1.Simple Assignment(products\_dim)

CREATE TABLE retail\_demo.products\_dim

(

product\_id TEXT,

category\_id TEXT,

price TEXT,

product\_name TEXT

)

1. Get the product between 2 price range(500-1000)
2. Find out all the categories that the stores sell.
3. Find out all the categories and the total product on that categories.

categoryId,categoryName,totalProducts,products details

36,sports,20,{(),()}

1. Find out the product with maximum & minimum price on each categories
2. Find out all the product between 2 prices on a particular category.
3. Search a product based on a keyword
4. Average price of each category products

2.Based on ORder table

CREATE TABLE retail\_demo.orders\_hawq

(

order\_id TEXT,

customer\_id TEXT,

store\_id TEXT,

order\_datetime TEXT,

ship\_completion\_datetime TEXT,

return\_datetime TEXT,

refund\_datetime TEXT,

payment\_method\_code TEXT,

total\_tax\_amount TEXT,

total\_paid\_amount TEXT,

total\_item\_quantity TEXT,

total\_discount\_amount TEXT,

coupon\_code TEXT,

coupon\_amount TEXT,

order\_canceled\_flag TEXT,

has\_returned\_items\_flag TEXT,

has\_refunded\_items\_flag TEXT,

fraud\_code TEXT,

fraud\_resolution\_code TEXT,

billing\_address\_line1 TEXT,

billing\_address\_line2 TEXT,

billing\_address\_line3 TEXT,

billing\_address\_city TEXT,

billing\_address\_state TEXT,

billing\_address\_postal\_code TEXT,

billing\_address\_country TEXT,

billing\_phone\_number TEXT,

customer\_name TEXT,

customer\_email\_address TEXT,

ordering\_session\_id TEXT,

website\_url TEXT

)

1. find out total orders by each customer.
2. find out all order details by any customer in the given time range.

input customer id,time range

output: all order description

1. find all the total\_paid amount city\_wise in a particular month
2. find out the maximum sales happend in any city in last 1 year

Hyd which month has highest sales

Banglore which month has highest sales

1. Find out total business done and total tax to be paid till date
2. Find out top 20 product in good demand & top 20 product in less demand.
3. Find out all customer with frequent order cancellation

output

c001 8(cancelled) 10(total) order by percentage

c005 2(cancelled) 10(total)

1. Find out all the products returned after using order by the total number of returns.

1.Store all the data in pig relation

A = load ' '

1.Find out top 3 customers with highest no of Orders.

2.Find out top 5 customers with highest amount of order.

3.How many Female staying at "Some Pin Code location " have purchased "Jewelry" Items.

4.How many Male customers have purchased "Jewelry" Item.

5.How many people from zipcode= "80744" has purchased "Video Games".