



Introduction to the Trace Model

Catherine Vanderwaart
July 24, 2018

Washington Metropolitan Area Transit Authority

Tour of Trace

1. Background
2. Components of Trace: Origin, Destination, Transfer, and Path Inference
3. Sample Day
4. Rail Examples and Uses
5. Bus Examples and Uses
6. Trace Dashboard
7. Future Enhancements



What is the Trace Model?

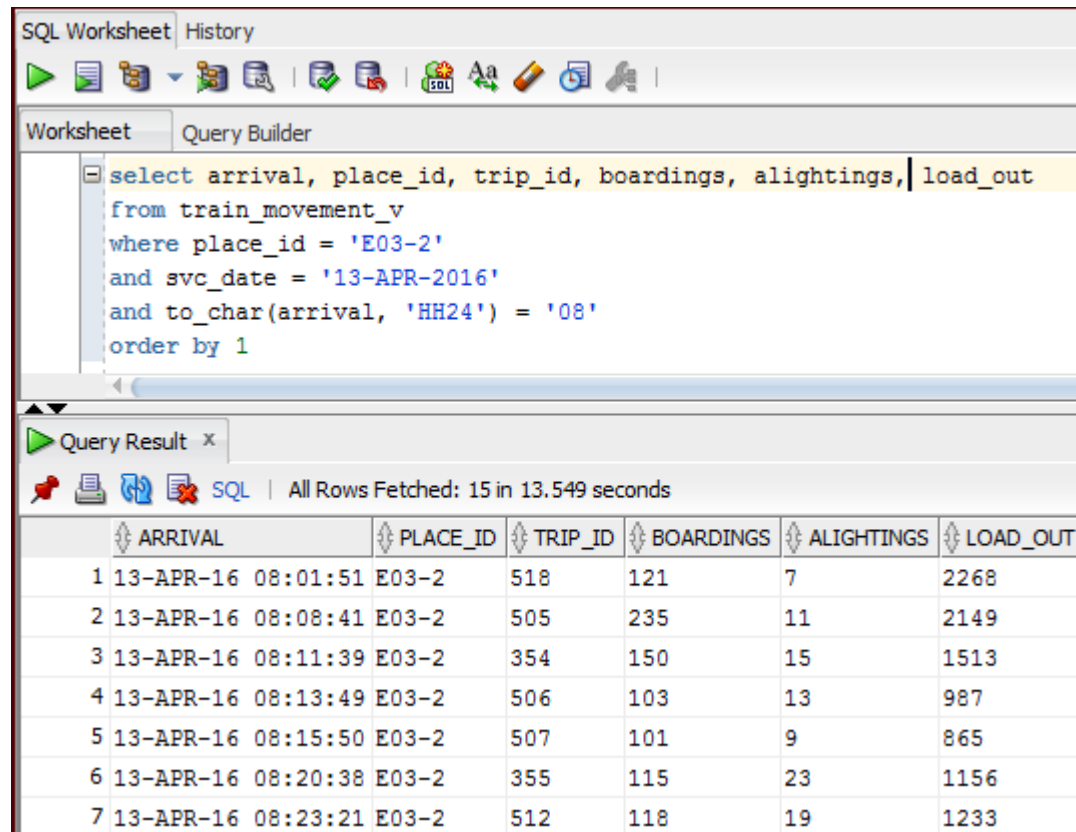
Why Are We Here? What We Don't Know:

Existing data sources don't give us a deep, robust understanding of how our passengers move through the system, and the conditions they experience.



Introducing the Trace Model

We now have very detailed data on passenger movements for bus and rail



The screenshot shows an SQL Worksheet interface. The top part contains a query editor with the following SQL code:

```
select arrival, place_id, trip_id, boardings, alightings, load_out
from train_movement_v
where place_id = 'E03-2'
and svc_date = '13-APR-2016'
and to_char(arrival, 'HH24') = '08'
order by 1
```

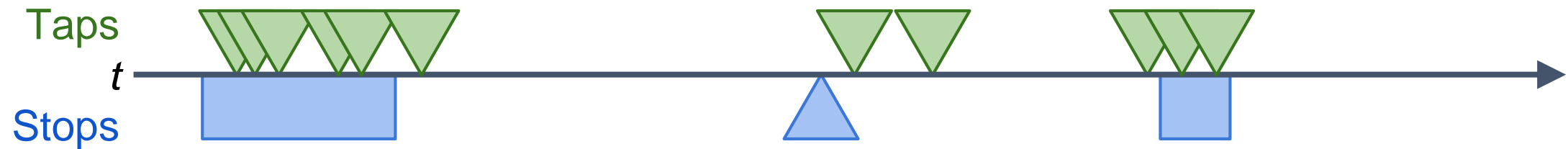
The bottom part of the screenshot shows the query results in a table format. The table has 7 columns: ARRIVAL, PLACE_ID, TRIP_ID, BOARDINGS, ALIGHTINGS, and LOAD_OUT. The results are as follows:

	ARRIVAL	PLACE_ID	TRIP_ID	BOARDINGS	ALIGHTINGS	LOAD_OUT
1	13-APR-16 08:01:51	E03-2	518	121	7	2268
2	13-APR-16 08:08:41	E03-2	505	235	11	2149
3	13-APR-16 08:11:39	E03-2	354	150	15	1513
4	13-APR-16 08:13:49	E03-2	506	103	13	987
5	13-APR-16 08:15:50	E03-2	507	101	9	865
6	13-APR-16 08:20:38	E03-2	355	115	23	1156
7	13-APR-16 08:23:21	E03-2	512	118	19	1233

- Data service with nightly updates
 - Vendor is Korbato
- We can query via SQL statements

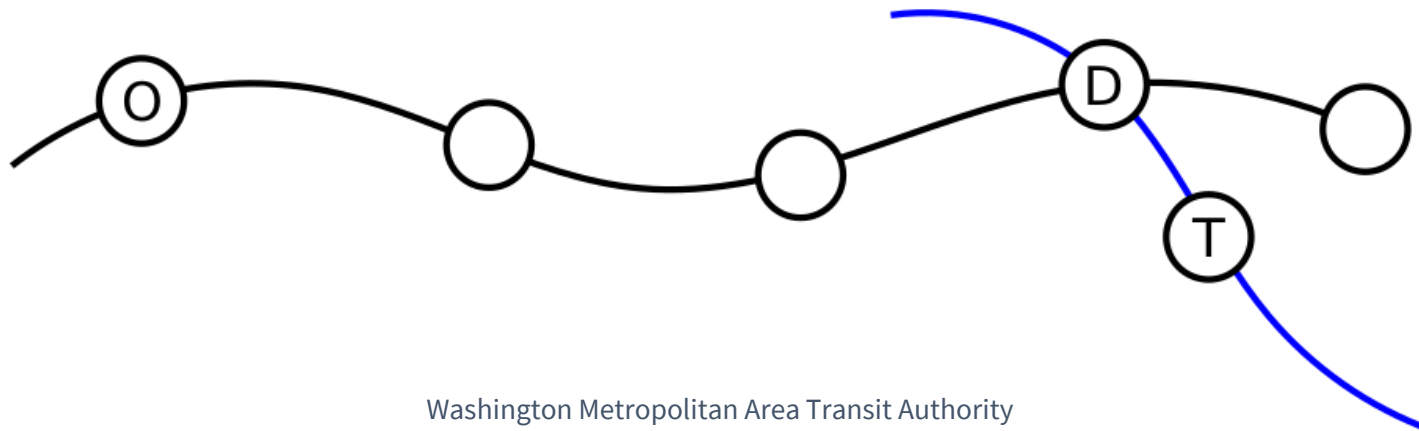
Origin Inference

- Bus: match fare transaction to stop visit by farebox ID and timestamp
- Rail: origin typically known, otherwise infer (next slide)

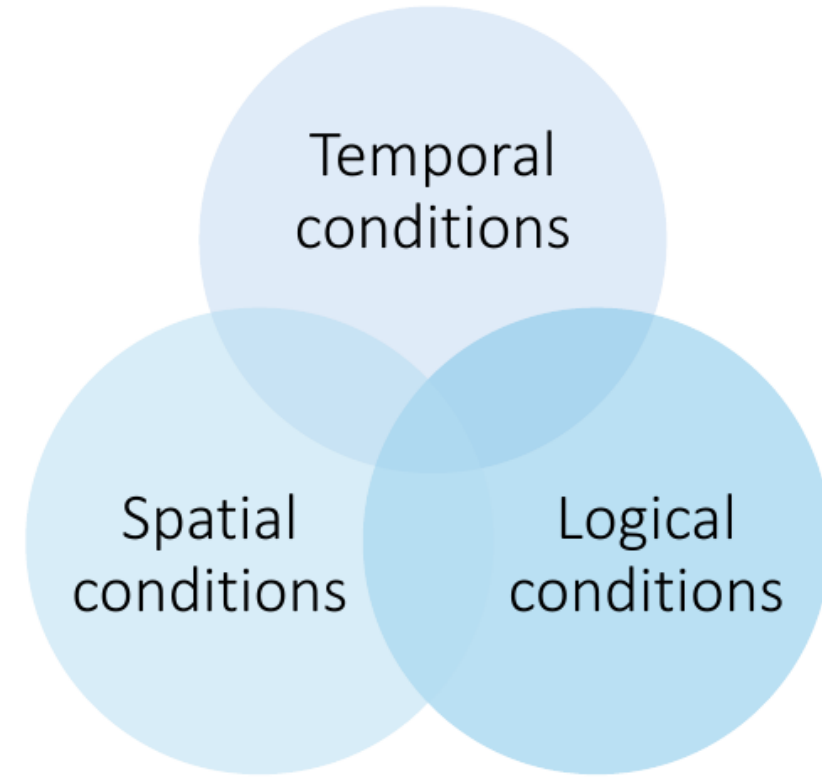
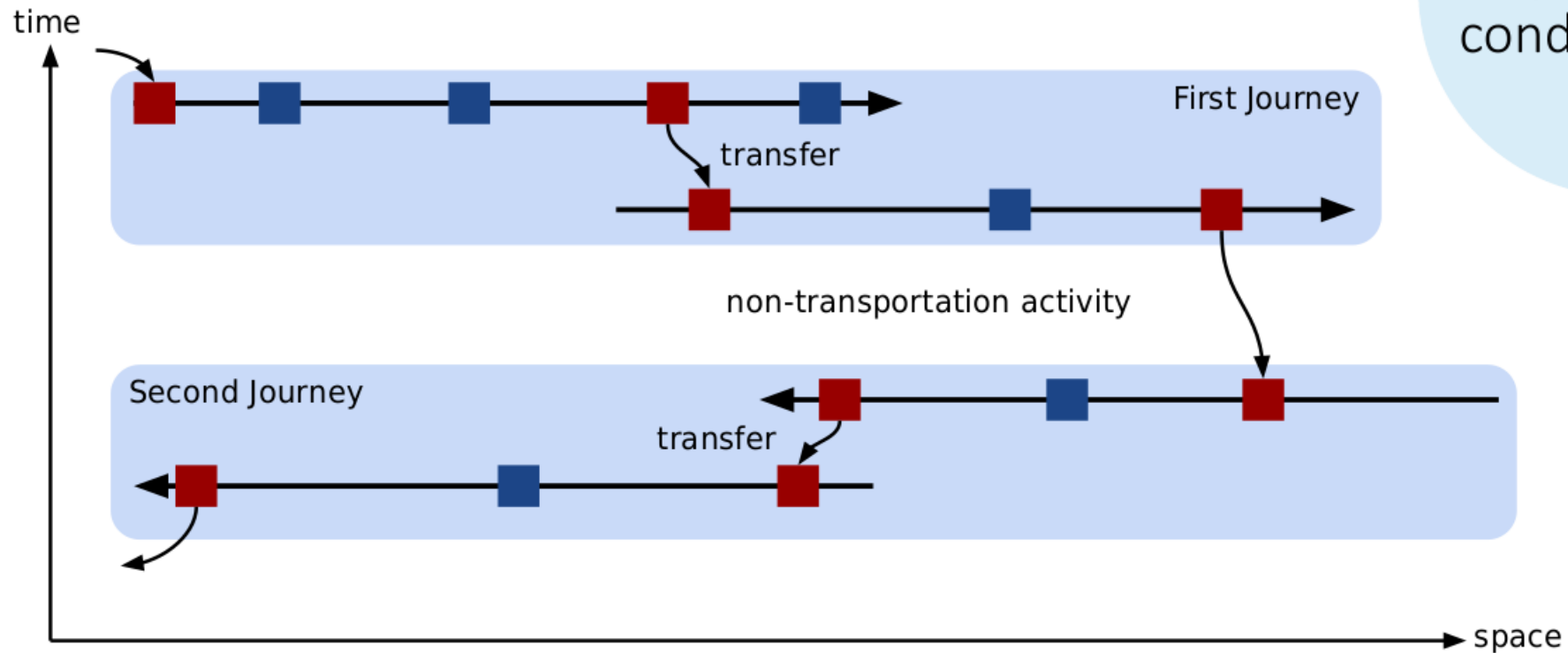


Destination Inference

- Bus: minimize generalized disutility from boarding stop to next tap's location, considering:
 - the destination of many trip segments is close to the origin of the following trip segment.
 - assume no intermediate non-AFC trip segment (private auto, rideshare, cash bus, etc.)
 - assume passengers will not walk a long distance
 - assume last trip of a day ends at the origin of the first trip of the day
- Rail: destination typically known, but missing D (or O) can be inferred like bus.

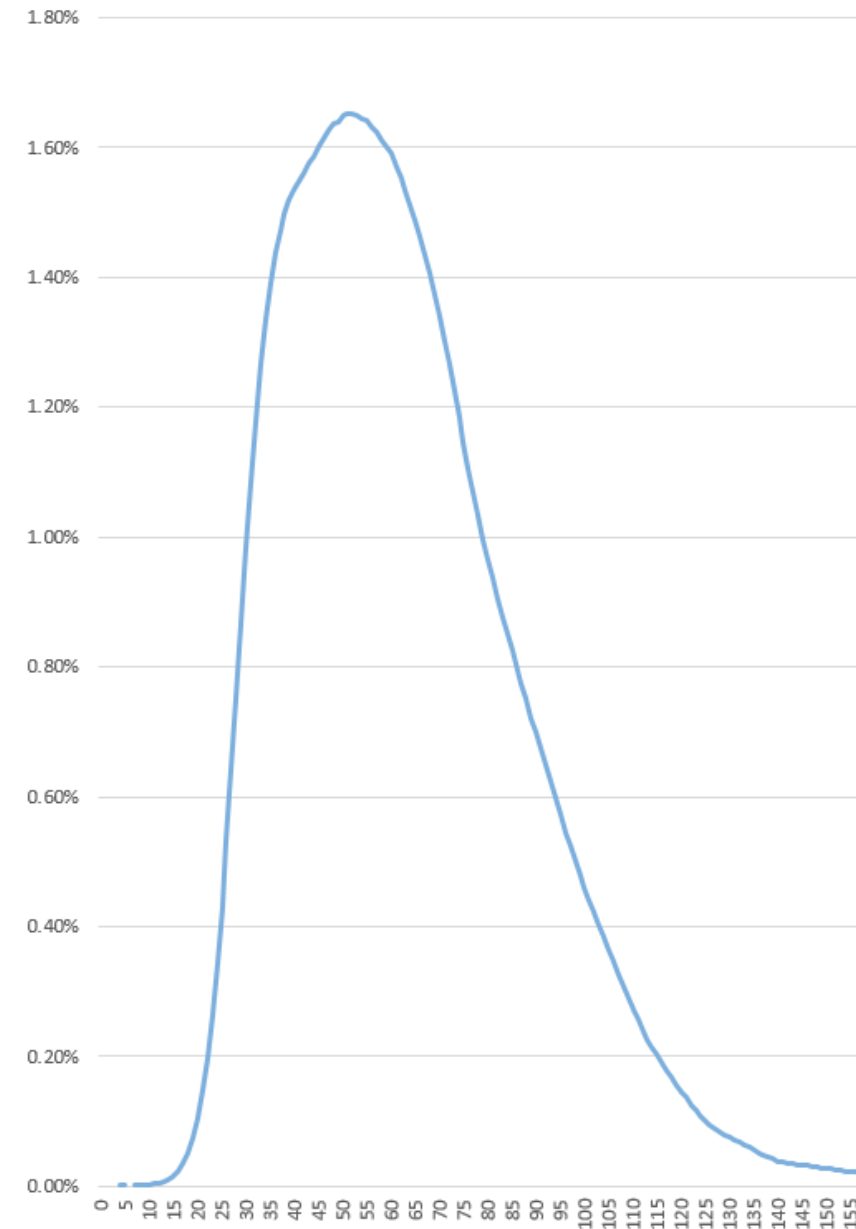


Transfer Inference



Path Inference

- Using AFC and cleaned rail AVL data, consider all feasible combinations of observed train rides (spacetime paths) that each cardholder might have taken.
- Estimate the probability of each feasible spacetime path using probability distributions for each gated pedestrian movement (access, transfer, egress).
 - Probability distributions are calculated for each mezz → platform, platform → platform, and platform → mezz pair, for various time periods, day types, and seasons.
- Inferred paths enable probabilistic train loads, denied boarding counts, platform occupancy, and vertical circulation



Example: One Rider's Day

Fare Stages

jny_seq	stage_seq	Route	origin	destination	origin_time	dest_time
1	1	29G	1795	14510	06:36:23	06:48:52
1	2	Yellow	Pentagon	L'Enfant, N	06:51:15	07:04:16
2	1	Yellow	L'Enfant N	Pentagon	16:01:12	16:08:52
2	2	29G	13320	1797	16:17:59	16:30:51

Morning Bus Ride Boarding

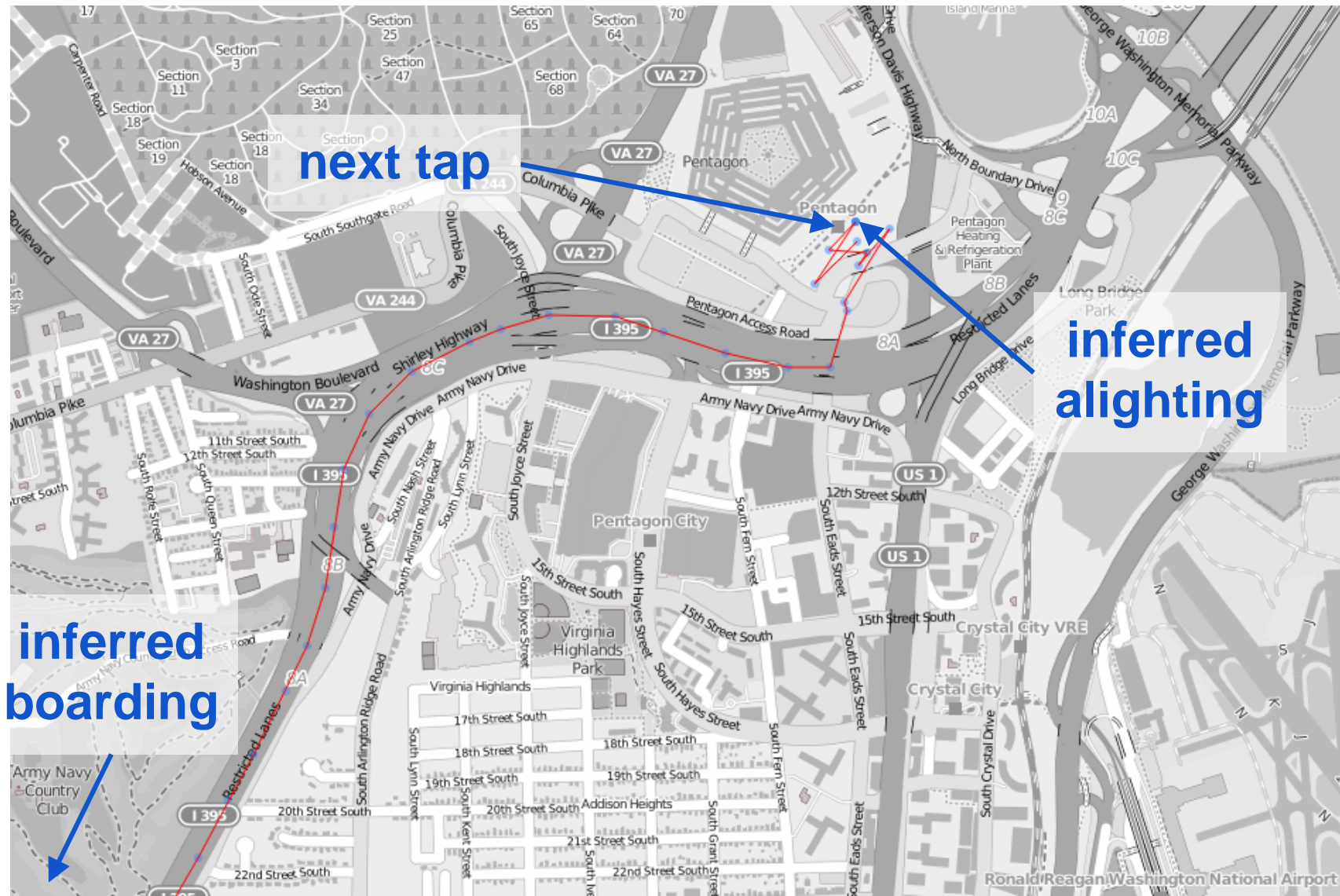
device_id	transaction_dtm	operator_id	bus_id
text	timestamp without time zone	integer	integer
FBX02596	2016-04-13 06:36:25	21	2596

← AFC

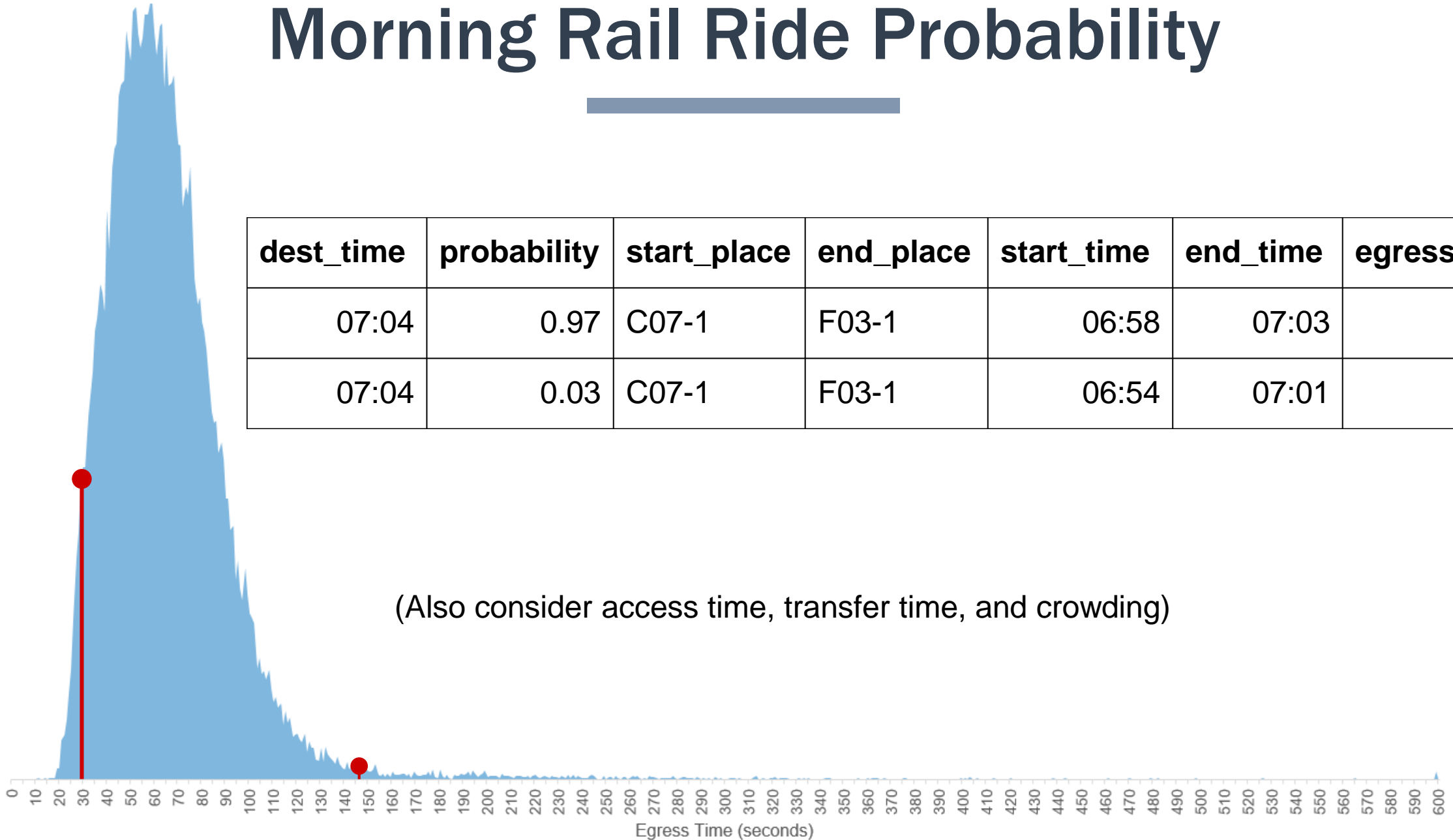
seq_in_day	pattern_id	seq_in_pattern	trip_id	place_id	arrival	departure
integer	text	integer	integer	text	timestamp with time zone	timestamp with time zone
24	29G01	23	1002562	5000684	2016-04-13 06:30:31-04	2016-04-13 06:30:31-04
25	29G01	24	1002562	5000669	2016-04-13 06:30:54-04	2016-04-13 06:30:54-04
26	29G01	25	1002562	5000659	2016-04-13 06:31:10-04	2016-04-13 06:31:10-04
27	29G01	26	1002562	5000653	2016-04-13 06:31:25-04	2016-04-13 06:31:31-04
28	29G01	27	1002562	5000637	2016-04-13 06:32:08-04	2016-04-13 06:32:08-04
29	29G01	28	1002562	5000618	2016-04-13 06:32:33-04	2016-04-13 06:32:33-04
30	29G01	29	1002562	5000605	2016-04-13 06:32:47-04	2016-04-13 06:32:47-04
31	29G01	30	1002562	5000593	2016-04-13 06:33:14-04	2016-04-13 06:33:14-04
32	29G01	31	1002562	5000587	2016-04-13 06:34:40-04	2016-04-13 06:34:51-04
33	29G01	32	1002562	5000583	2016-04-13 06:35:28-04	2016-04-13 06:35:34-04
34	29G01	33	1002562	5000578	2016-04-13 06:36:23-04	2016-04-13 06:36:28-04
35	29G01	34	1002562	[null]	2016-04-13 06:48:52-04	2016-04-13 06:48:52-04

AVL →

Morning Bus Ride Alighting



Morning Rail Ride Probability



dest_time	probability	start_place	end_place	start_time	end_time	egress_sec
07:04	0.97	C07-1	F03-1	06:58	07:03	30
07:04	0.03	C07-1	F03-1	06:54	07:01	146

(Also consider access time, transfer time, and crowding)



Afternoon Rail Ride

dest_time	probability	start_place	end_place	start_time	end_time	egress_sec
04:08 PM	1	F03-2	C07-2	4:02 PM	4:07 PM	54

Afternoon Bus Ride Boarding

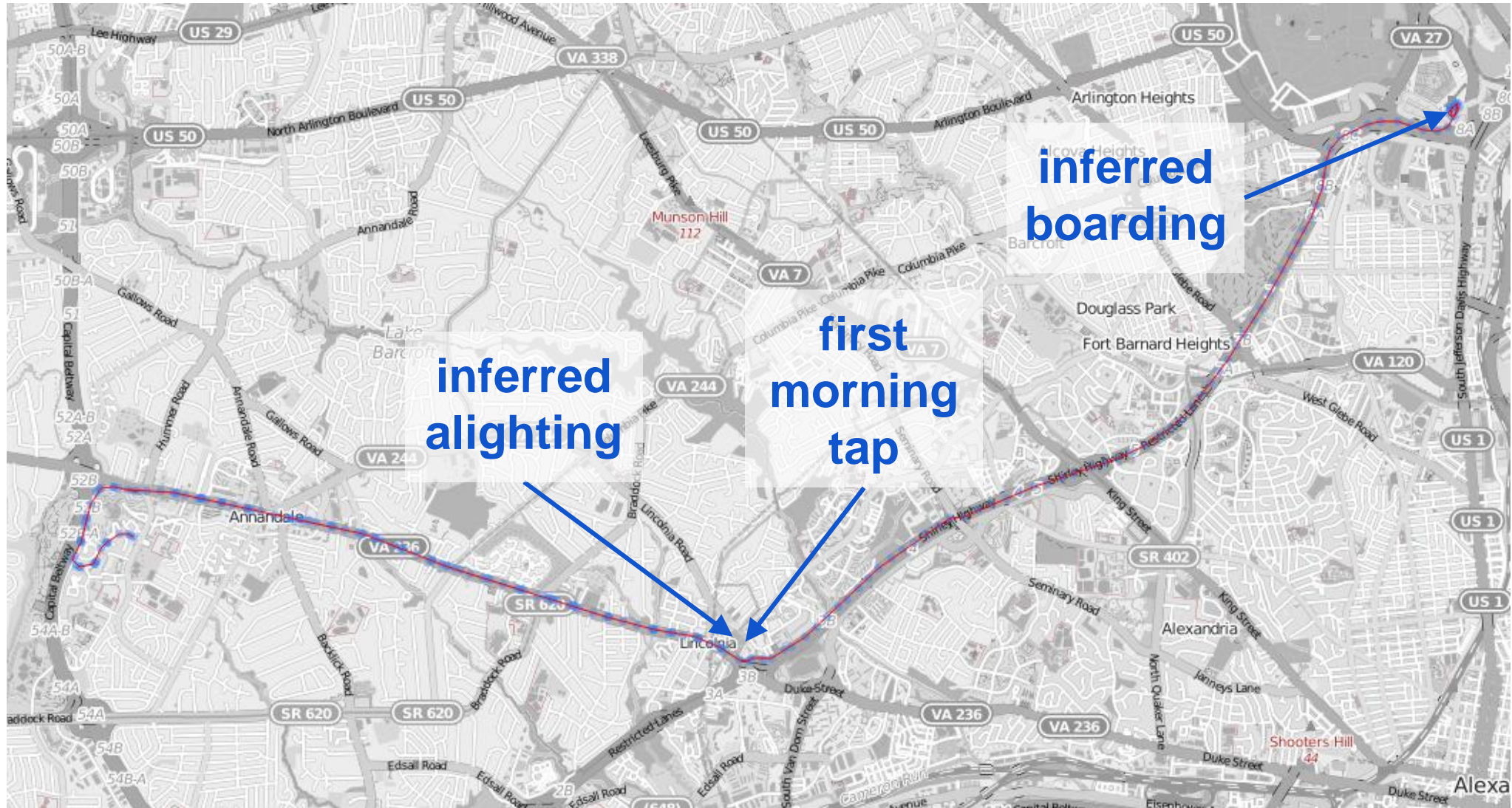
device_id text	transaction_dtm timestamp without time zone	operator_id integer	bus_id integer
FBX02534	2016-04-13 16:17:28	21	2534

← AFC

seq_in_day integer	pattern_id text	seq_in_pattern integer	trip_id integer	place_id text	arrival timestamp with time zone	departure timestamp with time zone
451	29G02	1	1001355	6000906	2016-04-13 16:17:59-04	2016-04-13 16:17:59-04
452	29G02	2	1001355	[null]	2016-04-13 16:18:59-04	2016-04-13 16:18:59-04
454	29G02	4	1001355	5000577	2016-04-13 16:30:51-04	2016-04-13 16:31:46-04
455	29G02	5	1001355	5002021	2016-04-13 16:32:26-04	2016-04-13 16:32:44-04
456	29G02	6	1001355	5000590	2016-04-13 16:33:23-04	2016-04-13 16:33:24-04
457	29G02	7	1001355	5000594	2016-04-13 16:35:19-04	2016-04-13 16:35:20-04
458	29G02	8	1001355	5000606	2016-04-13 16:36:11-04	2016-04-13 16:36:12-04
459	29G02	9	1001355	5000622	2016-04-13 16:36:57-04	2016-04-13 16:36:58-04
461	29G02	10	1001355	5000640	2016-04-13 16:38:27-04	2016-04-13 16:38:27-04
462	29G02	11	1001355	5000656	2016-04-13 16:39:36-04	2016-04-13 16:39:40-04

AVL →

Afternoon Bus Ride Alighting





Uses for Trace Data

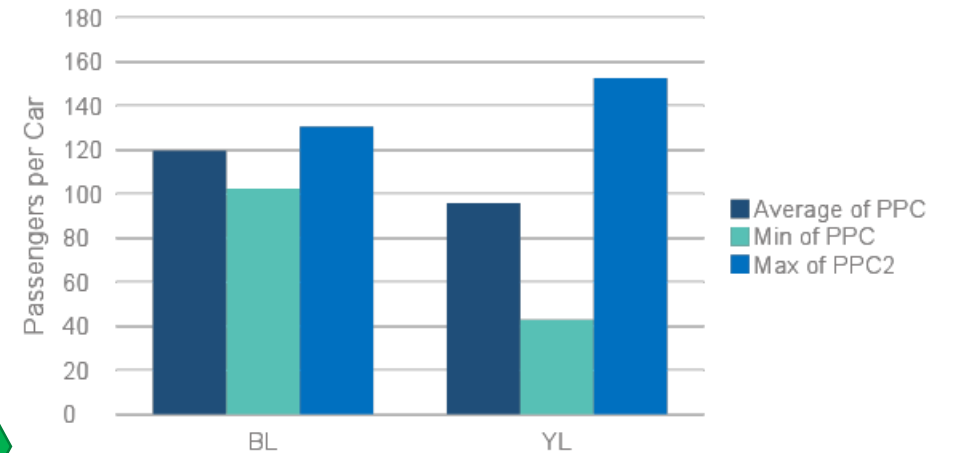
Rail Use Example: Link Loads

Trace Model Data:

1	place_id	arrival	TTI	train_id	boarding	alightings	load_out	Length	PPC	Color
2	C07-1	4/13/2016 8:01	193	305	214	70	520	6	87	YL
3	C07-1	4/13/2016 8:03	232	359	68	95	258	6	43	YL
4	C07-1	4/13/2016 8:08	159	309	143	219	599	6	100	YL
5	C07-1	4/13/2016 8:10	309	403	381	203	782	6	130	BL
6	C07-1	4/13/2016 8:13	58	306	238	52	660	6	110	YL
7	C07-1	4/13/2016 8:16	286	360	112	86	468	6	78	YL
8	C07-1	4/13/2016 8:20	172	310	247	197	754	6	126	YL
9	C07-1	4/13/2016 8:22	313	404	380	185	671	6	112	BL
10	C07-1	4/13/2016 8:25	161	313	180	64	679	6	113	YL
11	C07-1	4/13/2016 8:27	288	361	90	97	485	6	81	YL
12	C07-1	4/13/2016 8:31	45	302	138	216	560	6	93	YL
13	C07-1	4/13/2016 8:33	370	405	406	178	740	6	123	BL
14	C07-1	4/13/2016 8:38	37	311	293	83	916	6	153	YL
15	C07-1	4/13/2016 8:41	326	351	182	118	531	6	89	YL
16	C07-1	4/13/2016 8:44	270	303	143	175	564	6	94	YL
17	C07-1	4/13/2016 8:47	349	406	492	191	783	6	131	BL
18	C07-1	4/13/2016 8:52	317	312	253	67	691	6	115	YL
19	C07-1	4/13/2016 8:55	283	352	143	90	446	6	74	YL
20	C07-1	4/13/2016 8:57	271	304	86	170	499	6	83	YL
21	C07-1	4/13/2016 8:59	32	407	362	133	614	6	102	BL

PPC Statistics:

Rail Crowding at Pentagon, by Line
4/13/2016, 8am peak hour



Link Volumes:

Row Labels	Trains per Hour	Link Volume: (Sum of load_out)
C07-1	20	12,220
E02-2	16	6,959
F04-1	10	5,118
K01-1	19	11,395

Other Uses for Rail

- **DATA on the Customer Experience**
 - How crowded is a platform at peak times?
 - Vertical circulation - how utilized is an escalator? Are we violating our standards?
 - How many passengers experienced crowding conditions over X?
- **Behind-the-Gates Transfer Volumes:** what is causing platform crowding at a transfer station - entries and exits at that station, or within-station transfers?
- **Route Choice:** Are Yellow Line riders switching back to Blue since July?
- **Train-Level Ridership**
- **Cleaned Train Movement Data:** train runs
- **Platform for Future Research:** trains too full to board, dwell time analysis, linking delays to incidents, predicting crowding in real-time, effects of offloads, denied boardings, etc....

Bus Use Example: O-D Matrix

Route 16A East, 8 - 9 a.m. on April 13, 2016

	To	Annandale	Rte 236 & Medford	John Marr & Columbia Pk	Columbia Pk & Lincoln	Columbia Pk & Carlin	Columbia Pk & Fmr	Columbia Pk & Geo Mason	Columbia Pk & Walt Reed	Columbia Pk & S Orme	
From		Rte 236 & Medford	John Marr & Columbia Pk	Columbia Pk & Lincoln	Columbia Pk & Carlin	Columbia Pk & Fmr	Columbia Pk & Geo Mason	Columbia Pk & Walt Reed	Columbia Pk & S Orme	Pentagon	Total
Annandale	Rte 236 & Medford	1.5	0.0	1.2	5.7	1.1	0.0	8.7	0.0	0.7	19.0
Rte 236 & Medford	John Marr & Columbia Pk		0.0	0.0	1.2	0.0	1.1	2.4	0.0	3.0	7.8
John Marr & Columbia Pk	Columbia Pk & Lincoln			0.0	3.4	2.2	0.0	0.0	0.0	1.1	6.7
Columbia Pk & Lincoln	Columbia Pk & Carlin				3.7	0.0	0.0	2.2	0.0	11.6	17.5
Columbia Pk & Carlin	Columbia Pk & Fmr					1.4	0.0	4.0	0.1	6.7	12.2
Columbia Pk & Fmr	Columbia Pk & Geo Mason						0.0	0.0	0.0	2.2	2.2
Columbia Pk & Geo Mason	Columbia Pk & Walt Reed							0.2	1.2	0.8	2.2
Columbia Pk & Walt Reed	Columbia Pk & S Orme								0.0	12.3	12.3
Columbia Pk & S Orme	Pentagon									7.8	7.8
Total		1.5	0.0	1.2	14.0	4.8	1.2	17.6	1.3	46.2	87.7

Bus Use Example: O-D Matrix

Route 70 South, 8 - 9 a.m. on April 13, 2016

	To	Silver Spring	Georgia & Eastern	Georgia & Kennedy	Georgia & Petworth	Georgia & Irving	7th & Florida	7th & H	7th & Penn	
From		Georgia & Eastern	Georgia & Kennedy	Georgia & Petworth	Georgia & Irving	7th & Florida	7th & H	7th & Penn	Archives	Total
Silver Spring	Georgia & Eastern	13.1	8.5	17.6	0.0	1.7	4.9	1.1	0.0	46.9
Georgia & Eastern	Georgia & Kennedy		28.2	83.7	6.4	16.4	17.9	3.7	0.0	156.3
Georgia & Kennedy	Georgia & Petworth			37.0	3.2	23.0	32.0	12.0	0.0	107.2
Georgia & Petworth	Georgia & Irving				2.2	9.1	16.4	3.5	0.0	31.2
Georgia & Irving	7th & Florida					8.5	21.3	6.0	0.0	35.8
7th & Florida	7th & H						15.0	10.3	3.7	29.0
7th & H	7th & Pennsylvania							0.6	0.6	1.1
7th & Pennsylvania	Archives								0.0	0.0
Total		13.1	36.7	138.4	11.8	58.7	107.5	37.2	4.3	407.6

Other Uses for Metrobus

- **Directional transfer data:** when people transfer between the S2 and the H2, which way are they traveling on each route?
- **Boardings by stop and fare instrument:** where are people using UPass? Which stops do people with disabilities use most? Where do people pay in cash?
- **Transfers on passes:** pass transfers are not included in our current transfer counts
- **Follow people over time:** how do the same passengers use a local and a MetroExtra route over time, and at which stops? Do they ride more or less after a corridor is restructured?
- **Bus OD:** more accurate and granular than existing estimation and survey methods, including data by trip or hour, not just time period
- **Full system OD:** what are the ultimate destinations of people boarding at a specific stop? When people transfer from bus to rail at Fort Totten, are they transferring to Red or Green?



Visualizing Trace Data

Trace Dashboard: Bus Trips to Union Station

Tableau interface showing navigation and tool options. The breadcrumb trail reads: Home > Default > Origin-Destination Tools from Trace > Zone Ride OD Dashboard. The top navigation bar includes Undo, Redo, Revert, Refresh, and Pause. The right side of the bar contains View: Original, Alert, Share, Download, Comments, and Full Screen. Below the navigation bar, a series of tabs are visible: Zone Ride OD Dashboard (selected), Stop Ride OD Dashboard, Stop Ride OD Table, Zone Ride OD Table, Stop Journey OD Dashboard, and Zone Journey OD Dashboard.

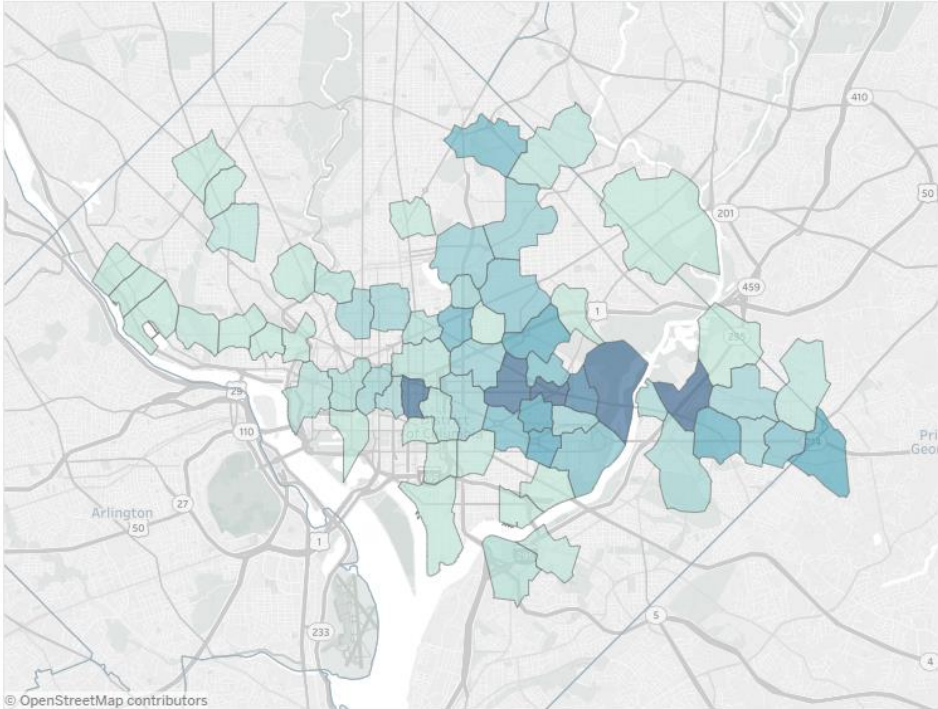
Zone Ride OD Dashboard

See where customers ended rides on their journey, given where they started them, or vice versa.

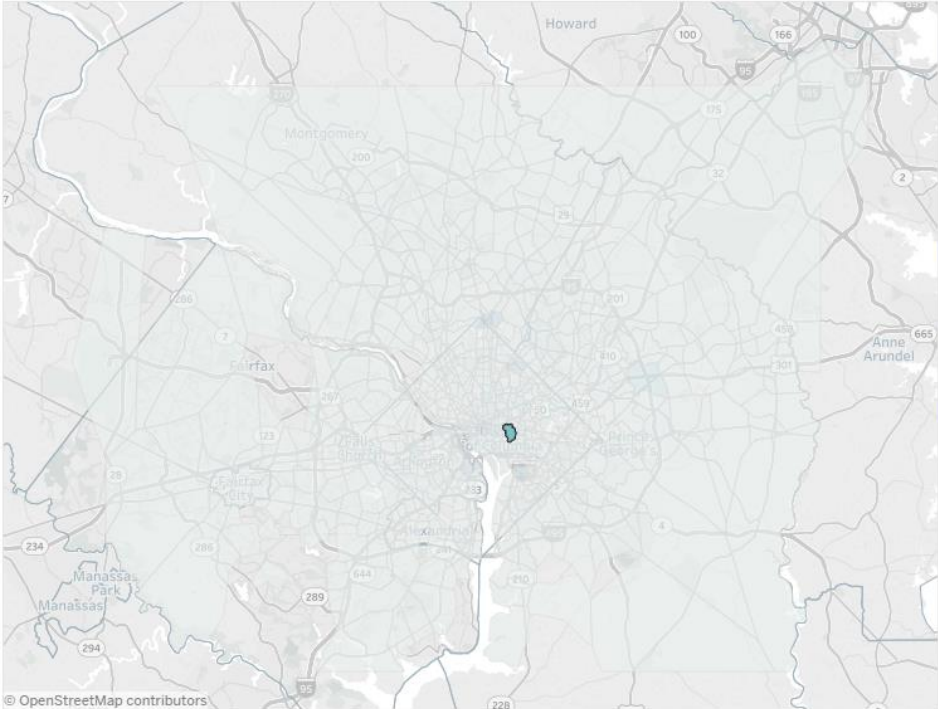
This dashboard summarizes passenger rides (unlinked rail and bus trips) for which both origin and destination are available. This means that the number of passengers travelling by bus will appear lower than it actually is.

Data from Trace Model v. 3.0, averaging across May 7 to 11, 2018.

Select an origin zone or a group of origin zones



Select a destination zone or a group of destination zones



Passengers
0.20 61.20

Filter by any combination of the following:

Subsystem

- (All)
- Metrobus
- Metrorail

Hour starting at

6 9

Bus Route or Rail Line

(All)

Direction

(All)



Trace Dashboard: All Journeys to Union Station

Home > Default > Origin-Destination Tools from Trace > Zone Journey OD Dashboard ☆ 1

← Undo → Redo ↶ Revert ↻ Refresh ⏸ Pause

View: Original Alert Share Download Comments Full Screen

Zone Ride OD Dashboard Stop Ride OD Dashboard Stop Ride OD Table Zone Ride OD Table Stop Journey OD Dashboard **Zone Journey OD Dashboard**

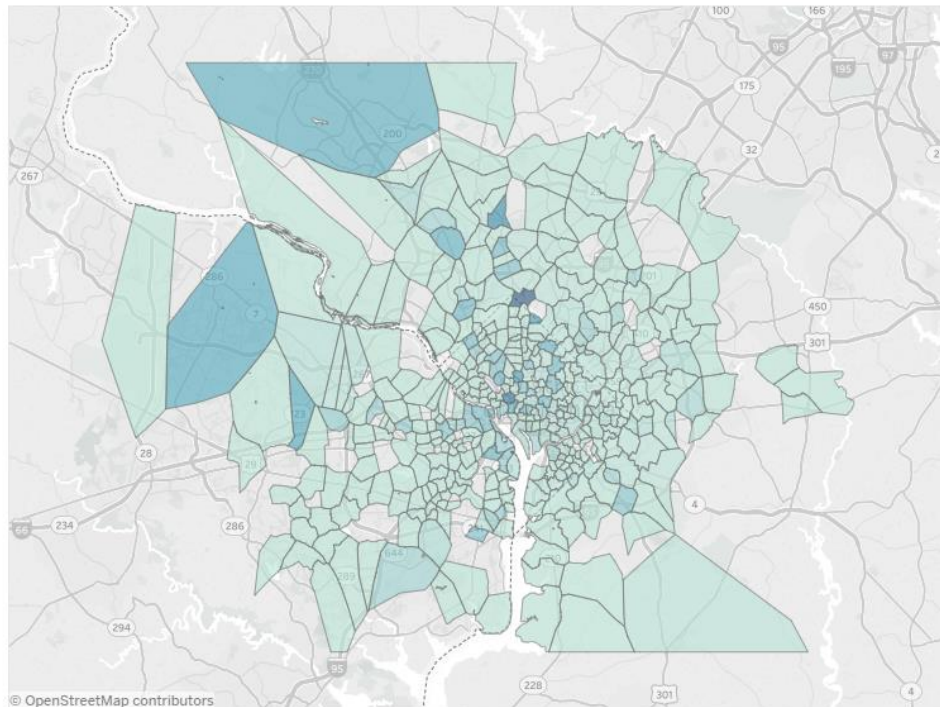
Zone Journey OD Dashboard

See passengers' ultimate destinations given their origin area, or vice versa.

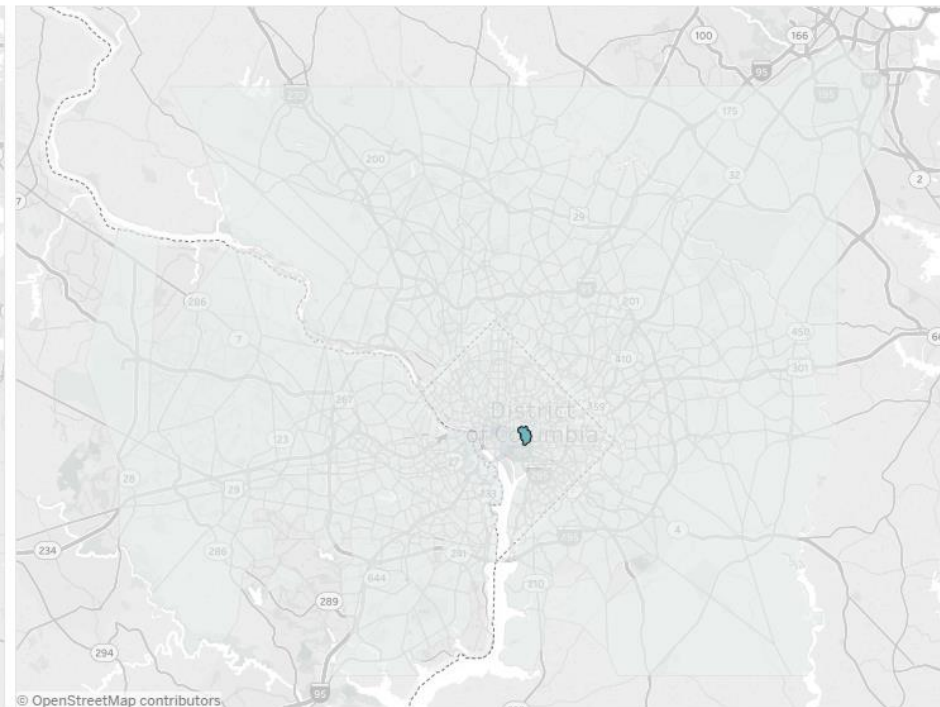
This dashboard summarizes journeys for which both origin and destination are available. This means that the number of passengers travelling by bus will appear lower than it actually is.

Data from Trace Model v. 3.0, averaging across May 7 to 11, 2018.

Select an origin zone or a group of origin zones



Select a destination zone or a group of destination zones



Trips per day
0.2 496.4

Filter by any combination of the following:

Origin Subsystem

- (All)
- Metrobus
- Metrorail

Destination Subsystem

- (All)
- Metrobus
- Metrorail

Hour starting at



Number of Stages





What's Next?

What's Next?

- Develop additional tools and visualizations to make the data more accessible
- Start using Trace data
 - Bus service planning
 - Example: If we moved the terminal of a route from one station to another, how would passengers' trips be affected?
 - Rail service planning and monitoring
 - Example: monitor platform crowding during disruptions due to track work
- Eventually, expand to include data from regional partners

