Market Assessment

Methodology Memo

August 2023

CONTENTS

Geomarket Methodology

Background

As part of the initial LOCUS project, WMATA created planning areas to achieve a balance between the need for granularity while being cognizant of the large-scale nature of the data. WMATA began with neighborhood/area definitions identified by local jurisdictional planning departments. These were then joined to Census Block Groups so that Metro could aggregate the LBS data to geographic areas with more recognizable names that would be useful in engagement and other technical discussions about the results. This exercise led to the creation of 207 planning areas. However, given the objectives of the Better Bus Network Redesign to understand demand at a market level, there was a need to aggregate these planning areas into a more manageable number of geographies (geomarkets) where the project team can assess the convenience and accessibility of transit with the goal to provide insights into the network design process.

Methodology

The geomarket development process, as detailed below, was an iterative process to ensure that the aggregated planning areas are both internally consistent and useful in the network redesign process. Interstates and water bodies serve as natural boundaries and no geomarket will straddle either of these.

- **Step 1:** The process started with identifying regions within which the planning areas will be aggregated. The reason to break it by regions is because the densities of trips and routes and trip patterns in Downtown DC are different from those in the rest of DC and these again are different from the trip patterns in Arlington, VA, Alexandria VA, Montgomery and Prince George's Counties in Maryland and Fairfax County in Virginia. We used the following regions for aggregating planning areas into geomarkets:
  - Downtown Washington DC
  - NE Washington DC
  - NW Washington DC
  - SE Washington DC
  - SW Washington DC
  - Alexandria, VA
  - Arlington County, VA
  - Montgomery County, MD
  - Prince George's County, MD
  - Fairfax County, VA
  - Fairfax City , VA
  - Falls Church, VA
- **Step 2:** For each of the regions mentioned in Step 1, a cluster analysis of the planning areas was conducted based on identifying adjacent planning areas with similar demographic characteristics such as:
  - Percent of population that is low income (2019 Census)
  - Percent of population that are people of color (2019 Census)
  - Total trip flow (in and out) (2019 LOCUS flows)

Scores based on the above three characteristics are computed based on the percent of low income and people of color, and the total trips in the planning area. It was determined that planning areas that had large percentages of low income residents and people of color and large total trip flows would have high scores, while planning areas with small percentages of low income residents and people of color and small total trip flows would have low scores. Given that the magnitude of these three characteristics are so disparate, there is a need to normalize them to the same scale. Normalized scores for each field are calculated as follows.

For each characteristic in each planning area, the following equations were applied to calculate the normalized score of the characteristic in the area. The number of characteristics used in the geomarket development (low-income percentage, percentage people of color, and total trips in and out) are used to compute the largest distance in the jurisdiction.

$$LargestDistance = \frac{MAX(jurisdiction\ latitude\ range, jurisdiction\ longitude\ range)}{no.of\ characteristics}$$

$$Factor = \frac{max.value\ of\ characteristic\ in\ jurisdiction - min.value\ of\ characteristic\ in\ jurisdiction}{LargestDistance}$$

$$AreaScore = \frac{value\ of\ characteristic\ in\ area}{Factor}$$

The total score for each planning area is the sum of the normalized scores and the maximum possible score for a planning area in a jurisdiction is equal to the range of the latitude or longitude, whichever is larger, of the planning area's jurisdiction. This normalization was applied to ensure that the demographic characteristics and locations of planning areas have equal weight in the clustering process, as both are equally important for the bus network redesign. **Table 1** shows the attributes and clustering score for Aspen Hill planning area.

Table 1: Example Metrics for Aspen Hill in Montgomery County, MD

Attribute	Metric
Low Income (Percent)	9%
People of Color (Percent)	59%
Number of Trips	7,258,054
Score: Percent Low Income	0.076
Score: Percent People of Color	0.087
Score: Number of Trips	0.002
Total Score	0.16

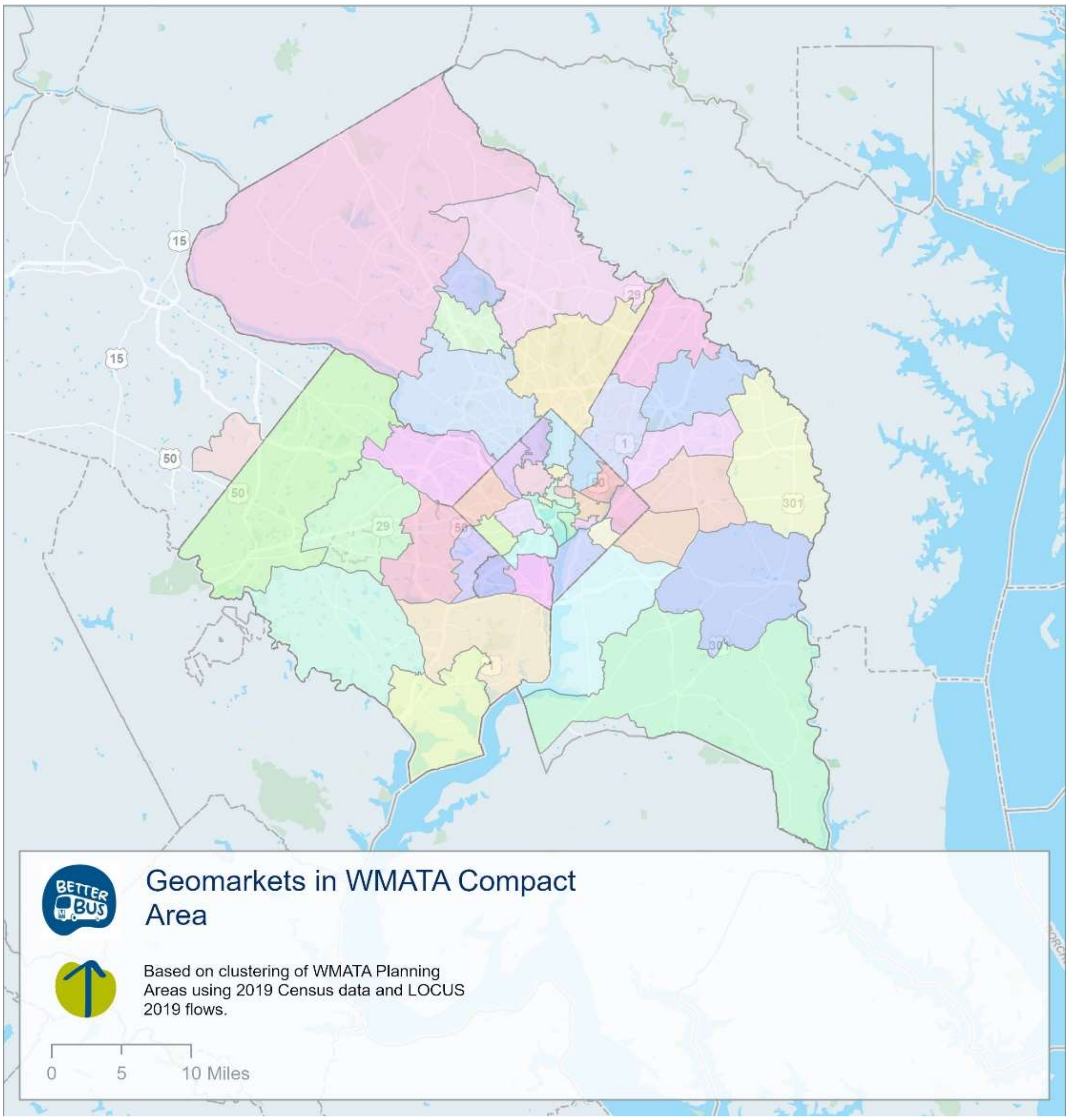
- **Step 3:** K-means clustering using the latitude, longitude, and total score was performed within each jurisdiction. First, the sum of squared distances was plotted for each jurisdiction to determine the ideal number of clusters, then each jurisdiction was clustered according to that number. The k-means clustering algorithm was applied to each of the jurisdictions separately.

- **Step 4:** Results from the k-means clustering process were refined using professional judgement to make sure the planning areas are reasonably segmented into distinct transit markets. Judgment was applied to consider input from the project team, natural barriers (Interstates, water bodies), major trip generators and neighborhood characteristics.

- **Step 5:** To further refine the Prince George's County geomarkets, consideration was given to the existing local bus network and route groupings. Additional geomarkets were defined to support a more discrete analysis for local bus.

Figure 1 shows the 49 Geomarkets generated as a part of this process.

Figure 1: Geomarkets

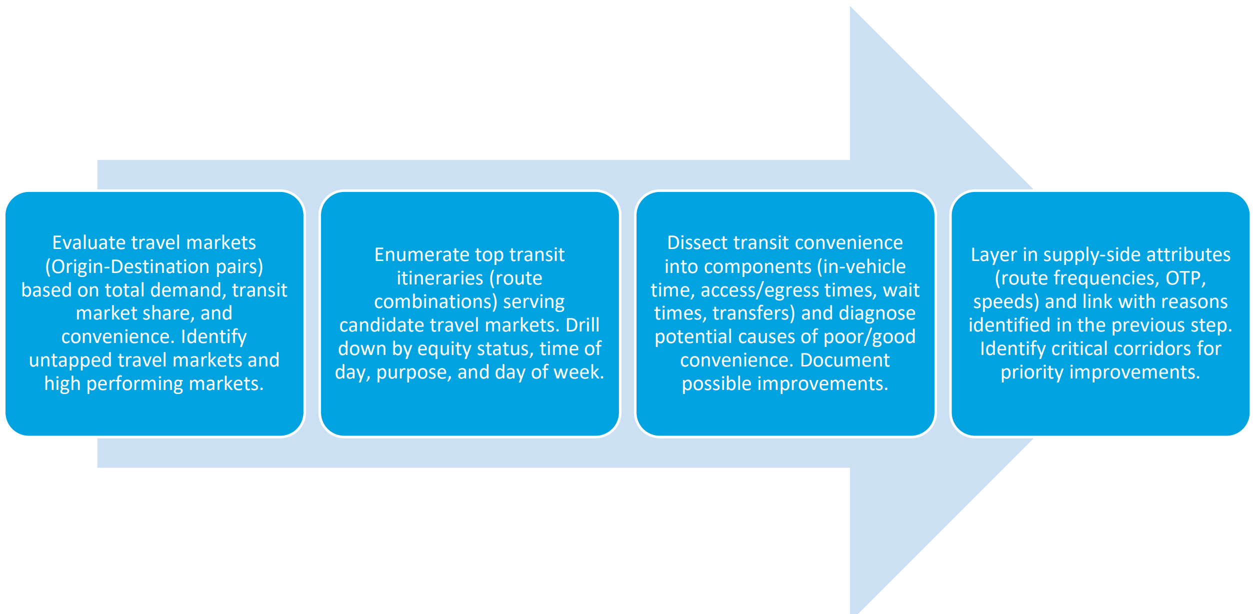


Travel Time Convenience Analysis

Background

One of the key tasks in the WMATA Bus Network Redesign is to understand the demand for transit and how effective the current transit network is in meeting that demand. The objective of the transit convenience task is to classify travel markets based on existing demand and existing quality of transit travel, prioritize markets for improvements, and then drill down to the itinerary level to diagnose why certain route itineraries are convenient or not by decomposing the transit travel times into access/egress, wait, transfer, and in-vehicle times. This is eventually used to inform the bus network redesign process. **Figure 2** shows the process. Key summaries from this analysis have been compiled in the **Market Assessment Appendix**.

Figure 2: Transit Convenience Analysis



Methodology

To understand how people are choosing to move throughout the Washington DC region, an in-depth analysis was conducted on transit mode share. This was paired with an analysis comparing transit travel time to auto travel time, to create a travel time ratio (TTR) that was used to assess the various types of travel markets. Travel time ratios (transit travel time/drive time) are used as an indicator of relative convenience of transit – the longer it takes to accomplish a given trip on transit compared to driving, is perceived as less convenient by the traveler.

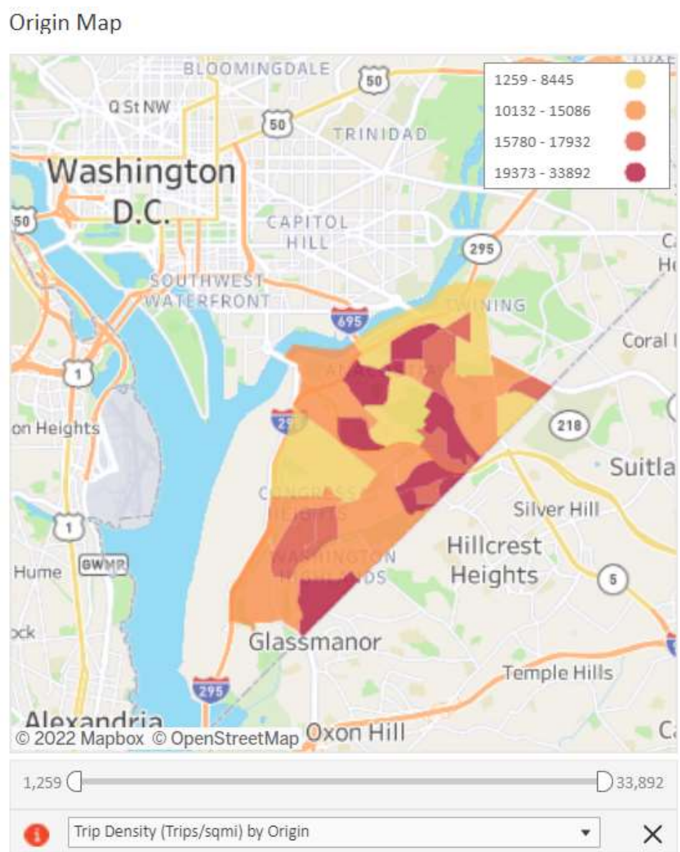
The data used for this analysis includes LOCUS flows (for overall travel) and WMATA Trace data (for transit travel) enriched with routing attributes of the "best" transit itineraries (see LOCUS Methods memo for more details). To accommodate the large-scale nature of the LOCUS and Trace datasets, and to make it easier to understand travel demand and provide market assessments, the analysis is conducted at the geomarkets level. The analysis not only includes the observed (actual transit trips), but also potential transit trips that could have been accomplished on transit (all trips which have a feasible itinerary).

Illustration

In order to make the methodology clear to the reader, it will be explained using a sample geomarket pair, Southeast DC and Downtown DC. The final analysis is packed into an interactive Tableau workbook that lets the users conduct end to end analysis in a streamlined fashion.

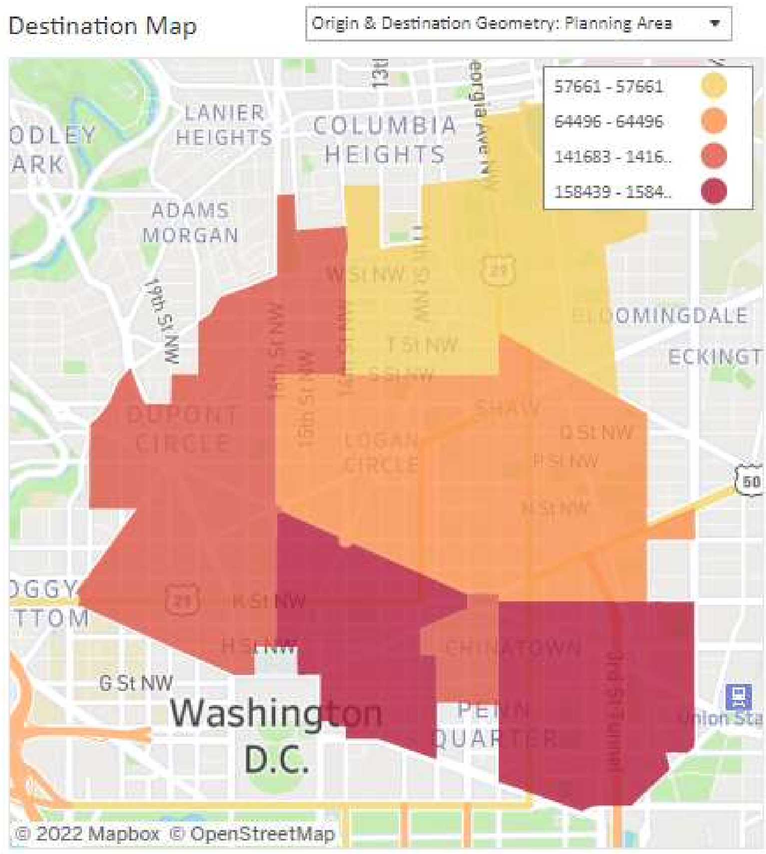
The origin Geomarket, Southeast DC includes the following planning areas: Congress Heights/Bellevue/Washington Highlands, Douglas/Shipley Terrace, Fairfax Village/Naylor Gardens/Hillcrest/Summit Park, Historic Anacostia, Saint Elizabeths, Sheridan/Barry Farm/Buena Vista, Twining/Fairlawn/Randle Highlands/Penn Branch, and Woodland/Ft Stanton/Garfield Heights/Knox Hill and shown in **Figure 3**.

Figure 3: Origin Geomarket



The Destination geomarket includes the following planning areas: Downtown, Chinatown, Penn Quarters, Mount Vernon Square, North Capitol Street, Dupont Circle, Connecticut Avenue/K Street, Shaw, Logan Circle, Howard University, Le Droit Park, Cardozo/Shaw and shown in **Figure 4**.

Figure 4: Destination Geomarket



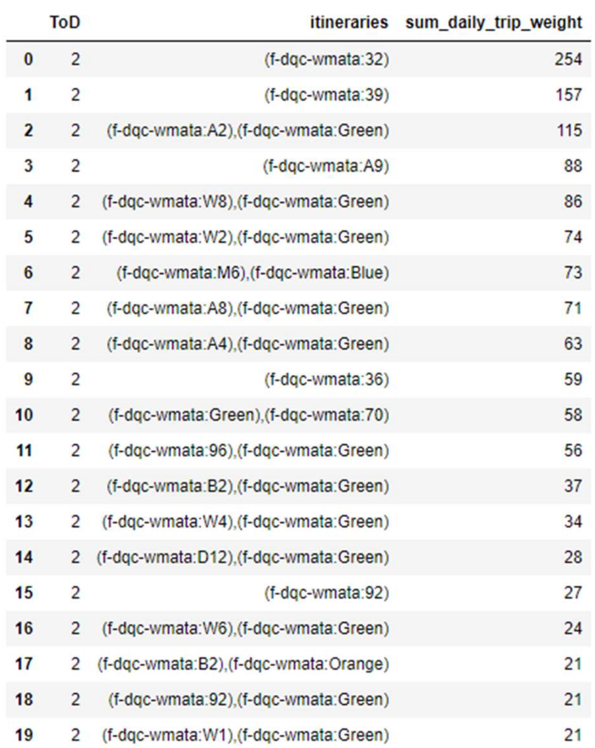
Once the geomarkets are identified, the next step is to determine the total and transit trips, the share of transit, the average travel time ratio (Transit Time/Auto Time), and the percent of peak and commute trips for each geomarket pair. This information comes directly from the routed LOCUS data and is used to provide an overall picture of convenience in the region. **Table 2**: Travel Market Profileshows these values for our sample travel market.

Table 2: Travel Market Profile

Attribute	Value
Geomarket Pair	Southeast DC to Downtown DC
Total Trips	10,383
Transit Trips	2,614
Transit Share	25.7%
Average Travel Time ratio	3.4
Percent of total trips in Peak Period	47.0%
Percent total trips that are Commute trips	32.6%

Following the overall assessment of a geomarket pair's convenience (as represented as TTR) and travel characteristics, the next step is to identify the top transit itineraries between geomarkets. This is achieved by parsing the mode summary field in the routing data, dropping "walk" and "auto" modes in the route records, and extracting the transit itineraries for all trips. Since the focus is on WMATA bus, those transit itineraries that are rail only are dropped from the analysis. The top 20 transit itineraries with either bus only or bus and rail for different time of day segments are identified. **Figure 5** and **Figure 6** show the top weekday itineraries for the AM Peak and Midday time periods respectively.

Figure 5: Top transit itineraries with bus for AM Peak (6:00 a.m. to 9:00 a.m.)



	ToD	itineraries	sum_daily_trip_weight
0	2	(f-dqc-wmata:32)	254
1	2	(f-dqc-wmata:39)	157
2	2	(f-dqc-wmata:A2),(f-dqc-wmata:Green)	115
3	2	(f-dqc-wmata:A9)	88
4	2	(f-dqc-wmata:W8),(f-dqc-wmata:Green)	86
5	2	(f-dqc-wmata:W2),(f-dqc-wmata:Green)	74
6	2	(f-dqc-wmata:M6),(f-dqc-wmata:Blue)	73
7	2	(f-dqc-wmata:A8),(f-dqc-wmata:Green)	71
8	2	(f-dqc-wmata:A4),(f-dqc-wmata:Green)	63
9	2	(f-dqc-wmata:36)	59
10	2	(f-dqc-wmata:Green),(f-dqc-wmata:70)	58
11	2	(f-dqc-wmata:96),(f-dqc-wmata:Green)	56
12	2	(f-dqc-wmata:B2),(f-dqc-wmata:Green)	37
13	2	(f-dqc-wmata:W4),(f-dqc-wmata:Green)	34
14	2	(f-dqc-wmata:D12),(f-dqc-wmata:Green)	28
15	2	(f-dqc-wmata:92)	27
16	2	(f-dqc-wmata:W6),(f-dqc-wmata:Green)	24
17	2	(f-dqc-wmata:B2),(f-dqc-wmata:Orange)	21
18	2	(f-dqc-wmata:92),(f-dqc-wmata:Green)	21
19	2	(f-dqc-wmata:W1),(f-dqc-wmata:Green)	21

Figure 6: Top transit itineraries with bus for Midday (9:00 a.m. to 3:00 p.m.)

	ToD	itineraries	sum_daily_trip_weight
0	3	(f-dqc-wmata:A8),(f-dqc-wmata:Green)	111
1	3	(f-dqc-wmata:A2),(f-dqc-wmata:Green)	79
2	3	(f-dqc-wmata:36)	78
3	3	(f-dqc-wmata:W4),(f-dqc-wmata:Green)	74
4	3	(f-dqc-wmata:W3),(f-dqc-wmata:Green)	62
5	3	(f-dqc-wmata:32)	62
6	3	(f-dqc-wmata:A4),(f-dqc-wmata:Green)	56
7	3	(f-dqc-wmata:W8),(f-dqc-wmata:Green)	53
8	3	(f-dqc-wmata:96),(f-dqc-wmata:Green)	52
9	3	(f-dqc-wmata:30S)	47
10	3	(f-dqc-wmata:V2),(f-dqc-wmata:Green)	46
11	3	(f-dqc-wmata:92)	44
12	3	(f-dqc-wmata:P6)	43
13	3	(f-dqc-wmata:A6),(f-dqc-wmata:Green)	39
14	3	(f-dqc-wmata:30N)	32
15	3	(f-dqc-wmata:52),(f-dqc-wmata:Blue)	29
16	3	(f-dqc-wmata:B2),(f-dqc-wmata:Green)	26
17	3	(f-dqc-wmata:W1),(f-dqc-wmata:Green)	25
18	3	(f-dqc-wmata:B2),(f-dqc-wmata:Silver)	22
19	3	(f-dqc-wmata:39)	22

The next step is to decompose the travel times associated with top transit itineraries into individual components of the transit journey. For illustration, a threshold of 70 daily trips for the itineraries is chosen for further analysis. The transit travel time components include access time, initial wait time, egress time, in-vehicle time, transfer wait time, transfer walk time. **Figure 7** shows the travel time components for itineraries in the midday and evening periods, along with the drive times (used for calculating the TTR).

Figure 7: Travel Time Decompositions for AM Peak and Midday Itineraries

AM Peak

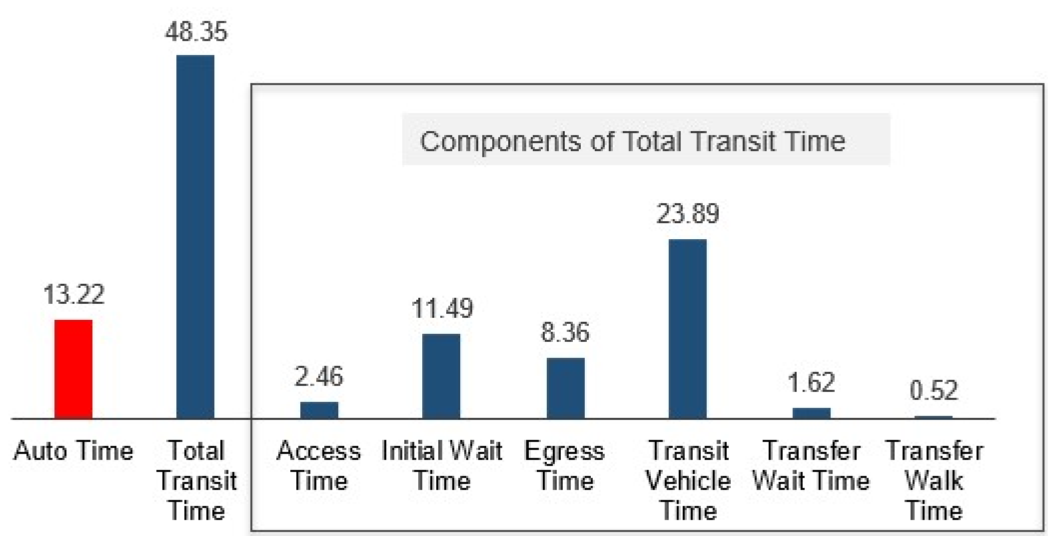
	itinerary	access_time	initial_wait_time	egress_time	transit_vehicle_time	transfer_wait_time	transfer_walk_time	auto_time
0	(f-dqc-wmata:32)	4.25	4.98	5.98	32.67	0.00	0.00	14.14
1	(f-dqc-wmata:39)	5.09	9.91	5.00	26.88	0.00	0.00	13.02
2	(f-dqc-wmata:A2),(f-dqc-wmata:Green)	2.46	11.49	8.36	23.89	1.62	0.52	13.22
3	(f-dqc-wmata:A9)	4.95	9.46	3.72	35.95	0.00	0.10	13.77
4	(f-dqc-wmata:W8),(f-dqc-wmata:Green)	2.90	10.38	6.77	24.12	1.14	1.48	13.25
5	(f-dqc-wmata:W2),(f-dqc-wmata:Green)	2.09	12.99	6.65	22.09	2.25	1.05	14.18
6	(f-dqc-wmata:M6),(f-dqc-wmata:Blue)	2.37	7.04	5.13	26.05	1.40	0.65	15.41
7	(f-dqc-wmata:A8),(f-dqc-wmata:Green)	3.00	7.78	8.10	26.83	2.17	0.80	14.30

Midday

	itinerary	access_time	initial_wait_time	egress_time	transit_vehicle_time	transfer_wait_time	transfer_walk_time	auto_time
0	(f-dqc-wmata:A8),(f-dqc-wmata:Green)	2.53	11.32	6.76	23.78	4.18	0.80	14.25
1	(f-dqc-wmata:A2),(f-dqc-wmata:Green)	3.06	12.82	6.18	24.61	2.58	1.22	14.37
2	(f-dqc-wmata:36)	4.26	13.50	8.41	26.80	0.00	0.00	11.23
3	(f-dqc-wmata:W4),(f-dqc-wmata:Green)	3.18	9.61	6.21	22.37	1.80	1.67	14.21

Following the breakdown by individual transit travel time components, a gap analysis is conducted to identify which components of transit travel time contribute to the inconvenience of a transit itinerary. **Figure 8** shows one example for the Metrobus A2 Anacostia-Washington Highlands Line during the weekday AM peak. This is an inconvenient service because transit takes 3.7 times as long as auto and the reasons are because the initial wait time and the transit vehicle time are so long. Therefore, the potential recommendations to make the itinerary more convenient are to increase frequency reducing initial wait time, or to make the route more direct reducing travel time.

Figure 8: A2 – AM Peak Gap Analysis



Finally, for each geomarket pair a table for the itineraries that have 70 or more daily trips showing the transit travel time components and highlighting the unreasonable ones as the potential causes for the lack of convenience of that itinerary was created. **Figure 9** shows the output for the sample geomarket.

Figure 9: Transit Competitiveness Output

OD Market	Travel Segment	Top Itineraries	Total Transit Travel Time	Travel Time Ratio	In Vehicle Travel Time Share	% OVTT – access time	% OVTT - initial wait time	% OVTT - transfer wait time	% OVTT - transfer walk time	% OVTT – egress time	Reason
Southeast Washington - Downtown DC *Total Demand*: 10,181 daily trips *Avg Transit Competitiveness*: 2.7 *Transit Market Share*: 25.7% *Avg Transfers*: 0.73 *Unique Transit Itineraries with Bus*: 981	AM Peak, Weekday	wmata:32	47.88	3.39	68%	28%	33%	0%	0%	39%	Bus is slow (mostly in-vehicle time)
		wmata:39	46.89	3.60	57%	25%	50%	0%	0%	25%	Out-of-vehicle time high, mostly due to high initial wait times
		wmata:A2-wmata:Green	48.35	3.66	49%	10%	47%	7%	2%	34%	Out-of-vehicle time high, mostly due to high initial wait times
		wmata:A9	54.17	3.93	66%	27%	52%	0%	1%	20%	Bus is slow (mostly in-vehicle time)
		wmata:W8-wmata:Green	46.79	3.53	52%	13%	46%	5%	7%	30%	Out-of-vehicle time high, mostly due to high initial wait times
		wmata:W2-wmata:Green	47.12	3.32	47%	8%	52%	9%	4%	27%	Out-of-vehicle time high, mostly due to high initial wait times
		wmata:M6-wmata:Blue	42.65	2.77	61%	14%	42%	8%	4%	31%	
		wmata:A8-wmata:Green	48.68	3.40	55%	14%	36%	10%	4%	37%	Out-of-vehicle time high, walk times are high (access/egress/transfer)
	Midday, Weekday	wmata:A8-wmata:Green	49.37	3.47	48%	10%	44%	16%	3%	26%	Out-of-vehicle time high, wait times (initial + transfer) high
		wmata:A2-wmata:Green	50.47	3.51	49%	12%	50%	10%	5%	24%	Out-of-vehicle time high, wait times (initial + transfer) high
		wmata:36	52.96	4.72	51%	16%	52%	0%	0%	32%	Out-of-vehicle time high, mostly due to high initial wait times
		wmata:W4-wmata:Green	44.83	3.15	50%	14%	43%	8%	7%	28%	Out-of-vehicle time high, wait times (initial + transfer) high

As previously stated, the actual analysis is packaged into a Tableau dashboard that allows users to dynamically adjust the thresholds used throughout the analysis as well as segment travel markets along several dimensions (such as Day of Week, time of day, Equity-focus Communities, and Travel Purpose). **Figure 10** and **Figure 11** showcase the geomarkets level and itinerary level analysis respectively.

Figure 10: Market Profiles and Classification

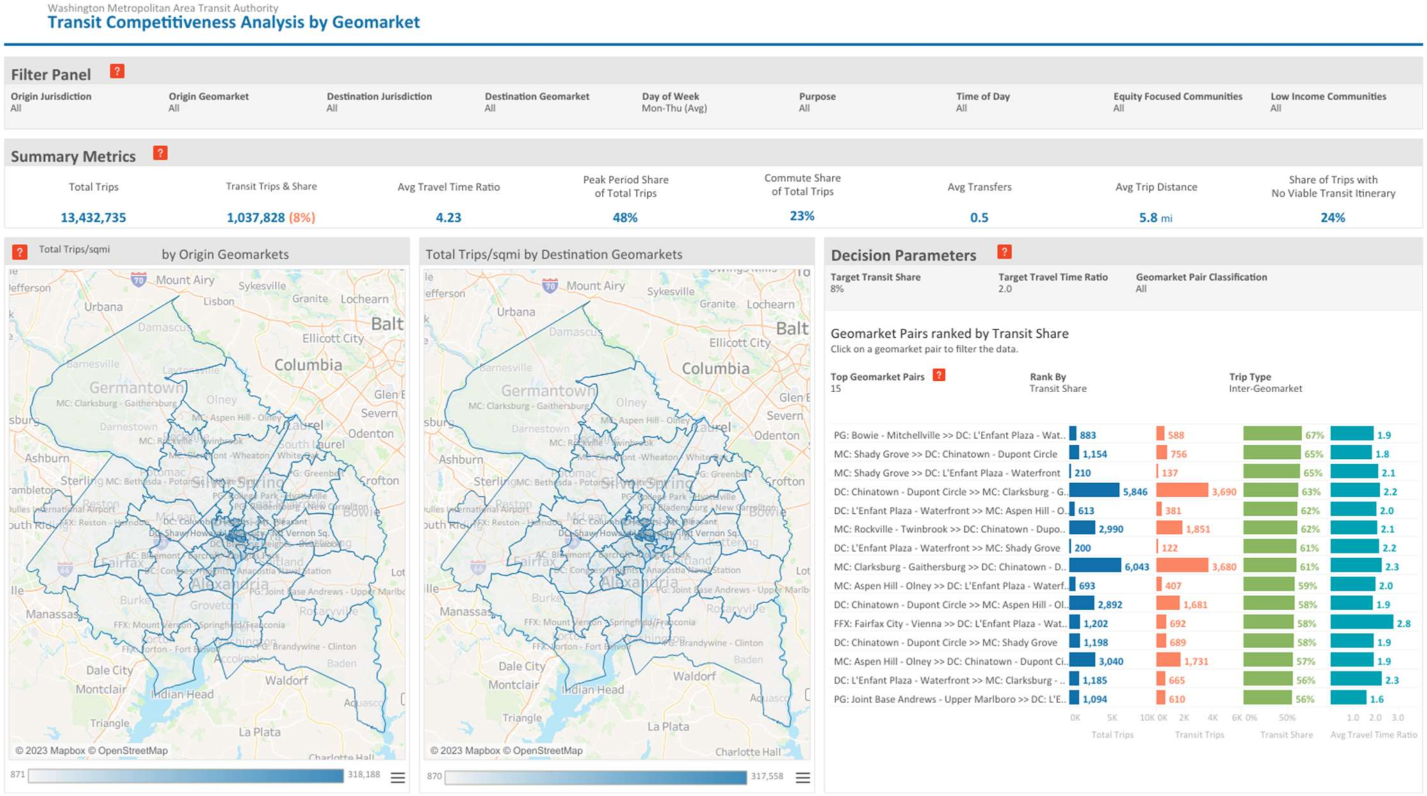
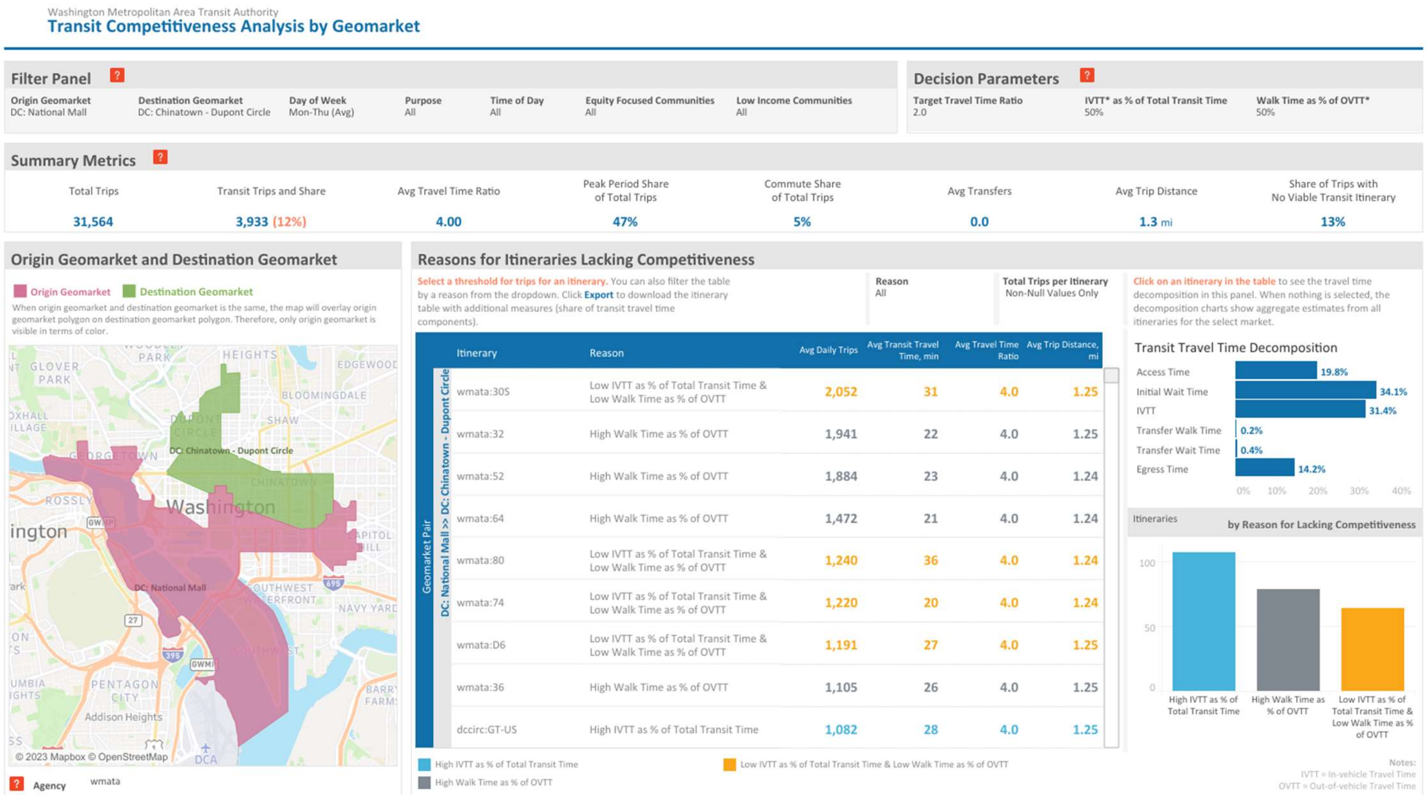


Figure 11: Itinerary Analysis



Accessibility Analysis

Background

The objective of the accessibility analysis is to quantify the accessibility to key destinations (identified by WMATA as destinations contributing to social welfare) and jobs by transit from different geomarkets under the existing transit network conditions. The unique aspect of this analysis is that it considers accessibility based on observed travel patterns (where people are actually traveling), as opposed to a track of accessibility analysis that looks at latent travel (where people can potentially travel). Key results have been presented in the Task 3 report appendix.

The analysis also splits the results for Equity-focused Communities, and Low-Income Communities to allow users to identify "accessibility deserts" for different population groups. Since transit convenience is a big driver for the network evaluation and improvements, accessibility measures based on "convenient transit options" (where TTR is below a certain threshold) are also included in the analysis.

Data Sources

Below are the key data sources leveraged for this analysis.

- 2019 LOCUS Data, with routing attributes
- Equity-focused Communities data (Census, compiled by WMATA)
- Census data
- Hospitals and Urgent Care Facilities: https://hifld-geoplatform.opendata.arcgis.com/
- Total Jobs: https://lehd.ces.census.gov/data/
- Grocery Store and Educational Facilities: Google

Figure 12 to **Figure 15** show the spatial distribution of these activity centers in the region.

Figure 12. Locations of Grocery Stores

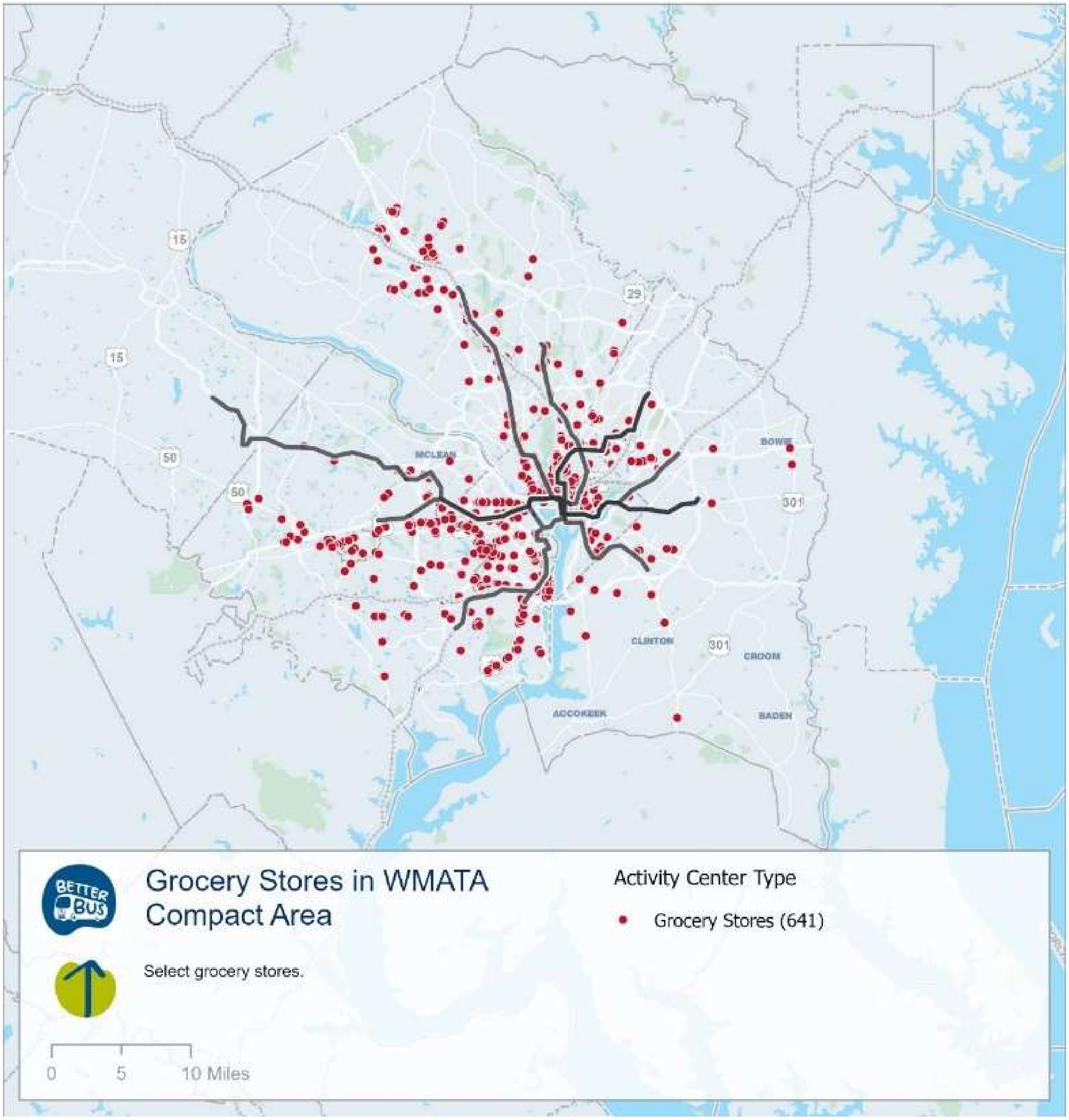


Figure 13. Locations of Educational Facilities

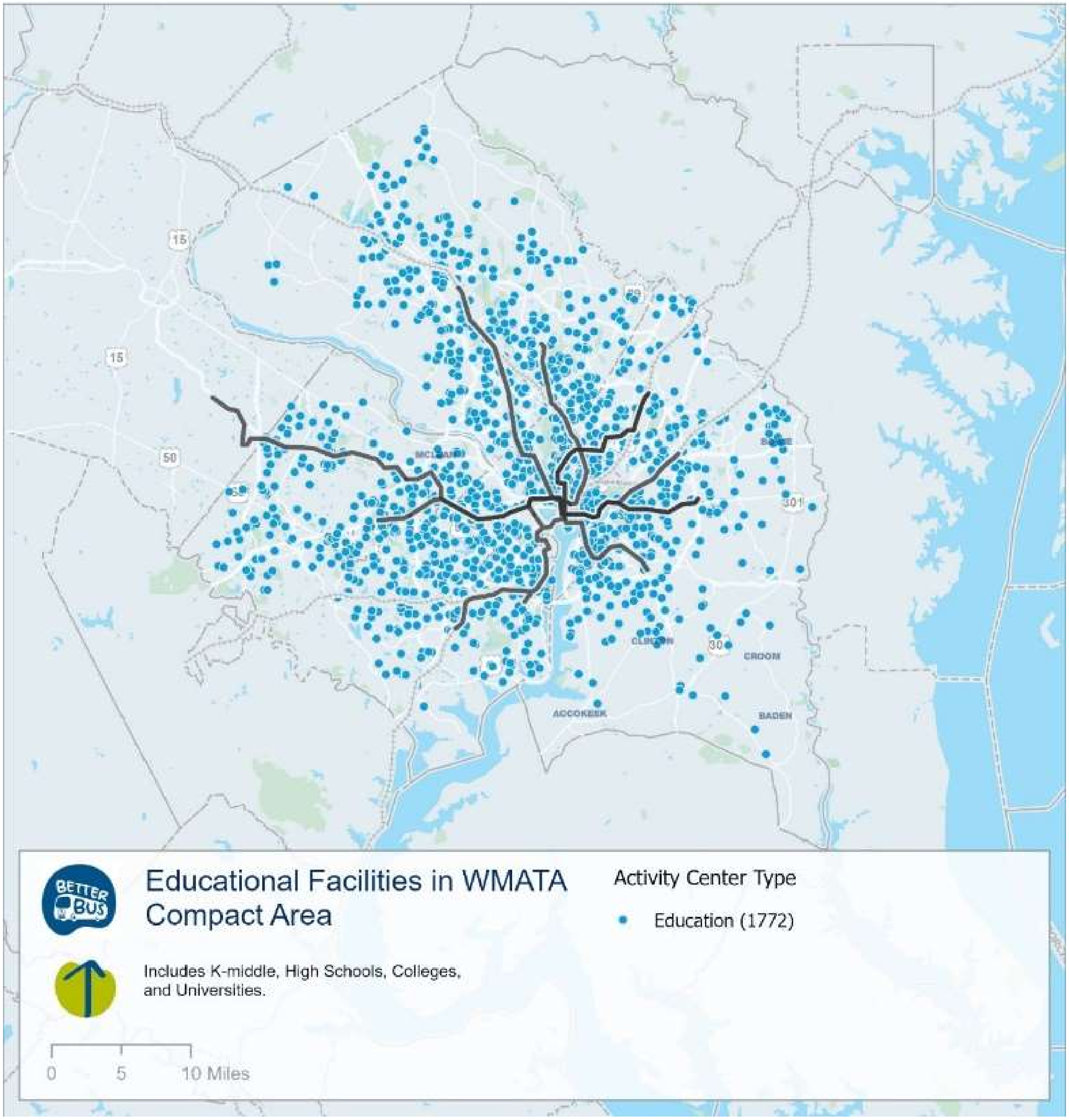


Figure 14. Locations of Medical Facilities

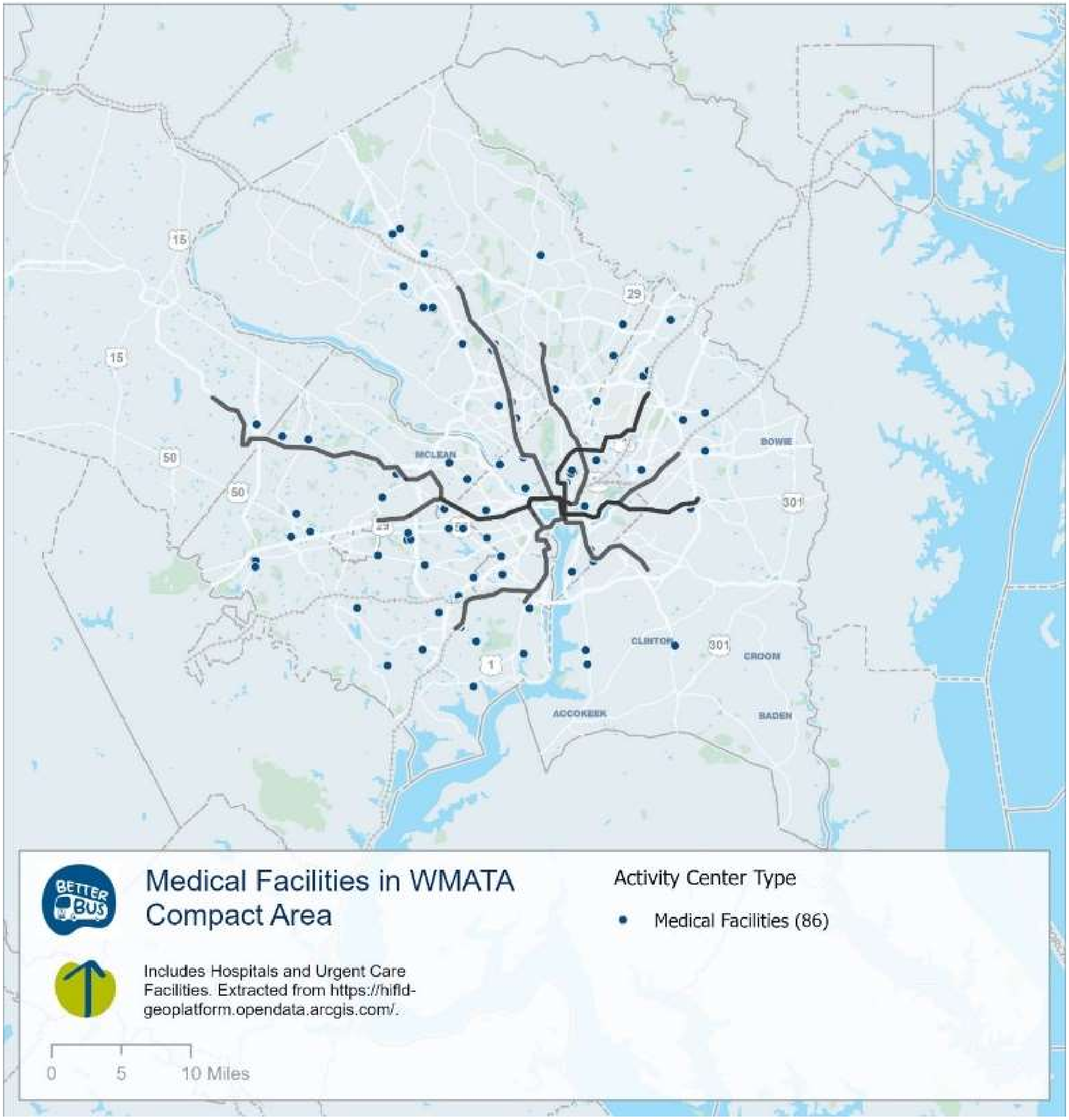
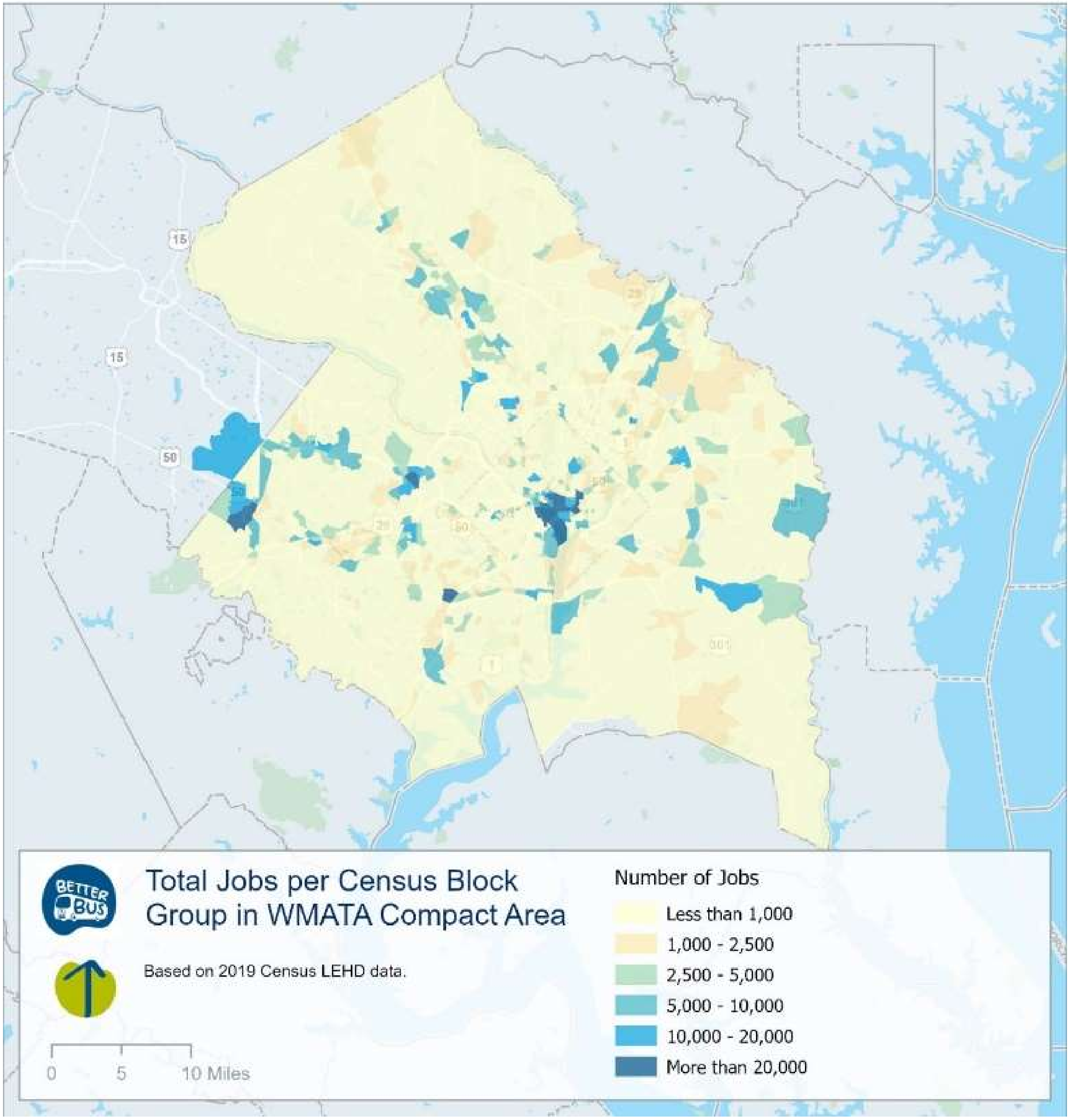


Figure 15. Total Jobs by Census Block Group



Methodology

The accessibility analysis begins by spatially intersecting points of each destination type (grocery stores, educational facilities, medical facilities, and jobs) to the Census Block Group (CBG) that contains it. Educational facilities used for this analysis include high schools, colleges, and universities; Medical facilities include hospitals and urgent care facilities; and jobs are total jobs by CBGs as reported by Census. The reason behind running the analysis at the BG aggregation level is to allow for smoothening of outliers (associated with point locations) and computation of metrics for activity centers not available as point locations (such as jobs).

Next, for each CBG ("origin CBG") in the region, the average travel times on transit to CBGs containing destinations of interest ("destinations CBG") are computed using the trips observed in the total flow dataset (so based on actual travel patterns, instead of latent demand).

Based on the selected transit travel time band (15 mins, 30 mins, or 60 mins), "destinations CBGs" accessible within the transit travel time from the given origin CBG are extracted and the destinations contained in the "destinations CBG "are summed. This yields the number of destinations of given type accessible within the selected transit travel time band per CBG. This number is then aggregated to the Geomarket level for summarization purposes – average values for CBGs contained in that geomarkets. The results are also compared against regionwide averages.

Additionally, the average travel time ratios between the "origin CBG" and "destination CBG" are also computed, and the analysis described above is regenerated with OD pairs with "convenient transit options". The implication here is that even though transit can get people to the activity centers in the given time bins, there might be faster and more convenient and competitive alternatives to transit. This is the reasoning behind using "convenient accessibility" as a measure, instead of just pure travel time accessibility because it can falsely give the impression that transit is truly a viable alternative.

Figure 16 shows the data format that feeds into the accessibility dashboards. The results of the analysis are packaged into Tableau dashboards, as exhibited in **Figure 17**.

Figure 16: Data Format for Accessibility Dashboard

start_bg	AC_ID	num	Type	15min_ttr	30min_ttr	60min_ttr
110010076015	240317010011	366	jobs_center	NaN	NaN	NaN
110010076012	510594222011	3	education	NaN	NaN	NaN
110010098011	110010073011	1331	jobs_center	NaN	2.342589	3.432966
110010098101	240317010021	2360	jobs_center	NaN	NaN	NaN
110010098021	110010019013	1	store	NaN	NaN	NaN
110010098072	240338036022	10129	jobs_center	NaN	NaN	2.898745
110010104001	515102004055	1327	jobs_center	NaN	NaN	NaN
110010074081	240317060123	1	store	NaN	NaN	NaN
110010098072	240317006081	194	jobs_center	NaN	NaN	NaN
110010098112	110010034002	3	education	NaN	NaN	2.479986
110010098072	240317050004	19539	jobs_center	NaN	NaN	NaN
110010074071	240338035251	804	jobs_center	NaN	NaN	NaN
110010098011	510131017021	3477	jobs_center	NaN	NaN	2.204324
110010075021	240338028042	1	store	NaN	NaN	2.081511
110010074092	110010074093	1	store	3.491588	3.509241	3.475198
110010074042	515102009003	2	education	NaN	NaN	NaN
110010074041	240338035203	195	jobs_center	NaN	NaN	NaN
110010075041	240338015002	4	store	NaN	NaN	2.708598
110010075032	516003002004	2	store	NaN	NaN	NaN
110010076015	110010111001	2	store	NaN	NaN	3.742044

Figure 17: Accessibility Analysis Dashboard