Zepto E-commerce SQL Data Analysis

Extracting business insights using SQL

Creating a table named zepto_data

Column Name	Description
sku_id	Unique identifier for each product entry (Synthetic Primary Key)
name	Product name as listed on the Zepto app
category	Category of the product (e.g., Fruits, Snacks, Beverages)
mrp	Maximum Retail Price (originally in paise, converted to ₹)
discountPercent	Discount applied on MRP (in %)
discountedSellingPrice	Final price after discount (in ₹)
availableQuantity	Units available in inventory
weightInGms	Product weight in grams
outOfStock	Boolean flag indicating if the product is out of stock
quantity	Units per package (may be combined with weight for loose produce)

```sql
create table zepto_data(
    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
);
```

sku_id	category	name	mrp	discountpercent	availablequantity	discountedsellingprice	weightir	outofsto	quantity
[PK] inte	character	characte	numeric	numeric (5,2)	integer	numeric (8,2)	integer	boolean	integer

Data Exploration

```sql
--data exploration

--count of rows

SELECT COUNT(*) FROM zepto_data;


--sample data



SELECT * FROM zepto_data
LIMIT 10;
```

count
bigint 🔒

3732

category character	name characte	mrp numeric (8,	discoun numeric	availabl integer	discounted numeric (8,	weightir integer	outofstock boolean	quantity integer
Fruits...	Onion	2500.00	16.00	3	2100.00	1000	false	1
Fruits...	Tom...	4200.00	16.00	3	3500.00	1000	false	1
Fruits...	Ten...	5100.00	15.00	3	4300.00	58	false	1
Fruits...	Cori...	2000.00	15.00	3	1700.00	100	false	100
Fruits...	Ladi...	1400.00	14.00	3	1200.00	250	false	250
Fruits...	Pota...	3500.00	17.00	3	2900.00	1000	false	1
Fruits...	Lem...	7500.00	16.00	3	6300.00	200	false	200
Fruits...	Wat...	5800.00	15.00	3	4900.00	58	false	1
Fruits...	Cap...	2300.00	17.00	3	1900.00	250	false	250
Fruits...	Chill...	1900.00	15.00	3	1600.00	100	false	100

Find Null Values

```sql
-- null values

SELECT * FROM zepto_data
WHERE name IS NULL
OR
category IS NULL
OR
mrp IS NULL
OR
discountPercent IS NULL
OR
discountedSellingPrice IS NULL
OR
weightInGms IS NULL
OR
availableQuantity IS NULL
OR
outOfStock IS NULL
OR
quantity IS NULL;
```

Distinct Product Categories

	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
10	Meats, Fish & Eggs
11	Munchies
12	Paan Corner
13	Packaged Food
14	Personal Care

Products
in stock vs out of stock

```sql
--products in stock vs out of stock

select outOfStock, COUNT(sku_id)
FROM zepto_data
GROUP BY outOfStock;
```

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

```sql
-- data cleaning

--profucts with price == 0

SELECT * FROM zepto_data
WHERE mrp = 0 OR discountedSellingPrice = 0;


DELETE FROM zepto_data
WHERE mrp = 0;


--convert paise to rupees

UPDATE zepto_data
SET mrp = mrp / 100.0,
discountedSellingPrice = discountedSellingPrice/100.0;

SELECT mrp, discountedSellingPrice FROM zepto_data;
```

mrp numeric (8,2) 🔒	discountedsellingprice numeric (8,2) 🔒
25.00	21.00
42.00	35.00
32.00	27.00
30.00	25.00
60.00	60.00

DATA CLEANING

DELETE PRODUCTS WITH PRICE ZERO.
CONVERT PAISE TO RUPEES.

```sql
SELECT DISTINCT name, mrp, discountPercent
FROM zepto_data
ORDER BY discountPercent DESC
LIMIT 10;
```

	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
10	Ceres Foods Nalli Nihari Instant Liquid Masala	220.00	50.00

Q1. Find the top 10 best-value products based on the discount percentage.

```sql
SELECT DISTINCT name, mrp
FROM zepto_data
WHERE outOfStock = TRUE and mrp >= 300
ORDER BY mrp DESC;
```

	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

Q2. What are the products with High MRP but out of stock?

Q3. Calculate Estimated Revenue for each category.

```sql
SELECT category,
SUM(discountedSellingPrice * availableQuantity) AS total_revenue
FROM zepto_data
GROUP BY category
ORDER BY total_revenue;
```

	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
10	Packaged Food	224385.00
11	Personal Care	270849.00
12	Paan Corner	270849.00
13	Munchies	337369.00
14	Cooking Essentials	337369.00

Q4. Find all the products where MRP is greater than Rs. 500 and discount is less than 10%.

```sql
SELECT DISTINCT name, mrp, discountPercent
FROM zepto_data
WHERE mrp > 500 AND discountPercent < 10
ORDER BY mrp DESC, discountPercent DESC;
```

	name character varying (150)	mrp numeric (8	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 (...	925.00	0.00

Q5. Identify the top 5 categories offering the highest average discount percentage.

```sql
SELECT category,
ROUND(AVG(discountPercent), 2) AS avg_discount
FROM zepto_data
GROUP BY category
ORDER BY avg_discount DESC
LIMIT 5;
```

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

```sql
SELECT DISTINCT name, weightInGms, discountedSellingPrice,
discountedSellingPrice/weightInGms AS price_per_gram
FROM zepto_data
WHERE weightInGms >= 100
ORDER BY price_per_gram;
```

name character varying (150)	weightingms integer	discountedsellingprice numeric (8,2)	price_per_gram numeric
Vicks Cough Drops Menthol	1160	20.00	0.0172413793100
Onion	3000	57.00	0.01900000000000
Aashirvaad Iodised Salt	1000	19.00	0.01900000000000
Onion	1000	21.00	0.02100000000000
Tata Salt	1000	24.00	0.02400000000000
Shubh kart - Nirmal sugand...	1160	28.00	0.024137931034
Beetroot	500	13.00	0.02600000000000
Shubh kart - Tejas Twisted ...	1000	28.00	0.02800000000000
Potato	3000	84.00	0.02800000000000
Potato	1000	29.00	0.02900000000000

Q6. Find the price per gram for products above 100g and sort by best value.

```sql
SELECT DISTINCT name, weightInGms,
CASE WHEN weightInGms < 1000 THEN 'Low'
    WHEN weightInGms < 5000 THEN 'Medium'
    ELSE 'Bulk'
    END AS weight_category
FROM zepto_data;
```

name character varying (150)	weightingms integer	weight_category text
Popular Essentials Poppy Seeds (Gasagase)	100	Low
Origami Kitchen Towels 2 In 1 Quilted 2ply	116	Low
Origami Good Karma Paper Serviettes - 100 napkins (1 ply)	58	Low
Johnson's Baby Top To Toe Bath Body Wash	210	Low
Harpic Hygienic Toilet Cleaner Rim Block Citrus	26	Low
Lizol Disinfectant Surface & Floor Cleaner Liquid Jasmine	500	Low
Sapota	348	Low
Britannia Treart Creme Wafers Chocolate	60	Low
Popular Essentials Dhania (Coriander Seeds)	200	Low
Popular Essentials Californian Almond	250	Low
Everest Shahi Biryani Masala	50	Low

Q7. Group the products into categories: Low, Medium, Bulk.

```sql
SELECT category,
SUM(weightInGms * availableQuantity) AS total_weight
FROM zepto_data
GROUP BY category
ORDER BY total_weight;
```

	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
10	Ice Cream & Desserts	490797
11	Chocolates & Candies	490797
12	Packaged Food	490797
13	Munchies	1404654
14	Cooking Essentials	1404654

Q8. What is the total inventory Weight Per Category.