**Name:**XinYue Liu

**Student ID:**1332-0443-43

**Email:**[liuxinyu@usc.edu](mailto:liuxinyu@usc.edu)

**Part 1:**

**Query #1**

#standardSQL

SELECT

name, count

FROM

`babynames.names\_2014`

WHERE

gender = 'F'

AND

name LIKE "\_i%"

ORDER BY count DESC;

**Query #2**

#standardSQL

SELECT

SUM(count)

FROM

`babynames.names\_2014`

WHERE

name LIKE "Xin%";

**result: 32**

**Part 2:**

**Query 2nd cell:**

%%bq query

#find the weekday of 1994/05/09

SELECT wday

FROM

`publicdata.samples.natality`

WHERE

year = 1994

AND

month = 5

AND

day = 9;

**Query 3rd cell:**

%%bq query --name total\_births

#find the number of people born on 05/09 in different years

SELECT CAST(source\_year AS string) AS year, COUNT(\*) AS birth\_count

FROM `publicdata.samples.natality`

WHERE

month = 5

AND

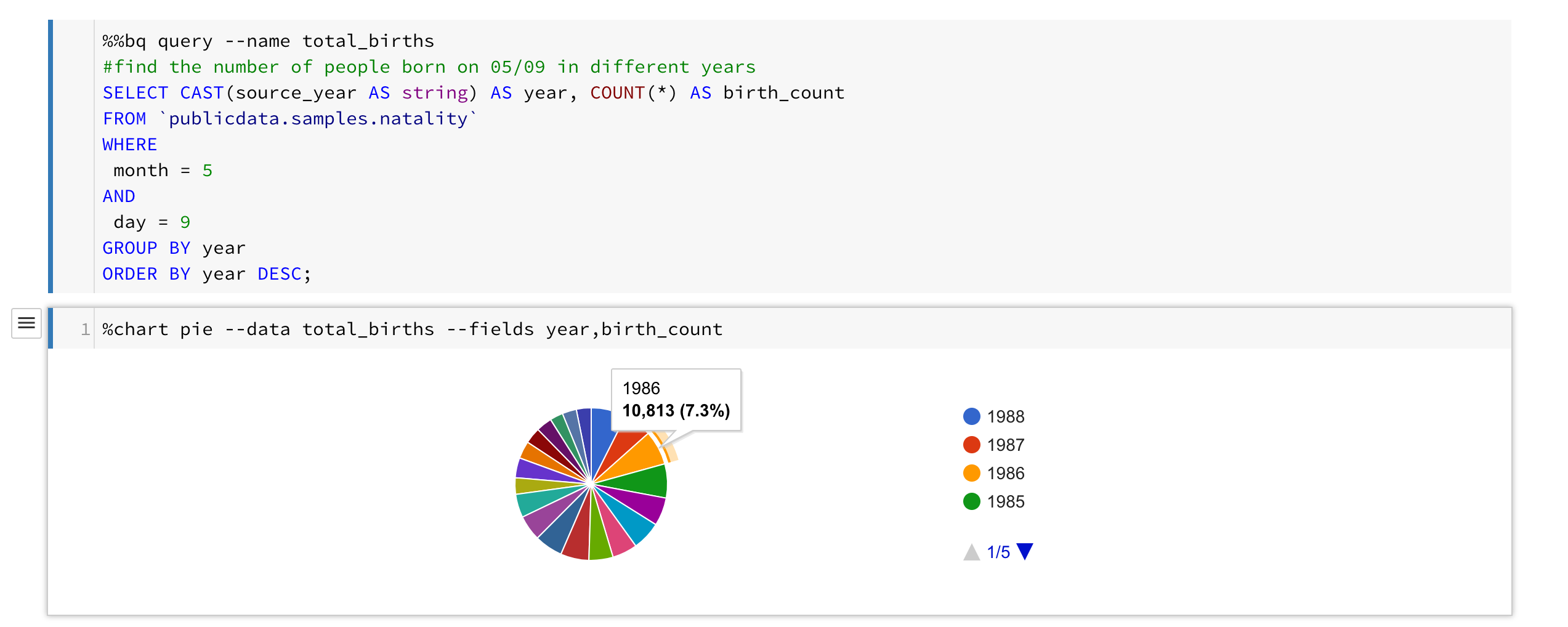
day = 9

GROUP BY year

ORDER BY year DESC;

%chart pie --data total\_births --fields year,birth\_count

**Snapshot of visualization:**

****

**Part 3:**

**1. BigQuery:**

SELECT DATE(pickup\_datetime) AS dateTime, SUM(passenger\_count) AS totalPassenger

FROM `nyc-tlc.yellow.trips`

WHERE DATE(pickup\_datetime) < '2015-01-01'

AND DATE(dropoff\_datetime) < '2015-01-01'

GROUP BY dateTime

ORDER BY dateTime;

**2.Datalab:**

%%bq query

SELECT DATE(pickup\_datetime) AS dateTime, SUM(passenger\_count) AS totalPassenger

FROM `nyc-tlc.yellow.trips`

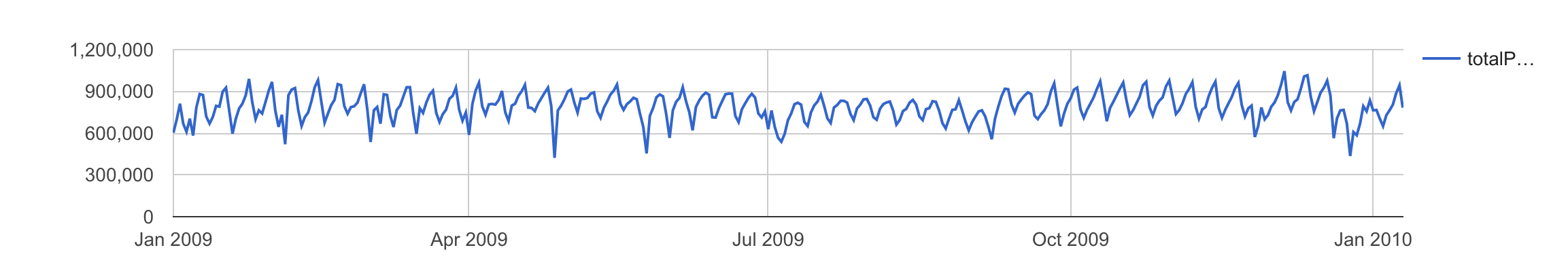
WHERE DATE(pickup\_datetime) < '2015-01-01'

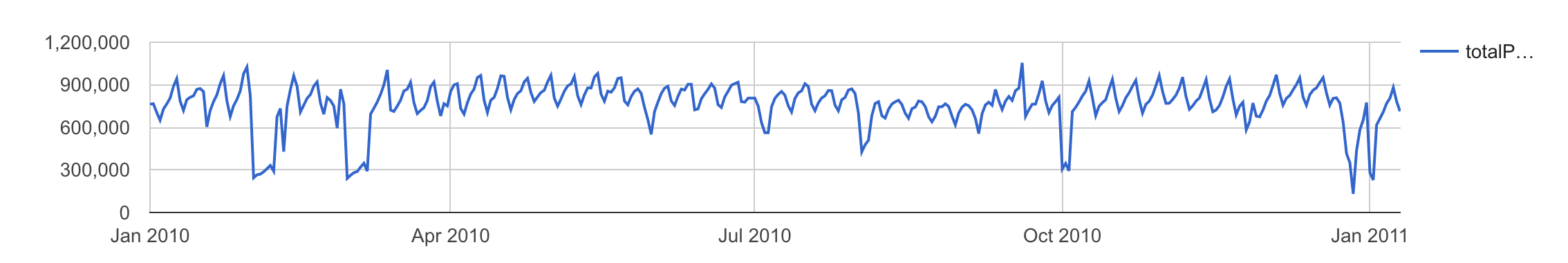
AND DATE(dropoff\_datetime) < '2015-01-01'

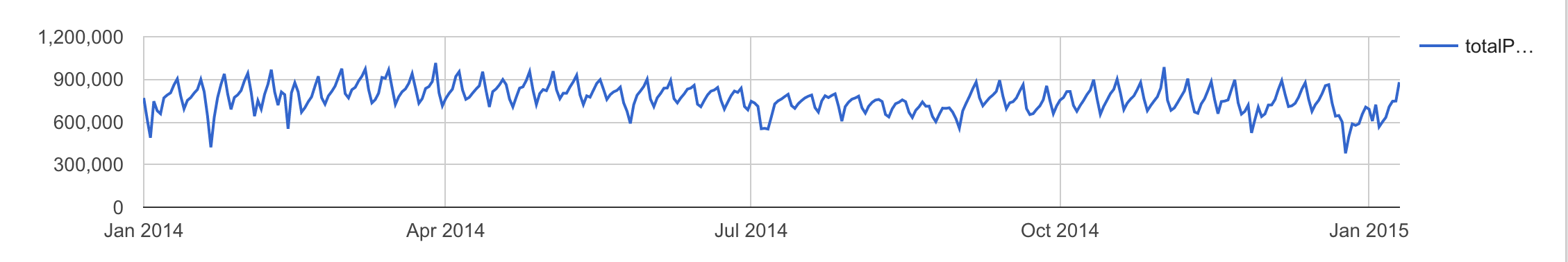
GROUP BY dateTime

ORDER BY dateTime;

**Visualization:**







**Semi-periodical pattern:**

Based on my observation, I think a peak of a week is Sunday.

**Query to prove:**

%%bq query

SELECT dateTime,EXTRACT(DAYOFWEEK FROM dateTime) dayOfweek

FROM

(SELECT DATE(pickup\_datetime) AS dateTime, SUM(passenger\_count) AS totalPassenger

FROM `nyc-tlc.yellow.trips`

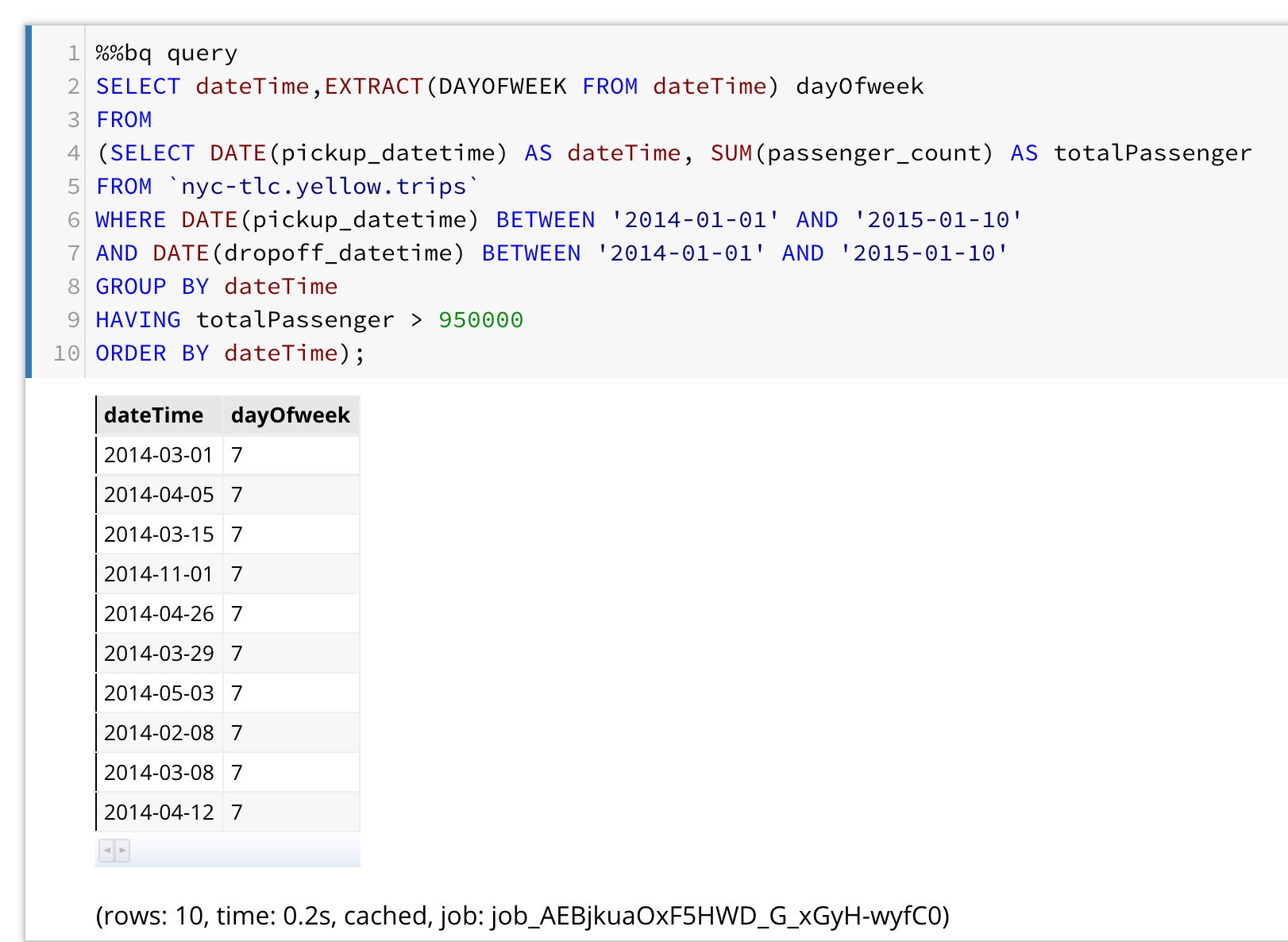
WHERE DATE(pickup\_datetime) BETWEEN '2010-01-01' AND '2011-01-10'

AND DATE(dropoff\_datetime) BETWEEN '2010-01-01' AND '2011-01-10'

GROUP BY dateTime

HAVING totalPassenger > 950000

ORDER BY dateTime);



**Two unusual patterns (two holidays):**

1. **Martin Luther King Jr. Day**
2. **Christmas Day**

**Explanation:** Due to those two holidays, people are more willing to stay home with their families. Therefore, there are big decreases in numbers around these two holidays.

**Minimum Point:**

### 2011-08-28 [Hurricane Irene hits New York](https://www.theguardian.com/world/blog/2011/aug/28/hurricane-irene-new-york-live-coverage1)

### 2012-10-29 [Time Lapse of Hurricane Sandy hitting New York City October 29th ...](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=0ahUKEwir04Cm1tXUAhVGHGMKHYGgBm4QtwIIMDAC&url=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3DOfnSIAVdfzU&usg=AFQjCNG8oTcZaY2lEfdQm7rVhAT1mkcRFQ)

### 2013-08-04 In August 2013 the City of New York introduced a restricted class of “boro” (or “green”) taxis. Boro taxis provide an equivalent service to that of “yellow” medallion taxis, but are restricted from pickups in Manhattan south of a boundary along East 96th St and West 110th St.

### Bonus:

### Query:

#legacySQL

SELECT pickup\_longitude,pickup\_latitude

FROM [nyc-tlc:yellow.trips]

WHERE DATE(pickup\_datetime) < '2013-01-01'

AND DATE(dropoff\_datetime) < '2013-12-31'

AND HOUR(pickup\_datetime) > 18

AND pickup\_longitude != 0

AND pickup\_latitude != 0

AND total\_amount >= 300.0

AND total\_amount <= 400.0;

**[Download as pickup.csv]**

#legacySQL

SELECT dropoff\_longitude,dropoff\_latitude

FROM [nyc-tlc:yellow.trips]

WHERE DATE(pickup\_datetime) < '2013-01-01'

AND DATE(dropoff\_datetime) < '2013-12-31'

AND HOUR(pickup\_datetime) > 18

AND dropoff\_longitude != 0

AND dropoff\_latitude != 0

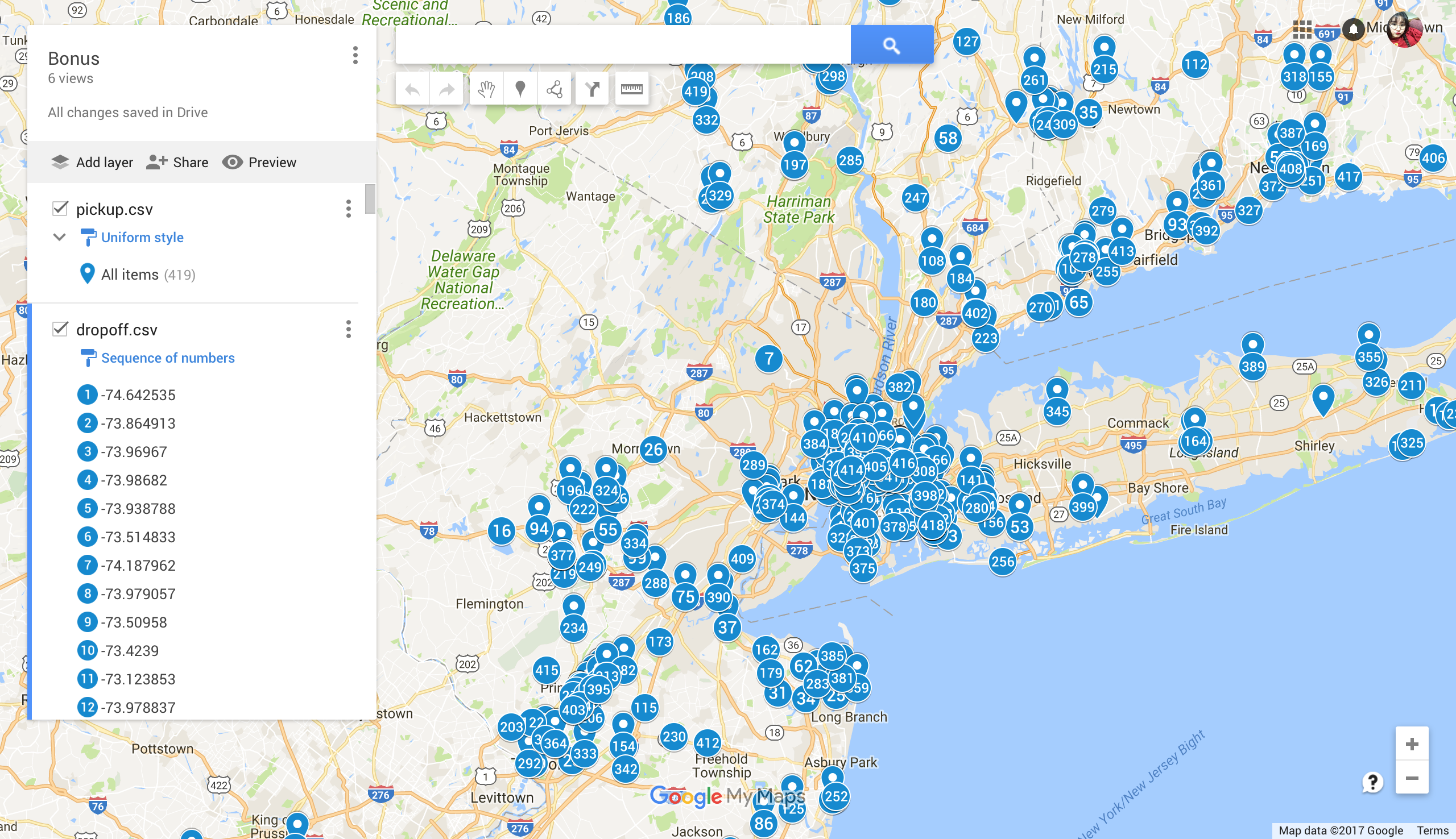
AND total\_amount >= 300.0

AND total\_amount <= 400.0;

**[Download as dropoff.csv]**

**Load pickup.csv and dropoff.csv into google my maps like below:**

[**https://drive.google.com/open?id=1\_aUWKbNzHknMP9w7jkke1TYdYT4&usp=sharing**](https://drive.google.com/open?id=1_aUWKbNzHknMP9w7jkke1TYdYT4&usp=sharing)

****