# Zepto Data Analysis Using SQL

Data Preprocessing:

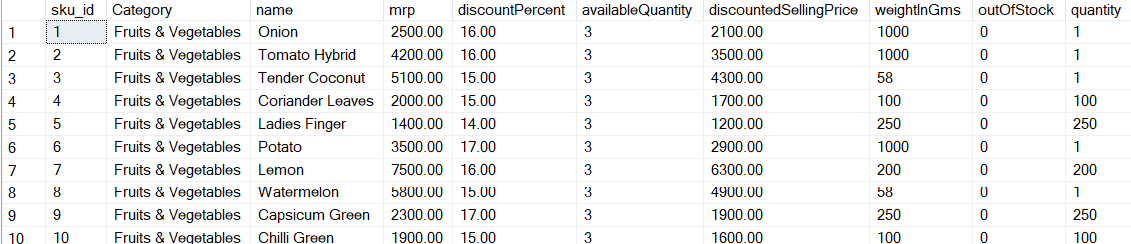
* Number of rows:

select count(\*) as number\_of\_rows from zepto\_v2;



* Displaying first few data:

select top 10 \* from zepto\_v2;



* Checking Null values:

select \* from zepto\_v2

where name is null

or category is null

or mrp is null

or discountPercent is null

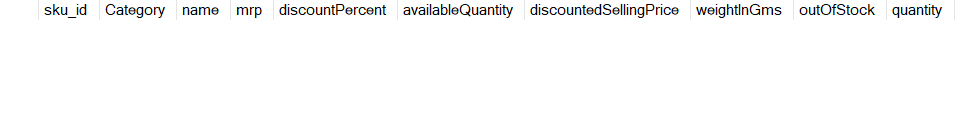
or availableQuantity is null

or discountedSellingPrice is null

or weightInGms is null

or outOfStock is null

or quantity is null;



* Different categories in the data:

select distinct category from zepto\_v2

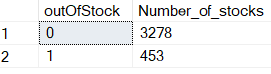
order by Category;



* Products in and out of stocks:

select outOfStock,count(\*) Number\_of\_stocks from zepto\_v2

group by outOfStock;



Data Cleaning:

select \* from zepto\_v2

where mrp=0 or discountedSellingPrice=0;

delete from zepto\_v2

where mrp=0;

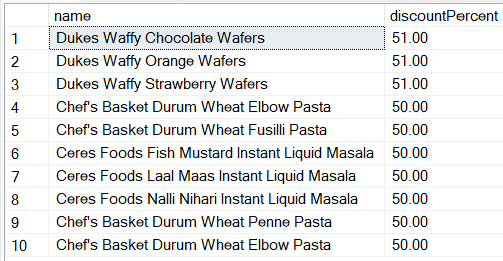
Analysis:

* Find the top 10 products with highest discount percentage?

select top 10 name,discountPercent

from zepto\_v2

order by discountPercent desc;

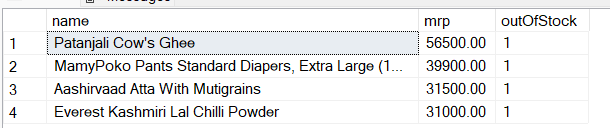


* Find products with high mrp but out of stock?

select distinct name,mrp,outOfStock from zepto\_v2

where outOfStock=1 and mrp> 30000

order by mrp desc;



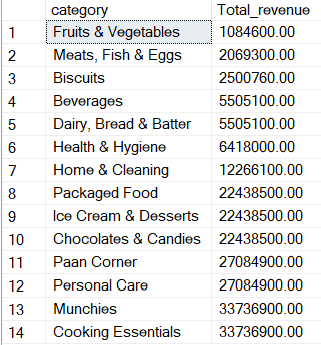
* Calculate estimate revenue for each category?

select category,sum(discountedsellingprice\*availablequantity) as Total\_revenue

from zepto\_v2

group by Category

order by Total\_revenue;



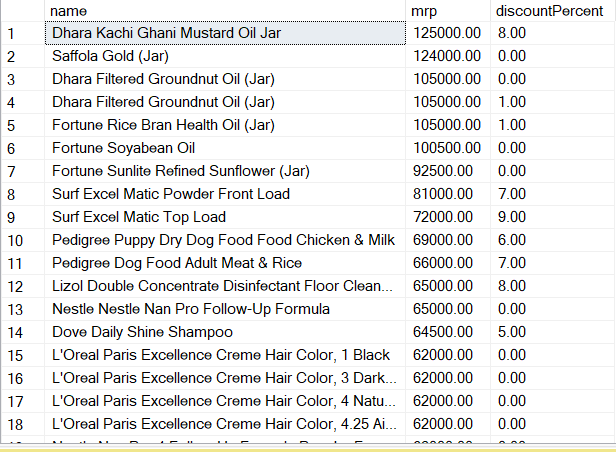
* Find all products where mrp less than 50000 and discount percentage less than 10 %?

select distinct name,mrp,discountPercent

from zepto\_v2

where mrp>50000 and discountPercent<10

order by mrp desc;



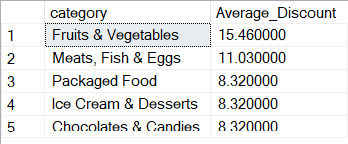
* Find top 5 categories offering highest average discount percentage?

select top 5 category,round(avg(discountPercent),2) as Average\_Discount

from zepto\_v2

group by Category

order by Average\_Discount desc;



* Find the price per gram of the products above 100 gms and sort best values?

select \* from zepto\_v2;

select distinct name,weightingms,discountedSellingPrice,

round(discountedSellingPrice/weightingms,2) as Price\_per\_Gram

from zepto\_v2

where weightInGms>100

order by Price\_per\_Gram desc;



* Group the proucts into categories like low,medium,bulk?

select distinct name,weightInGms,

case when weightInGms<1000 then 'Low'

when weightInGms<5000 then 'Medium'

else 'Bulk'

end as Weight\_Category

from zepto\_v2;



* What is the total inventory weight per category?

select category,

sum(weightingms\*availablequantity) as Total\_Weight

from zepto\_v2

group by Category

order by Total\_Weight;

