Section 3: Data Analysis with SQL

Business Report

The topic in this business report can best be treated under three headings: Understanding the problems, Finding the related business insights with SQL and recommendations.

Firstly, our users complained that customers frequently experience “empty” bike stations. I then using data analysis with SQL to find out empty stations and available ones. There are 758 available stations and 23 empty stations; as a result, we could expect that the proportion is merely **3%,** which is contrasting to our customers’ complaints.

There might be a risk that the available bikes are not working or in good condition. Instead of research empty (0) bikes, I look at the distribution of station has less than 5 bikes on the map of London. Unfortunately, there are many stations have **less than 5 bikes**. (Appendix 2 will provide a whole map)

Turning now to related business insights, we look at the busy stations, popular trips and usage peak (Appendix 4, 5 and 6). At first glance, the busiest stations are 191 Hyde Park, 307 Black Lion Gate, 303 Albert Gate, 785 Aquatic Centre, and 248 Triangle Car Park. It is an interesting business insight that the start stations and end stations are similar, we could assume that these are **tourism** trips; users ride around landmarks; moreover, they are also popular trips. Hence, we could expect that the majority of popular rides are for tourism. The third business insight relates to usage peak which is **8 AM** in **May** and **March**. May and March are considered as sweet months in London in which many events and outdoor activities (Reynolds, 2020).

Moving on now to consider recommendations, I strongly recommend reallocation of bikes and stations, according to SQL; I acknowledge that the company also has many free bikes which are located in various places; the map shows all the unused stations where the amount of **bikes equal to docks** (Appendix 3). If we could allocate them more appropriately; we could significantly reduce customer complaints. Additionally, our marketing plan should target on tourism users.

References

Reynolds, L., 2020. 44 Brilliant Things To Do In London In May 2019. [online] Londonist. Available at: <https://londonist.com/london/things-to-do/things-to-do-in-london-in-may> [Accessed 29 March 2020].

Appendix

1. ***Can you find any traces of empty stations?***

```
SELECT *
FROM `bigquery-public-data.london_bicycles.cycle_stations`
WHERE bikes_count = 0;
```

If yes, how big is this problem?

```
SELECT COUNT(*)  AS AVAILABLE_STATION,
 (SELECT COUNT(*)
  FROM `bigquery-public-data.london_bicycles.cycle_stations`
  WHERE bikes_count = 0)
  AS EMPTY_STATION,
 ROUND((SELECT COUNT(*)
  FROM `bigquery-public-data.london_bicycles.cycle_stations`
  WHERE bikes_count = 0) /
 (SELECT COUNT(*)
  FROM `bigquery-public-data.london_bicycles.cycle_stations`
  WHERE bikes_count != 0),3) AS EMPTY_OVER_AVAILABLE
FROM `bigquery-public-data.london_bicycles.cycle_stations` WHERE bikes_count != 0;
```

Query results SAVE RESULTS EXPLORE DATA

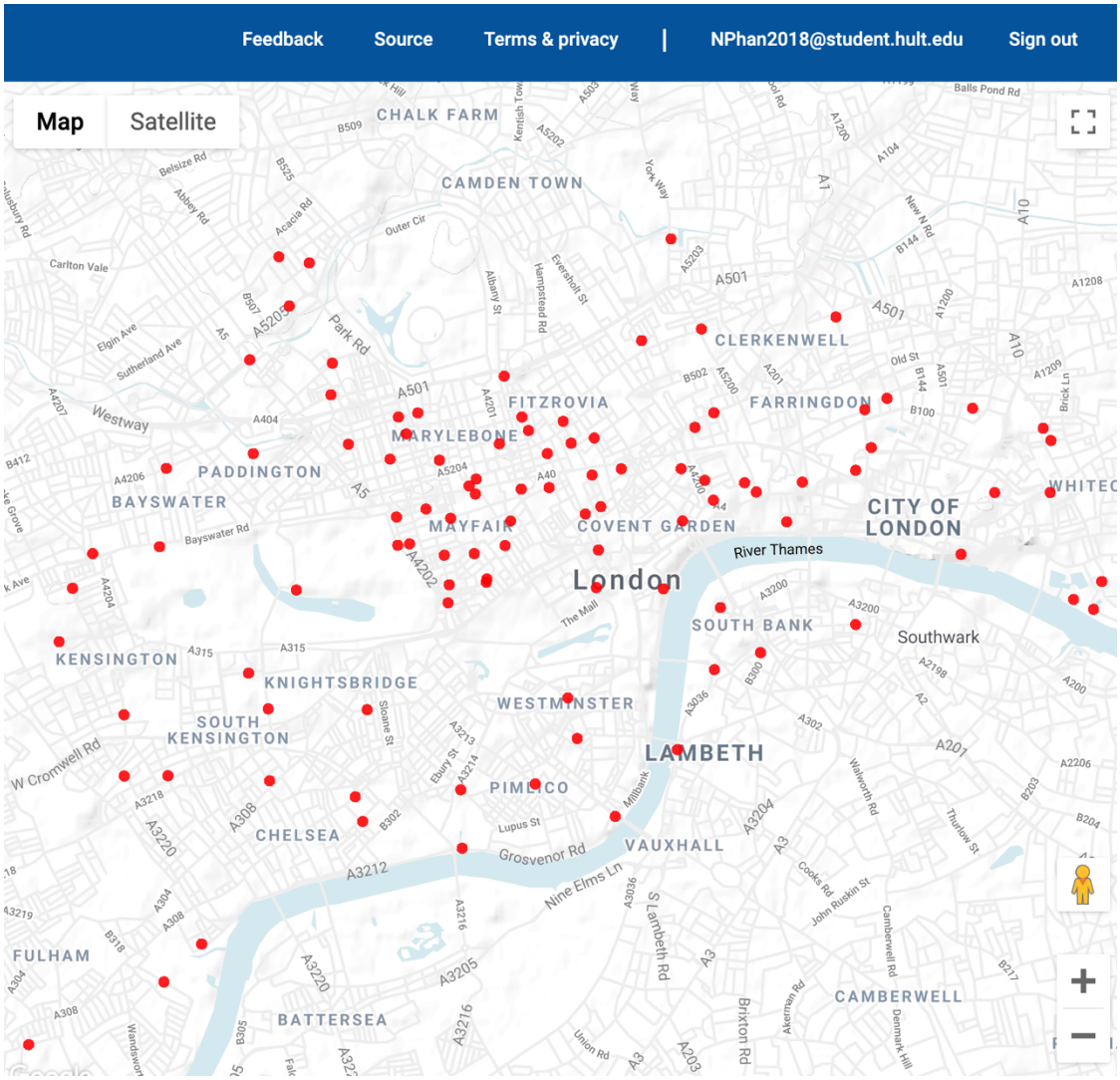
Query complete (0.4 sec elapsed, 6.1 KB processed)

Job information | Results | JSON | Execution details

Row	AVAILABLE_STATION	EMPTY_STATION	EMPTY_OVER_AVAILABLE
1	758	23	0.03

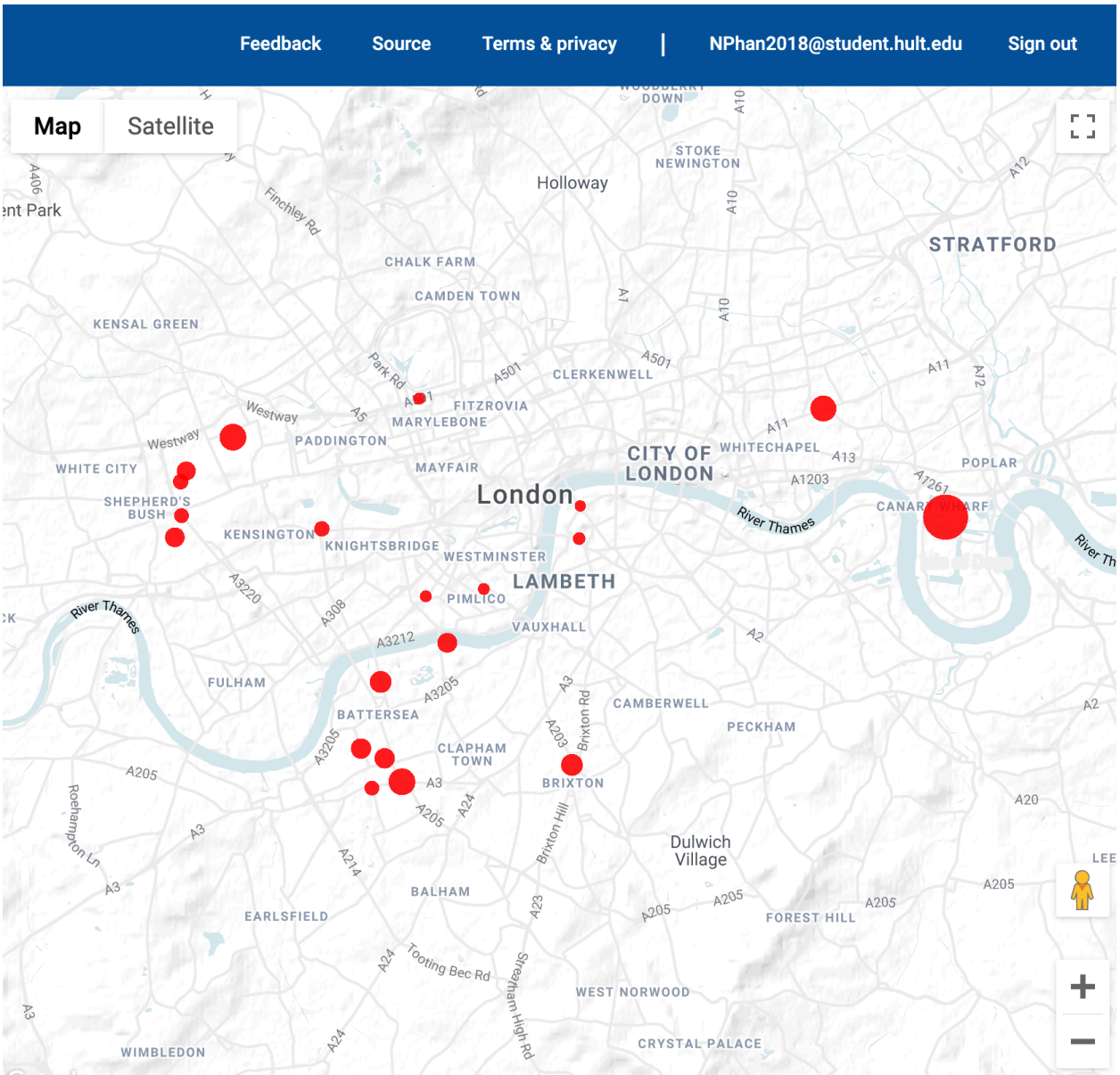
2. ***We could consider using map to find out the distribution of station less than 5.***

```
SELECT
  ST_GeogPoint(longitude, latitude)  AS WKT,
  bikes_count
FROM
  `bigquery-public-data.london_bicycles.cycle_stations`
WHERE bikes_count < 5;
```



3. ***A further business insights from distribution of free stations, where bikes are full.***

```
SELECT
  ST_GeogPoint(longitude, latitude)  AS WKT,
  bikes_count
FROM
  `bigquery-public-data.london_bicycles.cycle_stations`
WHERE bikes_count = docks_count;
```



4. What are the most popular stations in the network?

```
SELECT count(*) as frequency, end_station_id  , start_station_id
FROM `bigquery-public-data.london_bicycles.cycle_hire`
WHERE end_station_id is not null and start_station_id is not null
GROUP BY end_station_id, start_station_id
ORDER BY frequency DESC;
```

Query results SAVE RESULTS EXPLORE DATA

Query complete (8.3 sec elapsed, 368.3 MB processed)

Job information | Results | JSON | Execution details

Row	frequency	end_station_id	start_station_id
1	58269	191	191
2	31961	307	307
3	31218	303	303
4	23572	785	785
5	21273	248	248
6	15419	407	407

5. ***When does their usage peak?***

```
SELECT
COUNT(*) as Frequency,
EXTRACT (DAYOFWEEK FROM start_date) as Day,
EXTRACT (HOUR FROM start_date) as Hour,
EXTRACT (MONTH FROM start_date) as Month
FROM `bigquery-public-data.london_bicycles.cycle_hire`
Group by Day, Hour, Month
ORDER BY Frequency DESC;
```

Query results SAVE RESULTS EXPLORE DATA

Query complete (6.8 sec elapsed, 185.9 MB processed)

Job information | Results | JSON | Execution details

Row	Frequency	Day	Hour	Month
1	58466	4	8	5
2	58019	3	8	5
3	54982	3	8	3
4	54360	5	8	3
5	54002	5	8	5

6. *What are the most popular trips in the network?*

```
SELECT concat(end_station_id,' ',start_station_id) as Trip,
count(concat(end_station_id,' ',start_station_id)) as Number_of_Trip,
start_station_name,
end_station_name
FROM `bigquery-public-data.london_bicycles.cycle_hire`
WHERE end_station_id is not null and start_station_id is not null
Group by Trip, start_station_name, end_station_name
Order by Number_of_Trip DESC;
```

Query results SAVE RESULTS EXPLORE DATA

Query complete (9.9 sec elapsed, 1.7 GB processed)

Job information | Results | JSON | Execution details

Row	Trip	Number_of_Trip	start_station_name	end_station_name
1	191 191	58269	Hyde Park Corner, Hyde Park	Hyde Park Corner, Hyde Park
2	307 307	31961	Black Lion Gate, Kensington Gardens	Black Lion Gate, Kensington Gardens
3	303 303	31218	Albert Gate, Hyde Park	Albert Gate, Hyde Park
4	785 785	23572	Aquatic Centre, Queen Elizabeth Olympic Park	Aquatic Centre, Queen Elizabeth Olympic Park
5	248 248	21273	Triangle Car Park, Hyde Park	Triangle Car Park, Hyde Park
6	407 407	15419	Speakers' Corner 1, Hyde Park	Speakers' Corner 1, Hyde Park
7	404 404	14925	Palace Gate, Kensington Gardens	Palace Gate, Kensington Gardens
8	406 406	14232	Speakers' Corner 2, Hyde Park	Speakers' Corner 2, Hyde Park
9	111 111	12293	Park Lane , Hyde Park	Park Lane , Hyde Park
10	191 307	12000	Black Lion Gate, Kensington Gardens	Hyde Park Corner, Hyde Park
11	213 213	11912	Wellington Arch, Hyde Park	Wellington Arch, Hyde Park