```
-- Part 1: Usage Volume Overview

SELECT YEAR(start_date), COUNT(*) AS total_num_trips_2016
FROM bixit.trips
WHERE YEAR(start_date) = 2016
GROUP BY YEAR(start_date);


SELECT YEAR(start_date), COUNT(*) AS total_num_trips_2017
FROM bixit.trips
WHERE YEAR(start_date) = 2017
GROUP BY YEAR(start_date);


SELECT MONTH(start_date), COUNT(*) AS total_num_trips_by_month_2016
FROM bixit.trips
WHERE YEAR(start_date) = 2016
GROUP BY MONTH(start_date);


SELECT MONTH(start_date), COUNT(*) AS total_num_trips_by_month_2017
FROM bixit.trips
WHERE YEAR(start_date) = 2017
GROUP BY MONTH(start_date);


SELECT DAY(start_date),
AVG(total_num_trips_by_month_2017) OVER(PARTITION BY MONTH(start_date)
ORDER BY MONTH(start_date) DESC) AS total_num_trips_day
FROM bixit.trips
GROUP BY DAY(start_date);


SELECT YEAR(start_date) AS Year,
       MONTH(start_date) AS Month,
       ROUND(COUNT(*)/COUNT(DISTINCT DAY(start_date))) AS
Average_daily_trips
FROM trips
GROUP BY Year, Month
ORDER BY Average_daily_trips DESC;


SELECT is_member AS membership_status, YEAR(start_date) AS '2017',
COUNT(*) AS total_num_trips_2017
FROM bixit.trips
WHERE YEAR(start_date) = 2017
GROUP BY is_member;

SELECT is_member AS membership_status, YEAR(start_date) AS '2016',
COUNT(*) AS total_num_trips_2016
FROM bixit.trips
WHERE YEAR(start_date) = 2016
GROUP BY is_member;
```

```
SELECT MONTH(start_date) AS month, AVG(is_member) AS fraction_trips_2017
FROM bixit.trips
WHERE YEAR(start_date) = 2017
GROUP BY MONTH(start_date);

-- It seems that the demand for bicycles on July and August is at the peak
SELECT MONTH(start_date) AS month, COUNT(*) AS total_num_trips_2016
FROM trips
WHERE YEAR(start_date) = 2016
GROUP BY MONTH(start_date)
ORDER BY total_num_trips_2016 DESC
LIMIT 2;

SELECT MONTH(start_date) AS month, COUNT(*) AS total_num_trips_2017
FROM trips
WHERE YEAR(start_date) = 2017
GROUP BY month
ORDER BY total_num_trips_2017 DESC
LIMIT 2;


SELECT MONTH(start_date) AS month, COUNT(*) AS total_trips_non_membership
FROM trips
WHERE is_member = 0
GROUP BY month
ORDER BY total_trips_non_membership DESC;

SELECT MONTH(start_date) AS month, COUNT(*) AS total_trips_membership
FROM trips
WHERE is_member = 1
GROUP BY month
ORDER BY total_trips_membership DESC;

-- Part 2: Trip Characteristics

SELECT AVG(duration_sec) AS average_trip_time
FROM trips;


SELECT is_member, ROUND(AVG(duration_sec)) AS average_trip_time
FROM trips
GROUP BY is_member
ORDER BY average_trip_time DESC;

SELECT MONTH(start_date) as month, AVG(duration_sec) AS average_trip_time
FROM trips
GROUP BY month
ORDER BY average_trip_time DESC;

SELECT (CASE DAYOFWEEK(start_date)
      WHEN 1 THEN 'Monday'
    WHEN 2 THEN 'Tuesday'
    WHEN 3 THEN 'Wednesday'
```

```
    WHEN 4 THEN 'Thursday'
    WHEN 5 THEN 'Friday'
    WHEN 6 THEN 'Saturday'
    WHEN 7 THEN 'Sunday'
      END) AS day_of_the_week, AVG(duration_sec) AS average_trip_time
FROM trips
GROUP BY day_of_the_week
ORDER BY average_trip_time DESC;

-- station 'Métro Jean-Drapeau' has the longest trip on average. Average
trip time 1899 seconds
SELECT stations.name AS station_name, AVG(duration_sec) AS
average_trip_time
FROM trips
JOIN stations
ON stations.code = trips.start_station_code
GROUP BY station_name
ORDER BY average_trip_time DESC;

-- station Métro Georges-Vanier (St-Antoine / Canning) has the shortest
trip on average. Average trip time 499 seconds

SELECT stations.name AS station_name, AVG(duration_sec) AS
average_trip_time
FROM trips
JOIN stations
ON stations.code = trips.start_station_code
GROUP BY station_name
ORDER BY average_trip_time ASC;

-- use LIMIT to only show the first few rows


SELECT  is_member AS membership_status, AVG(start_station_code =
end_station_code) AS average_round_trip
FROM trips
GROUP BY membership_status;

-- Q 6.2
SELECT  (CASE DAYOFWEEK(start_date)
      WHEN 1 THEN 'Monday'
    WHEN 2 THEN 'Tuesday'
    WHEN 3 THEN 'Wednesday'
    WHEN 4 THEN 'Thursday'
    WHEN 5 THEN 'Friday'
    WHEN 6 THEN 'Saturday'
    WHEN 7 THEN 'Sunday'
      END) AS day_of_the_week, AVG(start_station_code = end_station_code)
AS average_round_trip
FROM trips
GROUP BY day_of_the_week
ORDER BY average_round_trip DESC;
```

```
-- Part 3: Popular Stations

SELECT stations.name AS station_name, COUNT(*) AS total_num_trips
FROM trips
JOIN stations
ON stations.code = start_station_code
GROUP BY station_name
ORDER BY total_num_trips DESC
LIMIT 5;

SELECT stations.name AS station_name, COUNT(*) AS total_num_trips
FROM trips
JOIN stations
ON stations.code = end_station_code
GROUP BY station_name
ORDER BY total_num_trips DESC
LIMIT 5;

SELECT stations.name AS station_name,
       (CASE
           WHEN HOUR(start_date) BETWEEN 7 AND 11 THEN "morning"
           WHEN HOUR(start_date) BETWEEN 12 AND 16 THEN "afternoon"
           WHEN HOUR(start_date) BETWEEN 17 AND 21 THEN "evening"
           ELSE "night"
        END) AS time_of_day,
        COUNT(*) AS num_trips_as_start_station
FROM stations
JOIN trips
ON stations.code = trips.start_station_code
WHERE stations.name = 'Mackay / de Maisonneuve'
GROUP BY station_name, time_of_day
ORDER BY num_trips_as_start_station DESC;



SELECT stations.name AS station_name,
       COUNT(*) AS num_of_trips,
       AVG(is_member) AS average_of_trips_done_by_members
FROM   (
        SELECT *
        FROM trips
        WHERE start_station_code = end_station_code
        ) AS round_trip
JOIN stations
ON round_trip.start_station_code = stations.code
GROUP BY station_name
HAVING num_of_trips >= 10
ORDER BY average_of_trips_done_by_members
LIMIT 1;

SELECT stations.name AS station_name,
       COUNT(*) AS num_of_trips,
       AVG(is_member) AS average_of_trips_done_by_members
```

```
FROM   (
        SELECT *
        FROM trips
        WHERE start_station_code = end_station_code
        ) AS round_trip
JOIN stations
ON round_trip.start_station_code = stations.code
GROUP BY station_name
HAVING num_of_trips >= 10
ORDER BY average_of_trips_done_by_members DESC
LIMIT 1;

SELECT stations.name AS station_name, COUNT(*) AS total_trips,
start_station_code
FROM trips
JOIN stations
ON stations.code = trips.start_station_code
GROUP BY station_name
ORDER BY total_trips DESC;

SELECT stations.name AS station_name, COUNT(trips.start_station_code =
trips.end_station_code) AS total_round_trips
FROM trips
JOIN stations
ON stations.code = trips.start_station_code
GROUP BY station_name
ORDER BY total_round_trips DESC;


SELECT stations.name AS station_name,
         COUNT(trips.start_station_code) AS number_of_trips,
         AVG(trips.start_station_code = trips.end_station_code) AS
average_of_round_trips
FROM trips
JOIN stations
ON trips.start_station_code = stations.code
GROUP BY station_name
HAVING average_of_round_trips >= 0.1 AND number_of_trips >= 50
ORDER BY average_of_round_trips DESC;
```

```
-- Part 1: Usage Volume Overview

SELECT YEAR(start_date), COUNT(*) AS total_num_trips_2016
FROM bixit.trips
WHERE YEAR(start_date) = 2016
GROUP BY YEAR(start_date);


SELECT YEAR(start_date), COUNT(*) AS total_num_trips_2017
FROM bixit.tripsa
WHERE YEAR(start_date) = 2017
GROUP BY YEAR(start_date);


SELECT MONTH(start_date), COUNT(*) AS total_num_trips_by_month_2016
FROM bixit.trips
WHERE YEAR(start_date) = 2016
GROUP BY MONTH(start_date);


SELECT MONTH(start_date), COUNT(*) AS total_num_trips_by_month_2017
FROM bixit.trips
WHERE YEAR(start_date) = 2017
GROUP BY MONTH(start_date);


SELECT DAY(start_date),
AVG(total_num_trips_by_month_2017) OVER(PARTITION BY MONTH(start_date) ORDER BY
MONTH(start_date) DESC) AS total_num_trips_day
```

```
-- Part 1: Usage Volume Overview

SELECT YEAR(start_date), COUNT(*) AS total_num_trips_2016
FROM bixit.trips
WHERE YEAR(start_date) = 2016
GROUP BY YEAR(start_date);


SELECT YEAR(start_date), COUNT(*) AS total_num_trips_2017
FROM bixit.trips
WHERE YEAR(start_date) = 2017
GROUP BY YEAR(start_date);


SELECT MONTH(start_date), COUNT(*) AS total_num_trips_by_month_2016
FROM bixit.trips
WHERE YEAR(start_date) = 2016
GROUP BY MONTH(start_date);


SELECT MONTH(start_date), COUNT(*) AS total_num_trips_by_month_2017
FROM bixit.trips
WHERE YEAR(start_date) = 2017
GROUP BY MONTH(start_date);


SELECT DAY(start_date),
AVG(total_num_trips_by_month_2017) OVER(PARTITION BY MONTH(start_date) ORDER BY
MONTH(start_date) DESC) AS total_num_trips_day
```

```
FROM bixit.trips
GROUP BY DAY(start_date);



SELECT YEAR(start_date) AS Year,
    MONTH(start_date) AS Month,
    ROUND(COUNT(*)/COUNT(DISTINCT DAY(start_date))) AS Average_daily_trips
FROM trips
GROUP BY Year, Month
ORDER BY Average_daily_trips DESC;



SELECT is_member AS membership_status, YEAR(start_date) AS '2017', COUNT(*) AS
total_num_trips_2017
FROM bixit.trips
WHERE YEAR(start_date) = 2017
GROUP BY is_member;

SELECT is_member AS membership_status, YEAR(start_date) AS '2016', COUNT(*) AS
total_num_trips_2016
FROM bixit.trips
WHERE YEAR(start_date) = 2016
GROUP BY is_member;



SELECT MONTH(start_date) AS month, AVG(is_member) AS fraction_trips_2017
FROM bixit.trips
WHERE YEAR(start_date) = 2017
GROUP BY MONTH(start_date);
```

```
-- It seems that the demand for bicycles on July and August is at the peak
SELECT MONTH(start_date) AS month, COUNT(*) AS total_num_trips_2016
FROM trips
WHERE YEAR(start_date) = 2016
GROUP BY MONTH(start_date)
ORDER BY total_num_trips_2016 DESC
LIMIT 2;

SELECT MONTH(start_date) AS month, COUNT(*) AS total_num_trips_2017
FROM trips
WHERE YEAR(start_date) = 2017
GROUP BY month
ORDER BY total_num_trips_2017 DESC
LIMIT 2;


SELECT MONTH(start_date) AS month, COUNT(*) AS total_trips_non_membership
FROM trips
WHERE is_member = 0
GROUP BY month
ORDER BY total_trips_non_membership DESC;

SELECT MONTH(start_date) AS month, COUNT(*) AS total_trips_membership
FROM trips
WHERE is_member = 1
GROUP BY month
ORDER BY total_trips_membership DESC;
```

```
-- Part 2: Trip Characteristics

SELECT AVG(duration_sec) AS average_trip_time
FROM trips;



SELECT is_member, ROUND(AVG(duration_sec)) AS average_trip_time
FROM trips
GROUP BY is_member
ORDER BY average_trip_time DESC;

SELECT MONTH(start_date) as month, AVG(duration_sec) AS average_trip_time
FROM trips
GROUP BY month
ORDER BY average_trip_time DESC;

SELECT (CASE DAYOFWEEK(start_date)
        WHEN 1 THEN 'Monday'
   WHEN 2 THEN 'Tuesday'
   WHEN 3 THEN 'Wednesday'
   WHEN 4 THEN 'Thursday'
   WHEN 5 THEN 'Friday'
   WHEN 6 THEN 'Saturday'
   WHEN 7 THEN 'Sunday'
        END) AS day_of_the_week, AVG(duration_sec) AS average_trip_time
FROM trips
GROUP BY day_of_the_week
ORDER BY average_trip_time DESC;
```

```
-- station 'Métro Jean-Drapeau' has the longest trip on average. Average trip time 1899 seconds

SELECT stations.name AS station_name, AVG(duration_sec) AS average_trip_time

FROM trips

JOIN stations

ON stations.code = trips.start_station_code

GROUP BY station_name

ORDER BY average_trip_time DESC;


-- station Métro Georges-Vanier (St-Antoine / Canning) has the shortest trip on average. Average trip time 499 seconds


SELECT stations.name AS station_name, AVG(duration_sec) AS average_trip_time

FROM trips

JOIN stations

ON stations.code = trips.start_station_code

GROUP BY station_name

ORDER BY average_trip_time ASC;


-- use LIMIT to only show the first few rows



SELECT  is_member AS membership_status, AVG(start_station_code = end_station_code) AS average_round_trip

FROM trips

GROUP BY membership_status;


-- Q 6.2

SELECT  (CASE DAYOFWEEK(start_date)

        WHEN 1 THEN 'Monday'
```

```
   WHEN 2 THEN 'Tuesday'
   WHEN 3 THEN 'Wednesday'
   WHEN 4 THEN 'Thursday'
   WHEN 5 THEN 'Friday'
   WHEN 6 THEN 'Saturday'
   WHEN 7 THEN 'Sunday'
        END) AS day_of_the_week, AVG(start_station_code = end_station_code) AS
average_round_trip
FROM trips
GROUP BY day_of_the_week
ORDER BY average_round_trip DESC;



-- Part 3: Popular Stations

SELECT stations.name AS station_name, COUNT(*) AS total_num_trips
FROM trips
JOIN stations
ON stations.code = start_station_code
GROUP BY station_name
ORDER BY total_num_trips DESC
LIMIT 5;

SELECT stations.name AS station_name, COUNT(*) AS total_num_trips
FROM trips
JOIN stations
ON stations.code = end_station_code
GROUP BY station_name
ORDER BY total_num_trips DESC
```

```
LIMIT 5;


SELECT stations.name AS station_name,
      (CASE
         WHEN HOUR(start_date) BETWEEN 7 AND 11 THEN "morning"
         WHEN HOUR(start_date) BETWEEN 12 AND 16 THEN "afternoon"
         WHEN HOUR(start_date) BETWEEN 17 AND 21 THEN "evening"
         ELSE "night"
       END) AS time_of_day,
       COUNT(*) AS num_trips_as_start_station
FROM stations
JOIN trips
ON stations.code = trips.start_station_code
WHERE stations.name = 'Mackay / de Maisonneuve'
GROUP BY station_name, time_of_day
ORDER BY num_trips_as_start_station DESC;




SELECT stations.name AS station_name,
      COUNT(*) AS num_of_trips,
      AVG(is_member) AS average_of_trips_done_by_members
FROM   (
       SELECT *
       FROM trips
       WHERE start_station_code = end_station_code
       ) AS round_trip
JOIN stations
```

```
ON round_trip.start_station_code = stations.code
GROUP BY station_name
HAVING num_of_trips >= 10
ORDER BY average_of_trips_done_by_members
LIMIT 1;

SELECT stations.name AS station_name,
      COUNT(*) AS num_of_trips,
      AVG(is_member) AS average_of_trips_done_by_members
FROM   (
       SELECT *
       FROM trips
       WHERE start_station_code = end_station_code
       ) AS round_trip
JOIN stations
ON round_trip.start_station_code = stations.code
GROUP BY station_name
HAVING num_of_trips >= 10
ORDER BY average_of_trips_done_by_members DESC
LIMIT 1;

SELECT stations.name AS station_name, COUNT(*) AS total_trips, start_station_code
FROM trips
JOIN stations
ON stations.code = trips.start_station_code
GROUP BY station_name
ORDER BY total_trips DESC;
```

```
SELECT stations.name AS station_name, COUNT(trips.start_station_code = trips.end_station_code) AS total_round_trips
FROM trips
JOIN stations
ON stations.code = trips.start_station_code
GROUP BY station_name
ORDER BY total_round_trips DESC;
```

```
SELECT stations.name AS station_name,
        COUNT(trips.start_station_code) AS number_of_trips,
        AVG(trips.start_station_code = trips.end_station_code) AS average_of_round_trips
FROM trips
JOIN stations
ON trips.start_station_code = stations.code
GROUP BY station_name
HAVING average_of_round_trips >= 0.1 AND number_of_trips >= 50
ORDER BY average_of_round_trips DESC;
```