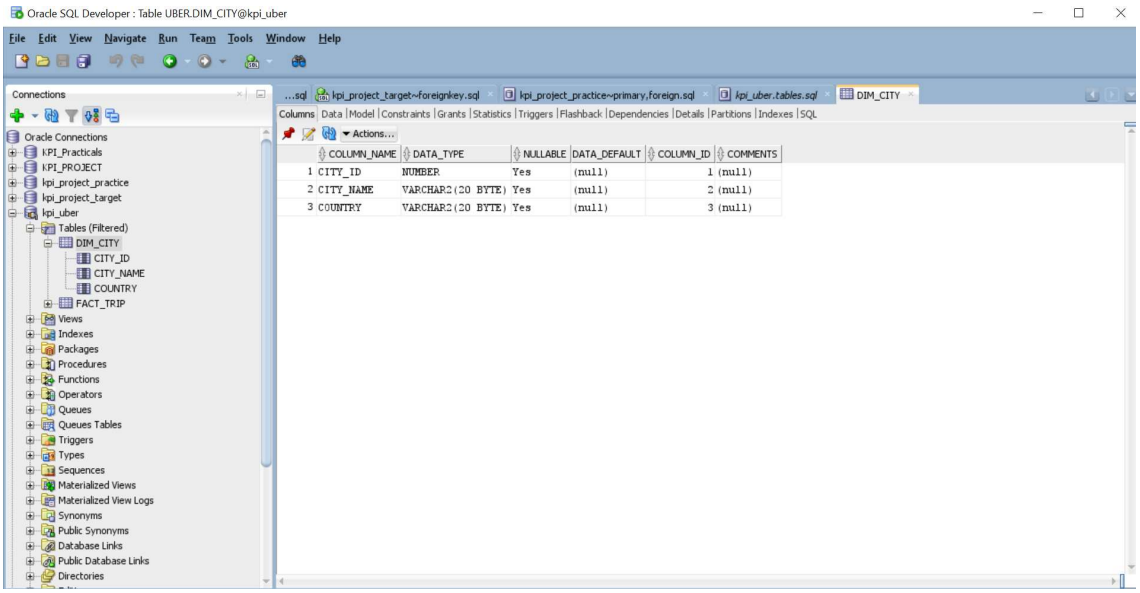
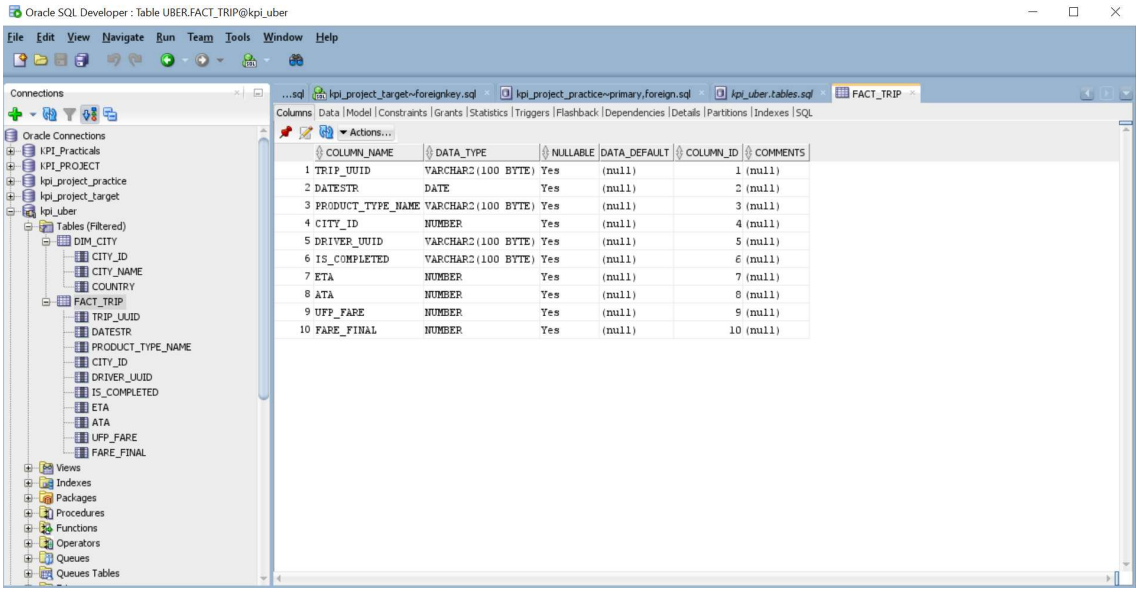
INTRODUCTION:

Uber is a prominent Taxi Aggregator that caters to commuters needs. Commuters can use 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:

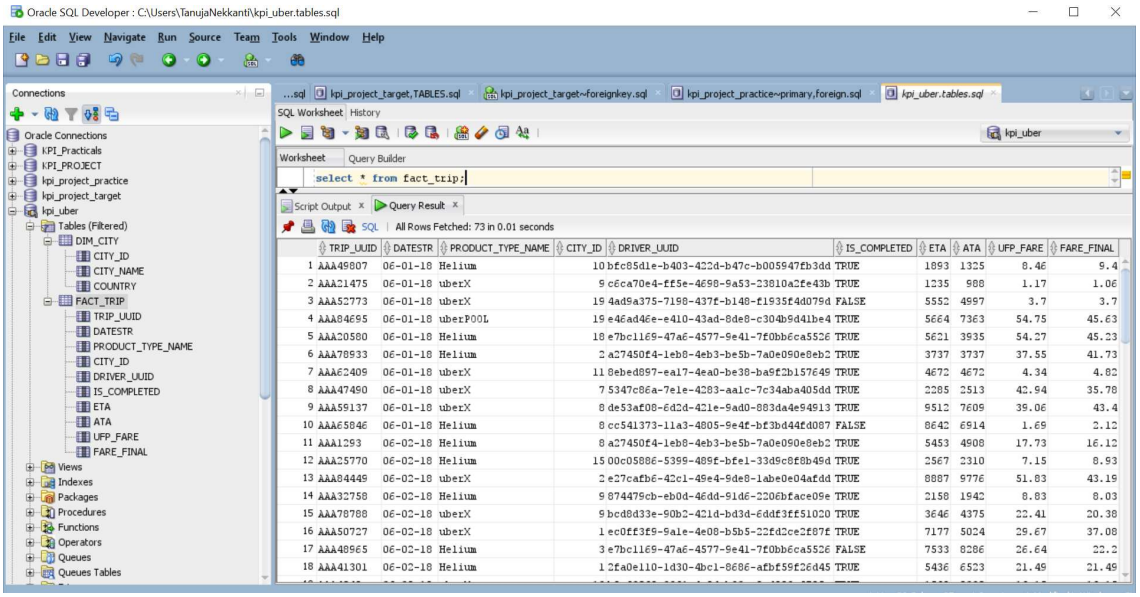
Here we are having two data sets 1. Dim_city and 2. fact_trip where Dim_city is a dimension which lists all the cities that Uber provides services to. Fact_trip provides details of all the trip transactions. In the dim_city we are having 3 columns City_id, city_name, country. And in the fact_trip we 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.

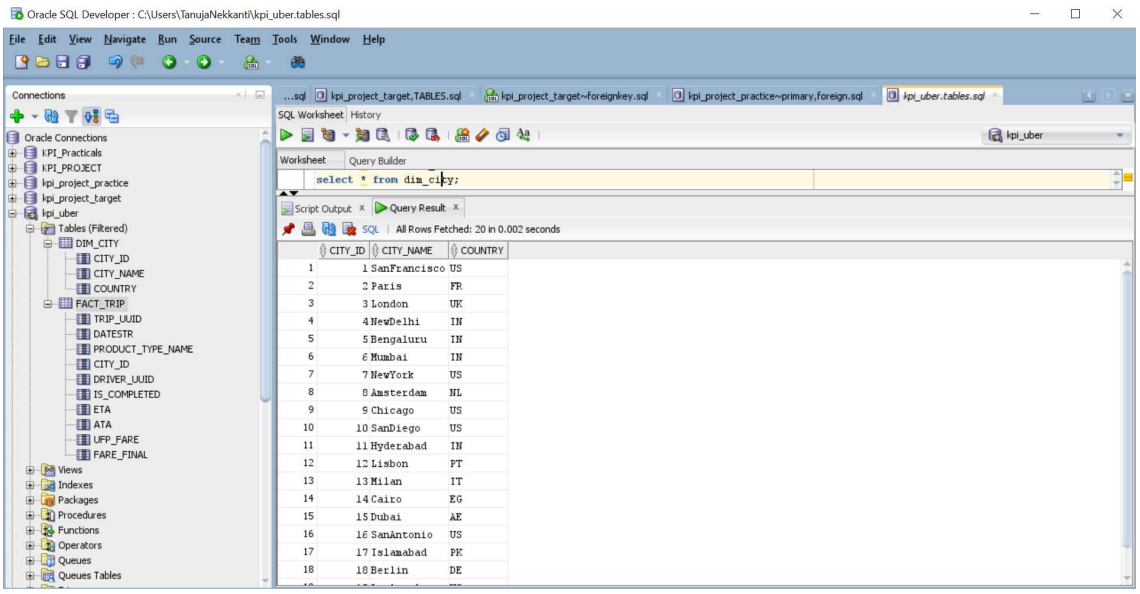




ANOMALIES:

In the given data set I didn't get any anomalies.

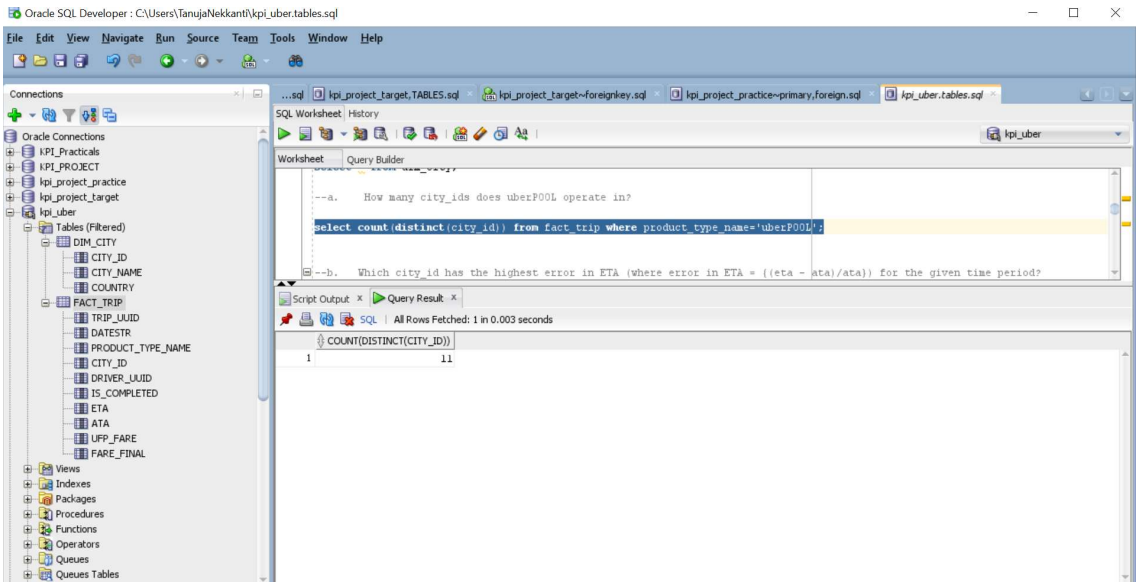




Query's and Outputs:

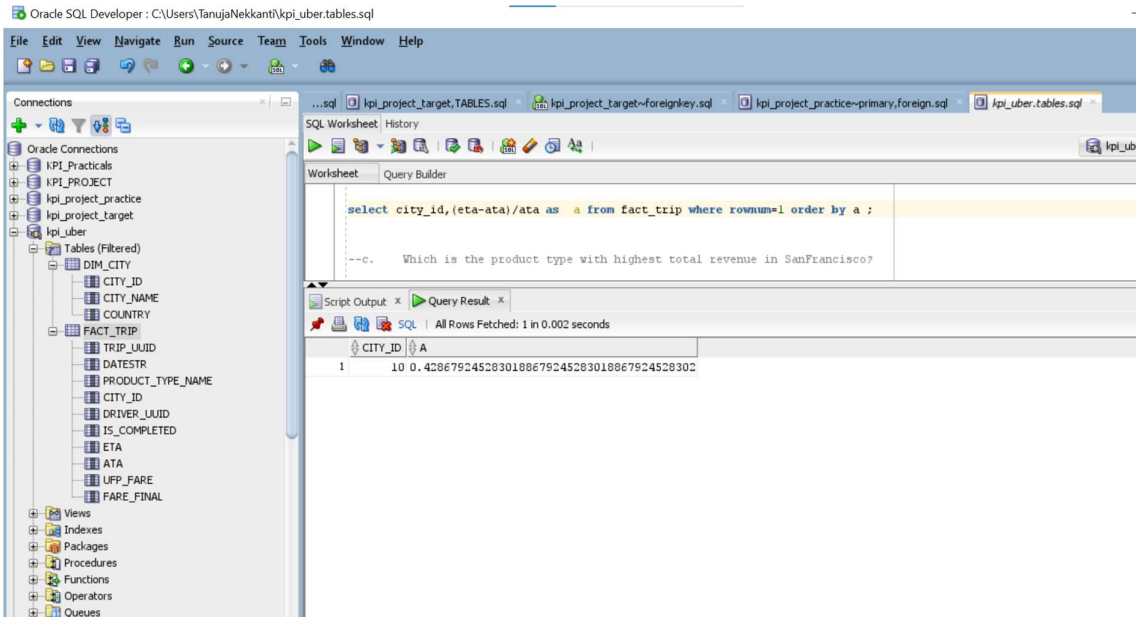
a. A brief description of your understanding of data.

```
select count(distinct(city_id)) from fact_trip where product_type_name = 'uberPOOL';
```



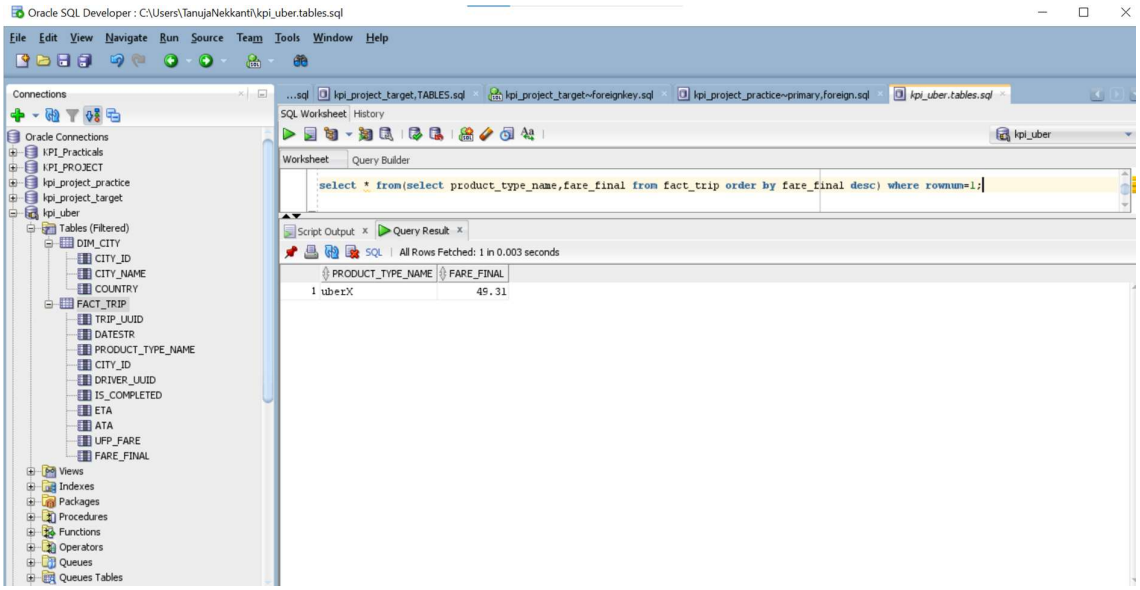
b. Which city_id has the highest error in ETA (where error in ETA = {(eta - ata)/ata}) for the given time period?

```
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;

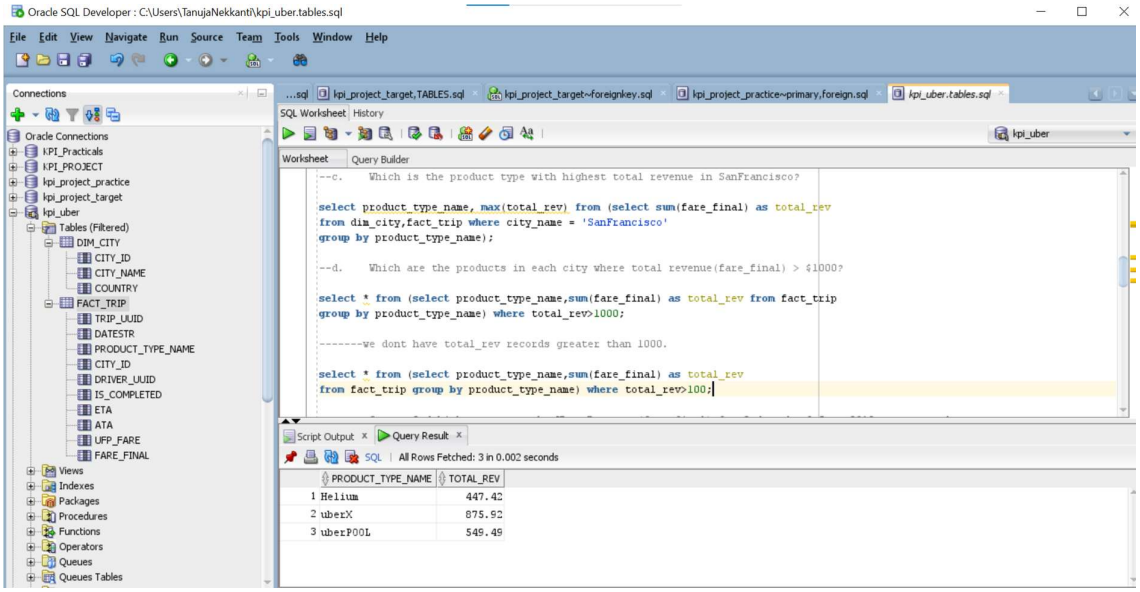


d. Which are the products in each city where total revenue(fare_final) > $1000?

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

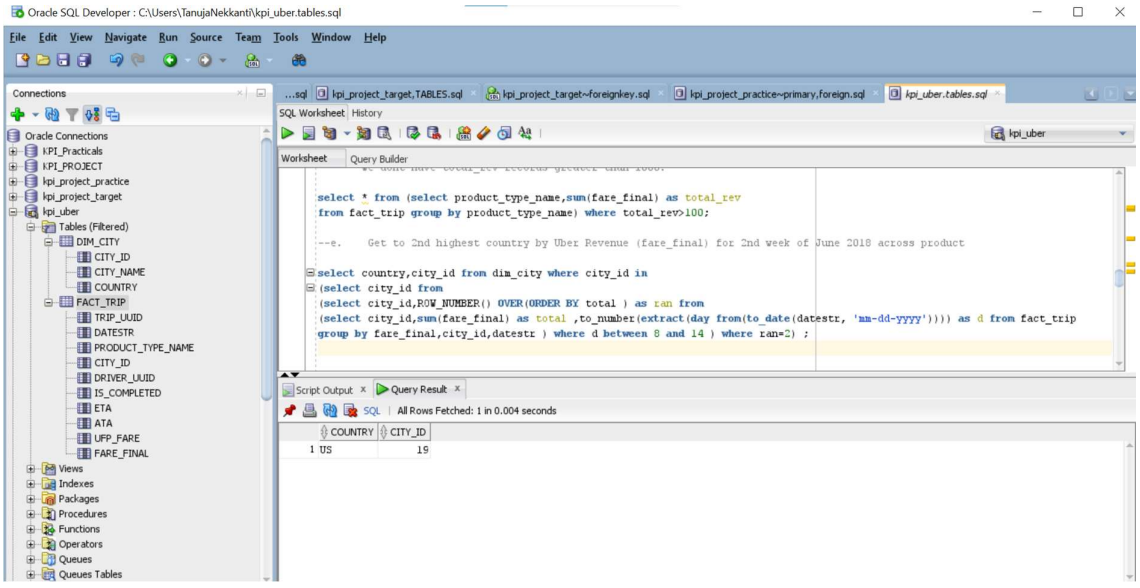
----from the given data sets we don't have the total_rev records greater than 1000 we have greater than 100. So we modified and calculating for greater than 100.

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

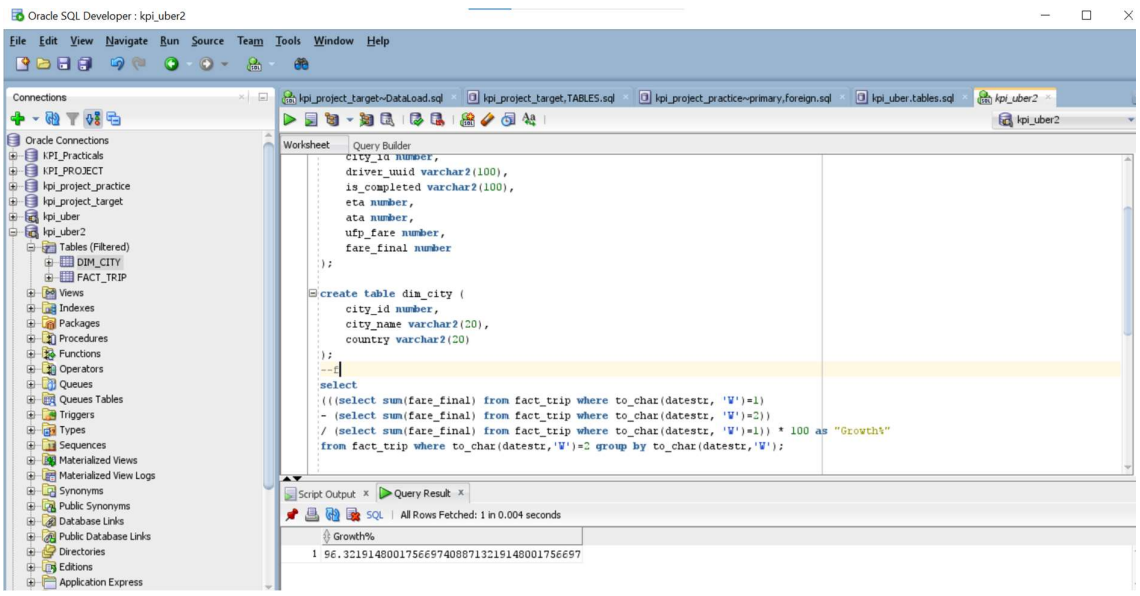
select country,city_id from dim_city where city_id in (select city_id from
(select city_id,ROW_NUMBER() OVER(ORDER BY total) as ran from
(select city_id,sum(fare_final) as total ,to_number(extract(day from(to_date(datestr, 'mm-dd-yyyy')))) as d from fact_trip group by fare_final,city_id,datestr) where d between 8 and 14) where ran=2) ;



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;