Name: Sheethal Melnarse
Dataset: London Bicycles
Purpose: Data Analysis with SQL
Source: BigQuery
Date: 27 March 2020

 # What is the size of this dataset? How many trips?

SELECT Count(*) FROM `bigquery-public-data.london_bicycles.cycle_hire` LIMIT 1000

 # When was the first and last trip in this dataset?

SELECT min(start_date), max(end_date) FROM `bigquery-public-data.london_bicycles.cycle_hire`;

 # How many bikes are there?

SELECT COUNT(DISTINCT bike_id) FROM `bigquery-public-data.london_bicycles.cycle_hire`;

 # How many stations are there?

SELECT COUNT(DISTINCT id) FROM `bigquery-public-data.london_bicycles.cycle_stations`;

 # How many trips are in the morning, and how many are in the afternoon?

SELECT CASE WHEN EXTRACT(hour FROM start_date) BETWEEN 6 and 11 THEN "morning"

WHEN EXTRACT(hour FROM start_date) BETWEEN 12 and 18 THEN "afternoon"

ELSE Null END as time_of_day,

COUNT(distinct rental_id) FROM `bigquery-public-data.london_bicycles.cycle_hire` GROUP BY time_of_day;

 # Percentage of empty bikes in every station

SELECT nbEmptyDocks/docks_count*100 as percentage, name FROM `bigquery-public-data.london_bicycles.cycle_stations`

ORDER BY percentage desc LIMIT 1000;

Total number of bikes in station and total docks

SELECT SUM(bikes_count) as total_bikes, SUM(docks_count) as docks_count FROM `bigquery-public-data.london_bicycles.cycle_stations` LIMIT 1000;

Percentage of rides where start and end station is same

SELECT (SELECT count(rental_id) as rent FROM `bigquery-public-data.london_bicycles.cycle_hire`

where end_station_id = start_station_id)/count(*)*100 FROM `bigquery-public-data.london_bicycles.cycle_hire` ;

Busiest hour in a day

SELECT count(*), EXTRACT(HOUR FROM start_date) as hour

FROM `bigquery-public-data.london_bicycles.cycle_hire`

group by hour

order by count(*) desc;

Busiest start station

SELECT count(*), start_station_name, EXTRACT(HOUR FROM start_date) as hour

FROM `bigquery-public-data.london_bicycles.cycle_hire`

group by start_station_name, hour

order by count(*) desc;

Customer who use bicycles for longer duration

SELECT start_station_name, end_station_name

FROM `bigquery-public-data.london_bicycles.cycle_hire`

where duration>900

GROUP BY start_station_name,end_station_name;

Visitors during peak hours in the evening

```
SELECT count(*), start_station_name, end_station_name, EXTRACT(HOUR FROM
start_date) as hour

FROM `bigquery-public-

data.london_bicycles.cycle_hire` where duration

> 900

group by

start_station_name,end_station_name, hour

order by count(*) desc;
```

Empty docks by station

```
SELECT nbEmptyDocks as empty , name

FROM `bigquery-public-

data.london_bicycles.cycle_stations` order by empty

desc;
```