**SQL USE CASE 3**

**LOADING FILES AND WORKING WITH DATA**

**Question 1**

**Upload the CSV data via CLI into the database**

CREATE TABLE USECASE3 (

TIME datetime NOT NULL,

DEVICE char(2) NOT NULL,

CONSUMPTION decimal(4,3) NOT NULL

);

LOAD DATA

LOCAL INFILE '/home/supraja/Downloads/usecase3.csv'

INTO TABLE USECASE3

FIELDS TERMINATED BY ','

IGNORE 1 LINES ;

**Question 2**

**Find all the duplicate time entries for respective devices.**

SELECT DISTINCT T1.DEVICE,

T1.TIME,

T1.CONSUMPTION

FROM USECASE3 T1 INNER JOIN

( SELECT TIME,DEVICE FROM USECASE3

GROUP BY TIME,DEVICE

HAVING COUNT(\*) >1 ) T2

ON T1.TIME = T2.TIME AND T1.DEVICE = T2.DEVICE;

**Question 3**

**Find all the missing timestamps for respective devices.**

WITH MT AS (

SELECT DEVICE,

TIME ,

ROW\_NUMBER() over (

partition by DEVICE,HOUR(TIME) order by TIME)-MINUTE(TIME) as rowdif from USECASE3

) SELECT DEVICE,

TIMESTAMPADD(MINUTE,rowdif-1,MT.TIME)

FROM MT WHERE rowdif!=1;

**Question 4**

**Find the hour-wise cumulative consumption of each device.**

SELECT DEVICE,

HOUR(TIME),

SUM(CONSUMPTION) AS HOURLY\_CONSUMPTION FROM USECASE3

GROUP BY DEVICE,HOUR(TIME)

ORDER BY DEVICE,HOUR(TIME);

**Question 5**

**Find the peak consumption reached in a given time range for all the devices.**

SELECT DEVICE,

MAX(CONSUMPTION) AS PEAK

FROM USECASE3

WHERE TIME >= '2020-01-01 10:00:00' AND

TIME <= '2020-01-01 10:05:00'

GROUP BY DEVICE ORDER BY DEVICE ;