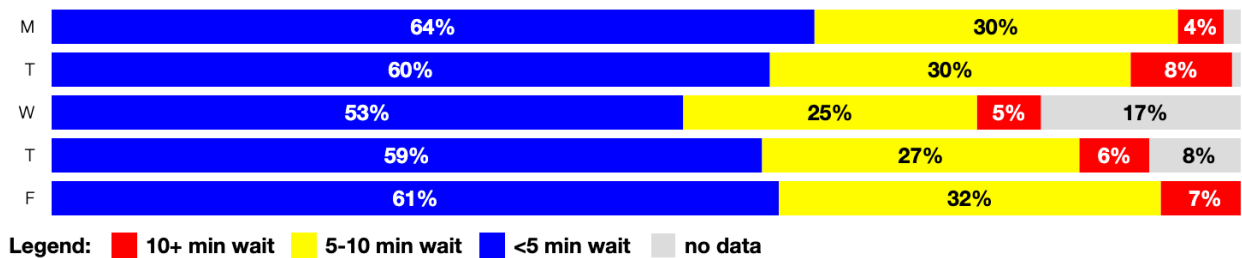**05.07.19**
Jeremy Steele
Work and Interests

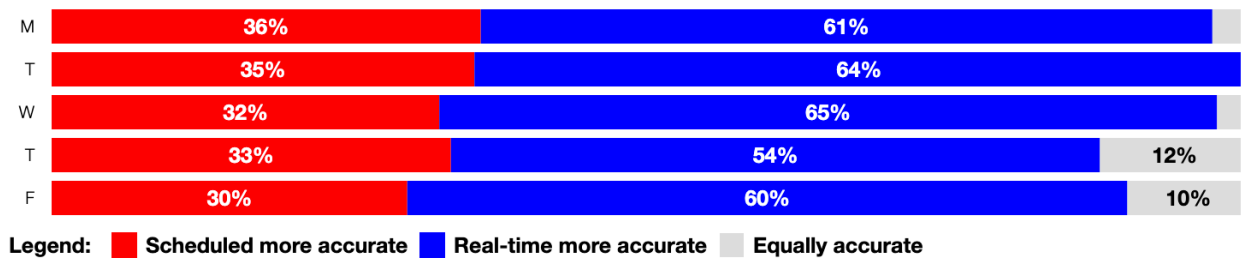**Transit Data Analysis and Visualization**

Bus Reliability on my Home Route

This small project was borne out of frustration with my bus route to campus. In theory, it runs every five to seven minutes on weekdays between 6am and 9pm, but my experience riding it for two years did not agree with this. In the absence of access to GTFS-realtime data to back my complaints up, I scraped real-time bus arrivals at my home stop off of the agency's website for a week, from 6am to 9pm. I estimated the real arrivals by interpolating the scraped estimations.

Figure 1 in the presentation provides a robust illustration of the wait time at the stop at each moment during the week. In the morning rush-hour, I rarely have to wait for more than five minutes. However, as demonstrated below, at off-peak periods, the wait time was above ten minutes between 5-10% of the time.

M	64%	30%	4%	
T	60%	30%	8%	
W	53%	25%	5%	17%
T	59%	27%	6%	8%
F	61%	32%	7%	

Legend: ■ 10+ min wait ■ 5-10 min wait ■ <5 min wait ■ no data

The second figure in the presentation shows how often the real-time data itself was accurate during the week. The results are disappointing, with estimations being inaccurate by more than two minutes nearly 20% of the time on Tuesday. Though the benefit over scheduled data is clear, following the schedule instead was still the more accurate option 30-40% of the time throughout the week, as evidenced below.

M	36%	61%	
T	35%	64%	
W	32%	65%	
T	33%	54%	12%
F	30%	60%	10%

Legend: ■ Scheduled more accurate ■ Real-time more accurate ■ Equally accurate

Just recently, I was finally able to access the GTFS-realtime data on the agency's website. The logical continuation of this project is to use the additional data provided by the VehiclePositions feed to fine-tune the interpolated arrivals, and to continue the analysis across all stops on the route to figure out where exactly buses become delayed.

Ridership Along a Redesigned Route

Upon completion of the BART extension to San Jose, the Valley Transportation Authority plans to overhaul its bus network. A major component of the overhaul is the Rapid 523 route, slated to replace and upgrade service on the San Carlos/Stevens Creek and Sunnyvale-Saratoga corridors. In an effort to understand the implications of the planned service change, I looked at existing ridership patterns on routes currently serving the areas.

The figure displayed looks at the most complicated area of service the 523 is slated to serve. As demonstrated, three primary routes connect Cupertino and Sunnyvale: the 53, 54, and 55. The 523 alignment makes use of Sunnyvale-Saratoga road, with existing route 54 planned for deletion. But existing ridership is largely similar on both existing alignments, and given the important distance between the two and the low walkability of the surrounding area, the trade-off does not appear to be overwhelmingly positive.

The other two sections of the route showed more clear-cut benefits to changing the service to a limited-stop model, with the exception of the current route 23 detour serving Valley Fair Transit Center not included in the 523 alignment, as demonstrated below.

23-Only Stops: 2298 b/wd
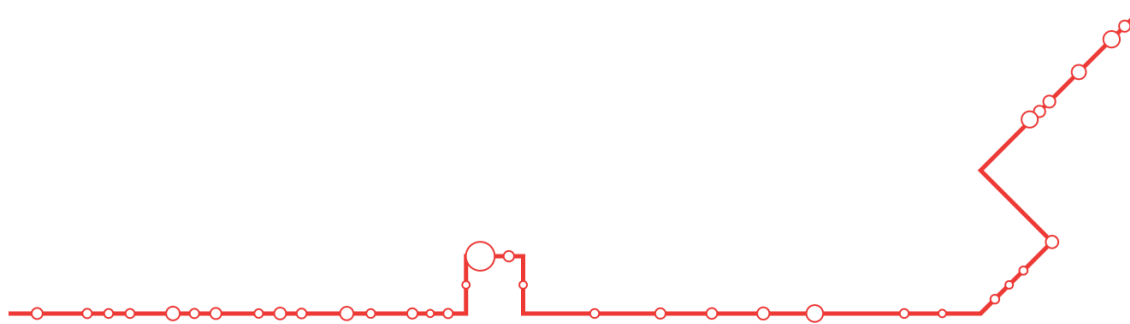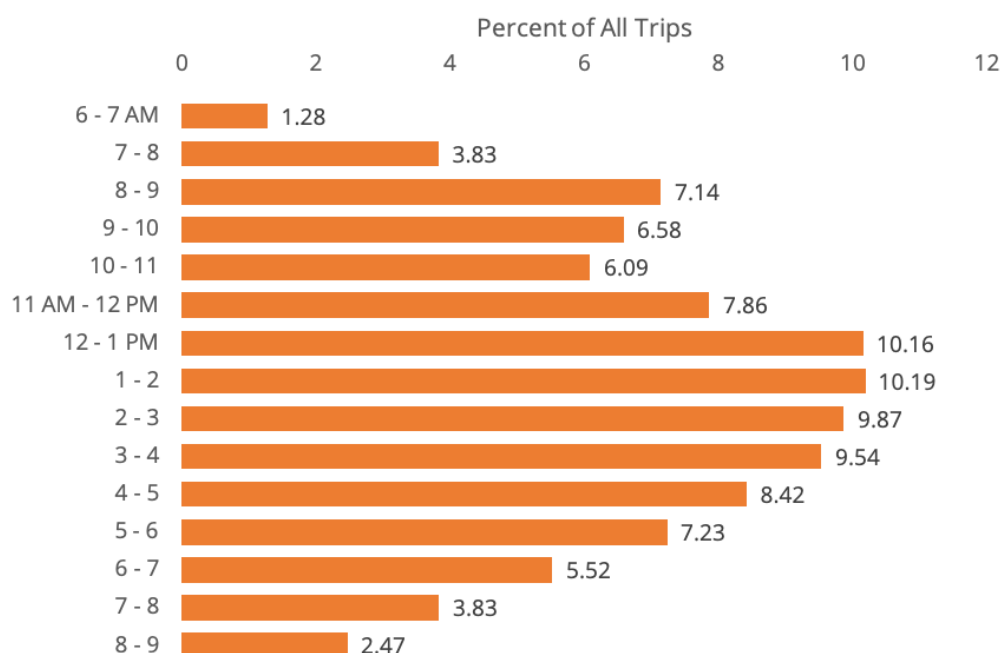

23/323/523 Stops: 3837 b/wd

There are limitations that come with the data source: the ridership numbers are partial, and this analysis only looks at boardings, not alightings. With better ridership data, and this consideration in mind, such a project would better evaluate trade-offs in new service.

**Analyzing New Modes of Transportation
(and occasionally trying to substantiate my skepticism of them)**

Travel Patterns of Scooters in DC

For my undergraduate research project, I was given the rather open-ended task of doing something interesting with the origins and destinations of nearly a million scooter trips over a four-month period in Washington, DC. Given the ridership decline that WMATA is undergoing, I wanted to see if these trips were replacing 'transitable' trips at a significant rate, and if scooter travel patterns could provide planning insight into where transit coverage was lacking.

The results were mixed. First, I found that scooter trips followed a much different temporal pattern than transit trips, as demonstrated below.

Percent of All Trips

Time	Percent
6 - 7 AM	1.28
7 - 8	3.83
8 - 9	7.14
9 - 10	6.58
10 - 11	6.09
11 AM - 12 PM	7.86
12 - 1 PM	10.16
1 - 2	10.19
2 - 3	9.87
3 - 4	9.54
4 - 5	8.42
5 - 6	7.23
6 - 7	5.52
7 - 8	3.83
8 - 9	2.47

I also found that most scooter trips were simply too short to be effectively completed on public transportation. However, I found that a large portion of trips, regardless of length, mirrored frequent transit routes. Below is a map showing the proportion of trips mirrored by frequent transit for each traffic analysis zone (TAZ) in the Washington, DC area.

Metro Stations

Bus Lines

Water

Percent of trips mirrored
by a frequent transit route

None

≤ 20

≤ 40

≤ 60

≤ 80
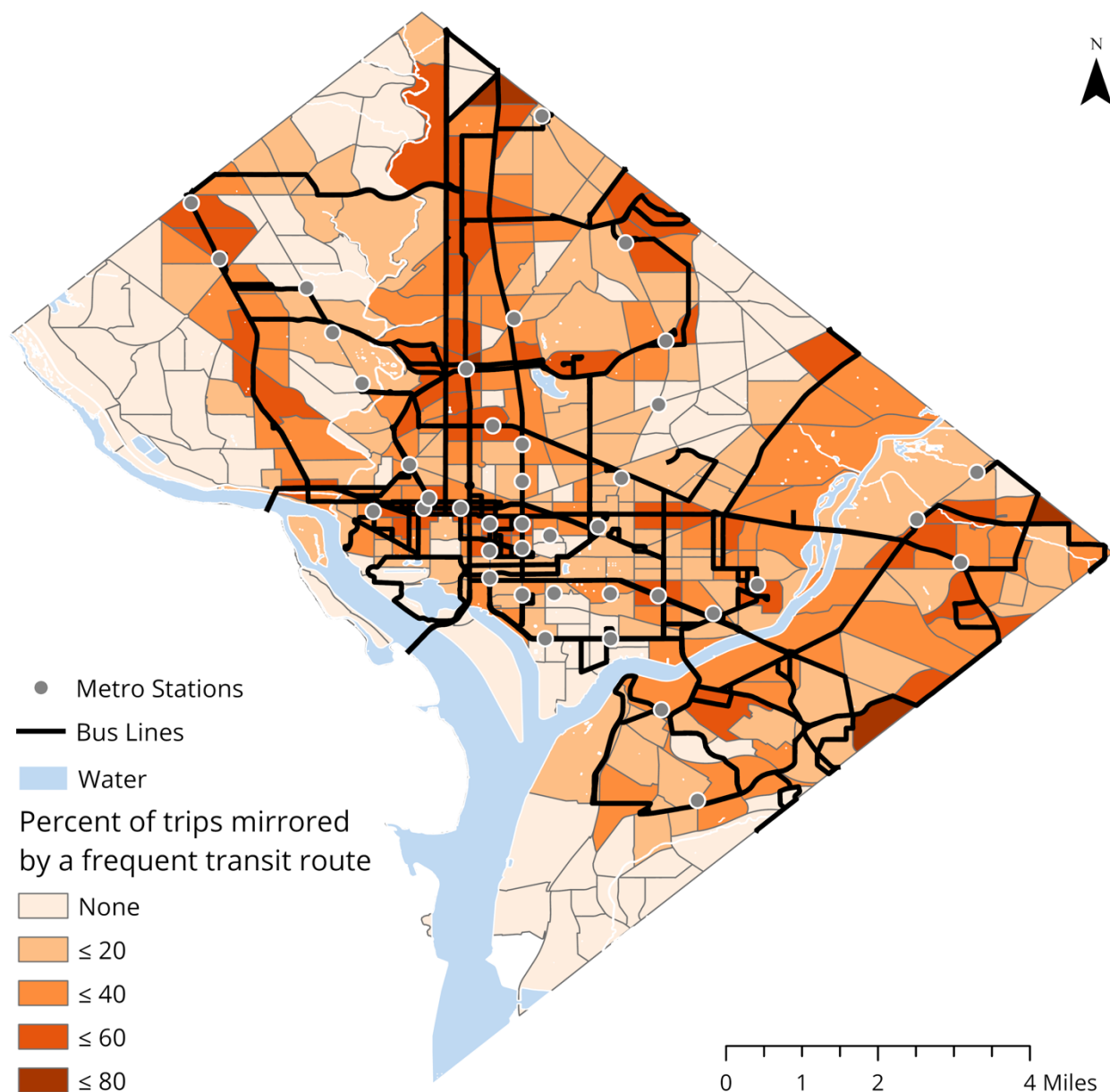
Overall, more than a quarter of trips started and ended within two minutes' walk of the same frequent transit route. Extended to a five minute walk, it becomes just under 65% of trips mirroring a frequent transit route.

Analyzing the results at the TAZ level allowed for an analysis of which areas, and more specifically which corridors, were lacking in transit service. Those areas and corridors with a relatively high number of scooter trips but low mirroring rates could be singled out, as demonstrated in the presentation.

Like the previous project, this one is largely exploratory and suffers from low data resolution. Creating a taxonomy of scooter trips based on how long the trip would have taken on transit, not how close to transit it was, would better compare the two modes.

Improving Transit Mapping and Information Design

VTA Next Network Diagram

Transit maps are often confusing and extremely cluttered. Harry Beck's London Underground diagram demonstrated that they don't have to be almost a century ago, yet North American transit agencies still choose to provide customer information at two levels: the system level, and the route level (and sometimes at the mode level), both using geographic illustrations. But presenting service based on this hierarchy fails to accommodate real-life travel patterns.

In this project, I attempt to touch on some of these issues. To best illustrate a service overhaul aimed at increasing frequency on the network as a whole, I chose to diagram those routes running at a fifteen-minute or better frequency, which is a more accurate depiction of the transit services available at any given moment. Abandoning geography for a simplified schematic allows for the corridors and stations to jump out of the map, which ironically ends up allowing for a faster and clearer understanding of where routes go when compared to a geographic display.
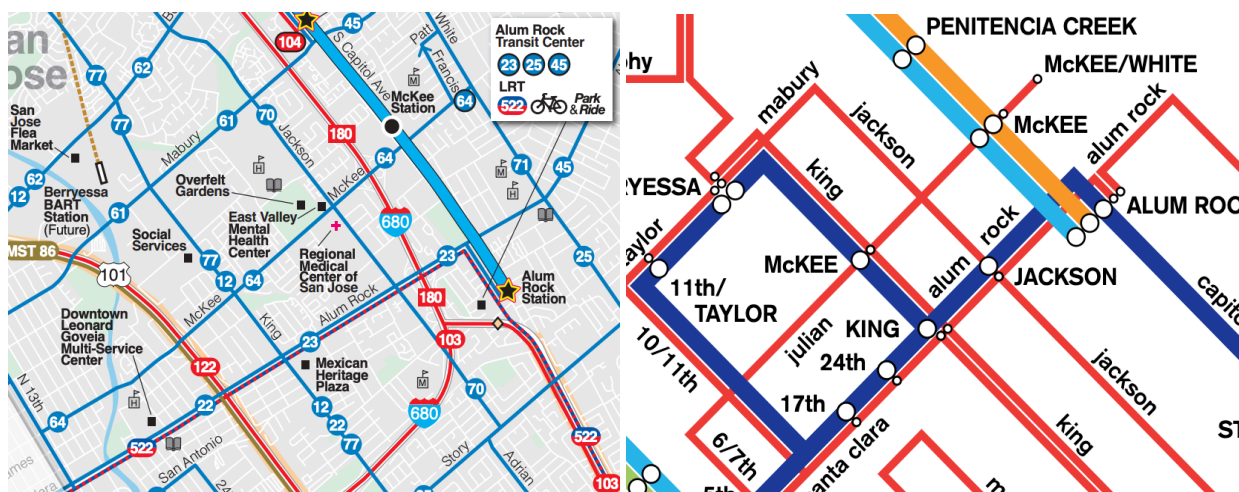
Simplifying map elements helps prioritize necessary information.

This project has a lot of room for improvement. Clicking on a stop should display all routes in the area, creating a connection between low-frequency routes and the 'core' network, in an effort to reduce the hierarchization of service I decried above. Accessibility data is currently not included. Also, there is room for much-improved integration between agencies in outlying areas (for example including BART service).

I also envision a wide range of potential uses for such a diagram. Using a simplified schematic like this one to display directions for riders would simplify the user interface and clarify the correct route to take. Additionally, a simplified schematic like this would be of great use in transit kiosks or information displays at important stations in the network.