census tract data was used for part 1 of the analysis.

Limitations: The accessible data covers a short period of time; limited data makes it difficult to conduct a proper analysis to determine the fare reinstatement's true impact on ridership.

Part 2: Since our assessment sought out to focus on different impacts on riders from different socioeconomic levels, we identified median household income within the American Communities Survey as our variable of interest. To measure the socioeconomic impacts of fare reinstatement in King County, this report merged the data from the American Communities Survey with the King County Metro Transit data to be able analyze ridership trends at the census-tract level. To do this, we used the available shape files of bus stops within King County, which contained the locations of each stop and matched them within their recorded census tracts. This enabled us to use the American Communities Survey data as a proxy for the socioeconomic and demographic characteristics of riders boarding and departing buses at stops within census tracts.

Limitations: Data for various confounding factors affecting ridership were not available. Alternative factors, especially Pandemic related reasoning like fear or remote working options, likely also affected residents' willingness/need to use this particular mode of transportation. In addition, since no individual-level socioeconomic or demographic data was collected on bus riders in King County, it cannot be assumed that all riders that were counted as boarding and departing at a given stop truly live within that census tract. In order to improve the accuracy, we created a true ridership variable described below.

Technical Appendix
Public Transit Fare Reinstatement Assessment for King County

2. Ridership

Part 1: Ridership was measured using average alightings (departures) per census tract. The team created a variable called "true ridership" to accurately aggregate ridership to census tracts. True ridership was created by making a dummy variable for individuals who exited the metro between 3pm and 8pm. Our assumption is that individuals exiting the metro at this time are returning from work or returning from their daily activities to their homes. To avoid double counting, we focused only on alightings and not boardings. After finding the alightings per stop, we aggregated the data by census tract and date, then found the average number of true alightings per census tract. This data is represented in figure 1 of the memo.

Part 2: To answer the second question of whether the fare reinstatement affected certain socioeconomic groups, we separated median household income into quintiles with each group representing 20% of the data. Census tracts were separated by Low-Income (below $90,179), Medium-Low Income ($90,170 - $111,705), Medium Income ($111,705-$125,313), Medium-High Income ($125,313 - $152,969), and High Income (above $152,969). We then found average daily ridership by taking daily average true alightings per income group. These findings are represented in figure 2 of the memo.

3. Data Adjustments

Outliers: Weekend dates were excluded from our analysis due to the variability in weekend ridership and the report's focus on the economic impacts of metro ridership in King County. Within our analysis, there were 13 weekend days. Particular weekend days showed extreme levels of ridership at a particular stop on certain Sundays. Based on correlations with the Seattle Seahawks' schedule home games, we hypothesize that these increases are due to their football games occurring on these Sundays. These types of irregularities and the increased economic implications of higher weekday ridership led the team to exclude weekend days from our ridership analysis. There was no exclusion of any outliers during our weekday analysis.

Missing & Incomplete Data: To answer part one, regarding overall ridership, the analysis used data from all bus stops regardless of whether there was census-tract demographic data linked to the stop or not. However, for part two's analysis of socio-economic trends, we did exclude bus stops that did not have census tract demographic data from the American Communities Survey linked to the bus stop. This limited the amount of stops included in our analysis, but provided more accurate information with regards to our socioeconomic analysis.