WALMART ANALYSIS

AIM:

This Project aims to explore the Walmart Sales data to understand top Performing branches and products, sales Trend of different products, customer behaviour. The aims is to study how sales strategies can be improved and optimized

PURPOSE OF THE PROJECT:

The major aim of the project is to gain insights into the sales data of Walmart to understand the different factors that affect sales of the different branches

ABOUT DATA:

Column	Description	Data Type
invoice_id	Invoice of the sales made	VARCHAR(30)
branch	Branch at which sales were made	VARCHAR(5)
city	The location of the branch	VARCHAR(30)
customer_type	The type of the customer	VARCHAR(30)
gender	Gender of the customer making purchase	VARCHAR(10)
product_line	Product line of the product solf	VARCHAR(100)
unit_price	The price of each product	DECIMAL(10, 2)
quantity	The amount of the product sold	INT
VAT	The amount of tax on the purchase	FLOAT(6, 4)
total	The total cost of the purchase	DECIMAL(10, 2)
date	The date on which the purchase was made	DATE
time	The time at which the purchase was made	TIMESTAMP
payment_method	The total amount paid	DECIMAL(10, 2)
cogs	Cost Of Goods sold	DECIMAL(10, 2)
gross_margin_percentage	Gross margin percentage	FLOAT(11, 9)
gross_income	Gross Income	DECIMAL(10, 2)
rating	Rating	FLOAT(2, 1)

STEPS HAS BEEN TAKEN:

I have completely understood all the important functions in SQL and also I have understand to solve difficult errors which has came while execution the problems also I have understand there are many way to solve each and every problems

1. How many unique product lines does the data have?

```
select count(distinct(Product_line)) as “product_line”
 from Walmart;
```

```
16  18:10:00  select count(distinct(Product_line)) from Walmart  1 row(s) returned
```

	product_line
▶	6

2. What is the most common payment method?

```
select max((Payment)) as "most common use"
 from Walmart;
```

```
17  18:10:17  select max((Payment)) as "most common use" from Walmart  1 row(s) returned
```

	most common use
▶	Ewallet

3. What is the most selling product line?

```
select max(Product_line) as maxmium_usage
from Walmart;
```

```
18  18:10:25  select max(Product_line) as maxmium_usage from Walmart  1 row(s) returned
```

	maxmium_usage
▶	Sports and travel

4. What is the total revenue by month?

```
with c1 as
            (select extract(month from Date) as month ,
             sum(Total) as total
             from Walmart
            group by 1)
select * from c1;
```

19 18:10:30 with c1 as (select extract(month from Date) as month , sum(Total) as total from Walmart group by 1) select * from c1 1 row(s) returned

month	total
1	30777.873
2	25393.168499999996
3	32543.763
4	27970.225499999997
5	30924.589499999987
6	32242.41299999999
7	29998.206000000002
8	25973.450999999994
9	30275.700000000004
10	17916.853499999997
11	16966.750500000002
12	21983.755500000003

5. What month had the largest COGS?

```
select extract(month from Date) as month, sum(cogs) as COGS
from Walmart
group by 1
order by month desc
LIMIT 1;
```

```
54  15:06:54  select extract(month from Date) as month,sum(cogs) as COGS from Walmart group by 1 order by month desc LIMIT 1  1 row(s) returned
```

month	COGS
12	20936.909999999993

6. What product line had the largest revenue?

```
select Product_line, sum(Total) as total
from Walmart
group by Product_line
order by total desc
LIMIT 1;
```

```
56  15:11:21  select Product_line,sum(Total) as total  from Walmart  group by Product_line  order by total desc  LIMIT 1  1 row(s) returned
```

Product_line	total
Food and beverages	56144.844000000005

7. What is the city with the largest revenue?

```
select City, sum(Total) as total
from Walmart
group by City;
```

```
58  15:14:06  select City,sum(Total) as total from Walmart group by City order by total desc LIMIT 1  1 row(s) returned
```

City	total
Naypyitaw	110568.70649999994

8. What product line had the largest VAT?

```
select Product_line, ROUND(max(Tax*cogs) as VAT
from Walmart
group by Product_line
order by VAT desc
limit 1;
```

```
17  11:40:06  select Product_line,ROUND(max(Tax*cogs),2) as VAT from Walmart group by Product_line order by VAT desc limit 1  1 row(s) returned
```

Product_line	VAT
Fashion accessories	49302.45

9. Fetch each product line and add a column to those product line showing "Good", "Bad". Good if its greater than average sales

```
select *,
      case
       when Total> (select avg(Total) from Walmart) then "good"
       else "bad"
end as suggestion
from Walmart;
```

```
24  18:11:31  select *,  case  when Total> (select avg(Total) from Walmart) then "good"  else "bad" end as suggestion from Walmart
```

Invoice_ID	Branch	City	Customer_type	Gender	Product_line	Unit_price	Quantity	Tax	Total	Date	Time	Payment
750-67-8428	A	Yangon	Member	Female	Health and beauty	74.69	7	26.1415	548.9715	2019-01-05	13:08:00	Ewallet
226-31-3081	C	Naypyitaw	Normal	Female	Electronic accessories	15.28	5	3.82	80.22	2019-01-06	10:29:00	Cash
631-41-3108	A	Yangon	Normal	Male	Home and lifestyle	46.33	7	16.2155	340.5255	2019-01-07	13:23:00	Credit card

cogs	gross_margin_percentage	gross_income	Rating	suggestion
522.83	4.761904762	26.1415	9.1	good
76.4	4.761904762	3.82	9.6	bad
324.31	4.761904762	16.2155	7.4	good

10. Which branch sold more products than average product sold?

```
select Branch, sum(Quantity) as sum_quantity
from Walmart
group by Branch
having sum_quantity > (select avg(Quantity) as avg_quantity from
Walmart);
```

```
62 15:20:34 select Branch,sum(Quantity) as sum_quantity from Walmart group by Branch having sum_quantity > (select avg(Quantity) as avg_quantity from Walmart) order by sum_quantity desc LIMIT 1
```

Branch	sum_quantity
A	1859

11. What is the most common product line by gender?

```
select Gender, Product_line, count(Gender) total_cnt
from Walmart
group by Product_line, Gender
order by total_cnt desc
limit 1;
```

```
64 15:23:00 select Gender,Product_line ,count(Gender) total_cnt from Walmart group by Product_line,Gender order by total_cnt desc limit 1 1 row(s) returned
```

Gender	Product_line	total_cnt
Female	Fashion accessories	96

12. What is the average rating of each product line?

```
select Product_line, avg(Rating) as rating
from Walmart
group by Product_line;
```

```
27 18:12:18 select Product_line,avg(Rating) as rating from Walmart group by Product_line
```

	Product_line	rating
▶	Health and beauty	7.003289473684212
	Electronic accessories	6.92470588235294
	Home and lifestyle	6.8375
	Sports and travel	6.916265060240964
	Food and beverages	7.113218390804598
	Fashion accessories	7.029213483146067

13. How many unique customer types does the data have?

```
select distinct(Customer_type) from Walmart;
```

	Customer_type
▶	Member
	Normal

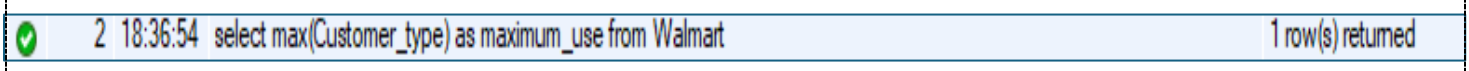
14. How many unique payment methods does the data have?

```
select distinct(Payment) from Walmart;
```

	Payment
▶	Ewallet
	Cash
	Credit card

15. Which customer type buys the most?

```
select max(Customer_type) as maximum_use
from Walmart;
```



16. What is the gender of most of the customers?

```
select max(Gender) as maximum_gender_buy
from Walmart;
```

```
3  18:36:57  select max(Gender) as maximum_gender_buy from Walmart  1 row(s) returned
```

	maximum_use
▸	Normal

17. What is the gender distribution per branch?

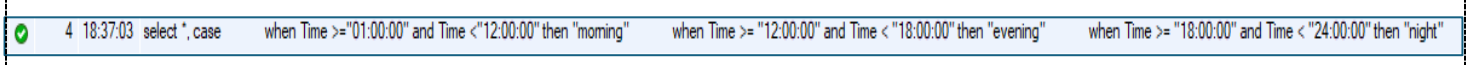
```
select
  Branch,
  sum(case when Gender = "Male" then 1  else 0 end)
  as Male_count,
  sum(case when Gender = "Female" then 1 else 0 end)
   as Female_count
from Walmart
group by Branch;
```

```
7  18:38:54  select  Branch,  sum(case when Gender = "Male" then 1 else 0 end) AS Male_count,  sum(case when Gender = "Female" then 1 else 0 end) AS Female_count from Walmart group by Branch
```

	Branch	Male_count	Female_count
▸	A	179	161
	C	150	178
	B	170	162

18. Which time of the day do customers give most ratings?

```
    select
      case
       when Time >="01:00:00" and
        Time <"12:00:00" then "morning"
       when Time >= "12:00:00" and
        Time < "18:00:00" then "evening"
       when Time >= "18:00:00" and
        Time < "24:00:00" then "night"
       else "nill"
end as time_of_rating,
extract(hour from Time) as time ,
avg(Rating) as count
from Walmart
group by 1,2
order by count desc
limit 1;
```



time_of_rating	time	count
evening	12	7.299999999999998

19. Which time of the day do customers give most ratings per branch?

```
select
    case
      when Time >="01:00:00" and
      Time <"12:00:00" then "morning"
      when Time >= "12:00:00" and
      Time < "18:00:00" then "evening"
      when Time >= "18:00:00" and
      Time < "24:00:00" then "night"
```

```
        else "nill"
end as time_of_rating,
extract(hour from Time) as time,
Branch,
count(Rating) as count
from Walmart
group by 1,2,3
order by count desc
limit 1;
```

```
7  18:38:54  select  Branch,  sum(case when Gender = "Male" then 1 else 0 end) AS Male_count,  sum(case when Gender = "Female" then 1 else 0 end) AS Female_count from Walmart group by Branch
```

time_of_rating	time	Branch	count
night	19	B	50

20. Which of the customer types brings the most revenue?

```
select Customer_type, max(total) as total_revenue
from Walmart
group by Customer_type
order by total_revenue desc
limit 1;
```

```
12  18:41:06  select Customer_type,max(total) as total_revenue from Walmart group by Customer_type order by total_revenue desc limit 1
```

Customer_type	total_revenue
Member	1042.65

21. Which city has the largest tax percent/ VAT (**Value Added Tax**)?

```
select City, ROUND(max(Tax*cogs),2) as VAT
from Walmart
group by City
order by VAT desc
limit 1;
```

```
18  11:45:07  select City,ROUND(max(Tax*cogs),2) as VAT from walmart group by City order by VAT desc limit 1  1 row(s) returned
```

City	VAT
Naypyitaw	49302.45

22. Which customer type pays the most in VAT?

```
select Customer_type, ROUND(max(Tax*cogs),2) as total_revenue
from Walmart
group by Customer_type
order by total_revenue desc
limit 1;
```

```
14  18:41:13  select Customer_type,max(Tax) as total_revenue from Walmart group by Customer_type order by total_revenue desc limit 1
```

Customer_type	total_revenue
Member	49302.45

23. Which day for the week has the best avg ratings?

```
select extract(week from Date) as Week, avg(Rating) as Rating
from walmart
group by 1
order by Rating desc
LIMIT 1;
```

```
78  15:38:26  select extract(week from Date) as Week, avg(Rating) as Rating from walmart group by 1 order by Rating desc LIMIT 1  1 row(s) returned
```

Week	Rating
0	8.028571428571428

24. Which day of the week has the best average ratings per branch?

```
select extract(week from Date) as Week, Branch, avg(Rating) as
Rating
from walmart
group by 1,2
order by Rating desc
LIMIT 1;
```

```
79  15:45:22  select extract(week from Date) as Week, Branch, avg(Rating) as Rating from walmart group by 1,2 order by Rating desc LIMIT 1   1 row(s) returned
```

	Week	Branch	Rating
▶	0	B	8.9

25. Number of sales made in each time of the day per weekday?

```
select extract(week from Date) as Week,count(Invoice_ID)
as "NO OF SALES"
from walmart
group by 1;
```

```
84  15:56:25  select extract(week from Date) as Week,count(Invoice_ID) as "NO OF SALES" from walmart group by 1   53 row(s) returned
```

	Week	NO OF SALES
▶	0	7
	1	21
	2	21
	3	21
	4	21
	5	21
	6	21
	7	21
	8	21
	9	21
	10	21

PROJECT DONE BY:

KIRIT P.S

30-05-2024