



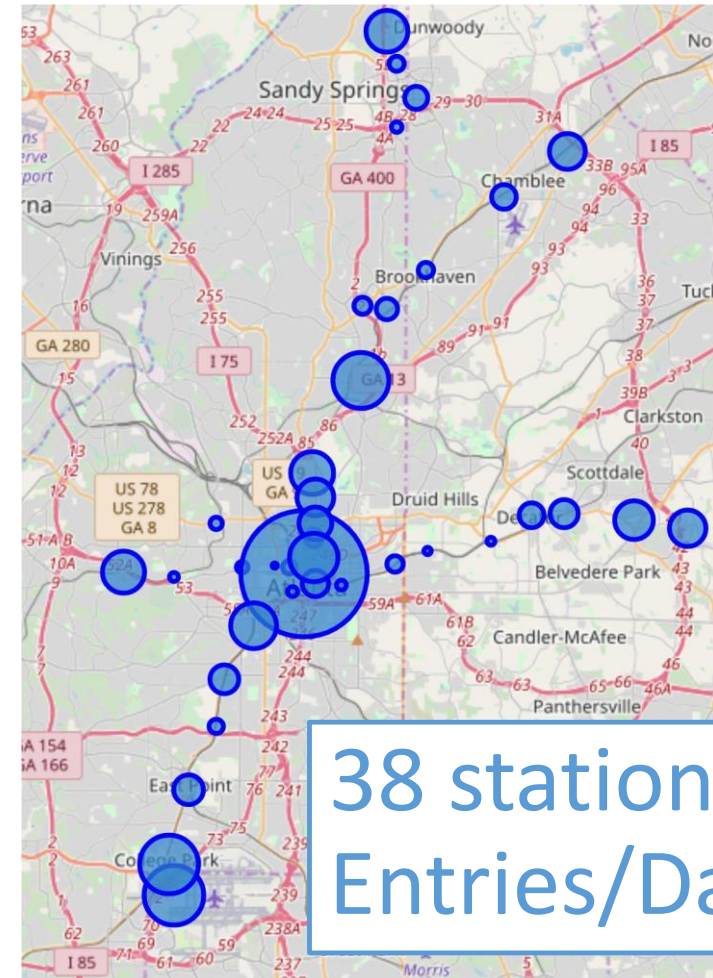
Segmenting and Labeling MARTA Rail Stations in Atlanta

Nelli Fedorova

<https://www.linkedin.com/in/nelli-fedorova-7710b01a/>

Data: https://en.wikipedia.org/wiki/MARTA_rail

Station	Entries/Day	Latitude	Longitude
Airport	9173	33.64076	-84.4463
Arts Center	6605	33.78971	-84.3878
Ashby	1791	33.75635	-84.4176
Avondale	4327	33.77528	-84.2819
Bankhead	1903	33.77189	-84.4288
Brookhaven/...	2357	33.86071	-84.3400
Buckhead	2643	33.84841	-84.3670
Chamblee	3785	33.88619	-84.3070
Civic Center	2692	33.76631	-84.3872
College Park	9026	33.65167	-84.4488
Decatur	3821	33.77472	-84.2956
Dome/...	2107	33.75629	-84.3978
...



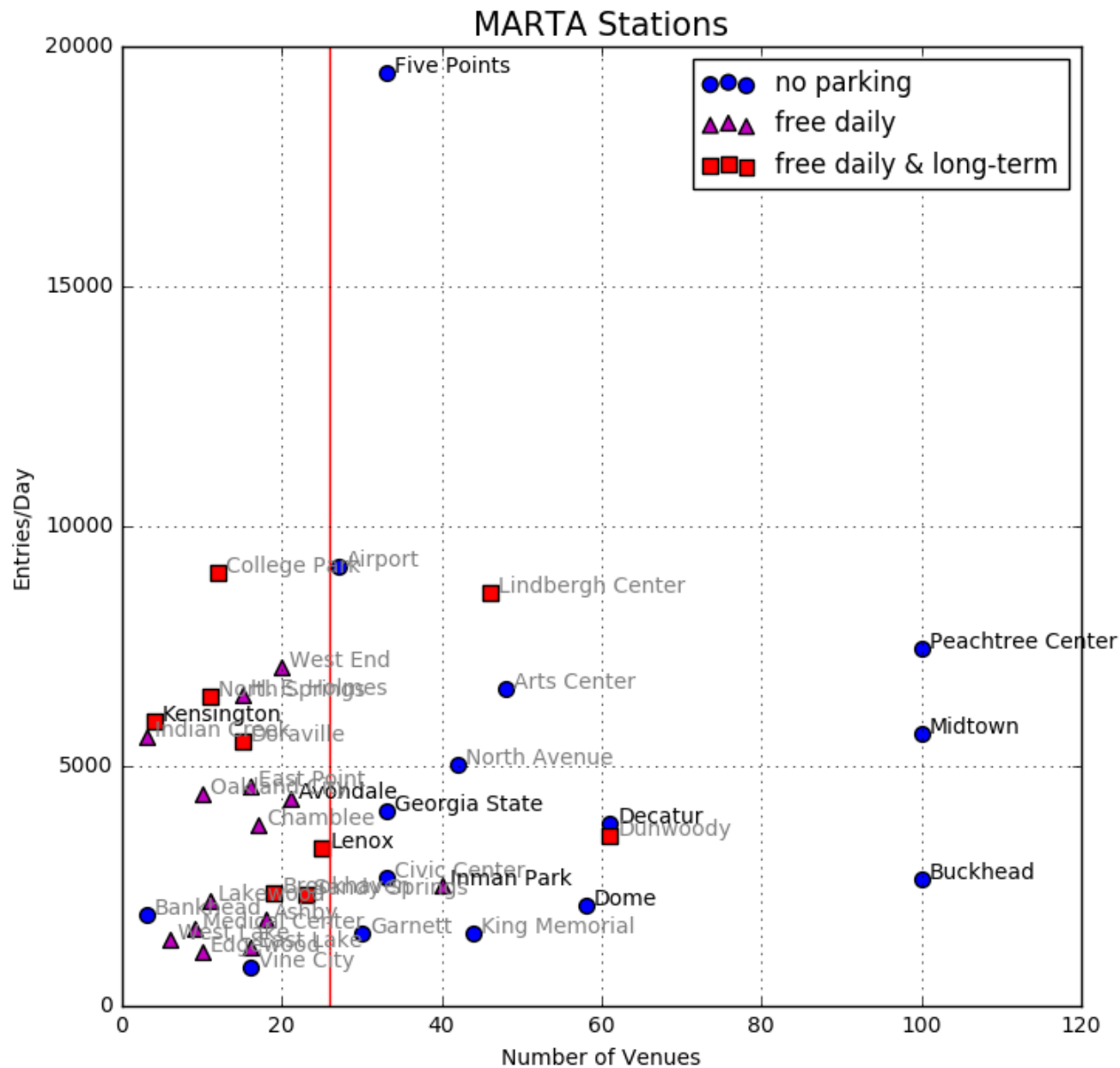
Data: Foursquare/explore

Top 100 venues with their primary categories within a radius of 500 meters from each station:

1156 venues
223 categories

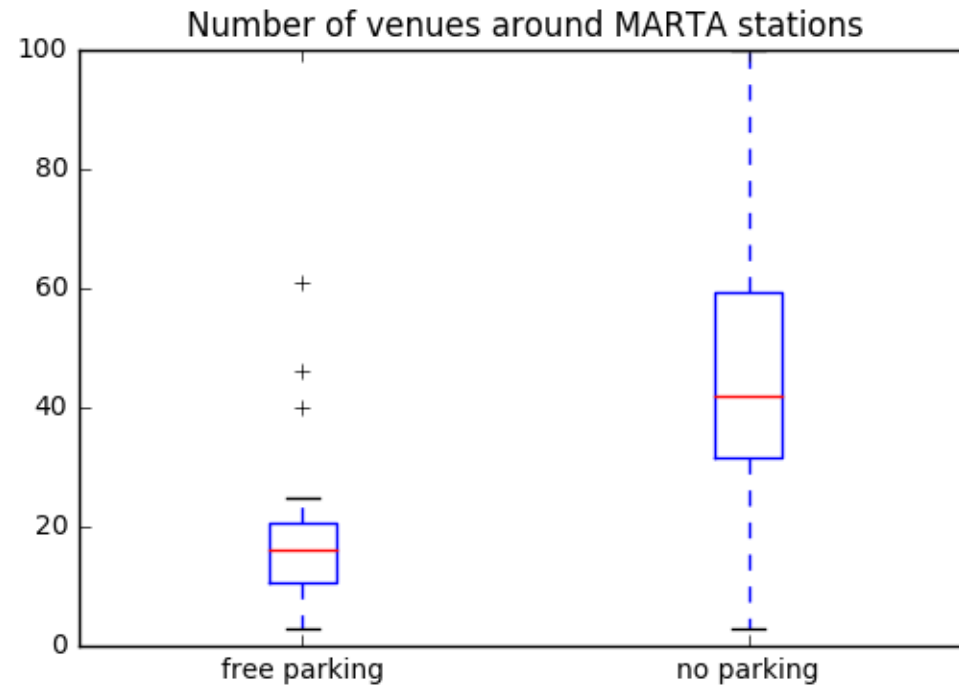
<https://api.foursquare.com/v2/venues/explore?&radius=500&limit=100...>

	Station	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Airport	33.640758	-84.446341	Starbucks	33.640891	-84.444769	Coffee Shop
1	Airport	33.640758	-84.446341	Metal Penguin	33.640059	-84.444530	Public Art
2	Airport	33.640758	-84.446341	Zimbabwe: A Tradition in Stone	33.640745	-84.442539	Art Gallery
3	Airport	33.640758	-84.446341	TSA PreCheck North	33.641276	-84.444029	Airport Service
4	Airport	33.640758	-84.446341	Starbucks	33.641678	-84.446620	Coffee Shop



▲ ■ free parking ● no free parking

Low business activity correlates with free parking



['Arts & Entertainment', 'Museum', 'Art Museum']

Nightclub

Data: Foursquare/categories

<https://developer.foursquare.com/docs/resources/categories>

General categories: **root parent** node for each **primary category**

Food → **Asian Restaurant**

Food → Asian Restaurant → **Japanese Restaurant**

Shop & Service → Clothing Store → **Women's Store**

Shop & Service → **Pharmacy**

Shop & Service → **Gas Station**

Nightlife Spot → **Nightclub**

Nightlife Spot → Bar → **Whisky Bar**

Arts & Entertainment → **Aquarium**

Arts & Entertainment → Museum → **Art Museum**

Outdoors & Recreation → **Park**

Outdoors & Recreation → Athletics & Sports → Gym / Fitness Center → **Yoga Studio**

Professional & Other Places → **Event Space**

Professional & Other Places → Medical Center → **Chiropractor**

Venues: One Hot Encoding for 223 primary categories

	Station	Accessories Store	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Aquarium	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Automotive Shop	BBQ Joint	Bagel Shop	B
0	Airport	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Airport	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Airport	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
3	Airport	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
4	Airport	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Venues: One Hot Encoding for 9 general categories

	Station	Arts & Entertainment	College & University	Food	Nightlife Spot	Outdoors & Recreation	Professional & Other Places	Residence	Shop & Service	Travel & Transport
0	Airport	0	0	1	0	0	0	0	0	0
1	Airport	1	0	0	0	0	0	0	0	0
2	Airport	1	0	0	0	0	0	0	0	0
3	Airport	0	0	0	0	0	0	0	0	1
4	Airport	0	0	1	0	0	0	0	0	0

Station: Bag of Categories

No matter where
people go, they
eat and drink

1. Group venues by station and compute average occurrence of each **primary** category
2. Extract 10 most frequent **primary** categories for each station

	Station	1st Most Common Category	2nd Most Common Category	3rd Most Common Category	4th Most Common Category	5th Most Common Category	6th Most Common Category	7th Most Common Category	8th Most Common Category	9th Most Common Category	10th Most Common Category
0	Airport	Airport Service	Coffee Shop	Airport Lounge	Rental Car Location	Bakery	Food Stand	Bus Station	Clothing Store	Fast Food Restaurant	Sandwich Place
1	Arts Center	Coffee Shop	Sandwich Place	Music Venue	Performing Arts Venue	Hotel	Pizza Place	American Restaurant	Theater	Comedy Club	Exhibit
6	Buckhead	Clothing Store	Cosmetics Shop	Hotel	Steakhouse	Furniture / Home Store	Women's Store	Department Store	American Restaurant	Bakery	Jewelry Store
17	Five Points	Sandwich Place	Pharmacy	Burger Joint	Event Space	Lounge	Falafel Restaurant	Coffee Shop	Bubble Tea Shop	Sports Bar	Theater
19	Georgia State	Sandwich Place	Coffee Shop	Breakfast Spot	Fast Food Restaurant	Asian Restaurant	Shoe Store	Bubble Tea Shop	Burger Joint	Mexican Restaurant	Movie Theater
27	Lindbergh Center	Coffee Shop	Sandwich Place	Shopping Plaza	Fast Food Restaurant	Gym	Nightclub	Shipping Store	Salon / Barbershop	Shopping Mall	Smoke Shop

Station: Bag of TF-IDF Categories

Station Label:
10 most frequent
categories

1. Group venues by station and compute TF-IDF frequency of each **primary** category
2. Extract 10 most frequent **primary** categories for each station

	Station	1st Most Common Category	2nd Most Common Category	3rd Most Common Category	4th Most Common Category	5th Most Common Category	6th Most Common Category	7th Most Common Category	8th Most Common Category	9th Most Common Category	10th Most Common Category
0	Airport	Airport Service	Airport Lounge	Airport Terminal	Food Stand	Betting Shop	Taxi Stand	Public Art	Caribbean Restaurant	Electronics Store	Bus Station
1	Arts Center	Performing Arts Venue	Music Venue	Theater	Exhibit	Indie Theater	Print Shop	Coffee Shop	Concert Hall	Art Museum	Lawyer
6	Buckhead	Clothing Store	Department Store	Women's Store	Cosmetics Shop	Steakhouse	Hotel	Furniture / Home Store	Jewelry Store	Shopping Plaza	Gift Shop
17	Five Points	Hookah Bar	Event Space	Cuban Restaurant	Deli / Bodega	Sandwich Place	Korean Restaurant	Falafel Restaurant	Bubble Tea Shop	Poke Place	Mediterranean Restaurant
19	Georgia State	College Basketball Court	Cuban Restaurant	Event Space	Asian Restaurant	Sandwich Place	Tapas Restaurant	Bubble Tea Shop	Movie Theater	Convenience Store	Sporting Goods Shop
27	Lindbergh Center	Shopping Plaza	Salon / Barbershop	Shipping Store	Nightclub	Garden Center	Big Box Store	Hardware Store	Tex-Mex Restaurant	Residential Building (Apartment / Condo)	Video Game Store

Station: Bag of Categories

Food and Shop &
Service dominate

1. Group venues by station and compute average occurrence of each **general** category
2. Sort **general** categories for each station

	Station	1st Most Common Category	2nd Most Common Category	3rd Most Common Category	4th Most Common Category	5th Most Common Category	6th Most Common Category	7th Most Common Category	8th Most Common Category	9th Most Common Category
0	Airport	Food	Travel & Transport	Shop & Service	Arts & Entertainment	-	-	-	-	-
1	Arts Center	Food	Arts & Entertainment	Shop & Service	Nightlife Spot	Travel & Transport	Outdoors & Recreation	-	-	-
6	Buckhead	Shop & Service	Food	Travel & Transport	Outdoors & Recreation	Nightlife Spot	Professional & Other Places	-	-	-
17	Five Points	Food	Arts & Entertainment	Nightlife Spot	Shop & Service	Travel & Transport	Professional & Other Places	-	-	-
19	Georgia State	Food	Shop & Service	Outdoors & Recreation	Arts & Entertainment	Travel & Transport	Professional & Other Places	College & University	-	-
27	Lindbergh Center	Shop & Service	Food	Nightlife Spot	Outdoors & Recreation	Residence	-	-	-	-

Station: Bag of TF-IDF Categories

Station Label:
3 most frequent
categories

1. Group venues by station and compute TF-IDF frequency of each **general** category
2. Sort TF-IDF **general** categories for each station

	Station	1st Most Common Category	2nd Most Common Category	3rd Most Common Category	4th Most Common Category	5th Most Common Category	6th Most Common Category	7th Most Common Category	8th Most Common Category	9th Most Common Category
0	Airport	Travel & Transport	Arts & Entertainment	Food	Shop & Service	-	-	-	-	-
1	Arts Center	Arts & Entertainment	Nightlife Spot	Food	Travel & Transport	Outdoors & Recreation	Shop & Service	-	-	-
6	Buckhead	Nightlife Spot	Professional & Other Places	Travel & Transport	Food	Outdoors & Recreation	Shop & Service	-	-	-
17	Five Points	Nightlife Spot	Arts & Entertainment	Professional & Other Places	Food	Travel & Transport	Shop & Service	-	-	-
19	Georgia State	College & University	Professional & Other Places	Food	Arts & Entertainment	Outdoors & Recreation	Travel & Transport	Shop & Service	-	-
27	Lindbergh Center	Nightlife Spot	Residence	Food	Outdoors & Recreation	Shop & Service	-	-	-	-

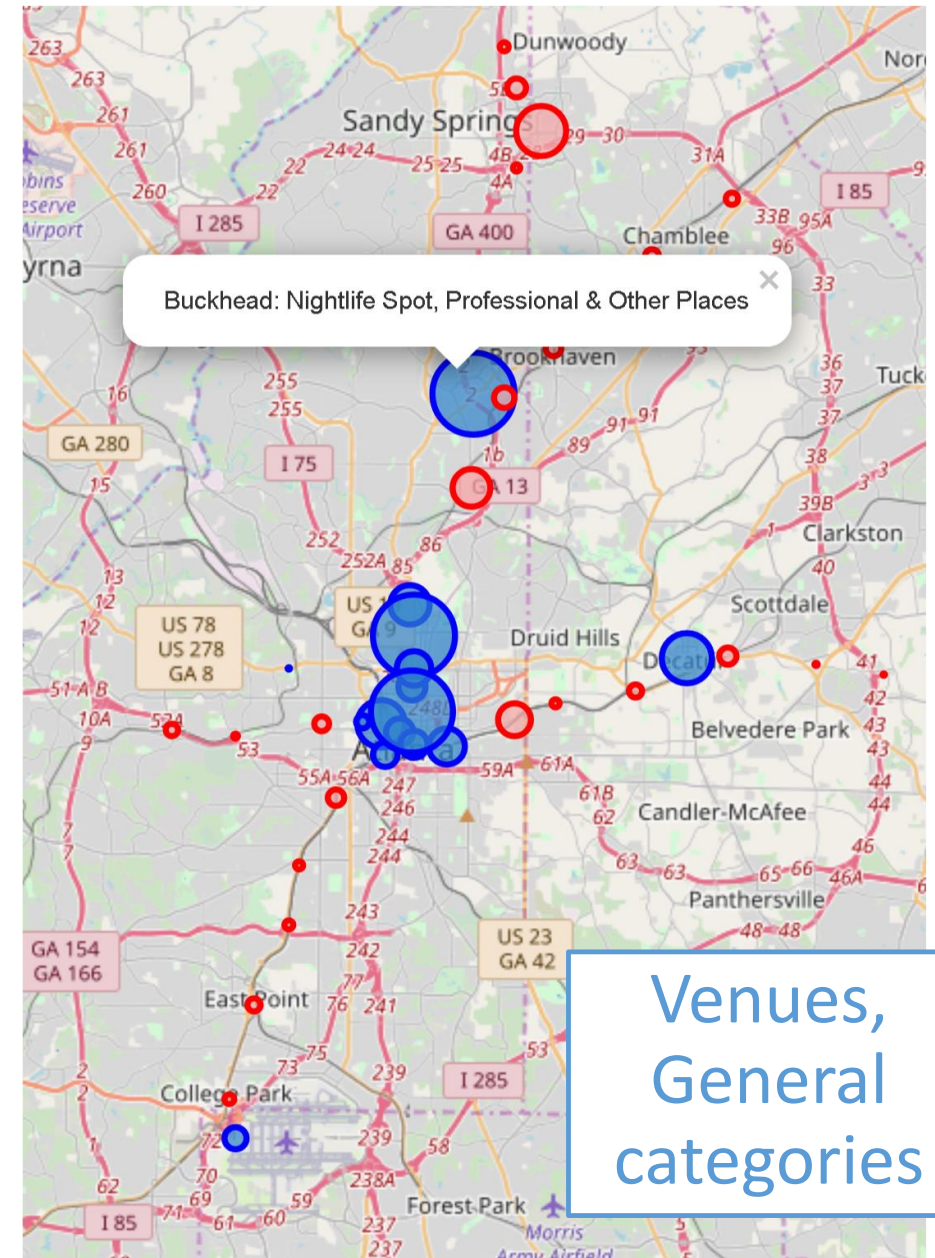
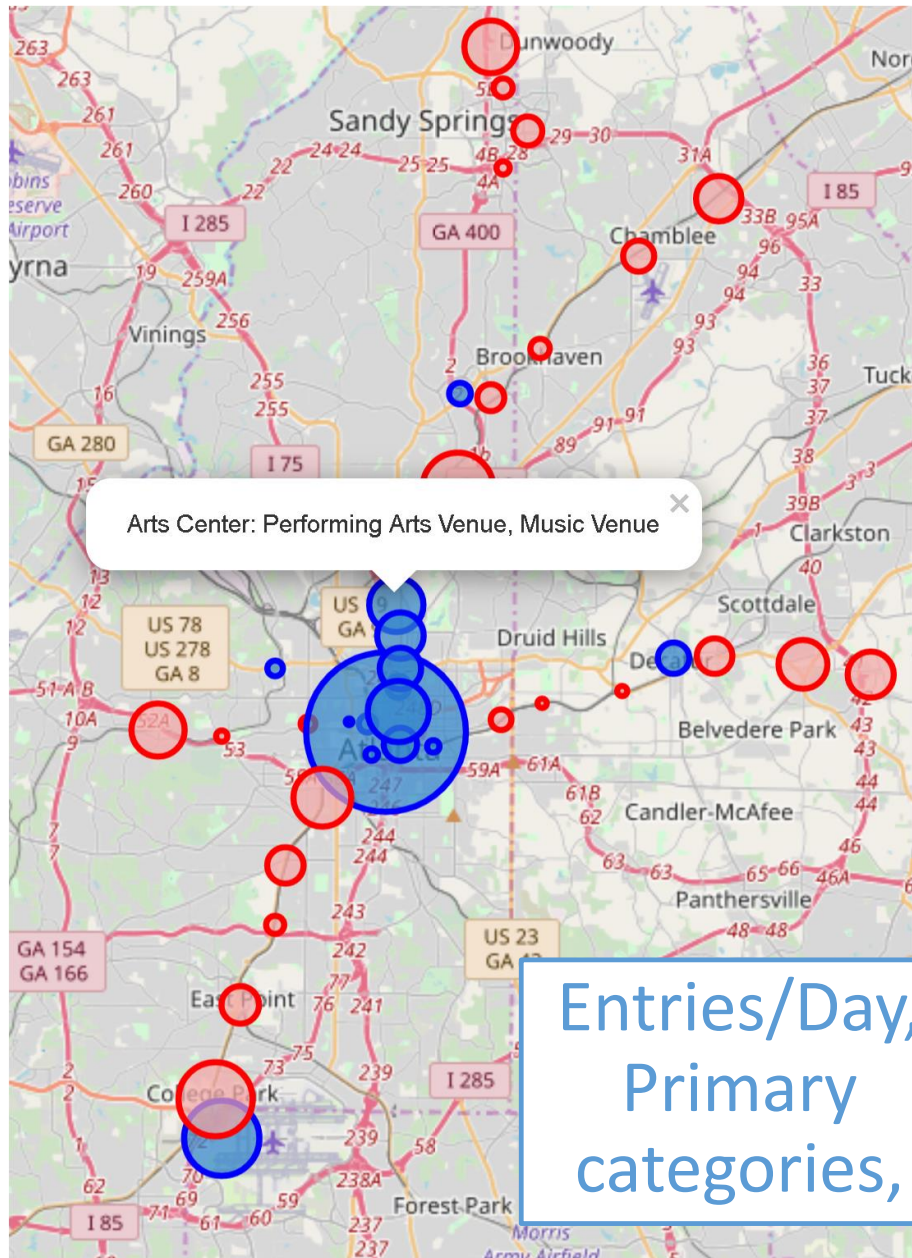
Station: LDA TF-IDF Topics

Station Label:
Top topics

1. Group venues by station
2. Use LDA TF-IDF to represent each document as weighted mixture of 10 latent topics

	Topic	Stemmed Words
0	store	0.020*hotel + 0.019*store + 0.017*chiropractor + 0.015*furnitur + ...
1	restaur	0.024*conveni + 0.014*restaur + 0.014*vegan + 0.014*vegetarian +...
2	station	0.036*station + 0.032*rail + 0.032*light + 0.026*cream + ...
3	airport	0.032*airport + 0.015*pharmaci + 0.014*depart + 0.013*sport + ...
4	colleg	0.019*basketbal + 0.015*stadium + 0.014*court + 0.012*fast + ...
5	metro	0.018*metro + 0.014*gastropub + 0.013*construct + 0.013*line + ...
6	fast	0.019*lawyer + 0.017*cosmet + 0.015*discount + 0.014*station + ...
7	place	0.021*hotel + 0.019*place + 0.016*sandwich + 0.015*theater + ...
8	galleri	0.015*galleri + 0.013*hall + 0.013*concert + 0.013*restaur + ...
9	pharm	0.027*pharmaci + 0.021*flower + 0.021*platform + 0.019*smoke + ...

Station	Weighted Sum of Topics
0 Airport	0.595* place + 0.334* airport + 0.057* store
1 Arts Center	0.989* place
6 Buckhead	0.7148* store + 0.213* place + 0.016* pharm
17 Five Points	0.985* place
19 Georgia State	0.831* place + 0.101* colleg + 0.057* restaur
27 Lindbergh Center	0.595* place + 0.395* store



• free parking • no free parking

Conclusion and Future Directions

- Low business activity correlates with free parking:
 - Stations with < 26 venues have free daily parking
 - Stations with > 26 venues, except Lindbergh Center, Dunwoody and Inman Park, have no free parking
- Most venues near MARTA stations are fast food restaurants and coffee shops
 - Add foot traffic, secondary categories and other detailed venue-related information to deepen our analysis
- TF-IDF normalization and TF-IDF LDA provide reasonably good labels for MARTA stations
 - Play with the number of topics in LDA
 - Trim Foursquare tree at a certain depth instead of using only 9 root nodes as general categories