SCREENSHOTS OF OUTPUT

Table creation:

- A table named zepto is created with columns like sku_id, category, name, mrp, discountPercent, availableQuantity, discountedSellingPrice, weightInGms, outOfStock, quantity.
- This is meant to store product information similar to what you see on the Zepto app.

```
drop table if exists zepto;

create table zepto (
    sku_id SERIAL PRIMARY KEY,
    category VARCHAR(120),
    name VARCHAR(150) NOT NULL,
    mrp NUMERIC(8,2),
    discountPercent NUMERIC(5,2),
    availableQuantity INTEGER,
    discountedSellingPrice NUMERIC(8,2),
    weightInGms INTEGER,
    outOfStock BOOLEAN,
    quantity INTEGER
);
```

CREATE TABLE

Query returned successfully in 58 msec.

Data exploration:

1. **Count of rows** → To check how many records are present in the table.

```
--data exploration

--count of rows
select count(*) from zepto;
```

count bigint
3732

2. **Sample data** → Display the first 10 rows for a quick look.

```
--sample data
SELECT * FROM zepto
LIMIT 10;
```

Showing rows: 1 to 10   Page No:

sku_id [PK] integer	category character varying (120)	name character varying (150)	mrp numeric (8,2)	discountpercent numeric (5,2)	availablequantity integer	discountedsellingprice numeric (8,2)	weightingms integer	outofstock boolean	quantity integer
1	1	Fruits & Vegetables	Onion	2500.00	16.00	3	2100.00	1000	false	1
2	2	Fruits & Vegetables	Tomato Hybrid	4200.00	16.00	3	3500.00	1000	false	1
3	3	Fruits & Vegetables	Tender Coconut	5100.00	15.00	3	4300.00	58	false	1
4	4	Fruits & Vegetables	Coriander Leaves	2000.00	15.00	3	1700.00	100	false	100
5	5	Fruits & Vegetables	Ladies Finger	1400.00	14.00	3	1200.00	250	false	250
6	6	Fruits & Vegetables	Potato	3500.00	17.00	3	2900.00	1000	false	1
7	7	Fruits & Vegetables	Lemon	7500.00	16.00	3	6300.00	200	false	200
8	8	Fruits & Vegetables	Watermelon	5800.00	15.00	3	4900.00	58	false	1
9	9	Fruits & Vegetables	Capsicum Green	2300.00	17.00	3	1900.00	250	false	250

3. **Null values** → Identify rows where important columns have missing data.

```
25    --null values
26  ∨ SELECT * FROM zepto
27    WHERE name IS NULL
28    OR
29    category IS NULL
30    OR
31    mrp IS NULL
32    OR
33    discountPercent IS NULL
34    OR
35    discountedSellingPrice IS NULL
```

Data Output   Messages   Notifications

sku_id [PK] integer	category character varying (120)	name character varying (150)	mrp numeric (8,2)	discountpercent numeric (5,2)	availablequantity integer	discountedsellingprice numeric (8,2)	weightingms integer	outofstock boolean	quantity integer

4. **Different product categories** → Show all unique categories.

```
45       --different product categories
46  ∨  SELECT DISTINCT category
47     FROM zepto
48     ORDER BY category;
49
```

Data Output   Messages   Notifications

	category character varying (120) 🔒
1	Beverages
2	Biscuits
3	Chocolates & Candies
4	Cooking Essentials
5	Dairy, Bread & Batter
6	Fruits & Vegetables
7	Health & Hygiene
8	Home & Cleaning
9	Ice Cream & Desserts

Total rows: 14   Query complete 00:00:00.091

5. **Products in stock vs out of stock** → Group products by stock availability.

```
50     --products in stock vs out of stock
51  ∨ SELECT outOfStock, COUNT(sku_id)
52    FROM zepto
53    GROUP BY outOfStock;
```

Data Output   Messages   Notifications

	outofstock boolean 🔒	count bigint 🔒
1	false	3279
2	true	453

6. **Product names present multiple times** → Find duplicate product names with multiple SKUs.

```
55       --product names present multiple times
56  ∨  SELECT name, COUNT(sku_id) AS "Number of SKUs"
57     FROM zepto
58     GROUP BY name
59     HAVING count(sku_id) > 1
60     ORDER BY count(sku_id) DESC;
61
```

Data Output   Messages   Notifications

	name character varying (150) 🔒	Number of SKUs bigint 🔒
1	Sunfeast Yippee! Pasta Treat - Sour Cream Onion	10
2	Arden Eggs White	10
3	Quaker Oats	10
4	Saffola Veggie Twist Masala Oats	10
5	Amul Delicious Fat Spread - Cholesterol Free	10
6	Mother's Recipe Tamarind Paste	10
7	Kellogg's Real Almond & Honey Corn Flakes	9
8	Amul Fresh Cream	8
9	iD Idli & Dosa Batter	7

Total rows: 1214   Query complete 00:00:00.111

Data cleaning:

1. **Products with price = 0** → Identify and delete invalid entries where mrp or discountedSellingPrice is zero.

```
63
64    --products with price = 0
65  ∨ SELECT * FROM zepto
66    WHERE mrp = 0 OR discountedSellingPrice = 0;
67
68  ∨ DELETE FROM zepto
69    WHERE mrp = 0;
70
```

Data Output   Messages   Notifications

```
DELETE 0

Query returned successfully in 52 msec.
```

2. **Convert paise to rupees** → Divide mrp and discountedSellingPrice by 100 to correct scaling issues.

```
71    --convert paise to rupees
72  ∨ UPDATE zepto
73    SET mrp = mrp / 100.0,
74    discountedSellingPrice = discountedSellingPrice / 100.0;
75
76    SELECT mrp, discountedSellingPrice FROM zepto;
77
```

Data Output   Messages   Notifications

	mrp numeric (8,2)	discountedsellingprice numeric (8,2)
1	25.00	21.00
2	42.00	35.00
3	30.00	29.00
4	50.00	44.00
5	60.00	60.00
6	425.00	383.00
7	91.00	82.00
8	97.00	97.00
9	48.00	47.00

Total rows: 3731   Query complete 00:00:00.076

Data analysis:

1. **Top 10 best-value products** → Find products with the highest discount percent.

```
78    --data analysis
79
80    -- Q1. Find the top 10 best-value products based on the discount perce
81  ∨ SELECT DISTINCT name, mrp, discountPercent
82    FROM zepto
83    ORDER BY discountPercent DESC
84    LIMIT 10;
```

Data Output   Messages   Notifications

	name character varying (150)	mrp numeric (8,2)	discountpercent numeric (5,2)
1	Dukes Waffy Orange Wafers	45.00	51.00
2	Dukes Waffy Chocolate Wafers	45.00	51.00
3	Dukes Waffy Strawberry Wafers	45.00	51.00
4	RRO Mozzarella Block Cheese	295.00	50.00
5	Ceres Foods Laal Maas Instant Liquid Masala	220.00	50.00
6	Chef's Basket Durum Wheat Elbow Pasta	160.00	50.00
7	RRO Mozzarella Pizza Cheese	275.00	50.00
8	Ceres Foods Fish Mustard Instant Liquid Ma...	220.00	50.00
9	Dukes Waffy Strawberry Roll	150.00	50.00

Total rows: 10   Query complete 00:00:00.099

2. High MRP but out of stock → Products costing more than ₹300 but currently unavailable.

```
85
86    --Q2.What are the Products with High MRP but Out of Stock
87
88  ∨ SELECT DISTINCT name,mrp
89    FROM zepto
90    WHERE outOfStock = TRUE and mrp > 300
91    ORDER BY mrp DESC;
92
```

Data Output   Messages   Notifications

	name character varying (150)	mrp numeric (8,2)
1	Patanjali Cow's Ghee	565.00
2	MamyPoko Pants Standard Diapers, Extra Large (12 - 17 kg)	399.00
3	Aashirvaad Atta With Mutigrains	315.00
4	Everest Kashmiri Lal Chilli Powder	310.00

3. Estimated revenue per category → discountedSellingPrice × availableQuantity summed by category.

```
93    --Q3.Calculate Estimated Revenue for each category
94  ∨ SELECT category,
95    SUM(discountedSellingPrice * availableQuantity) AS total_revenue
96    FROM zepto
97    GROUP BY category
98    ORDER BY total_revenue;
99
```

Data Output   Messages   Notifications

	category character varying (120)	total_revenue numeric
1	Fruits & Vegetables	10846.00
2	Meats, Fish & Eggs	20693.00
3	Biscuits	25007.60
4	Beverages	55051.00
5	Dairy, Bread & Batter	55051.00
6	Health & Hygiene	64180.00
7	Home & Cleaning	122661.00
8	Ice Cream & Desserts	224385.00
9	Chocolates & Candies	224385.00

Total rows: 14   Query complete 00:00:00.069

4. MRP > 500 and discount < 10% → Expensive products with very little discount.

```
99
100   -- Q4. Find all products where MRP is greater than ₹500 and discount is less than 10%.
101 ∨ SELECT DISTINCT name, mrp, discountPercent
102   FROM zepto
103   WHERE mrp > 500 AND discountPercent < 10
104   ORDER BY mrp DESC, discountPercent DESC;
```

Data Output   Messages   Notifications

	name character varying (150)	mrp numeric (8,2)	discountpercent numeric (5,2)
1	Dhara Kachi Ghani Mustard Oil Jar	1250.00	8.00
2	Saffola Gold (Jar)	1240.00	0.00
3	Dhara Filtered Groundnut Oil (Jar)	1050.00	1.00
4	Fortune Rice Bran Health Oil (Jar)	1050.00	1.00
5	Dhara Filtered Groundnut Oil (Jar)	1050.00	0.00
6	Fortune Soyabean Oil	1005.00	0.00
7	Fortune Sunlite Refined Sunflower (Jar)	925.00	0.00
8	Surf Excel Matic Powder Front Load	810.00	7.00
9	Surf Excel Matic Top Load	720.00	9.00

5. Top 5 categories with highest avg discount → Using AVG(discountPercent).

```
105
106    -- Q5. Identify the top 5 categories offering the highest average discount percentage.
107 ∨  SELECT category,
108    ROUND(AVG(discountPercent),2) AS avg_discount
109    FROM zepto
110    GROUP BY category
111    ORDER BY avg_discount DESC
112    LIMIT 5;
113
```

Data Output   Messages   Notifications

	category character varying (120)	avg_discount numeric
1	Fruits & Vegetables	15.46
2	Meats, Fish & Eggs	11.03
3	Ice Cream & Desserts	8.32
4	Chocolates & Candies	8.32
5	Packaged Food	8.32

6. Price per gram → Calculate discountedSellingPrice ÷ weightInGms for products above 100g.

```
113
114    -- Q6. Find the price per gram for products above 100g and sort by best value.
115 ∨  SELECT DISTINCT name, weightInGms, discountedSellingPrice,
116    ROUND(discountedSellingPrice/weightInGms,2) AS price_per_gram
117    FROM zepto
118    WHERE weightInGms >= 100
119    ORDER BY price_per_gram;
120
```

Data Output   Messages   Notifications

	name character varying (150)	weightingms integer	discountedsellingprice numeric (8,2)	price_per_gram numeric
1	Tata Salt	1000	24.00	0.02
2	Onion	1000	21.00	0.02
3	Onion	3000	57.00	0.02
4	Shubh kart - Nirmal sugandhi mogra wet dhoop zipper 20 sticks	1160	28.00	0.02
5	Vicks Cough Drops Menthol	1160	20.00	0.02
6	Aashirvaad Iodised Salt	1000	19.00	0.02
7	Carrot	500	15.00	0.03
8	Potato	1000	29.00	0.03
9	Beetroot	500	13.00	0.03

7. Group products into Low / Medium / Bulk → Based on weight:

- ○ <1000 g = Low

- ○ <5000 g = Medium

- ○ ≥5000 g = Bulk

```
121    --Q7.Group the products into categories like Low, Medium, Bulk.
122 ∨  SELECT DISTINCT name, weightInGms,
123    CASE WHEN weightInGms < 1000 THEN 'Low'
124        WHEN weightInGms < 5000 THEN 'Medium'
125        ELSE 'Bulk'
126        END AS weight_category
127    FROM zepto;
128
```

Data Output   Messages   Notifications

	name character varying (150)	weightingms integer	weight_category text
1	Popular Essentials Poppy Seeds (Gasagase)	100	Low
2	Origami Kitchen Towels 2 In 1 Quilted 2ply	116	Low
3	Origami Good Karma Paper Serviettes - 100 napkins (1 ply)	58	Low
4	Johnson's Baby Top To Toe Bath Body Wash	210	Low
5	Harpic Hygienic Toilet Cleaner Rim Block Citrus	26	Low
6	Lizol Disinfectant Surface & Floor Cleaner Liquid Jasmine	500	Low
7	Sapota	348	Low
8	Britannia Treart Creme Wafers Chocolate	60	Low
9	Popular Essentials Dhania (Coriander Seeds)	200	Low

Total rows: 1782   Query complete 00:00:00.097

8. Total inventory weight per category →

Calculate SUM(weightInGms × availableQuantity) to see how heavy the stock is for each category.

```
129    --Q8.What is the Total Inventory Weight Per Category
130 ∨  SELECT category,
131    SUM(weightInGms * availableQuantity) AS total_weight
132    FROM zepto
133    GROUP BY category
134    ORDER BY total_weight;
```

Data Output   Messages   Notifications

	category character varying (120)	total_weight bigint
1	Meats, Fish & Eggs	48016
2	Biscuits	84431
3	Fruits & Vegetables	91794
4	Health & Hygiene	142904
5	Dairy, Bread & Batter	143735
6	Beverages	143735
7	Paan Corner	348187
8	Personal Care	348187
9	Home & Cleaning	373161

Total rows: 14   Query complete 00:00:00.118