1. SQL_DAY3_HW1

```
-- 1. has more than 1000 trips
--
2. has average trip time less than the average trip time of all trips in san Franciso
-- 3. does not belong to the san Francisco region

select station_id, station_name, cnt_trip_station, duration as avg_trip_time_station
from `gg-data-bangsa.gda01221_irawansyah.sql_day2_task2_2`
where cnt_trip_station > 1000 AND duration < (
    select
        avg(trips.duration_sec)
    from `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips` trips
    INNER JOIN `bigquery-public-
data.san_francisco_bikeshare.bikeshare_station_info` info
    on trips.start_station_id = info.station_id
    INNER JOIN `bigquery-public-
data.san_francisco_bikeshare.bikeshare_regions` regions
    on  regions.region_id = info.region_id
    where regions.name = 'San Francisco')
    and region_name != 'San Francisco'
```

Row	station_id	station_name	cnt_trip_station	avg_trip_time_station
1	208	4th St at San Carlos St	1342	907.8159463487326
2	280	San Fernando St at 7th St	2429	768.496088925484
3	296	5th St at Virginia St	2448	658.8153594771239
4	305	Ryland Park	2434	821.1047658175845
5	310	San Fernando St at 4th St	2894	835.0283344851409
6	311	Paseo De San Antonio at 2nd St	1772	896.6822799097059
7	312	San Jose Diridon Station	4040	690.235148514851
8	317	San Salvador St at 9th St	1698	847.7444051825679
9	18	Telegraph Ave at Alcatraz Ave	1155	1016.1220779220788
10	158	Shattuck Ave at Telegraph Ave	1668	696.959232613909
11	160	West Oakland BART Station	3407	986.7848547108906
12	162	Franklin St at 9th St	2381	870.3716925661486
13	163	Lake Merritt BART Station	5695	799.3041264266899
14	167	62nd St at Claremont Ave	1332	913.2454954954949
15	171	Rockridge BART Station	1564	704.1317135549871

2. SQL_DAY3_HW2

```
-- 1. First find out the time gap/duration in days between one trip and the next trip
-- 2. Next get the average of these time gaps for each region

WITH
X as (
SELECT
   trip_id,
   start_date,
   start_station_name,
   end_station_name,
   bike_number,
```

```
   LAG(start_date) OVER(PARTITION BY start_station_name ORDER BY start_date) prev_destination,
   LEAD(start_date) OVER(PARTITION BY start_station_name ORDER BY start_date) next_destination
,
   FROM `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips`),
Y as (
SELECT trip_id,
    start_station_name,
    TIMESTAMP_DIFF(next_destination, prev_destination, DAY) as Day
    FROM X
    WHERE prev_destination is not null and next_destination is not null
    )

SELECT b.region_id, b.name AS region_name, AVG(trips.Day) AS avg_day_diff
FROM Y as trips
JOIN `bigquery-public-data.san_francisco_bikeshare.bikeshare_station_info` a
  ON a.name=trips.start_station_name
LEFT JOIN  `bigquery-public-data.san_francisco_bikeshare.bikeshare_regions` b
  ON b.region_id=a.region_id
GROUP BY 1,2
```

Row	region_id	region_name	avg_day_diff
1	12	Oakland	0.07481559536354059
2	14	Berkeley	0.04848055154878617
3	13	Emeryville	0.23467259835951632
4	3	San Francisco	0.003438376335738076
5	5	San Jose	0.2761470830662149