E-Commerce Products Sales Analysis Using PostgreSQL

The dataset used in this project contains 8000 observations of e-commerce products sales including product details and customer demographics. Sourced from Kaggle, it provides insights into consumer behavior and sales trend in products. I used SQL queries to answer 12 questions, uncovering valuable insights. It is an ideal resource for understanding E-commerce sales patterns.

Author: Dawood Shah

Google certified Data analyst

Verify my Data Analytics certificate at:
https://coursera.org/verify/professional-cert/HPS5H6DXPZPZ

IMPORTING DATA TO SQL

```
Product_ID VARCHAR(50) PRIMARY KEY,
Product_Name VARCHAR(100) NOT NULL,
Category VARCHAR(100) NOT NULL,
Sub_Category VARCHAR(100) NOT NULL,
Price NUMERIC(10,2),
Customer_Age INT,
Customer_Gender VARCHAR(50),
Purchase_History INT,
Review_Rating INT,
Review_Sentiment VARCHAR(50),
Total_price NUMERIC(10,2)
);
SELECT * FROM products;
```

--Q1) Retrieve average price of each product.

```
SELECT product_name, avg(price) AS avg_price
FROM products
GROUP BY (product_name);
```

--Q2) Retrieve the top 20 product with highest review rating.

```
SELECT product_name, review_rating
FROM products
ORDER BY review_rating DESC LIMIT 20 ;
```

--Q3) What is the total sales for each product?

```
SELECT product_name, SUM(total_price) AS total_sales
FROM products
GROUP BY product_name;
```

--Q4) Which customers (by age group) have purchase the most products?

```
SELECT
CASE
WHEN customer_age BETWEEN 15 AND 30 THEN '15-30'
WHEN customer_age BETWEEN 31 AND 40 THEN '31-40'
WHEN customer_age BETWEEN 41 AND 50 THEN '41-50'
WHEN customer_age BETWEEN 51 AND 60 THEN '51-60'
ELSE 'other'
END AS age_group,
SUM(purchase_history) AS purchased_products FROM products
GROUP BY age_group
ORDER BY purchased_products DESC LIMIT 1;
```

--Q5) Retrieve the total number of customers in each age group.

```
SELECT
CASE
WHEN customer_age BETWEEN 15 AND 30 THEN '15-30'
WHEN customer_age BETWEEN 31 AND 40 THEN '31-40'
WHEN customer_age BETWEEN 41 AND 50 THEN '41-50'
WHEN customer_age BETWEEN 51 AND 60 THEN '51-60'
ELSE 'other'
END AS age_group,
COUNT(*) AS number_of_customers FROM products
GROUP BY age_group;
```

--Q6) Retrieve the total number of products purchased.

```
select sum(purchase_history) from products;
```

--Q7) Retrieve the total number of customers.

```
SELECT COUNT(product_id) FROM products;
```

--Q8) What is the distribution of products by subcategory?

```
SELECT sub_category,COUNT(*) AS number_of_products

FROM products

GROUP BY sub_category;
```

--Q9) Which product subcategory has the highest average price?

```
SELECT  product_name,sub_category, AVG(price) AS avg_price

FROM products

GROUP BY product_name,sub_category

ORDER BY avg_price DESC LIMIT 1;
```

--Q10) What is the average review rating for products purchased by male vs. female customers?

```
SELECT customer_gender, AVG(review_rating) AS avg_review_ratimg

FROM products

GROUP BY customer_gender;
```

--Q11) Which products have a price above the average price of all products?

```
SELECT product_name,price FROM products

WHERE price>(SELECT AVG(price)FROM products);
```

--Q12) Which product category has the highest total sales?

```
SELECT category,SUM(total_price) AS total_sales

FROM products

GROUP BY category

ORDER BY total_sales LIMIT 1;
```

THE END