ATLANTA RAPID MINERS

Emory Shuttle Simulation: C Route

October 15th, 2019

EMORY UNIVERSITY

Agenda

- Our Goal
- Data Capture and Data Processing
- Modeling Process
- Model Validation
- Future Improvement

Our Goal

Analyze C Route Shuttle

- Identify arrival times between each 2 pairs of stops
- Calculate the passenger arrival pattern across 11 stops (and fit it to a distribution)
- Simulate this process for Monday-Thursday on Simio

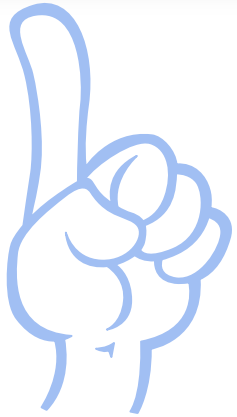
Pressure Test System

- What happens if demand surges? That is, how does throughput & utilization change if number of people at each stop increases?
- What happens if number of buses on a particular day decreases?

Our Goal

- ✓ Analyze the performance of Emory's Route C system as it stands today
- ✓ Identify key gaps that lead to inefficiency
- ✓ Suggest pro-active solutions for anticipated or commonly experienced problems

Our Process: 3 Key Steps



Data Capture & Processing

- Understand and analyze data provided by the Emory Shuttles
- Clean dataset by filtering out non-identified values
- Calculate parameters for each stop (such as passenger arrival pattern, riders getting off, service time)

Model Induction

Simulate Route C shuttle with the clean dataset on Simio

Validation and Implication

Validate results by comparing the model outputs with the data

Data Preparation

1

Missing Value

Delete Unidentified Data Points, such as those with unknown stops

2

Specific Days

Filter data to keep only weekdays (Mon-Thu) for simplicity

3

Time Buckets

4 buckets: 6-11 AM, 11-3 PM, 3-7 PM, 7PM- Midnight
Assumes that the system operation (demand, travel times) varies at different times of the day

4

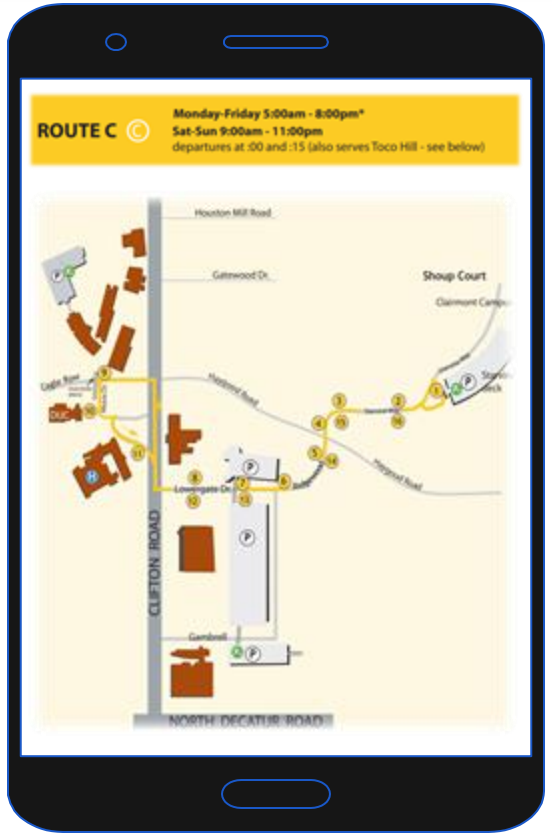
Regroup Data

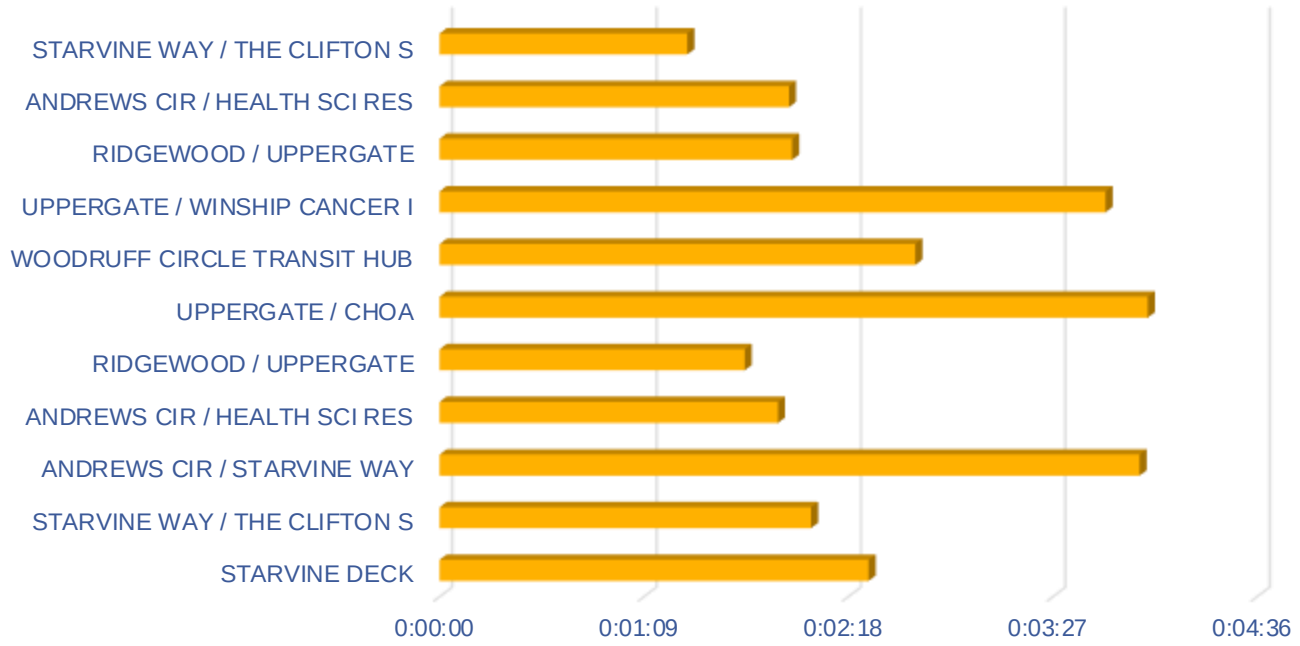
Calculate Average Travel Time between every 2 stops

Calculate # of Passengers on the bus at each stop

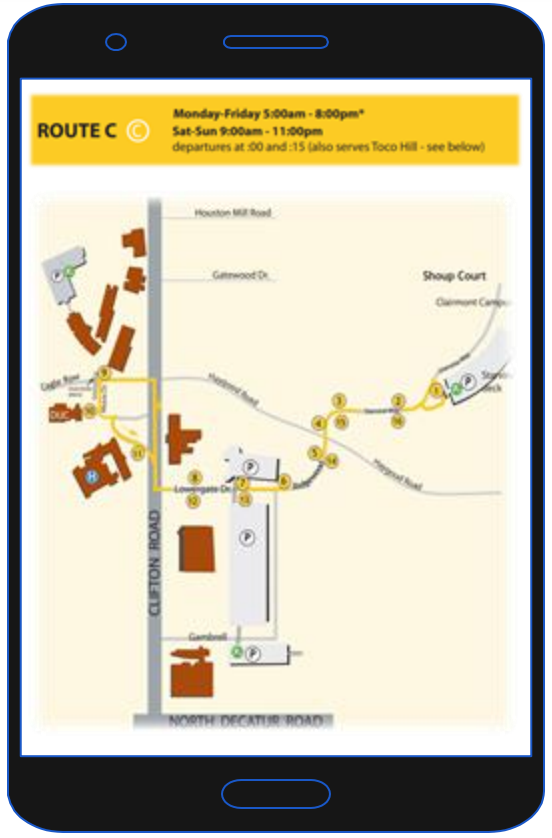
Calculate # and % of passengers getting off at each stop

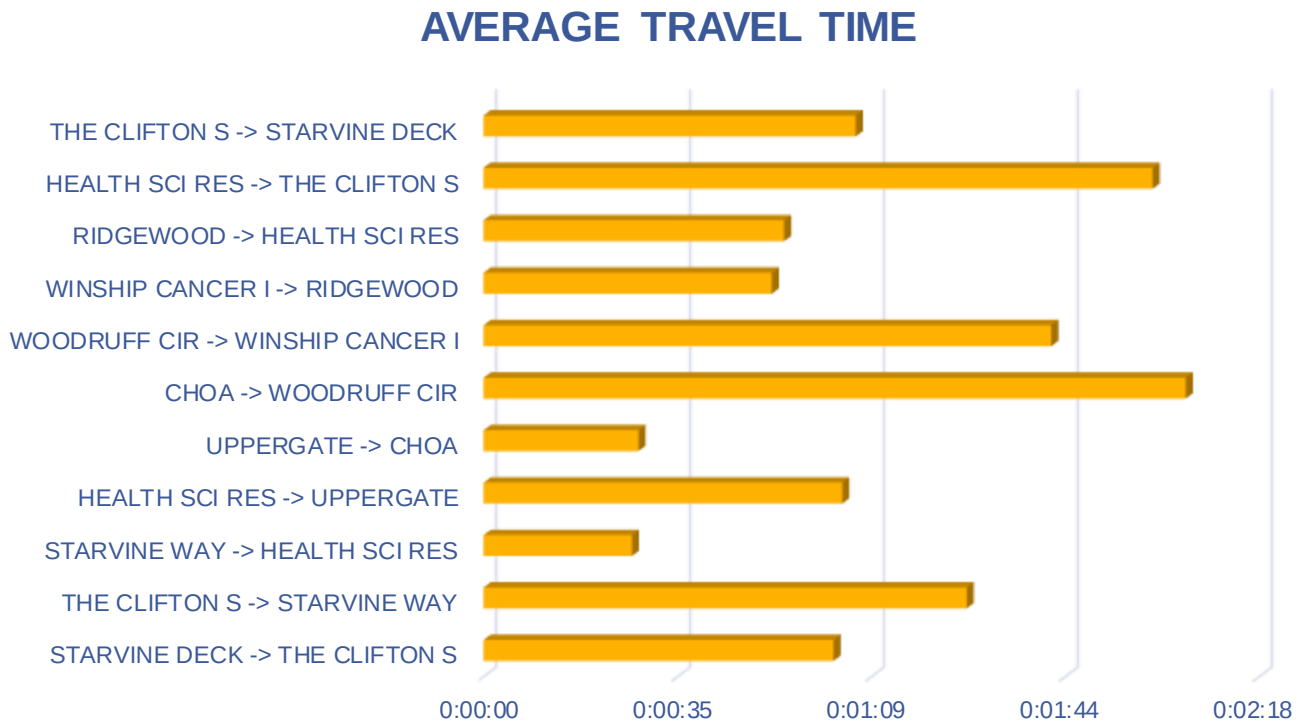
Data Insights – Average Time Interval at Each Stop





Data Insights – Average Travel Time between Stops

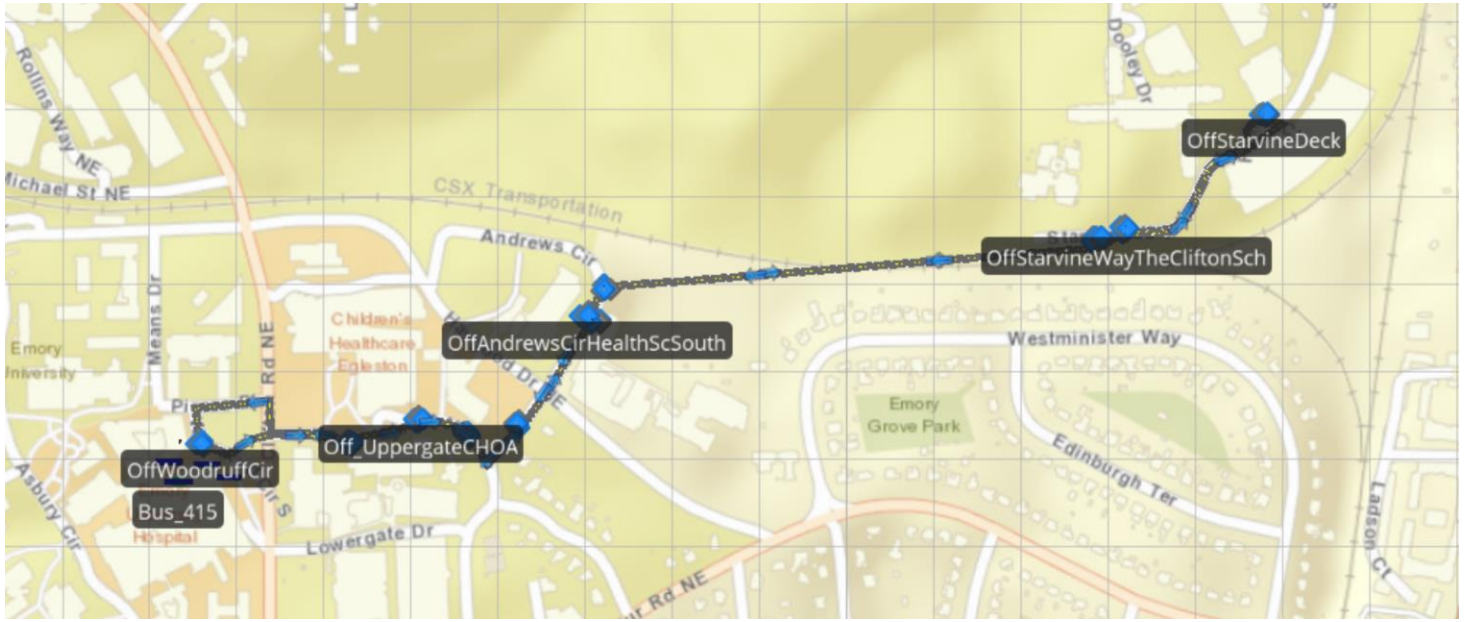




Route structure and Model Preview

- **Start** • Starvine Deck
- **Stop 1** • Starvine Way
- **Stop 2** • Andrews Cir @ Health Sci Research Building
- **Stop 3** • Andrews Cir @ Haygood
- **Stop 4** • Ridgewood @ Uppergate
- **Stop 5** • Uppergate @ CHOA
- **Stop 6** • Woodruff Cir
- **Stop 7** • Uppergate @ Whinship Cancer Institution
- **Stop 8** • Uppergate @ Ridgewood
- **Stop 9** • Andrews Cir @ Starvine Way
- **Stop 10** • Starvine Way @ The Clifton Sch
- **Stop 11** • Starvine Deck

Simulated Model – First look



Model Process

Add 11 Stops

- We used 11 pairs of Source-Server-Sink nodes
- **Source Node**: generating passengers with interarrival time following an **Exponential Distribution** *(the mean -* $\frac{1}{Lamda}$ *was calculated using the data for each stop in each time bucket)*
- **Server Node**: a second server was added at each stop to deal with the problem of all passengers departing at the next step

Connecting Stops

- Time-paths were used to connect all the stops
- A **normal distribution** is used with travel time mean and std calculated between every 2 pair of stops

4 Buses on Route C

- Buses depart from Starvine deck **every 5 minutes, starting 6 AM**
- Based on the bus patterns in real-time in the Rider App, we chose 4 buses for Route C
- While the data suggests there were over 30 buses, each day only saw 3-5 buses run. This was another reason to choose 4 buses

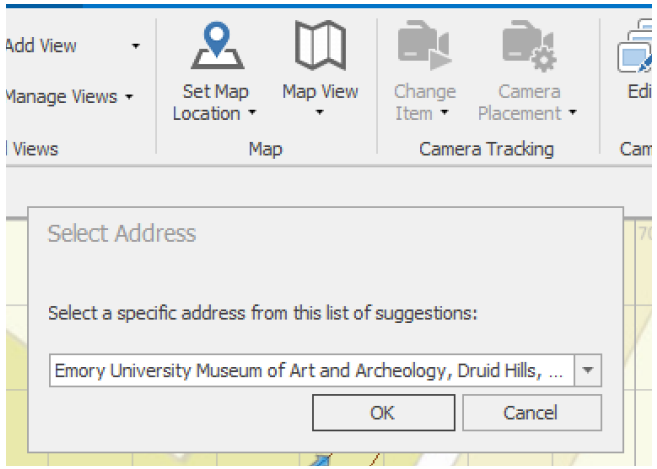
Set Maximum Capacity

- The bus capacity was set at 65 for all four buses
- This value is the maximum observed LOAD at any point in the data, and we assume the utilization was 100% at this point

Model: Map View

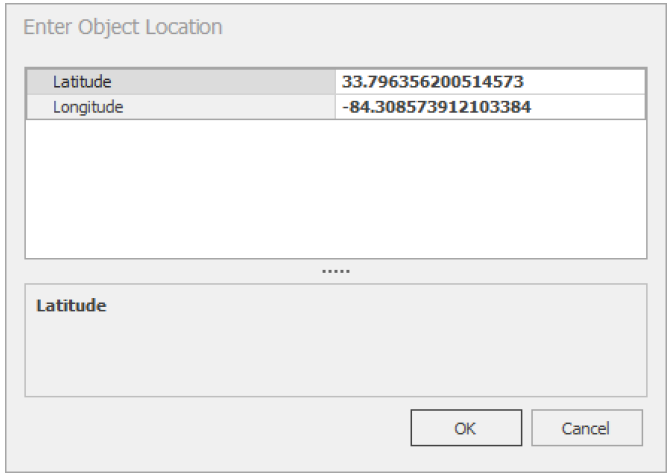
Select A Specific Address

To set up the map view, we first selected a specific location for Route C



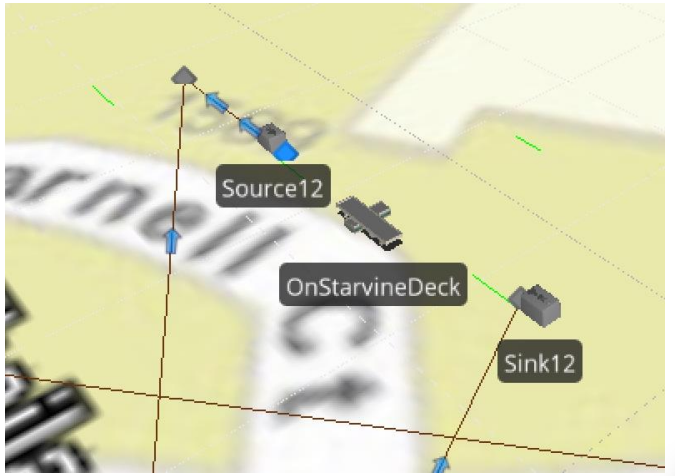
Stop Location

To set the stop location, we then used longitude and latitude for each bus stop



Connecting Stops

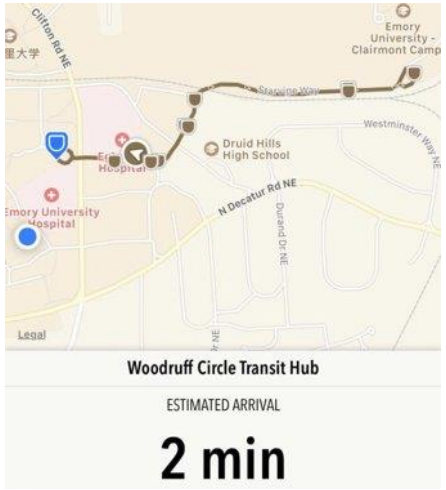
After setting up the map view and locating each stop, we use Time Path function to connect the stops



Model: Stop Sequence

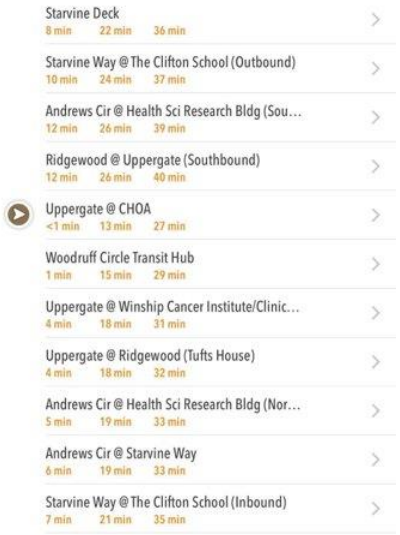
Design the route

We design the map view of the route based on the Rider App.



Stop Sequence

From the Rider App, we were also able to acquire the stop sequence.



Build the Sequence Table

Based on the information acquired from Rider, we built a sequence table and let the bus travel sequentially between stops

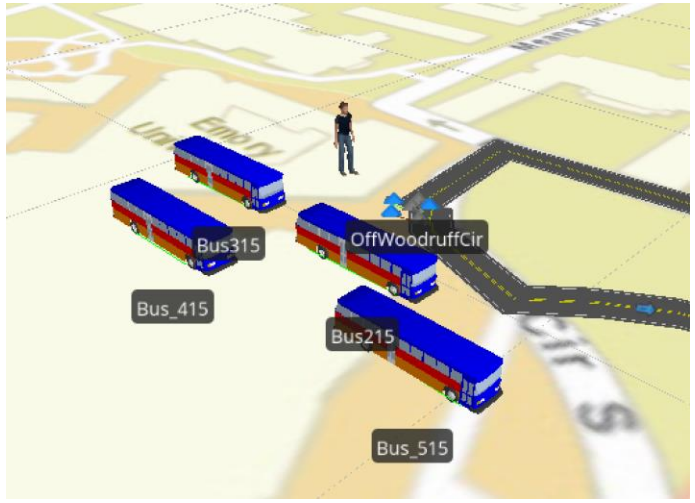
	Sequence	Stop Time
1	Output@OnStarvineDeck	1
2	Input@OffStarvineWayTheCliftonSch	1
3	Output@OnStarvineWayTheCliftonSch	1
4	Input@OffAndrewsCirStarvineWay	1
5	Output@OnAndrewsCirStarvineWay	1
6	Input@OffAndrewsCirHealthScSouth	1
7	Output@OnAndrewsCirHealthScSouth	1
8	Input@OffRidgewoodUppergate	1
9	Output@OnRidgewoodUppergate	1
10	Input@OffUppergateCHOA	1
11	Output@OnUppergateCHOA	1
12	Input@OffWoodruffCir	1
13	Output@OnWoodruffCir	1
14	Input@OffUppergateWinshipCancerInst	1
15	Output@OnUppergateWinshipCancerInst	1
16	Input@OffUppergateRidgewood	1
17	Output@OnUppergateRidgewood	1
18	Input@OffAndrewsHealthScienceNorth	1
19	Output@OnAndrewsHealthScienceNorth	1
20	Input@OffStarvineWayTheCliftonSch2	1
21	Output@OnStarvineWayTheCliftonSch2	1
22	Input@OffStarvineDeck	1

Model: Bus Schedule

Buses with capacity 65

Four buses operate simultaneously

Depature every 5 mins



Transporter
Bus215
Bus315
Bus_415
Bus_515

Name	Description
StandardDay	Standard 8-5 Work Day

Work Periods

Start Time	Duration	End Time	Value	Cost Multiplier	Description
11:00 AM	4 hours	3:00 PM	1	1	

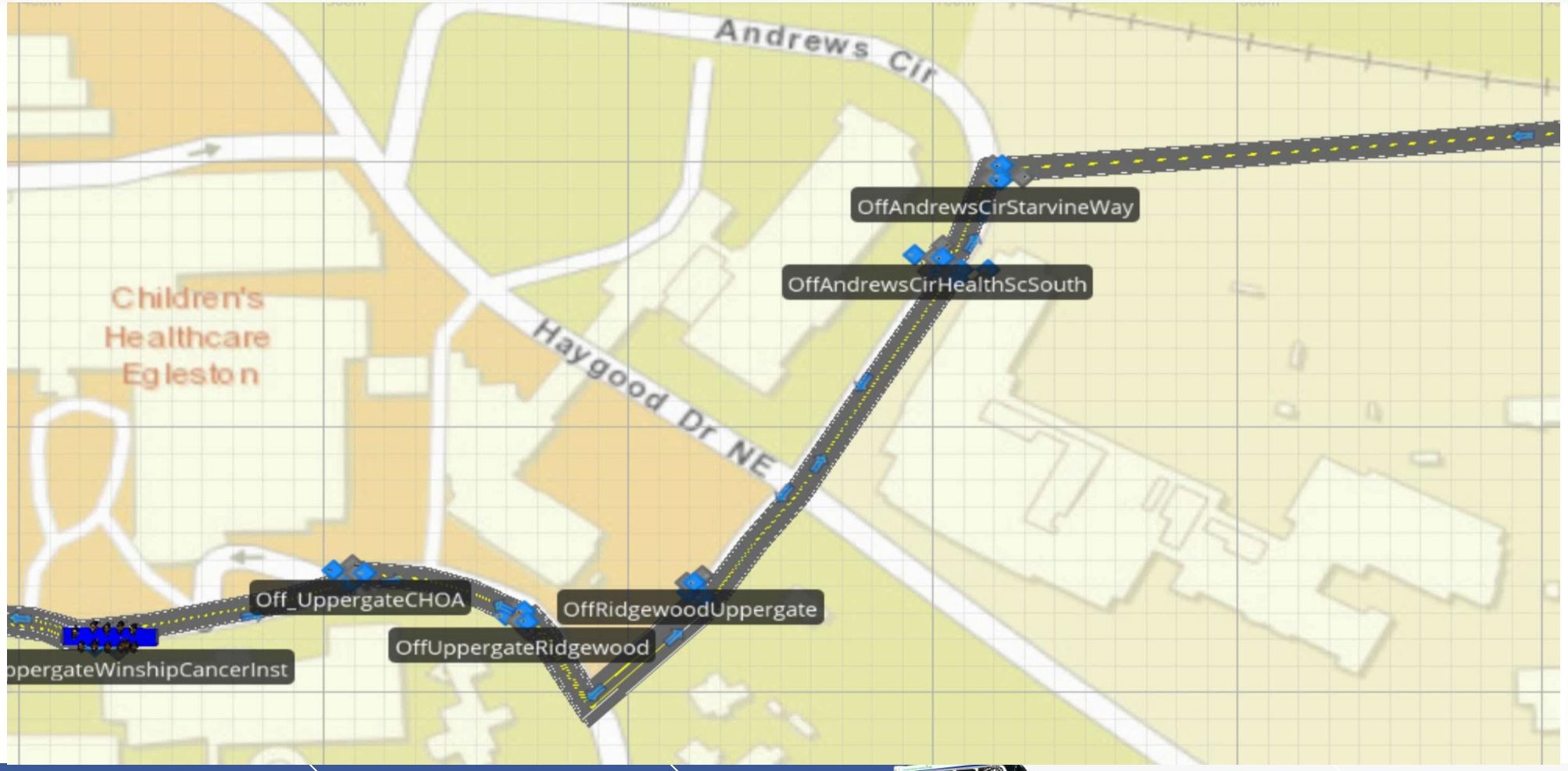
Name	Description
DayPattern1	

Work Periods

Start Time	Duration	End Time	Value	Cost Multiplier	Description
11:05 AM	235 minutes	3:00 PM	1	1	

Model: Simulation Video (1 of 2)

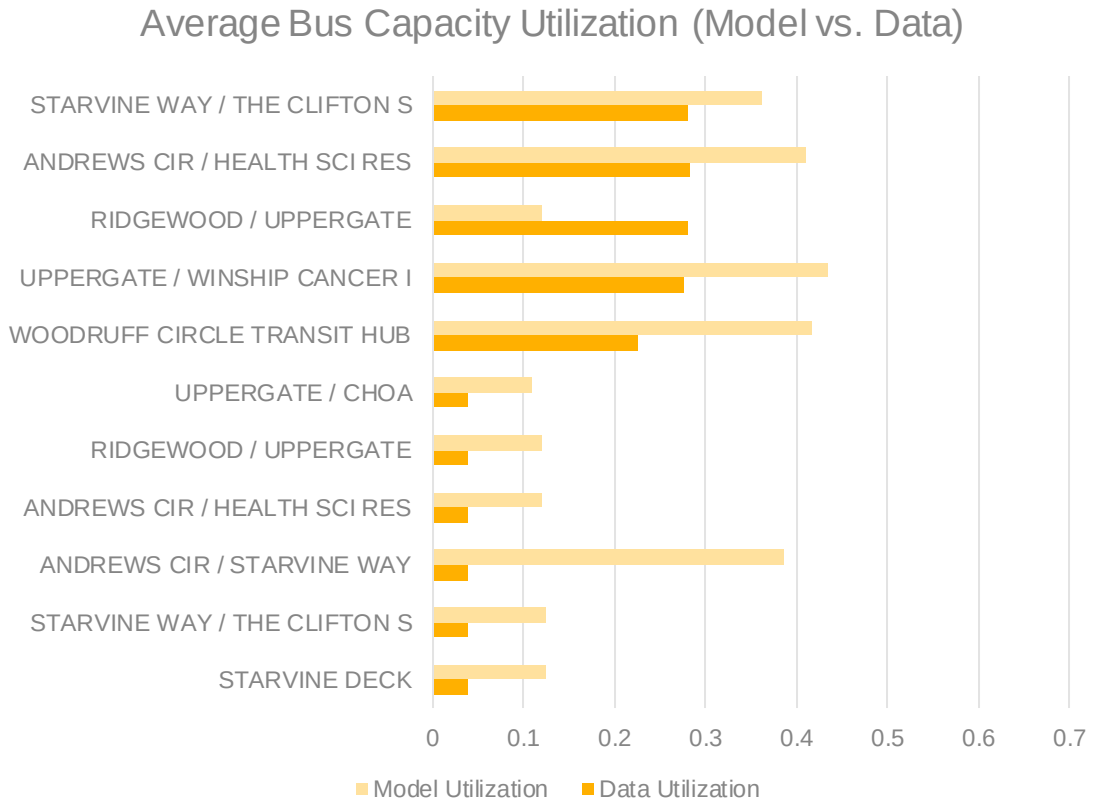
Model: Simulation Video (2 of 2)



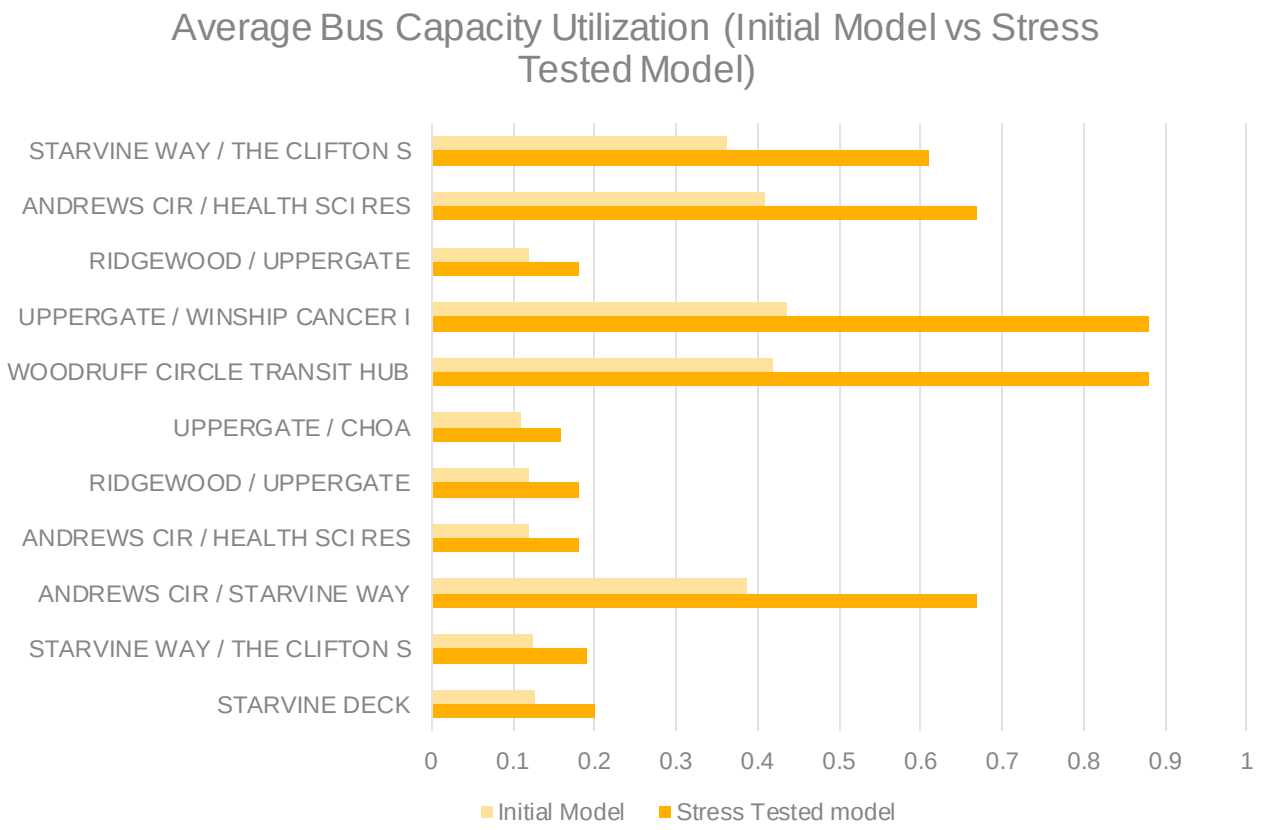
Comparing model results with actual data

Total # of passengers in Simulation vs. Data

	Model	Data
6-11 AM	3303	1618
11- 3 PM	1101	767
3-7 PM	2204	1491
7-11:59 PM	691	479



Model: Stress Test



What happens if passengers increase?

- We modified the inter-arrival rate at each stop (by 30%) to analyze demand increase
- The bus utilization after arriving at Woodruff Circle and Winship approached **90%** after this modification
- While this is a drastic change from the initial model (which was already somewhat over-estimating passenger demand on few stops), we still do not see the buses reach ~100% utilization
- We can conclude that a **full load** is probably not going to be a big program on Route C, if 4 buses keep running on weekdays
- However, any further increase in demand OR decrease in buses can be a problem at Route C

Other Considerations

Stress Testing

Demand Surge

- During a special event, the passenger at a specific stop will increase dramatically, and this might lead to a full-load situation
- In this case some passengers would end up unable to get on the bus and would have to wait for the next bus

Shortage of Bus

- In our model, there are four buses operating simultaneously.
- We would like to test if the system will operate normally when there's a shortage of bus
- This will be especially important because there might be some days where bus availability is less than usual

Accidents/ Road blocks

- When a road is influenced by construction or an accident, the travel time between stops might increase
- The model should be stress tested to see how throughput and utilization varies if travel time between certain stops were to increase

Future Improvements

Current limitations and probable modifications

Reducing simplifications

- Our model is currently based on a lot of assumptions
  - Number of buses is constant at 4
  - The model assumes certain time buckets, which may or may not be the most optimal choices
  - Model assumes similar behavior across Mon-Thu and excludes Fri-Sat, where surge could be expected
  - Assumes certain distributions (e.g. for time travel b/w stops, passenger arrivals) which may or may not be the best fit
- While some of these make sense to keep, others could be modified. For example, we can probably estimate # of buses running at each day, rather than using a constant
  - Further, we saw that number of buses running also varies by bucket (more buses at surge hours). This can be incorporated in the model

Model expansion

- We have 4 models, one for each bucket
- This is inconvenient, especially because each model starts from the initial stop. This would not be practical, since buses don't pause operations during the day
- A possible modification could be to use a single model for all buckets, and use rate tables to incorporate varying behavior across buckets
- Further, weekends can also be taken into consideration, but weekend model behavior would likely be different from our current weekday model, and may require two separate models
- Finally, The model is currently based on data for two months
- This is not ideal, especially to see how operations change seasonally. Using year long data can possibly improve model reliability over longer periods

Model Accuracy

- We compared our model with the actual data based on capacity utilization and throughput (number o passengers serviced)
- Our parameters are based on average values from the data, which is vulnerable to skewness
- This is especially true for some observations that may be recorded incorrectly, and they should be removed using outlier detection. Next, some levers can be modified to get results that more closely match the data.
- Finally, we would like to compare the model and the data based on more parameters, such as flowtime

ATLANTA RAPID MINERS

Thank You!
Questions?

EMORY UNIVERSITY

Appendix 1: Data Understanding

Data Source and Description

- Emory Shuttle Data from the Emory Transportation Department
- Includes 264214 instances of shuttle stop names, record date, record day, arrival and departure times, passenger on and off headcounts, bus number, longtitude and latitude. Each instance records a bus activity at a certain stop.
- Contains 13 attributes: 8 numerical and 5 categorical.
- There are no missing values or anomalies (such as not identified data)

Independent Variables

Bus Information (Categorical)

Route
Bus Number

Stop-Related Information

Stop Name (Categorical)
Bus Arrival/Departure Time for Each Stop
Passenger On/Off for Each Stop

Time Information

Date (Categorical)
Day (Categorical)
Date/Time (Categorical)

Bus-Related Information

Longtitude
Latitude

New Variables Needed

- To better understand the operating system, we need to do some data manipulation and generate some new variables for the model building process...

Appendix 2: Summary Statistics for Different Time Buckets

Time Bucket #1: 6:00 – 11:00

Row Labels	Average of Travel Time	Average of Percentage	Average of Time Interval	Average of # of passengers
1	0:00:53	0.071849234	0:00:48	14
2	0:01:20	0.006786575	0:01:14	0.094265515
3	0:00:26	0.000266986	0:02:38	0.007880221
4	0:01:13	0.079978658	0:01:19	0.123563218
5	0:00:32	0.045185558	0:01:12	0.042260961
6	0:02:00	0.374908304	0:02:24	0.163650878
7	0:01:41	0.724910606	0:01:52	0.854521625
8	0:00:52	0.057982567	0:02:35	0.116090147
9	0:00:59	0.018443452	0:01:03	0.028951487
10	0:01:58	0.073879216	0:01:07	0.034518828
11	0:01:49	0.028501983	0:00:54	0.002085506
Grand Total	**0:01:16**	**0.134235632**	**0:01:33**	**1.395324527**

Time Bucket #2: 11:00 – 15:00

Row Labels	Average of Travel Time	Average of Percentage	Average of Time Interval	Average of # of passengers
1	0:01:01	0.432473964	0:01:51	4.827681027
2	0:01:25	0.003371816	0:01:44	0.068121104
3	0:00:27	0	0:03:36	0.001353791
4	0:01:10	0.036837711	0:01:46	0.148839326
5	0:00:30	0.018750502	0:01:41	0.012974977
6	0:02:04	0.081392189	0:03:35	0.040201005
7	0:01:41	0.419770344	0:02:10	3.408602151
8	0:00:56	0.016555961	0:03:30	0.81839294
9	0:01:03	0.006672671	0:01:39	0.131313131
10	0:01:59	0.02771073	0:01:42	0.186057248
11	0:00:56	0.011995467	0:01:23	0.0087236
Grand Total	**0:01:13**	**0.094810547**	**0:02:14**	**0.863012003**

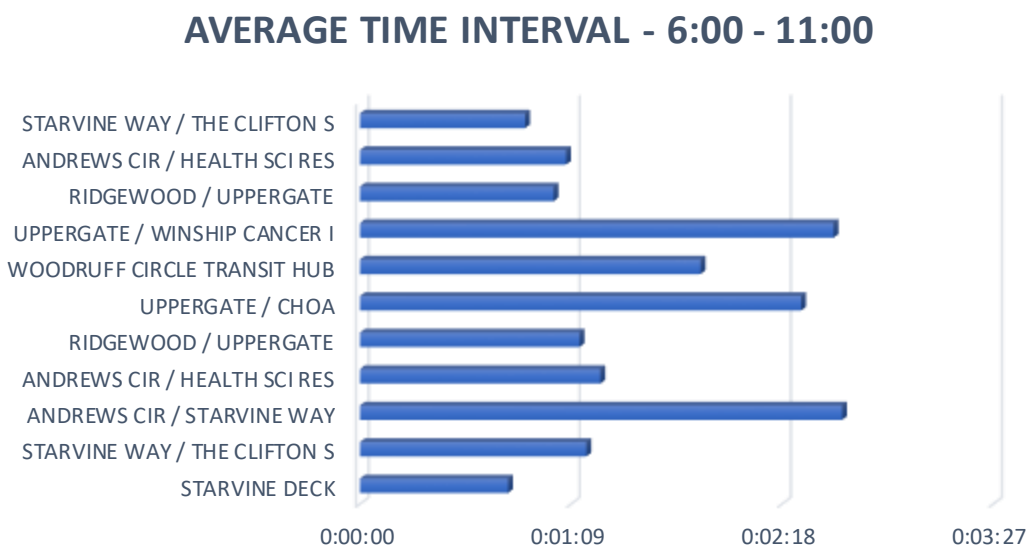
Time Bucket #3: 15:00 – 19:00

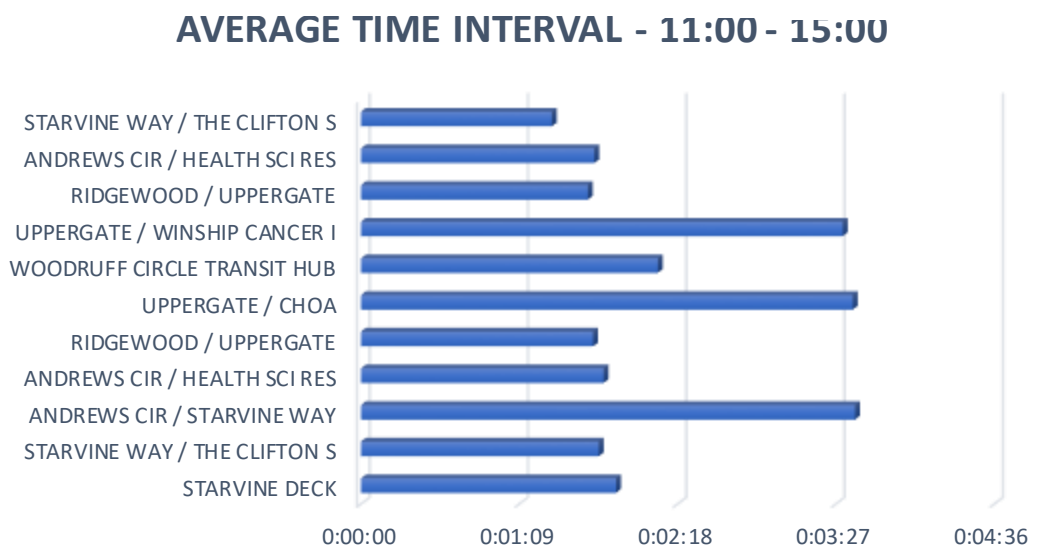
Row Labels	Average of Travel Time	Average of Percentage	Average of Time Interval	Average of # of passengers
1	0:01:06	0.787205649	0:01:32	1.887903512
2	0:01:29	0.002391023	0:01:11	0.120902362
3	0:00:27	0.000175685	0:02:49	0.00035137
4	0:01:13	0.012962834	0:01:18	0.158898305
5	0:00:26	0.008312525	0:01:12	0.021048744
6	0:02:31	0.02543551	0:02:57	0.042442293
7	0:02:03	0.212514188	0:01:27	7.574749642
8	0:00:51	0.005653226	0:02:16	3.984216065
9	0:01:01	0.001873369	0:01:12	0.454225352
10	0:02:01	0.0073806	0:01:22	0.675017643
11	0:00:53	0.007775709	0:01:07	0.05334728
Grand Total	**0:01:19**	**0.098508082**	**0:01:40**	**1.362620668**

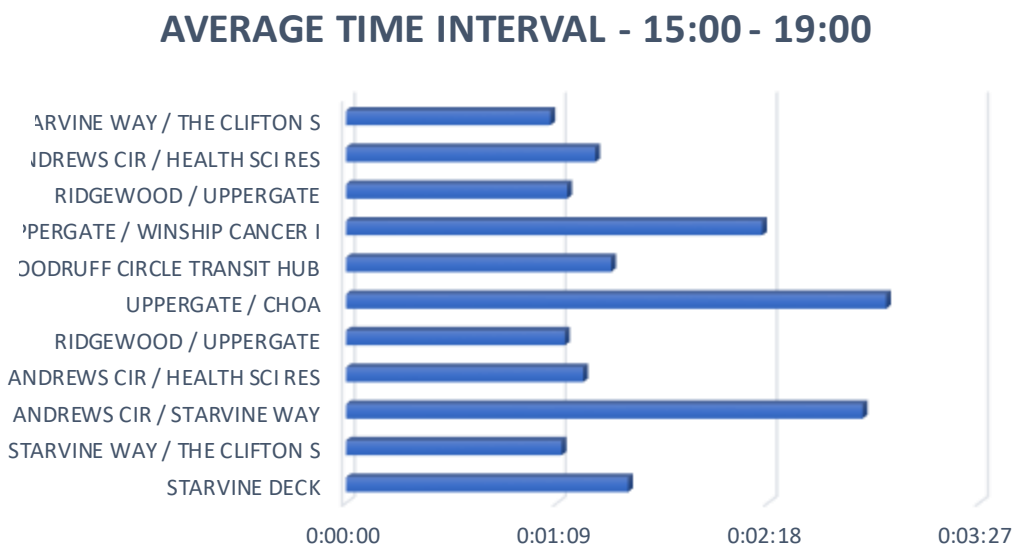
Time Bucket #4: 19:00 – 0:00

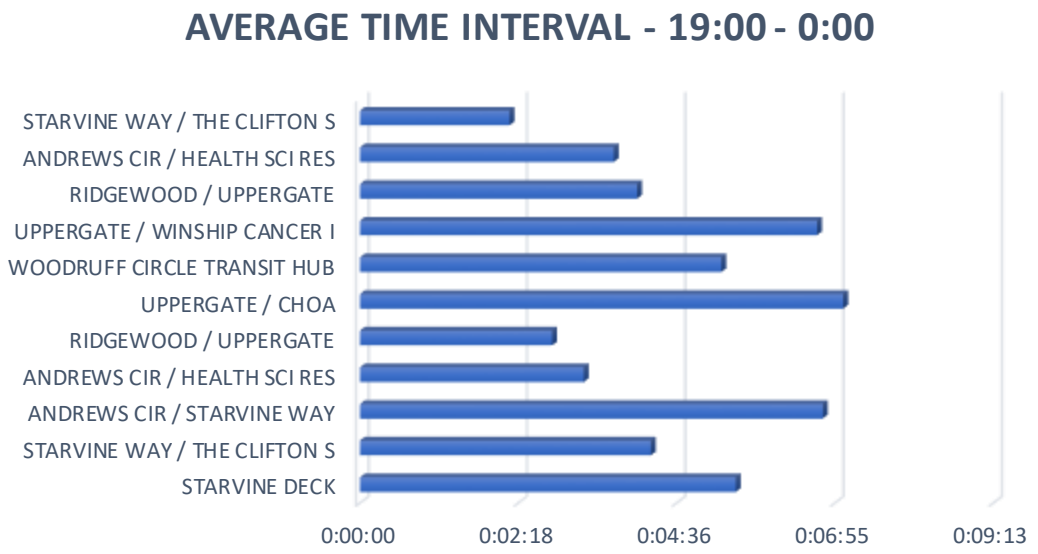
Row Labels	Average of Travel Time	Average of Percentage	Average of Time Interval	Average of # of passengers
1	0:01:10	0.798681627	0:05:28	1.573405073
2	0:01:30	0.00148258	0:04:14	0.002965159
3	0:00:26	0.000372301	0:06:44	0.000744602
4	0:00:41	0.012020817	0:03:15	0.025894897
5	0:00:23	0.002401922	0:02:48	0.008006405
6	0:01:46	0.008758963	0:07:02	0.027027027
7	0:01:19	0.15970926	0:05:15	8.78960195
8	0:00:47	0.003266596	0:06:39	2.047768207
9	0:00:32	0.000696881	0:04:02	0.125796178
10	0:01:58	0.003137474	0:03:42	0.211145997
11	0:00:48	0.000444283	0:02:11	0.001529052
Grand Total	**0:01:03**	**0.089138932**	**0:04:37**	**1.093790295**

Appendix 3: Average Time Interval at Each Stop for Each Time Bucket









Appendix 4: Average Travel Time at Each Stop for Each Time Bucket