Lecture 7

Activity 1

```
CREATE TABLE gps_data (
    datapoint_id varchar(50),
    journey_id varchar(50),
    latitude float,
    longitude float,
    month int,
    day int,
    hour int,
    Primary key (datapoint_id)
);
```

```
CREATE TABLE vehicle_data (
    datapoint_id varchar(50),
    geohash varchar(50),
    speed float,
    make char(50),
    model char(50),
    route_id char(50),
    segment_start_measure float,
    Primary key (datapoint_id)
);
```

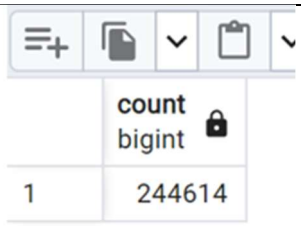
```sql
-- Step 1: Drop the table if it exists
DROP TABLE IF EXISTS gps_data;

-- Step 2: Recreate the table
CREATE TABLE gps_data (
    datapoint_id varchar(50),
    journey_id varchar(50),
    latitude float,
    longitude float,
    month int,
    day int,
    hour int,
    PRIMARY KEY (datapoint_id)
);

-- Step 3: Import data from the CSV file
-- Duplicate key violations will cause the offending row to be skipped
COPY gps_data(datapoint_id, journey_id, latitude, longitude, month, day, hour)
FROM 'C:\Users\Public\gps_data.csv'
DELIMITER ','
CSV HEADER;
```

Activity #2

```sql
SELECT *
FROM gps_data
WHERE month = 10 AND
      day = 3 AND
      hour = 4;
```

```sql
SELECT COUNT(*)
FROM gps_data;
```

Answer:



	count bigint
1	244614

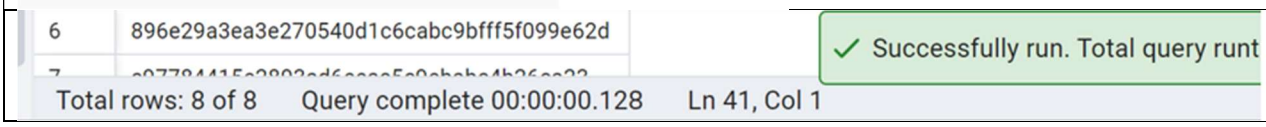
```
SELECT COUNT(DISTINCT journeyid)
FROM gps_data
WHERE month = 10 AND
      day = 3 AND
      hour = 4;
```

Answer :

	count bigint
1	7

```
SELECT *
FROM gps_data
WHERE month = 10 AND
      day BETWEEN 1 AND 3 AND
      hour BETWEEN 3 AND 6;
```


```
SELECT DISTINCT journeyid
FROM gps_data
WHERE journeyid LIKE '%abc%';
```



Activity #3

```
SELECT MAX(speed) AS max_speed
FROM vehicle_data;
```

	max double precision
1	162.43

```sql
SELECT
    AVG(speed) AS avg_speed,
    MIN(speed) AS min_speed,
    MAX(speed) AS max_speed,
    STDDEV(speed) AS stddev_speed
FROM vehicle_data
WHERE make = 'CHEVROLET';
```

Data Output Messages Notifications

SQL

	avg_speed double precision	min_speed double precision	max_speed double precision	stddev_speed double precision
1	79.31768078013866	0	157.82	47.02569582299691

```sql
SELECT *,
    CONCAT(route_id, '-', segment_start_measure) AS seg_id
FROM vehicle_data;
```

Total rows: 1000 of 244614 Query complete 00:00:00.377 Ln 56, Col 19

```sql
SELECT
    seg_id,
    PERCENTILE_CONT(0.85) WITHIN GROUP (ORDER BY speed) AS speed_85th_percentile
FROM vehicle_data
GROUP BY seg_id;
```

	seg_id character varying (255)	speed_85th_percentile double precision
1	C000141000N-2.997	13.469
2	C000141000N-3.011	35.477
3	C000141000N-3.061	42.62
4	C000141000N-3.086	48.035000000000004
5	C000141000N-3.161	48.38
6	C000141000N-3.198	44.92
7	C000141000N-3.241	42.735

Total rows: 1000 of 8745 Query complete 00:00:00.

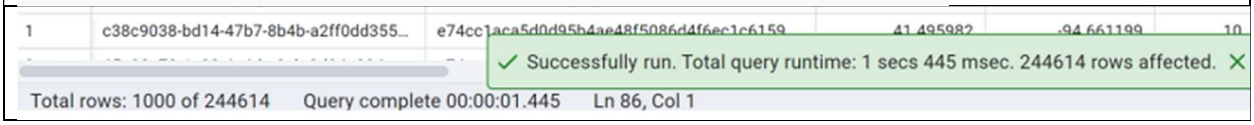
```sql
SELECT
    model,
    PERCENTILE_CONT(0.95) WITHIN GROUP (ORDER BY speed) AS speed_95th_percentile
FROM vehicle_data
GROUP BY model;
```

	model character (50)	speed_95th_percentile double precision
1	ACADIA	126.71
2	ATS	118.65
3	AVEO_SONIC	130.17
4	BLAZER	126.71
5	CAMARO	127.87
6	CANYON	80.64
7	COLORADO	129.02

Total rows: 34 of 34 Query complete 00:00:00.228 Ln 78

Activity #4

```sql
SELECT
    gps_data.*,
    vehicle_data.*
FROM
    gps_data
JOIN
    vehicle_data
ON
    gps_data.datapoint_id = vehicle_data.datapoint_id;
```

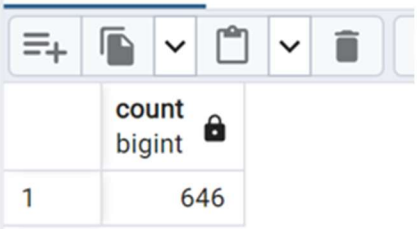


Total rows: 1000 of 244614 Query complete 00:00:01.445 Ln 86, Col 1

```sql
SELECT
    gps_data.datapoint_id,
    gps_data.journey_id,
    gps_data.speed,
    vehicle_data.seg_id
FROM
    gps_data
JOIN
    vehicle_data
ON
    gps_data.datapoint_id = vehicle_data.datapoint_id;
```

Total rows: 1000 of 244614 Query complete 00:00:01.074 Ln 92, Col 4

```sql
SELECT
    COUNT(DISTINCT gps_data.journey_id) AS unique_journeys_stopped
FROM
    gps_data
JOIN
    vehicle_data
ON
    gps_data.datapoint_id = vehicle_data.datapoint_id
WHERE
    vehicle_data.speed = 0;
```



	count bigint
1	646