1. SELECT

tripduration,

CURRENT\_DATE() as currentdate,

CURRENT\_TIME() as currnettime,

CURRENT\_TIMESTAMP() as currentimestamp,

CAST(starttime AS date) as start\_time,

CAST(stoptime as date) as stop\_time,

cast(starttime as STRING) as test,

count(distinct bikeid) as num\_bikes,

CONCAT(usertype,'-',gender) as CONCAT\_test,

substr(gender,1,1) as gender,

substr(usertype,1,3) as subscriber,

ARRAY\_AGG(usertype order by  usertype) as avg\_type,

#ARRAY\_CONCAT\_AGG(bikeid) as bikeid or ARRAY\_CONCAT(bikeid), #must be array type need in here

cast(concat(cast(extract(year from stoptime) as STRING),'-','01','-01') as date) as date,

date\_diff(current\_date(),cast(stoptime as date),DAY) as day\_stop\_time\_to\_current\_time,

generate\_timestamp\_array(TIMESTAMP(DATE\_SUB(CURRENT\_DATE("GMT"),INTERVAL 2 DAY)),TIMESTAMP(current\_date("Australia/Sydney")),INTERVAL 1 HOUR) as days\_hour

FROM `bigquery-public-data.new\_york\_citibike.citibike\_trips`

where gender='female'

group by tripduration, starttime,bikeid,CONCAT\_test,gender,subscriber,stoptime limit 100;



SELECT \* FROM `bigquery-public-data.google\_analytics\_sample.ga\_sessions\_20170801` LIMIT 1000;

SELECT \* FROM `bigquery-public-data.new\_york\_citibike.citibike\_trips` limit 100;

select timestamp(current\_datetime("Australia/Sydney"))as now;

SELECT date\_diff(current\_date(),cast(stoptime as date),DAY) as day\_stop\_time\_to\_current\_time

FROM `bigquery-public-data.new\_york\_citibike.citibike\_trips`;

1. Conver String date to normal date format :

SELECT channelGrouping,

CAST(CONCAT(SUBSTR(date, 0 , 4),'-',SUBSTR(date, 5 , 2),'-',SUBSTR(date, 7 , 2))as DATE )AS date,

pageview,

sum(pageview) over (PARTITION BY date) totalpageViews

FROM(

SELECT channelGrouping,

  date,

 sum(totals.pageViews) as pageview ,

  FROM `bigquery-public-data.google\_analytics\_sample.ga\_sessions\_20170801` group by channelGrouping,date LIMIT 1000)

1. Array :

1)SELECT generate\_timestamp\_array(TIMESTAMP(DATE\_SUB(CURRENT\_DATE("GMT"),INTERVAL 2 DAY)),TIMESTAMP(current\_date("Australia/Sydney")),INTERVAL 1 HOUR) as days\_hour

2) SELECT ARRAY<FLOAT64>[1,2,3] AS float;

3)SELECT tripduration, ARRAY\_AGG(birth\_year) AS birth\_year, FROM `bigquery-public-data.new\_york\_citibike.citibike\_trips` GROUP BY tripduration LIMIT  10;

1. SELECT  GENERATE\_DATE\_ARRAY(CAST(starttime AS date),CAST(stoptime AS date), INTERVAL 1 DAY) AS date\_array FROM   `bigquery-public-data..new\_york\_citibike.citibike\_trips` LIMIT   1000;
2. WITH

  last\_3\_days\_datetime AS (

  SELECT

    generate\_timestamp\_array(TIMESTAMP(DATE\_SUB(CURRENT\_DATE("GMT"),INTERVAL 2 DAY)),

      TIMESTAMP\_TRUNC(TIMESTAMP(current\_datetime("GMT")),HOUR),

      INTERVAL 1 HOUR) days\_hours),

  hours\_list AS (

  SELECT

    datetime(timestamp\_hour) starttime

  FROM

    last\_3\_days\_datetime,UNNEST(days\_hours) as timestamp\_hour

  WHERE

    datetime\_diff(current\_datetime("GMT"),

      datetime(timestamp\_hour),

      HOUR)>=2),

  hours\_list\_by AS (

  SELECT

    usertype,

    cast(starttime as date) as starttime

  FROM

    hours\_list,

    UNNEST(['subscriber','Consumer']) AS usertype)

SELECT

  \*

FROM

  hours\_list\_by;

1. select date\_sub(current\_date("GMT"),INTERVAL 1 DAY) as IND\_DAY,date\_sub(current\_datetime("GMT"),INTERVAL 1 HOUR) as IND\_HOUR
2. SELECT

  starttime,

  stoptime

FROM

  `bigquery-public-data.new\_york\_citibike.citibike\_trips`

WHERE

  UPPER(gender) LIKE UPPER(CONCAT("%",gender,"%"))

  OR LOWER(gender) LIKE LOWER(CONCAT("%",gender,"%"))

LIMIT

  1000;

1. SELECT

  UPPER(gender)

FROM

  `bigquery-public-data.new\_york\_citibike.citibike\_trips`

WHERE

  gender="male"

LIMIT

  100;

1. SELECT

  FORMAT\_DATE("%A",starttime) AS week\_days,

  FORMAT\_DATE("%C",starttime)AS week\_date,

  FORMAT\_DATE("%B",starttime) AS Month

FROM

  `bigquery-public-data.new\_york\_citibike.citibike\_trips`

LIMIT

  100;

1. WITH

  aa AS (

  SELECT

    FORMAT\_DATE("%A",starttime) AS week\_days,

    COUNT(1) AS numsub,

    SUM(

    IF

      (starttime BETWEEN DATE\_SUB(CURRENT\_DATE("GMT"),INTERVAL 7 day)

        AND DATE\_SUB(CURRENT\_DATE("GMT"),INTERVAL 1 day),

        1,

        0)) AS weeks

  FROM

    `bigquery-public-data.new\_york\_citibike.citibike\_trips` group by starttime),

  bb AS (

  SELECT

    ARRAY\_AGG(STRUCT(week\_days))

  FROM

    aa

  LIMIT

    100 )

    select \* from bb

1. SELECT EXTRACT(DAY FROM current\_date('GMT')) AS the\_day, EXTRACT(Month FROM current\_date('GMT')) AS the\_month, EXTRACT(Year FROM current\_date('GMT')) AS the\_year;
2. SELECT extract(DAY FROM CURRENT\_DATE("GMT")) AS DAY, FORMAT\_DATE('%A', CURRENT\_DATE("GMT")) AS TodaysDay,FORMAT\_DATE("%B",CURRENT\_DATE("GMT")) AS Month
3. select FORMAT\_DATE('%A', CURRENT\_DATE("GMT")) AS TodaysDay

BigQuery nested and repeated

1. create or replace table `test.demo` as(select 'cafe pacific' as name ,'new york' as location,['indian','marathi'] as food
2. union all
3. select 'boston pizza' as name,'toranto' as location,['indian','chines','maggi'] as food);
4. select \* from `test.demo`;
5. select name,location,food from `test.demo`, unnest(food) as food;
6. select name,ARRAY\_LENGTH(food) as food from `test.demo`;
7. select \* from  `test.demo` where 'marathi' in unnest(food);
8. create or replace table `test.struct\_demo` as (select 'cafe pacific' as name ,'new york' as location,STRUCT(['indian','marathi'] as food\_country,'100$' as price) as food
9. union all
10. select 'boston pizza' as name,'toranto' as location,Struct(['indian','chines','maggi'] as food\_country,'200$' as price) as food);
11. select \* from `test.struct\_demo`;
12. select name,location,food.price as price\_range,country from `test.struct\_demo`, unnest(food.food\_country) as country;

BigQuery nested and repeated inside nested and repeated query read data

1. SELECT  \* FROM `bigquery-public-data.google\_analytics\_sample.ga\_sessions\_20170801` LIMIT 1000;
2. select hits[offset(0)] as hits, hits[offset(0)].page.pageTitle as pageTitle from `bigquery-public-data.google\_analytics\_sample.ga\_sessions\_20170801` LIMIT 1000;
3. Array out of index error got in offset then use safe\_offset:

select hits[offset(0)] as hits, hits[offset(0)].page.pageTitle as pageTitle ,hits[offset(0)].product[SAFE\_OFFSET(0)].productSKU as productSKU from `bigquery-public-data.google\_analytics\_sample.ga\_sessions\_20170801` LIMIT 1000;

1. select date\_sub(current\_date(),interval 1 day) as date;

#Boook Bigqueries for practice

#300 and 600 second means 5 and 10 minutes

1. select \* from `bigquery-public-data.new\_york\_citibike.citibike\_trips`;
2. select gender,tripduration from `bigquery-public-data.new\_york\_citibike.citibike\_trips` where  gender='female' or gender='male' and tripduration>=300 and tripduration < 600 limit 10;
3. select gender,tripduration/60 as minutes from `bigquery-public-data.new\_york\_citibike.citibike\_trips` where (tripduration/60) < 10 limit 10;
4. select \* from `bigquery-public-data.new\_york\_citibike.citibike\_stations` limit 100;
5. select \* from `bigquery-public-data.new\_york\_citibike.citibike\_stations` where name like '%Riverside%';
6. select \* except(short\_name,last\_reported) from `bigquery-public-data.new\_york\_citibike.citibike\_stations` where name like '%Riverside%';
7. with all\_trip as (select gender,tripduration/60 as minutes from `bigquery-public-data.new\_york\_citibike.citibike\_trips`)
8. select \* from all\_trip where minutes<10 limit 10;

#Order by Clause

1. select gender,tripduration/60 as minutes from `bigquery-public-data.new\_york\_citibike.citibike\_trips` where  gender='female' order by minutes DESC;
2. select gender,avg(tripduration/60) as minutes from `bigquery-public-data.new\_york\_citibike.citibike\_trips` where  tripduration is not null  group by gender order by minutes;

#Count clause

1. select gender,count(\*) as rides, avg(tripduration/60) as minutes from `bigquery-public-data.new\_york\_citibike.citibike\_trips` where tripduration is not null  group by gender order by minutes;

#Having  clause

1. select gender,avg(tripduration/60) as minutes from `bigquery-public-data.new\_york\_citibike.citibike\_trips` group by gender having minutes >14 order by minutes;

#Distinct Clause

1. select DISTINCT gender from `bigquery-public-data.new\_york\_citibike.citibike\_trips`;
2. select count(DISTINCT gender) as gender\_count,gender,usertype from `bigquery-public-data.new\_york\_citibike.citibike\_trips` where gender !=" " group by 2,3;

#Split and unnest

1. select city,split(city,' ') as parts from (select \* from unnest(['Seattle WA','New York','Singapore'])as city);

#Unnest : unnest must be array use after from

1. select date,channelGrouping,h.isEntrance,h.page.pagePath,totals.transactions,totals.transactionRevenue from `bigquery-public-data.google\_analytics\_sample.ga\_sessions\_20170801` cross join unnest(hits) as h where h.isEntrance=True limit 10;