



New York City Parking Tickets

Data Source

From NYC OpenData

Fiscal Year 2014 :

<https://data.cityofnewyork.us/City-Government/Parking-Violations-Issued-Fiscal-Year-2014/jt7v-77mi>

Fiscal Year 2015:

<https://data.cityofnewyork.us/City-Government/Parking-Violations-Issued-Fiscal-Year-2015/c284-tqph>

Fiscal Year 2016:

<https://data.cityofnewyork.us/City-Government/Parking-Violations-Issued-Fiscal-Year-2016/kiv2-tbuis>

Fiscal Year 2017:

<https://data.cityofnewyork.us/City-Government/Parking-Violations-Issued-Fiscal-Year-2017/2bnn-yakx>

Hardware Specifications - Google Dataproc

-Master node

Standard (1 master, N workers)

-Machine type

n1-standard-4 (4 vCPU, 15.0 GB memory)

-Primary disk type

pd-standard

-Primary disk size

500 GB

-Worker nodes

2

-Machine type

n1-standard-4 (4 vCPU, 15.0 GB memory)

-Primary disk type

pd-standard

-Primary disk size

500 GB

Hardware Specifications cont.

Image Version:1.3.45-debian9

Applications: Hadoop 2.9.2, Spark 2.3.4, Hive 2.3.5, Pig 0.17.0, Tez 0.9.2

Ticket Summary

```
CREATE EXTERNAL TABLE IF NOT EXISTS tickets_summary  
(summons_number INT, plate_id STRING, vehicle_make STRING, vehicle_body_type STRING, vehicle_color  
STRING, violation_code STRING, registration_state STRING)  
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','  
STORED AS TEXTFILE LOCATION '/output/tickets_summary'  
TBLPROPERTIES ('skip.header.line.count'='1');
```

```
INSERT OVERWRITE TABLE tickets_summary SELECT summons_number, plate_id, vehicle_make,  
vehicle_body_type, vehicle_color, violation_code, registration_state FROM tickets;
```

```
select * from tickets_summary LIMIT 10;
```

Results

```
0: jdbc:hive2://localhost:10000/default> CREATE EXTERNAL TABLE IF NOT EXISTS tickets_summary
. . . . .> (summons_number INT, plate_id STRING, vehicle_make STRING, vehicle_body_type STRING, vehicle_color STRING, violation_code STRING, registration_state STRING)
. . . . .> ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
. . . . .> STORED AS TEXTFILE LOCATION '/output/tickets_summary'
. . . . .> TBLPROPERTIES ('skip.header.line.count'='1');
No rows affected (2.573 seconds)
0: jdbc:hive2://localhost:10000/default> INSERT OVERWRITE TABLE tickets_summary SELECT summons_number, plate_id, vehicle_make, vehicle_body_type, vehicle_color, violation_code, registration_state FROM tickets;
No rows affected (140.807 seconds)
0: jdbc:hive2://localhost:10000/default> select * from tickets_summary LIMIT 10;
```

tickets_summary.summons_number	tickets_summary.plate_id	tickets_summary.vehicle_make	tickets_summary.vehicle_body_type	tickets_summary.vehicle_color	tickets_summary.violation_code	tickets_summary.registration_state
1366962000	63540MC	FRUEH	DELV	BRN	46	NY
1342296187	GCY4187	FORD	VAN	BLUE	21	NY
1342296199	95V6675	GMC		SILVR	21	TX
1342296217	FYM5117	NISSA	SUBN	WHITE	21	NY
1356906515	GFM1421	MAZDA	SDN	BLK	40	NY
1337077380	18972BB	INTER	BUS	YELLOW	14	NY
1364523796	WNJ4730	TOYOT	SDN	BLK	14	VA
1359914924	68091JZ	TOYOT	DELV	WH	46	NY
1355498326	EWV4127	ACURA	SUBN	GREY	21	NY
1361272259	P4ZO671	BMW	SDN	BK	21	99

Displaying certain information by plate id

```
SELECT * FROM tickets_summary WHERE plate_id LIKE 'GFM1421' LIMIT 1;
```

```
0: jdbc:hive2://localhost:10000/default> SELECT * FROM tickets_summary WHERE plate_id LIKE 'GFM1421' LIMIT 1;
```

tickets_summary.summons_number	tickets_summary.plate_id	tickets_summary.vehicle_make	tickets_summary.vehicle_body_type	tickets_summary.vehicle_color	tickets_summary.violation_code	tickets_summary.registration_state
1356906515	GFM1421	MAZDA	SDN	BLK	40	NY

How many tickets were given to Toyotas in NY

```
SELECT COUNT (*) as count, vehicle_make, registration_state
FROM tickets
WHERE vehicle_make LIKE 'TOYOT' AND registration_state LIKE 'NY'
GROUP BY vehicle_make, registration_state
LIMIT 1;
```

```
0: jdbc:hive2://localhost:10000/default> SELECT COUNT (*) as count, vehicle_make, registration_state
. . . . .> FROM tickets
. . . . .> WHERE vehicle_make LIKE 'TOYOT' AND registration_state LIKE 'NY'
. . . . .> GROUP BY vehicle_make, registration_state
. . . . .> LIMIT 1;
+-----+-----+-----+
| count | vehicle_make | registration_state |
+-----+-----+-----+
| 1955288 | TOYOT | NY |
+-----+-----+-----+
```


How many tickets were given because of expired meters in NY

```
SELECT COUNT (*) as count, violation_code, registration_state
FROM tickets
WHERE violation_code = 37 AND registration_state LIKE 'NY'
GROUP BY violation_code, registration_state
LIMIT 1;
```

```
0: jdbc:hive2://localhost:10000/default> SELECT COUNT (*) as count, violation_code, registration_state
. . . . .> FROM tickets
. . . . .> WHERE violation_code = 37 AND registration_state LIKE 'NY'
. . . . .> GROUP BY violation_code, registration_state
. . . . .> LIMIT 1;

+-----+-----+-----+
| count | violation_code | registration_state |
+-----+-----+-----+
| 1268920 | 37 | NY |
+-----+-----+-----+
```

How many Toyotas were given tickets because of expired meters in NY

```
SELECT COUNT (*) as count, vehicle_make, violation_code, registration_state
FROM tickets
WHERE vehicle_make LIKE 'TOYOT' AND violation_code = 37 AND registration_state LIKE 'NY'
GROUP BY vehicle_make, violation_code, registration_state
LIMIT 1;
```

```
0: jdbc:hive2://localhost:10000/default> SELECT COUNT (*) as count, vehicle_make, violation_code, registration_state
. . . . .> FROM tickets
. . . . .> WHERE vehicle_make LIKE 'TOYOT' AND violation_code = 37 AND registration_state LIKE 'NY'
. . . . .> GROUP BY vehicle_make, violation_code, registration_state
. . . . .> LIMIT 1;

+-----+-----+-----+-----+
| count | vehicle_make | violation_code | registration_state |
+-----+-----+-----+-----+
| 142420 | TOYOT        | 37             | NY                 |
+-----+-----+-----+-----+
```

Top 5 police squads that gave out tickets in NY

```
SELECT COUNT (*) as count, issuer_squad, registration_state
FROM tickets
WHERE registration_state LIKE 'NY'
GROUP BY issuer_squad, registration_state
ORDER BY count DESC
LIMIT 5;
```

```
0: jdbc:hive2://localhost:10000/default> SELECT COUNT (*) as count, issuer_squad, registration_state
. . . . .> FROM tickets
. . . . .> WHERE registration_state LIKE 'NY'
. . . . .> GROUP BY issuer_squad, registration_state
. . . . .> ORDER BY count DESC
. . . . .> LIMIT 5;

+-----+-----+-----+
| count | issuer_squad | registration_state |
+-----+-----+-----+
| 4020046 | 0000 | NY |
| 2214929 | A | NY |
| 894656 | M | NY |
| 888602 | B | NY |
| 791490 |  | NY |
+-----+-----+-----+
```

Top 10 Registration State by number

```
0: jdbc:hive2://localhost:10000/default> Select Count (*) as count, registration_state
- - - - - .> From tickets
- - - - - .> Group by registration_state
- - - - - .> Order by count desc
- - - - - .> limit 10;
```

count	registration_state
9081366	NY
1125070	NJ
313659	PA
177766	CT
155466	FL
102787	MA
85164	VA
67692	MD
64714	NC
61546	IN

Bottom 15 Registration State by number

```
0: jdbc:hive2://localhost:10000/default> Select Count (*) as count, registration_state
- - - - -> From tickets
- - - - -> Group by registration_state
- - - - -> Having Count (*) > 549
- - - - -> Order by count
- - - - -> limit 15;

+-----+-----+
| count | registration_state |
+-----+-----+
| 724   | WY                 |
| 746   | BC                 |
| 866   | NS                 |
| 974   | HI                 |
| 1150  | ND                 |
| 1632  | MT                 |
| 2246  | AK                 |
| 2368  | SD                 |
| 2404  | UT                 |
| 3254  | KS                 |
| 4136  | NE                 |
| 4552  | WV                 |
| 4656  | GV                 |
| 5166  | NV                 |
| 6724  | AR                 |
+-----+-----+
15 rows selected (154.377 seconds)
```

Top 15 Violation times

```
0: jdbc:hive2://localhost:10000/default> Select Count (*) as count, violation_time
- - - - -> From tickets
- - - - -> Group by violation_time
- - - - -> Order by count desc
- - - - -> limit 15;
```

count	violation_time
32230	0836A
28632	1136A
27829	1140A
27596	0936A
27409	0806A
26922	0840A
26557	0906A
26505	0940A
25956	1145A
25033	0945A
24937	1139A
24709	1137A
24683	1138A
24672	0838A
24630	0837A

Bottom 15 Violation time

```
O: jdbc:hive2://localhost:10000/default> Select Count (*) as count, violation_time
- - - - -> From tickets
- - - - -> Group by violation_time
- - - - -> Having Count(*) > 412
- - - - -> Order by count
- - - - -> limit 15;

+-----+-----+
| count | violation_time |
+-----+-----+
| 413   | 0004A         |
| 437   | 0002A         |
| 441   | 0003A         |
| 458   | 0429A         |
| 468   | 0501A         |
| 489   | 0504A         |
| 490   | 0419A         |
| 491   | 0434A         |
| 491   | 0416A         |
| 496   | 0439A         |
| 500   | 0359A         |
| 500   | 0449A         |
| 504   | 0421A         |
| 507   | 0441A         |
| 509   | 0424A         |
+-----+-----+
15 rows selected (133.781 seconds)
```

Top 15 Violations

```
0: jdbc:hive2://localhost:10000/default> Select Count (*) as count, violation_description
- - - - -> From tickets
- - - - -> Group by violation_description
- - - - -> Order by count desc
- - - - -> limit 15;
```

count	violation_description
3372231	
2631892	38-Failure to Display Muni Rec
2341312	21-No Parking (street clean)
2260668	FAILURE TO STOP AT RED LIGHT
1805166	PHTO SCHOOL ZN SPEED VIOLATION
1615408	14-No Standing
1559830	37-Expired Muni Meter
881972	71A-Insp Sticker Expired (NYS)
881250	20A-No Parking (Non-COM)
759048	40-Fire Hydrant
697904	BUS LANE VIOLATION
551998	69-Failure to Disp Muni Recpt
437642	19-No Stand (bus stop)
358706	31-No Stand (Com. Mtr. Zone)
358476	70A-Reg. Sticker Expired (NYS)

New York VS Non-New York Violations

count	violation_description
2214903	
2063544	38-Failure to Display Muni Rec
1870814	FAILURE TO STOP AT RED LIGHT
1751506	21-No Parking (street clean)
1572378	PHTO SCHOOL ZN SPEED VIOLATION
1256152	37-Expired Muni Meter
1121810	14-No Standing
881416	71A-Insp Sticker Expired (NYS)
649216	20A-No Parking (Non-COM)
576832	BUS LANE VIOLATION
544224	40-Fire Hydrant
431592	69-Failure to Disp Muni Recpt
358310	70A-Reg. Sticker Expired (NYS)
303314	19-No Stand (bus stop)
255174	46A-Double Parking (Non-COM)

(New York)

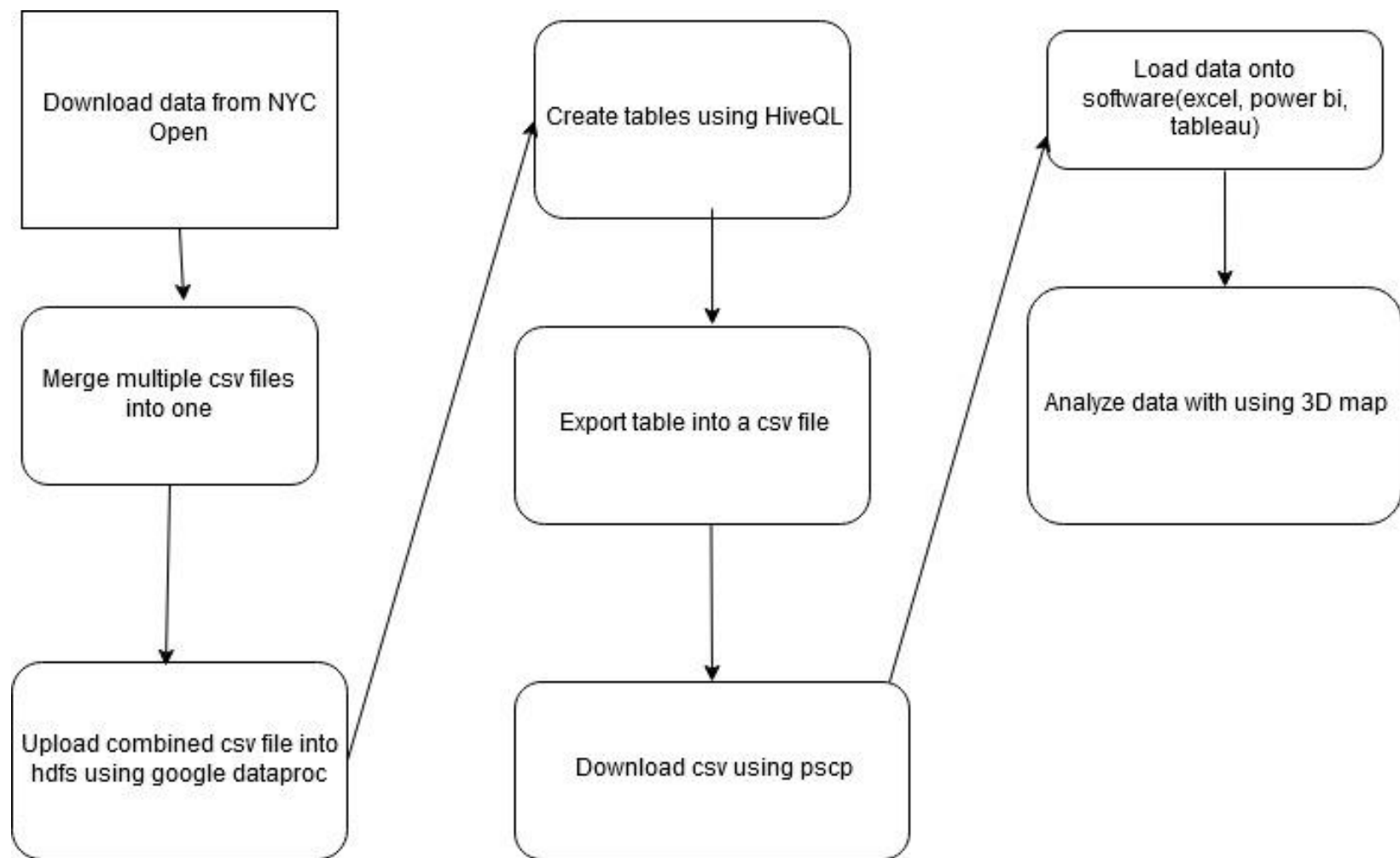
count	violation_description
1157328	
589806	21-No Parking (street clean)
568348	38-Failure to Display Muni Rec
493598	14-No Standing
389854	FAILURE TO STOP AT RED LIGHT
303678	37-Expired Muni Meter
232788	PHTO SCHOOL ZN SPEED VIOLATION
232034	20A-No Parking (Non-COM)
214824	40-Fire Hydrant
149872	31-No Stand (Com. Mtr. Zone)
134328	19-No Stand (bus stop)
121072	BUS LANE VIOLATION
120406	69-Failure to Disp Muni Recpt
83946	46A-Double Parking (Non-COM)
83648	84-Platform lifts in low posit

Not New York

Top 15 Vehicle Body Types

```
0: jdbc:hive2://localhost:10000/default> Select Count (*) as count, vehicle_body_type
- - - - -> From tickets
- - - - -> Group by vehicle_body_type
- - - - -> Order by count desc
- - - - -> limit 15;
```

count	vehicle_body_type
7326941	SUBN
6640760	4DSD
3236722	VAN
1749424	DELV
1185478	SDN
638836	2DSD
543162	PICK
269892	
183018	UTIL
154950	TAXI
145240	REFG
142568	4 DR
141416	TRAC
122268	TRLR
113736	BUS



Visualization

