```
create schema Uber_Supply_Demand_Gap;

create table user_request_data (
request_id int primary key,
pickup_point varchar(20),
driver_id int null,
status varchar(30),
time_of_request time,
date_of_request date,
drop_time time null,
drop_date date null
);

select count(drop_date) from user_request_data;
```

-- since the data was not uploaded completely - added the "2019-01-01" in blank spaces and "-1" for drive_id and after this all the data has been uploaded

-- **To Convert the 2019-01-01 and -1 into null**

```
update user_request_data
        set drop_date = NULL
        where drop_date = '2019-07-16';
```

-- ( *Error Code: 1175. You are using safe update mode and you tried to update a table without a WHERE that uses a KEY column. To disable safe mode, toggle the option in Preferences -> SQL Editor and reconnect.* )

```
SET SQL_SAFE_UPDATES = 0; -- To turn off the safe mode

update user_request_data
set drop_date = NULL
where drop_date = '2019-07-16'; -- Updated

update user_request_data
set drop_time = NULL
where drop_time = '0';

update user_request_data
set driver_id = NULL
where driver_id = '-1'; -- 3914 row(s) affected Rows matched: 3914  Changed: 3914  Warnings: 0

select count(drop_time) from user_request_data
where drop_time is not null;
```

----------**Analysis** ----------

**-- 1 - Identify Time slots face the highest unmet demand / demand-supply gap**

```
with demand_summary AS (
Select
        pickup_point,
                Case
    When hour(time_of_request) between 0 and 4 then 'Late Night'
                When hour(time_of_request) between 5 and 8 then 'Early Morning'
```

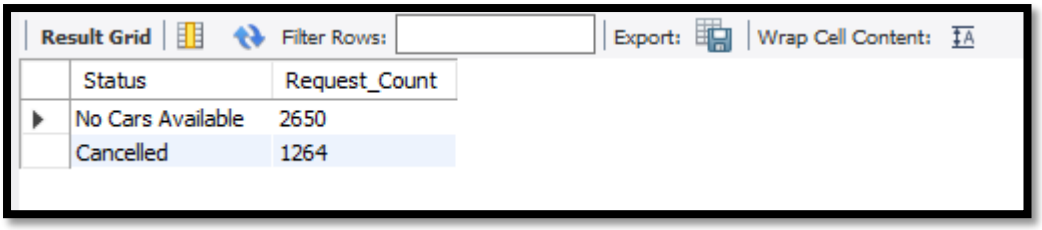
```
                When hour(time_of_request) between 9 and 12 then  'Morning'
                When hour(time_of_request) between 13 and 16 then 'Afternoon'
                When hour(time_of_request) between 17 and 20 then 'Evening'
                Else 'Night'
        End as Time_Slot,
   count(*) as Total_Request,
        Sum(case when status = 'Trip completed' then 1 else 0 End) as 'Completed',
        Sum(case when status = 'cancelled' then 1 else 0 End) as 'Cancelled',
        Sum(case when status = 'No Cars Available' then 1 else 0 End) as 'No_Cars'
From user_request_data
Group By pickup_point, Time_Slot
)
Select *,
        round((1 - Completed * 1.0/ Total_Request)* 100,2 ) AS Gap_Percentage
From demand_summary
order By pickup_point, Time_Slot;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

| pickup_point | Time_Slot | Total_Request | Completed | Cancelled | No_Cars | Gap_Percentage |
|---|---|---|---|---|---|---|
| Airport | Evening | 1457 | 312 | 78 | 1067 | 78.59 |
| City | Early Morning | 1335 | 373 | 653 | 309 | 72.06 |
| Airport | Night | 624 | 203 | 31 | 390 | 67.47 |
| City | Late Night | 325 | 111 | 63 | 151 | 65.85 |
| City | Morning | 714 | 286 | 239 | 189 | 59.94 |
| Airport | Late Night | 253 | 103 | 2 | 148 | 59.29 |
| City | Afternoon | 374 | 208 | 32 | 134 | 44.39 |
| City | Night | 323 | 196 | 33 | 94 | 39.32 |
| Airport | Afternoon | 252 | 162 | 40 | 50 | 35.71 |
| City | Evening | 436 | 330 | 46 | 60 | 24.31 |
| Airport | Morning | 315 | 239 | 32 | 44 | 24.13 |
| Airport | Early Morning | 337 | 308 | 15 | 14 | 8.61 |

---

```
-- 2 reasons for unmet demand
Select
        Status,
        Count(request_id) as Request_Count
From user_request_data
Where status in ('No Cars Available' , 'Cancelled')
Group By status
Order By Request_Count Desc;
```

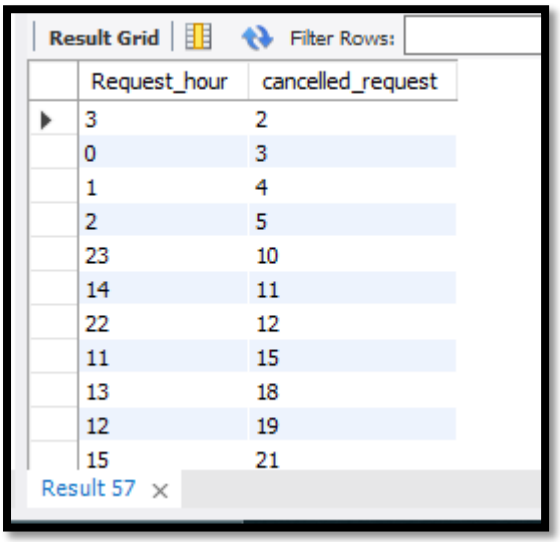


| Status | Request_Count |
|---|---|
| No Cars Available | 2650 |
| Cancelled | 1264 |

---

**-- 3 Count of most cancellations by Hour**

```
Select
        Hour(time_of_request) as Request_hour,
        Count(*) as cancelled_request
From user_request_data
Where status = 'Cancelled'
group by Request_hour
order by cancelled_request;
```



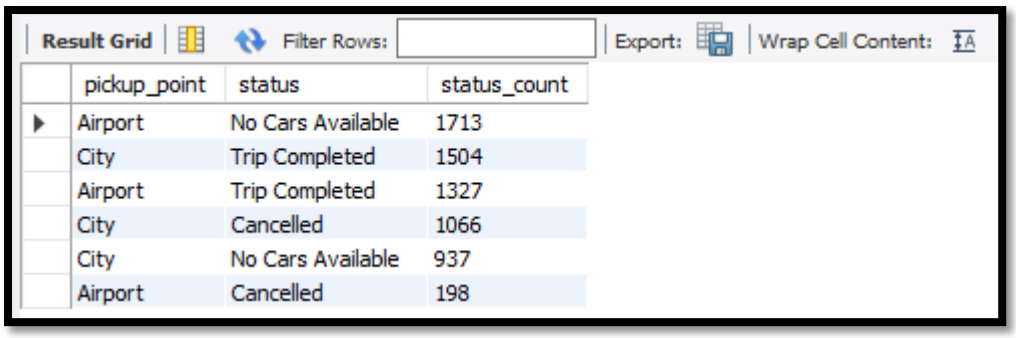
| Request_hour | cancelled_request |
|---|---|
| 3 | 2 |
| 0 | 3 |
| 1 | 4 |
| 2 | 5 |
| 23 | 10 |
| 14 | 11 |
| 22 | 12 |
| 11 | 15 |
| 13 | 18 |
| 12 | 19 |
| 15 | 21 |

Result 57

---

**-- 4 pickup point status distribution**

```
Select
        pickup_point,
   status,
   Count(*) as status_count
From user_request_data
group by pickup_point, status
Order By status_count desc;
```

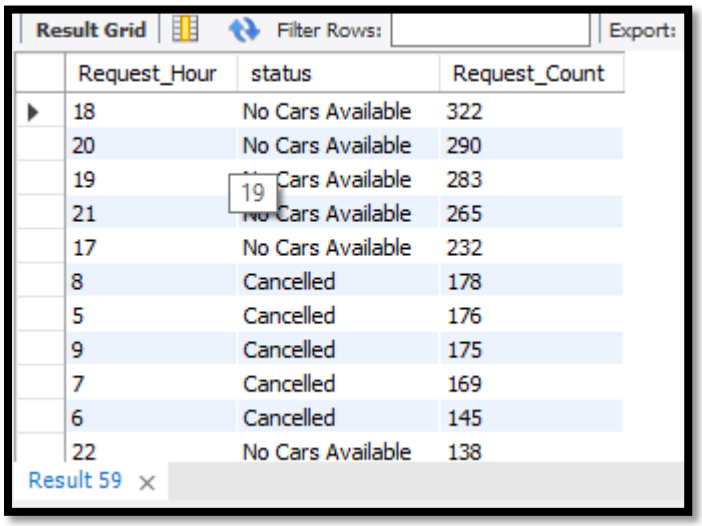


| pickup_point | status | status_count |
|---|---|---|
| Airport | No Cars Available | 1713 |
| City | Trip Completed | 1504 |
| Airport | Trip Completed | 1327 |
| City | Cancelled | 1066 |
| City | No Cars Available | 937 |
| Airport | Cancelled | 198 |

---

**-- 5 cancellations and unavailability peak**

```
Select
        hour(time_of_request) As Request_Hour,
   status,
   count(request_id) as Request_Count
From user_request_data
where status in ('No Cars Available', 'Cancelled')
Group By status, Request_Hour
Order By Request_Count Desc;
```



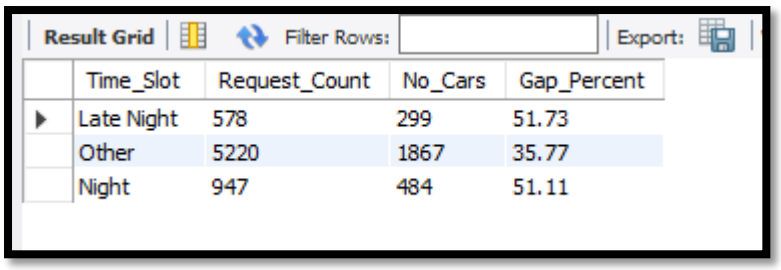
| Request_Hour | status | Request_Count |
|---|---|---|
| 18 | No Cars Available | 322 |
| 20 | No Cars Available | 290 |
| 19 | No Cars Available | 283 |
| 21 | No Cars Available | 265 |
| 17 | No Cars Available | 232 |
| 8 | Cancelled | 178 |
| 5 | Cancelled | 176 |
| 9 | Cancelled | 175 |
| 7 | Cancelled | 169 |
| 6 | Cancelled | 145 |
| 22 | No Cars Available | 138 |

Result 59

---

**-- 6 - % of 'No Cars Available' during night**

```
Select
        case
        when Hour(time_of_request) between 21 and 23 then 'Night'
   when Hour(time_of_request) between 0 and 4 then 'Late Night'
        Else 'Other'
End As Time_Slot,
Count(*) As Request_Count,
Sum(status = 'No Cars Available') as No_Cars,
Round(Sum(status = 'No Cars Available') * 100.0 / Count(*), 2 ) as Gap_Percent
From user_request_data
GROUP BY Time_Slot;
```



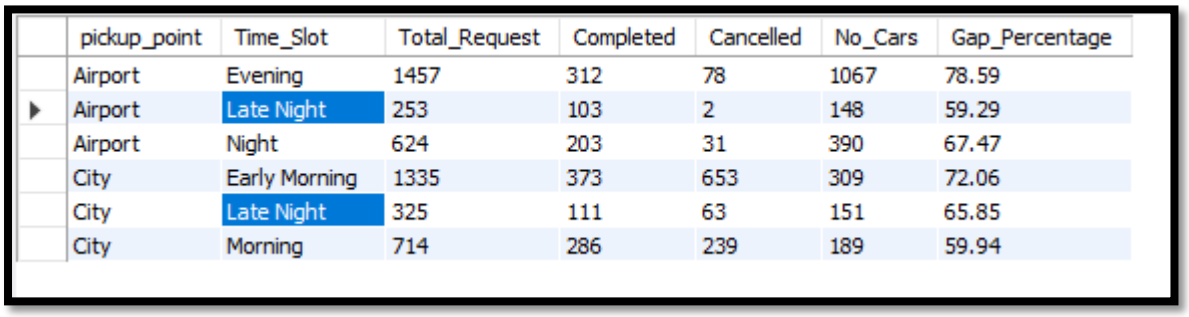
Result Grid | Filter Rows: | Export:

| Time_Slot | Request_Count | No_Cars | Gap_Percent |
|---|---|---|---|
| Late Night | 578 | 299 | 51.73 |
| Other | 5220 | 1867 | 35.77 |
| Night | 947 | 484 | 51.11 |

---

**-- 7 - demand-supply gap where Gap is more than 50%**

```
with demand_summary AS (
Select
        pickup_point,
                Case
    When hour(time_of_request) between 0 and 4 then 'Late Night'
                When hour(time_of_request) between 5 and 8 then 'Early Morning'
                When hour(time_of_request) between 9 and 12 then  'Morning'
                When hour(time_of_request) between 13 and 16 then 'Afternoon'
                When hour(time_of_request) between 17 and 20 then 'Evening'
                Else 'Night'
        End as Time_Slot,
  count(*) as Total_Request,
        Sum(case when status = 'Trip completed' then 1 else 0 End) as 'Completed',
        Sum(case when status = 'cancelled' then 1 else 0 End) as 'Cancelled',
        Sum(case when status = 'No Cars Available' then 1 else 0 End) as 'No_Cars'
From user_request_data
Group By pickup_point, Time_Slot
)
Select *,
        round((1 - Completed * 1.0/ Total_Request)* 100,2 ) AS Gap_Percentage
From demand_summary
Having Gap_Percentage > 50
order By pickup_point, Time_Slot;
```



| pickup_point | Time_Slot | Total_Request | Completed | Cancelled | No_Cars | Gap_Percentage |
|---|---|---|---|---|---|---|
| Airport | Evening | 1457 | 312 | 78 | 1067 | 78.59 |
| Airport | Late Night | 253 | 103 | 2 | 148 | 59.29 |
| Airport | Night | 624 | 203 | 31 | 390 | 67.47 |
| City | Early Morning | 1335 | 373 | 653 | 309 | 72.06 |
| City | Late Night | 325 | 111 | 63 | 151 | 65.85 |
| City | Morning | 714 | 286 | 239 | 189 | 59.94 |