**Check Point 1**

**Task 1.2 (SQL-Oracle)**

**Stage 1:**

* **Construct and ER-Diagram for the above-mentioned Requirement**

Diagram

Description automatically generated

* **Construct Tables has per the ER-Diagram.**

The tables according to our ER Diagrams are:

Customer Table:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **USER\_ID** | **GENDER** | **AGE** | **OCCUPATION** | **CITY\_CATEGORY** | **STAY-IN\_CURRENT\_CITY\_YEARS** | **MARITAL\_STATUS** |

Product Table:

|  |  |  |  |
| --- | --- | --- | --- |
| **PRODUCT\_ID** | **PRODUCT\_CATEGORY\_1** | **PRODUCT\_CATEGORY\_2** | **PRODUCT\_CATEGORY\_3** |

Purchase Table:

|  |  |  |
| --- | --- | --- |
| **USER\_ID** | **PRODUCT\_ID** | **PURCHASE** |

* **Identify the relationships between tables and use appropriate standards for the same where applicable**

Primary Key of Customer table is USER\_ID.

Primary Key of Product table is PRODUCT\_ID.

Purchase table have USER\_ID and PRODUCT\_ID as foreign keys and thus help in joining the Customer and Product table.

* **Insert the appropriate data into the identified tables from the sample dataset provided.**

Queries to insert the data into the Customer table from the given two csv files (Having table name customerproduct and userdemographic) are:

1. Insert into purchase

Select distinct user\_id, product\_id, purchase

From customerproduct;

1. Insert into product

Select distinct product\_id, product\_category\_1, product\_category\_2, product\_category\_3

From customerproduct ;

1. Insert into customer

Select distinct cp.user\_id, gender, age, occupation, city\_category, stay\_in\_current\_city\_years, marital\_status

From customerproduct cp, userdemographic ud

Where cp.user\_id = ud.user\_id;

**Stage 2. Write SQL queries for below**

1. **TOP 50 best selling products**

**QUERY:**

select \* from (Select product\_id, count(product\_id)

From purchase

Group by product\_id

Order by count(product\_id) desc)

where rownum<=50;

**SNAPSHOTS:**

**Table

Description automatically generatedTable

Description automatically generatedGraphical user interface, text, application, email

Description automatically generated**

1. **TOP best selling product categories**

As each of the product id is associated at least with product category 1, and the product category 2 and product category 3 have many null values, thuswe considered the product category 1 column to identify the subcategory associations.

**QUERY:**

Select product\_category\_1, count(p.product\_id)

From product p, purchase pr

Where p.product\_id = pr.product\_id

Group By product\_category\_1

Order By count(p.product\_id) desc;

**SNAPSHOTS:**

**Table

Description automatically generatedGraphical user interface, text, application

Description automatically generated**

**CONCLUSION:**

The top bestselling product categories are number 5, 1, 8, 11, 2.

1. **count\_sales by gender**

**QUERY:**

Select gender, count(\*) as Count\_Sales

From customer c, purchase pc, product p

Where c.user\_id = pc.user\_id and

pc.product\_id = p.product\_id

Group by Gender;

**SNAPSHOTS:**

**Graphical user interface

Description automatically generatedGraphical user interface, text, application

Description automatically generated**

1. **favourite products by gender(Female)**

**QUERY:**

Select p.product\_id , count(p.product\_id)

From customer c, purchase pc, product p

Where c.user\_id = pc.user\_id

and pc.product\_id = p.product\_id

and c.gender = 'F'

Group by p.product\_id

Order by count(p.product\_id) desc;

**SNAPSHOTS:**

**Table

Description automatically generatedGraphical user interface, application, Word

Description automatically generated**

**CONCLUSION:**

The top favourite products of females are P00265242, P00220442, P00058042, etc.

1. **favourite products by gender(male)**

**QUERY:**

Select p.product\_id , count(p.product\_id)

From customer c, purchase pc, product p

Where c.user\_id = pc.user\_id

and pc.product\_id = p.product\_id

and c.gender = 'M'

Group by p.product\_id

Order by count(p.product\_id) desc;

**SNAPSHOTS:**

**Table

Description automatically generatedGraphical user interface, text, application

Description automatically generated**

**CONCLUSION:**

The top favourite products of males are P00265242, P00025442, P00110742, etc.

1. **sales\_volume by occupation**

**QUERY:**

Select occupation, count(p.product\_id) as Sales\_Volume

From product p, purchase pr, customer c

Where p.product\_id = pr.product\_id

and c.user\_id = pr.user\_id

Group By occupation

Order By occupation;

**SNAPSHOTS:**

Table

Description automatically generatedGraphical user interface, text, application

Description automatically generated

1. **sales\_volume by age**

**QUERY:**

Select age, count(p.product\_id) as Sales\_Volume

From product p, purchase pr, customer c

Where p.product\_id = pr.product\_id

and c.user\_id = pr.user\_id

Group By age

Order By age;

**SNAPSHOTS:**

**Table

Description automatically generatedGraphical user interface, text, application

Description automatically generated**

1. **sales\_volume by city category**

**QUERY:**

Select city\_category, count(p.product\_id) as Sales\_Volume

From product p, purchase pr, customer c

Where p.product\_id = pr.product\_id

and c.user\_id = pr.user\_id

Group By city\_category

Order By city\_category;

**SNAPSHOTS:**

Table

Description automatically generated with medium confidence Graphical user interface, text, application

Description automatically generated

1. **sales\_volume by stay in city years**

**QUERY:**

Select stay\_in\_current\_city\_years, count(p.product\_id) as Sales\_Volume

From product p, purchase pr, customer c

Where p.product\_id = pr.product\_id

and c.user\_id = pr.user\_id

Group By stay\_in\_current\_city\_years

Order By stay\_in\_current\_city\_years;

**SNAPSHOTS:**

Table

Description automatically generatedGraphical user interface, text, application

Description automatically generated

1. **sales\_volume by Marital\_Status**

**QUERY:**

Select Marital\_Status, count(p.product\_id) as Sales\_Volume

From product p, purchase pr, customer c

Where p.product\_id = pr.product\_id

and c.user\_id = pr.user\_id

Group By Marital\_Status

Order By Marital\_Status;

**SNAPSHOTS:**

**Table

Description automatically generated with low confidenceGraphical user interface, text, application

Description automatically generated**

1. **purchase amount by age and gender**

**QUERY:**

Select age, gender, sum(purchase)

From product p, purchase pr, customer c

Where p.product\_id = pr.product\_id

and c.user\_id = pr.user\_id

Group By age, gender

Order By age, gender;

**SNAPSHOTS:**

Table

Description automatically generatedGraphical user interface, text, application, email

Description automatically generated

1. **purchase amount by age, gender and marital\_status**

**QUERY:**

Select age, gender, marital\_status, sum(purchase)

From product p, purchase pr, customer c

Where p.product\_id = pr.product\_id

and c.user\_id = pr.user\_id

Group By age, gender, marital\_status

Order By age, gender, marital\_status;

**SNAPSHOTS:**

**Table

Description automatically generatedGraphical user interface, text, application

Description automatically generated**