Data Analyst Assignment

Dataset Description:

For this assignment, you will be working with a dataset containing information about a retail company's sales and inventory transactions. The dataset includes the following tables:

Customers:

o customer_id: Unique identifier for each customer.

o name: Customer's name.

o age: Customer's age.

o gender: Customer's gender.

customer_id	cust_name	cust_age	cust_gender
1	John	30	M
2	Emily	25	F
3	Michael	40	M
4	Sophia	35	F
5	Jacob	28	M
6	Olivia	32	F
7	William	45	M
8	Ava	27	F
9	James	33	M
10	Mia	29	F
11	Benjamin	38	M
12	Harper	31	F
13	Samuel	42	M
14	Amelia	26	F
15	Joseph	34	M
16	Charlotte	30	F
17	David	39	M
18	Emma	24	F
19	Daniel	36	M
20	Grace	33	F

Products:

o product_id: Unique identifier for each product.

o name: Product name.

o category: Product category.

o price: Price of the product.

product_id	prod_name	prod_category	prod_price
1	T-Shirt	Clothing	20
2	Jeans	Clothing	50
3	Shoes	Footwear	80
4	Watch	Accessories	100
5	Laptop	Electronics	800
6	Backpack	Accessories	30
7	Dress	Clothing	60
8	Sneakers	Footwear	70
9	Sunglasses	Accessories	40
10	Smartphone	Electronics	1000
11	Shorts	Clothing	25
12	Sandals	Footwear	45
13	Headphones	Electronics	150
14	Hat	Accessories	20
15	Skirt	Clothing	35
16	Slippers	Footwear	25
17	Camera	Electronics	600
18	Belt	Accessories	15
19	Jacket	Clothing	70
20	Heels	Footwear	60

Sales:

o transaction_id: Unique identifier for each transaction.

o customer_id: Unique identifier for each customer.

o product_id: Unique identifier for each product.

o date: The date of the transaction.

o quantity: The quantity of products purchased in that transaction.

o amount: The total amount spent in that transaction.

transaction_id	customer_id	product_id	date	quantity	amount
1	1	3	01-06-2023	2	160
2	2	7	02-06-2023	1	60
3	3	9	03-06-2023	3	120
4	4	4	04-06-2023	1	100
5	5	1	05-06-2023	2	40
6	6	13	06-06-2023	1	150
7	7	10	07-06-2023	1	1000
8	8	8	08-06-2023	2	140
9	9	2	09-06-2023	1	50
10	10	5	10-06-2023	1	800
11	11	6	11-06-2023	2	60
12	12	15	12-06-2023	1	35
13	13	14	13-06-2023	1	20
14	14	19	14-06-2023	1	70
15	15	12	15-06-2023	2	90
16	16	11	16-06-2023	1	25
17	17	17	17-06-2023	1	600
18	18	18	18-06-2023	1	15
19	19	20	19-06-2023	2	120
20	20	16	20-06-2023	1	25

Inventory:

o product_id: Unique identifier for each product.

o stock_count: The current stock count for each product.

product_id	stock_count
1	8
2	15
3	10
4	5
5	3
6	12
7	6
8	10
9	4
10	2
11	9
12	7
13	6
14	14
15	10
16	11
17	4
18	20
19	8
20	5

➢ Calculate the total revenue generated by the company for each product category.

```
SELECT p.prod_category, SUM(s.amount) AS total_revenue
FROM product p
JOIN salesss s ON p.product_id = s.product_id
GROUP BY p.prod_category;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

prod_category	total_revenue
Footwear	535
Clothing	280
Accessories	315
Electronics	2550

- Determine the top 5 customers who have made the highest total purchases, considering the customer's age and gender.

```
SELECT c.customer_id, c.cust_name, c.cust_age, c.cust_gender, SUM(s.amount) AS total_purchases
FROM customer c
JOIN salesss s ON c.customer_id = s.customer_id
GROUP BY c.customer_id, c.cust_name, c.cust_age, c.cust_gender
ORDER BY total_purchases DESC
LIMIT 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

customer_id	cust_name	cust_age	cust_gender	total_purchases
7	William	45	M	1000
10	Mia	29	F	800
17	David	39	M	600
1	John	30	M	160
6	Olivia	32	F	150

- Identify the most profitable product category by calculating the average revenue per unit sold.

```
SELECT p.prod_category, AVG(s.amount / s.quantity) AS average_revenue_per_unit
FROM product p
JOIN salesss s ON p.product_id = s.product_id
GROUP BY p.prod_category
ORDER BY average_revenue_per_unit DESC
LIMIT 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

prod_category	average_revenue_per_unit
Electronics	637.50000000

- Analyze the inventory data and identify products that need restocking (stock count less than a specified threshold).

```
SELECT p.product_id, p.prod_name, p.prod_category, i.stock_count
FROM product p
JOIN inventoryy i ON p.product_id = i.product_id
WHERE i.stock_count < 10;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

product_id	prod_name	prod_category	stock_count
1	T-Shirt	Clothing	8
4	Watch	Accessories	5
5	Laptop	Electronics	3
7	Dress	Clothing	6
9	Sunglasses	Accessories	4
10	Smartphone	Electronics	2
11	Shorts	Clothing	9
12	Sandals	Footwear	7
13	Headphones	Electronics	6
17	Camera	Electronics	4
19	Jacket	Clothing	8
20	Heels	Footwear	5

- Write a SQL query to calculate the average age of customers for each product category

```sql
SELECT p.prod_category, AVG(c.cust_age) AS average_cust_age
FROM product p
JOIN salesss s ON p.product_id = s.product_id
JOIN customer c ON s.customer_id = c.customer_id
GROUP BY p.prod_category;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

prod_category	average_cust_age
Footwear	32.0000
Clothing	28.8333
Accessories	35.8000
Electronics	36.2500

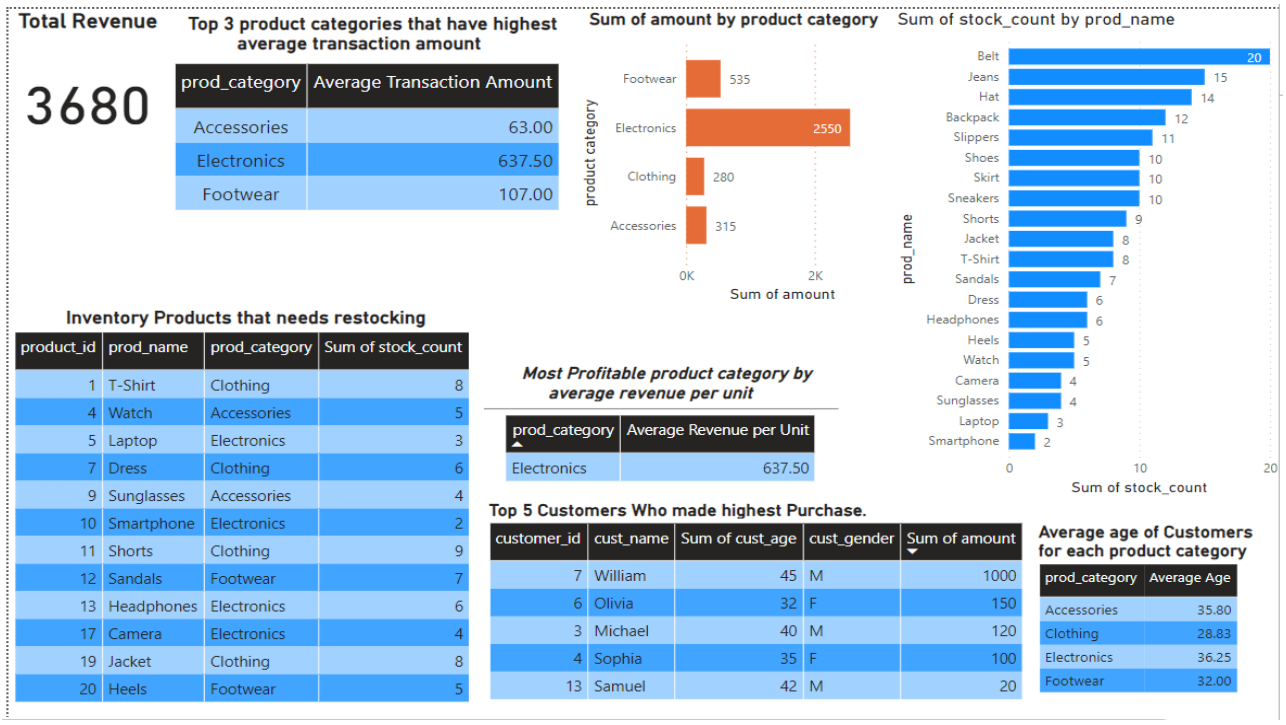
- Write a SQL query to retrieve the top 3 product categories that have the highest average transaction amount

```sql
SELECT p.prod_category, AVG(s.amount) AS average_transaction_amount
FROM product p
JOIN salesss s ON p.product_id = s.product_id
GROUP BY p.prod_category
ORDER BY average_transaction_amount DESC
LIMIT 3;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

prod_category	average_transaction_amount
Electronics	637.5000
Footwear	107.0000
Accessories	63.0000

Final Report :



Final Report based on the retail company's sale and inventory transaction.

- Where total revenue is 3680 (Revenue)
- Accessories , Electronics , Footwear are the top product that have highest average transaction amount
- Electronics are the most Profitable product with the average revenue per unit of 637 (Average Revenue Per Unit)

Report analysis based on Electronics :

- There are four products under Electronics namely Laptop , Smartphone, Headphones, Camera
- It's the most Profitable product category of 637.50 (Average revenue per unit)
- Average age of customers purchasing electronics items are 36.25 (Average Age)
- Finally , it's total revenue is 2550 (Revenue)