

a) Which top 10 BikeId has most of the trip duration

```
-- SELECT bikeid, tripduration
-- FROM `bigquery-public-data.new_york_citibike.citibike_trips`
-- ORDER BY tripduration DESC
-- LIMIT 10;
```

bikeid	tripduration
30503	19510049
30875	15962256
16777	15020934
16533	13931824
18504	13586276
30446	12479323
33252	11749576
19957	11699746
15336	11138807
24939	10283682

b) which two stations have the longest travel time

note: my interpretation of this question is the longest duration a bike is being used between two given stations

```
-- SELECT start_station_name,end_station_name ,MAX(tripduration) travel
-- FROM `bigquery-public-data.new_york_citibike.citibike_trips`
-- GROUP BY start_station_name,end_station_name
-- ORDER BY travel DESC
-- LIMIT 2;
```

Row	start_station_name	end_station_name	travel
1	Hope St & Union Ave	Nassau St & Navy St	19510049
2	Grand Army Plaza & Plaza St West	NYCBS DEPOT - DELANCEY	15962256

c) Which station is more popular (frequently used station)

```
---- Note: used station start and end
-- SELECT start_station_id,start_station_name, COUNT(start_station_name) AS visits
-- FROM `bigquery-public-data.new_york_citibike.citibike_trips`
-- WHERE start_station_id >= 1
-- GROUP BY start_station_name,start_station_id
-- ORDER BY visits DESC
-- LIMIT 4
---- Note: Excluding null values for location
```

Row	start_station_id	start_station_name	visits
2	497	E 17 St & Broadway	867,794

d) Which station has most number of bikes available for rent

```
-- SELECT station_id, name,MAX(num_bikes_available) bike_availability
-- FROM `bigquery-public-data.new_york_citibike.citibike_stations`
-- GROUP BY name,station_id,num_bikes_available
-- ORDER BY num_bikes_available DESC
-- LIMIT 1
---- Note: research top function
```

Row	station_id	name	bike_availability
1	445	E 10 St & Avenue A	79

e) What are the rush hours when most trips are taken

```
-- SELECT a.hour_of_day, COUNT((a.hour_of_day)) hour_count
-- FROM (
-- SELECT EXTRACT (hour FROM starttime) AS hour_of_day
-- FROM `bigquery-public-data.new_york_citibike.citibike_trips`
-- ) a
-- GROUP BY a.hour_of_day
-- ORDER BY a.hour_of_day DESC
```

```

69
70 hour_of_day    hour_count
71 17             5232747
72 18             5024723
73 8              4161676
74 16             3798414
75 19             3517025
76 9              3346692
77
78 ---- Another option: See the number of trips before or after a given hour
79 -- SELECT COUNT(*)
80 -- FROM `bigquery-public-data.new_york_citibike.citibike_trips`
81 -- WHERE EXTRACT (hour FROM starttime) > 12;
82 ----- WHERE EXTRACT (hour FROM starttime) <= 12;
83 ----- WHERE EXTRACT (hour FROM stoptime) > 12;
84 ----- WHERE EXTRACT (hour FROM stoptime) <= 12;
85
86

```

#### f) What are the most popular stations to start a trip

```

87
88
89 -- SELECT x.start_station_name, x.num_station
90 -- FROM (
91 -- SELECT start_station_name,COUNT(start_station_name) num_station
92 -- FROM `bigquery-public-data.new_york_citibike.citibike_trips`
93 -- WHERE start_station_id >= 1
94 -- GROUP BY start_station_name
95 -- ) x
96 -- GROUP BY x.start_station_name, x.num_station
97 -- ORDER BY x.num_station DESC
98 -- LIMIT 3
99

```

Row	start_station_id	start_station_name	visits
100	1 519	Pershing Square North	438077
101	2 497	E 17 St & Broadway	423334
102	3 435	W 21 St & 6 Ave	403795

#### g) What category (usertype) of users take most of the bike-trips

```

103
104
105
106
107 -- SELECT usertype,COUNT(usertype) user_trips
108 -- FROM `bigquery-public-data.new_york_citibike.citibike_trips`
109 -- GROUP BY usertype
110

```

Row	usertype	user_trips
112	1 Subscriber	46917572
113	2 Customer	6191149
114	3	5828994 -- Random Null results

#### h) What are the top 5 routes taken by women in the year 2013

```

115
116
117 ---- notes: route should calculated as start and end from same location.
118 -- SELECT x.route, COUNT((x.route)) route_taken
119 -- FROM (
120 -- SELECT CONCAT(start_station_name,'-',end_station_name) route, EXTRACT (year FROM
121 -- starttime) AS start_year
122 -- FROM `bigquery-public-data.new_york_citibike.citibike_trips`
123 -- WHERE gender = "female"
124 -- ) x
125 -- WHERE x.start_year = 2013
126 -- GROUP BY x.route, x.start_year
127 -- ORDER BY route_taken DESC
128

```

Row	route	route_taken
129	1 West Thames St-Vesey Pl & River Terrace	633
130	2 Vesey Pl & River Terrace-West Thames St	470
131	3 Central Park S & 6 Ave-Central Park S & 6 Ave	460
132	4 West St & Chambers St-West St & Chambers St	442
133	5 Washington Square E-University Pl & E 14 St	420

```
137
138 i) Write a query to extract Day,Hour,Year from Timestamp (use starttime column)
139
140 -- SELECT starttime,
141 --     EXTRACT(day FROM starttime) AS day_start,
142 --     EXTRACT(hour FROM starttime ) AS hour_start,
143 --     EXTRACT(year FROM starttime) AS year_start
144 -- FROM `bigquery-public-data.new_york_citibike.citibike_trips`
145 -- WHERE starttime = '2016-08-29T14:45:02';
146
147
148 Row starttime                day_start  hour_start  year_start
149 1    2016-08-29T14:45:02      29        14        2016
150
```