

My progression, the original "ER Diagram"

Station_entrances
Station_name **

Station_name, Station_lat, Station_lon**
Station_geom, Division, Line

Stations
Station_name**

Subway_ridership_2013_present
Station_name**

Complex_ID

Station_Complexes
Complex_ID is primary key

Borough_name

Route_ID

Routes
*route_ID is primary key here, 30 distinct rows

Trips
Trip_id is primary key

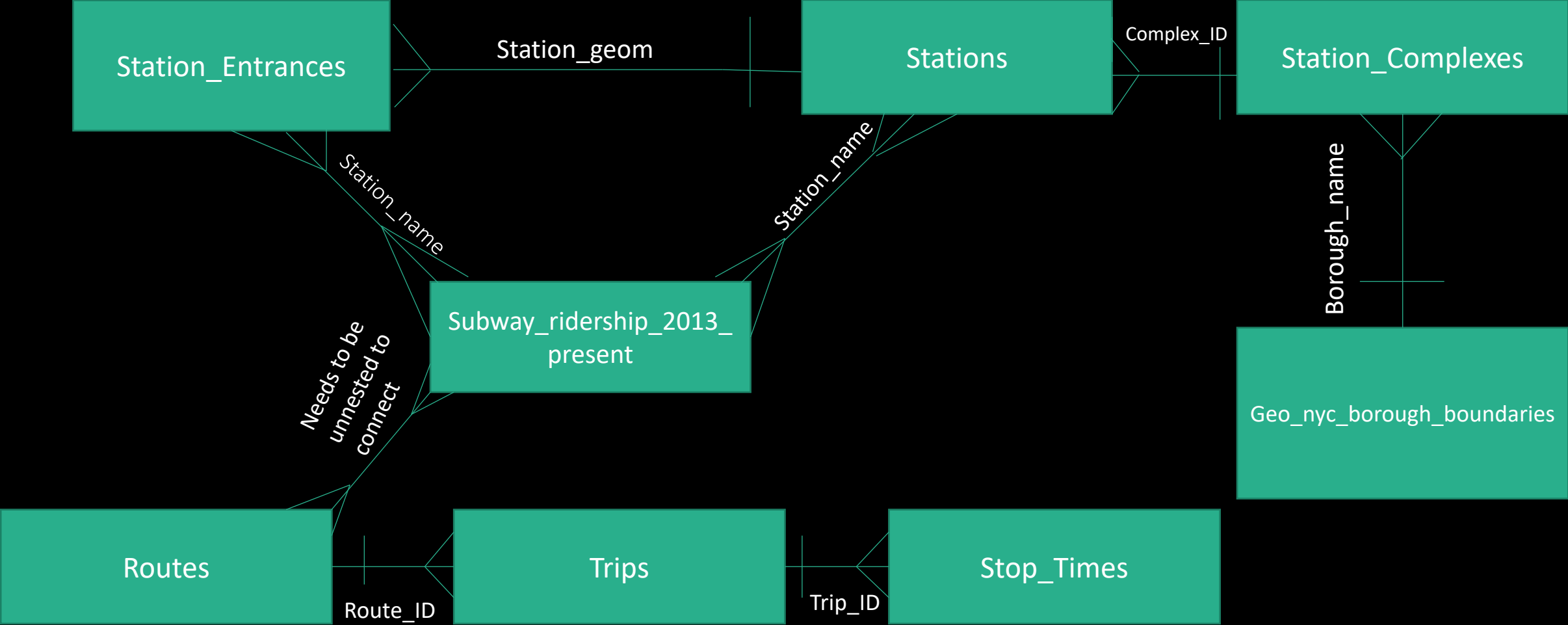
Trip_ID

geo_nyc_borough_boundaries
Borough_name is primary key

Stop_Times

RESTRICTED

ER Diagram for New York Subway Dataset



Query

🔍 *Editor

station_entrances

+

🏠

ℹ️

📄

🖨️

▶️ RUN

⚙️ MORE

💾 SAVE

👤 SHARE

🕒 SCHEDULE

✅ This query will process 131.96 KB when run.

```
1 Select *
2 From(
3 SELECT S.station_name, SE.staff, SE.entrance_type, SE.vending, SE.corner
4 FROM bigquery-public-data.new_york_subway.station_entrances SE
5 Join bigquery-public-data.new_york_subway.stations S on st_equals(SE.station_geom, S.station_geom))
6 Order by Station_name
7
```

Processing location: US ✕Press Alt+F1 for Accessibility Options.

Query results

📄 SAVE RESULTS

📊 EXPLORE DATA

⬆️

JOB INFORMATION

RESULTS

JSON

EXECUTION DETAILS

EXECUTION GRAPH

PREVIEW

Row	station_name	staff	entrance_type	vending	line	corner
5	103 St	FULL	Stair	true	Lexington Av	SE
6	103 St	FULL	Stair	true	8th Av - Fulton St	NW
7	103 St	FULL	Stair	true	Broadway - 7Av	NW
8	103 St	NONE	Stair	fal...	Broadway - 7Av	SE
9	103 St	FULL	Stair	true	Broadway - 7Av	NE
10	103 St	FULL	Stair	true	Lexington Av	SW

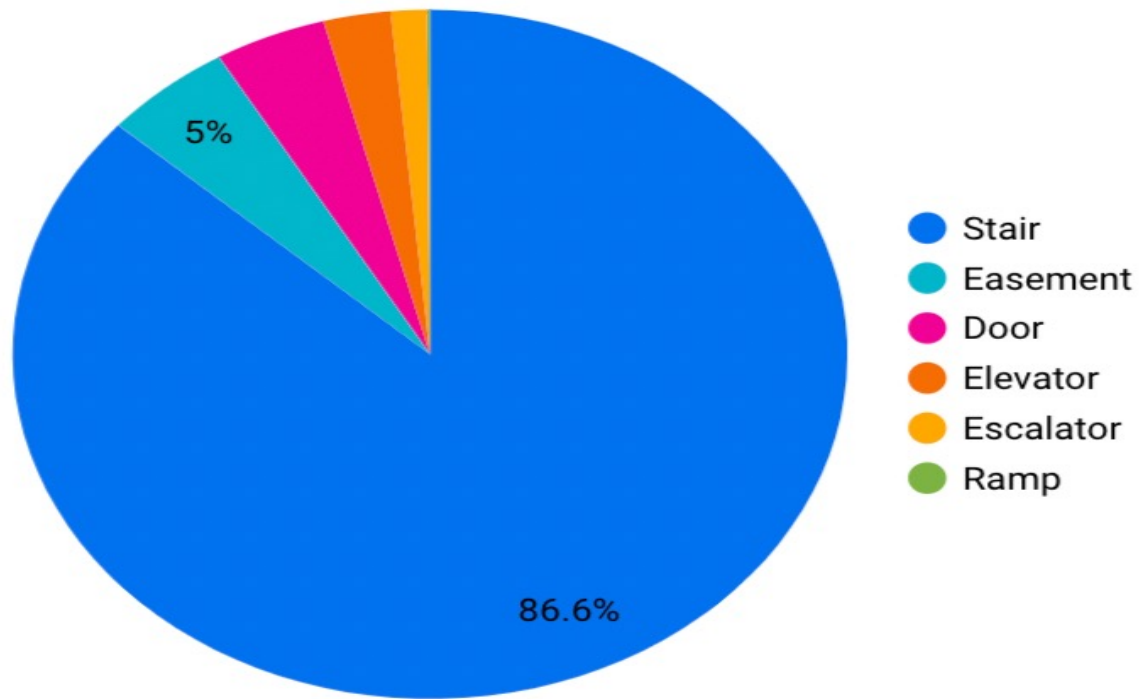
Results per page: 50 1 – 50 of 1847 |< < > >|

Interesting Conclusions(Cont.)

- On the southeast corner of 103 st station, there isn't vending or staff. Why is that? Maybe because the entire part of that station is digital?
- Also, it's important to notice that the station name itself isn't enough to draw conclusions from, or even the line. The corner seems to be key to seeing the real differences in figuring out what part of the station truly isn't staffed or doesn't have vending. This could be a decision maker for why someone make not take a certain route or station, maybe they don't like digital service.

Interesting Conclusions

As you can see, the most common entrance into a subway in NYC is through the stairs



Interesting Conclusions(Cont.)

	station_name	Number_Of_Entrances	Number_Of_Enterances_With_Staff ▾	Percent
1.	34th St	41	17	41.46%
2.	59th St	25	16	64%
3.	23rd St	40	15	37.5%
4.	125th St	19	15	78.95%
5.	50th St	20	15	75%
6.	86th St	22	13	59.09%
7.	57th St	16	12	75%
8.	7th Av	14	12	85.71%
9.	14th St	26	12	46.15%
10.	103rd St	12	11	91.67%
11.	Flushing-Main St	10	10	100%
12.	Jamaica-179th St	16	10	62.5%
13.	96th St	13	9	69.23%
14.	Yankee Stadium-161st St	16	9	56.25%
15.	Lafayette Av	9	9	100%

1 - 100 / 355 < >

Interesting Conclusions(Cont.)

- Someone could look at this dataset and say that they want to go to Flushing-Main St's station simply because they are 100% likely to have some sort of staff, whether it's part time, full-time, or a special event.
- In that same token, the data could be looked at in a way that someone could decide that they want to go to a station that is likely not staffed so that they could just use the digital services that the subways provide.

Connectivity Issues

routes				subway_riders...			
This is a partitioned table. Learn more				This is a partitioned table. Learn more			
SCHEMA DETAILS PREVIEW LINEAGE PREVIEW				SCHEMA DETAILS PREVIEW LINEAGE PREVIEW			
Row	route_id	route_short_name	route_long_name	Row	ion_name	routes	ridership_2013
1	GS	S	42 St Shuttle	1	St	D,N,R	12059
				2	St - Washington Sq	A,B,C,D,E,F,M	40798
				3	St - MetroTech	A,C,F,R	40183
				4	es Sq - 42 St / 42 St	N,Q,R,W,S,1,2,3,7 / A,C,E	197696
				5	/	F,G,R	13156
2	C	C	8 Avenue Local	6	mer St / Metropolitan Av	L/G	14004
				7	v Utrecht Av / 62 St	N / G	5159
				8	ington Av/53 St / 51 St	E,M/ 6	69973
				9	//53 St	E,M	25939
3	E	E	8 Avenue Local	10	St - Columbus Circle	A,B,C,D,1	72236
				11	tle - Wyckoff Avs	L,M	18167
				12	ington Av/59 St	N,W,R / 4,5,6	67841
				13	//59 St	N,W,R	18394
4	W	W	Broadway Local	14	St / 8 Av	A,C,E / L	42718
Results per page: 50 1 - 30 of 30				Results per page: 50 1 - 50 of 430			

Technical issues

- The bottom part of the data, routes, trips, and stop_times are essential to draw interesting conclusions from and there is a technical link through them. But, this is where I'm having trouble connecting the tables. Routes links to subway_ridership through routes and route_short_name but this link isn't a 1-1 link. There would have to be some sort of filter used to link them.
- The current technical problem is connecting the two tables, which would possibly lead to some interesting connections. For example, we would be able to find all the stop times for the stations.
- The solution to this would be to unnest the data