Submission

A document file containing

A brief description of your understanding of data

INTRODUCTION:

Uber is a prominent Taxi Aggregator that caters to baesd upon commuters needs. Commuters can uses Uber app to request a taxi for their commute needs. With ever increasing smart phones, Uber has become a go to option for most of the travellers.

A BRIEF DESCRIPTION OF THE DATA USED:

Here we are having two data sets Dim_city and and fact_trip where Dim_city is a dimension which lists all the cities that Uber provides services to. Fact_trip provides details of all the trip transactions. In the dim_city we are having 3 columns City_id, city_name, country. And in the fact_trip we are having trip_uuid, datastr, product_type _name, city_id, driver_uuid, is_completed, ETA, ATA, UFF_fare, fare_final this columns will provide all the data. By using this 2 data sets we can solve the customer requirements, and Uber provides services across lot of cities and there are various products catered to the traveller's needs. Uber seeks our help to understand which of the products are profitable and how many times were they able to meet the ETA so they can fine tune the service offerings.

B. Any anomalies you identified in the provided dataset and a brief description of how you identified them and why do you think they are anomalies

There is no anomalies

→ *Create the table structure with appropriate data types before loading with SQL Loader*

```
create table fact_trip (
trip_uuid varchar2(100),
datestr date,
product_type_name varchar2(100),
city_id number,
driver_uuid varchar2(100),
is_completed varchar2(100),
eta number,
ata number,
ufp_fare number,
fare_final number
);
```

```
create table dim_city (
city_id number,
city_name varchar2(20),
country varchar2(20)
);
```

Queries you have written including the DDLs

A. How many city_ids does uberPOOL operate in?

```
select count(d.city_id) from dim_city d,fact_trip f
where d.city_id=f.city_id and product_type_name='uberPOOL';
```

B. Which city_id has the highest error in ETA (where error in ETA = {(eta - ata)/ata}) for the given time period?

```
select city_id,(eta-ata)/ata as  a from fact_trip where rownum=1 order by a ;
```

C. Which is the product type with highest total revenue in SanFrancisco?

```
select * from(select product_type_name,fare_final from fact_trip order by fare_final desc) where
rownum=1;
```

D. Which are the products in each city where total revenue(fare_final) > $1000?

```
select * from (select product_type_name,sum(fare_final) as total_rev
from fact_trip
group by product_type_name) where total_rev>1000;

<<<<<<<<<<by total_rev>1000 there is no records>>>>>>>>>>
select * from (select product_type_name,sum(fare_final) as total_rev
from fact_trip
group by product_type_name) where total_rev>100;
```

E .Get to 2nd highest country by Uber Revenue (fare_final) for 2nd week of June 2018 across product

```
select * from(select d.country, f.fare_final, rownum as rank from dim_city d join fact_trip f on
d.city_id=f.city_id where
to_char(datestr, 'W')=2 order by 2) where mod(rank, 2)=0;
```

F. Get WOW growth % for US region for June Month. WOW- Week over week .

```
select
(((select sum(fare_final) from fact_trip where to_char(datestr, 'W')=1)
- (select sum(fare_final) from fact_trip where to_char(datestr, 'W')=2))
/ (select sum(fare_final) from fact_trip where to_char(datestr, 'W')=1)) * 100 as "Growth%"
from fact_trip where to_char(datestr,'W')=2 group by to_char(datestr,'W');
```

G. Growth % = ((Current week fare final - previous week fare final) / previous week fare final) * 100

```
select
(((select sum(fare_final) from fact_trip where to_char(datestr, 'W')='1')
- (select sum(fare_final) from fact_trip where to_char(datestr, 'W')='2'))
/ (select sum(fare_final) from fact_trip where to_char(datestr, 'W')='1') *100 ) as "Growth%"
from dual;
```

Results to the queries above

A

	COUNT(D.CITY_ID)
1	16

B.

	CITY_ID	A
1	10	0.428679245283018867924528301886792452830

C.

	PRODUCT_TYPE_...	FARE_FINAL
1	uberX	49.31

D.

BY USING total_rev>1000 THERE IS NO RECORDS IF WE USE 100 WE CAN GET SOME RECORDS

	PRODUCT_TYPE_NAME	TOTAL_REV
1	Helium	447.42
2	uberX	875.92
3	uberPOOL	549.49

E.

SQL | All Rows Fetched: 1 in 0.005 seconds

	COUNTRY
1	US

F.

	Growth%
1	96.321914800175669740887132191480017566 97

G.

	Growth%
1	96.321914800175669740887132191480017566 97