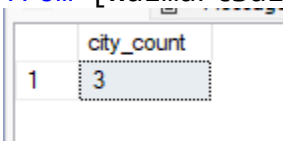
Walmart Sales

Common question

1. How many unique cities does the data have?

```
 select COUNT(distinct city) as city_count
from [WalmartSalesData.csv]
```

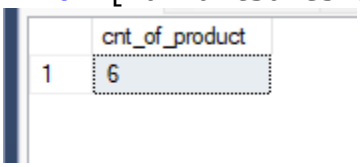


	city_count
1	3

Product Based question

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

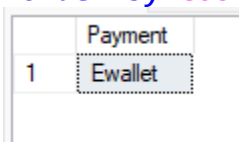
```
select COUNT(distinct product_line) as cnt_of_product
 from [WalmartSalesData.csv]
```



	cnt_of_product
1	6

2. What is the most common payment method

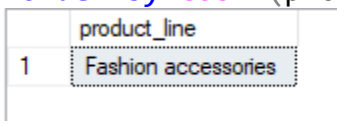
```
 select top 1 Payment
 from [WalmartSalesData.csv]
 group by Payment
 order by COUNT(payment) desc;
```



	Payment
1	Ewallet

3. What is the most selling product line?

```
 select top 1 product_line
 from [WalmartSalesData.csv]
 group by Product_line
 order by COUNT(product_line) desc
```



	product_line
1	Fashion accessories

4.What is the total revenue by month?

```
 SELECT
    month_day AS Month,
    CAST(ROUND(SUM(total) / 1000.0, 2) AS VARCHAR(20)) + 'k' AS
totalrevenue
FROM
    [WalmartSalesData.csv]
```

```
GROUP BY
    month_day
ORDER BY
    SUM(total) DESC;
```

	Month	totalrevenue
1	January	116.29k
2	March	109.46k
3	February	97.22k

\

5. What month had the largest COGS;

```
 SELECT
    month_day AS Month,
    CAST(ROUND(SUM(cogs) / 1000.0, 2) AS VARCHAR(20)) + 'k' AS
totalcogs
FROM
    [WalmartSalesData.csv]
GROUP BY
    month_day
ORDER BY
    SUM(cogs) DESC;
```

	Month	totalcogs
1	January	110.75k
2	March	104.24k
3	February	92.59k

6.What product line has the largest revenue

```
 select product_line,
     cast(round(SUM(total)/1000.0, 2) AS varchar(20)) + 'k' as
totalrev
from [WalmartSalesData.csv]
group by Product_line
order by SUM(total) desc;
```

	product_line	totalrev
1	Food and beverages	56.14k
2	Sports and travel	55.12k
3	Electronic accessories	54.34k
4	Fashion accessories	54.31k
5	Home and lifestyle	53.86k
6	Health and beauty	49.19k

7.What is the city with the largest revenue

```
 select City,
     cast(round(SUM(total)/1000.0, 2) AS varchar(20)) + 'k' as
totalrev
```

```
from [WalmartSalesData.csv]
group by city
order by SUM(total) desc;
```

	City	totalrev
1	Naypyitaw	110.57k
2	Yangon	106.2k
3	Mandalay	106.2k

8.What product line had highest VAT

```
 select  PRODUCT_LINE,
         cast(round(SUM(0.05* cogs)/1000.0,2) AS varchar(20)) + 'k' as
TOTAL_VAT
 FROM [WalmartSalesData.csv]
 GROUP by Product_line
 order by TOTAL_VAT desc
```

	PRODUCT_LINE	TOTAL_VAT
1	Fashion accessories	2.59k
2	Home and lifestyle	2.56k
3	Electronic accessories	2.59k
4	Health and beauty	2.34k
5	Food and beverages	2.67k
6	Sports and travel	2.62k

9. Which branch sold more products than avg product sold.

```
  select top 1 branch,
       SUM(quantity) as total_sold
from Walmart_Sales
group by Branch
having SUM(quantity)> (select AVG(quantity) as avg_sold from
Walmart_Sales
)
```

	branch	total_sold
1	A	1859

10.What is the most common product line by gender

```
WITH GenderProductCounts AS (
    SELECT
        gender,
        product_line,
        COUNT(*) AS count
    FROM
        walmart_sales
    GROUP BY
```

```
        gender,
        product_line
),
MaxCounts AS
    (SELECT
        gender,
        MAX(count) AS max_count
    FROM
        GenderProductCounts
    GROUP BY
        gender
)
SELECT
    gpc.gender,
    gpc.product_line,
    gpc.count
FROM
    GenderProductCounts gpc
JOIN
    MaxCounts mc
ON
    gpc.gender = mc.gender AND gpc.count = mc.max_count;
```

	gender	product_line	count
1	Male	Health and beauty	88
2	Female	Fashion accessories	96

11. What is Avg rating of each product line

```
  select product_line,
       AVG(rating) as avg_rating
from Walmart_Sales
group by Product_line
```

	product_line	avg_rating
1	Fashion accessories	7.02921346600136
2	Home and lifestyle	6.8375
3	Electronic accessories	6.92470588123097
4	Health and beauty	7.00328945799878
5	Food and beverages	7.11321838970842
6	Sports and travel	6.91626506253897

Sales based questions

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

```
 SELECT
    Datename(weekday,date) AS weekday,
    time_of_day,
```

```
    COUNT(*) AS sales_count
FROM
    walmart_sales
GROUP BY
    Datename(weekday,date),
    time_of_day
ORDER BY
    weekday;
```

	weekday	time_of_day	sales_count
1	Friday	Afternoon	58
2	Friday	Evening	52
3	Friday	Morning	29
4	Monday	Afternoon	48
5	Monday	Evening	56
6	Monday	Morning	21
7	Saturday	Afternoon	55
8	Saturday	Evening	81
9	Saturday	Morning	28
10	Sunday	Afternoon	53
11	Sunday	Evening	58
12	Sunday	Morning	22
13	Thursday	Afternoon	49
14	Thursday	Evening	56
15	Thursday	Morning	33
16	Tuesday	Afternoon	53
17	Tuesday	Evening	69
18	Tuesday	Morning	36
19	Wednesday	Afternoon	61

2. Which of the customer types bring the most revenue

```
 SELECT Customer_type,
       cast(round(SUM(total)/1000.0,2) AS varchar(20)) + 'k' as
totalrev
from walmart_sales
group by Customer_type
```

	Customer_type	totalrev
1	Member	164.22k
2	Normal	158.74k

3. Which city has the largest tax percent/VAT (Value Added Tax)

```
  SELECT top 1
    city,
   AVG(Tax_5) as avg_vat
FROM
    walmart_sales
group by City
```

```
ORDER BY
    avg_vat DESC;
```

	city	avg_vat
1	Naypyitaw	16.05236

4.Which customer type pays the most in VAT?

```
 select top 1
       customer_type,
       cast(round(SUM((tax_5 /100)* total)/1000.0,2)AS varchar(20)) +
'k' as vat_paid
        from Walmart_Sales
group by customer_type
order by vat_paid desc;
```

	customer_type	vat_paid
1	Member	40.28k

Customer based question

1.How many unique customer type does the data have?

```
select distinct Customer_type
from Walmart_Sales
```

	Customer_type
1	Member
2	Normal

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

```
 select distinct Payment
from Walmart_Sales
```

	Payment
1	Credit card
2	Cash
3	Ewallet

3. What is the most common customer type?

```
 select top 1 Customer_type
 from Walmart_Sales
group by customer_type
order by COUNT(*) desc
```

	Customer_type
1	Member

4. Which Customer type buy the most?

```
  select top 1 customer_type
from Walmart_Sales
group by Customer_type
order by SUM(total) desc;
```

	customer_type
1	Member

5. What is the gender of the most common customer ?

```
SELECT top 1
    gender
FROM
    walmart_sales
GROUP BY
    gender
ORDER BY
    COUNT(*) DESC
```

	gender
1	Female

6. What is gender distribution per branch?

```
SELECT
    branch,
    gender,
    COUNT(*) AS customer_count
FROM
    walmart_sales
GROUP BY
    branch,
    gender
ORDER BY
    branch,
    gender;
```

	branch	gender	customer_count
1	A	Female	161
2	A	Male	179
3	B	Female	162
4	B	Male	170
5	C	Female	178
6	C	Male	150

7. Which time of day do customers give most ratings?

```
select time_of_day,
       COUNT(rating) as ratings
from walmart_sales
group by
        time_of_day
order by
       COUNT(rating) desc
```

	time_of_day	ratings
1	Evening	432
2	Afternoon	377
3	Morning	191

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

```
 select branch, time_of_day,
       COUNT(rating) as ratings
from walmart_sales
group by
        time_of_day, Branch
order by
      Branch, COUNT(rating)
```

branch	time_of_day	ratings
A	Morning	73
A	Afternoon	126
A	Evening	141
B	Morning	59
B	Afternoon	125
B	Evening	148
C	Morning	59
C	Afternoon	126
C	Evening	143

9. Which day of the week has the best avg ratings?

```
select day_name,
     ROUND( AVG(rating),2) as avg_rating
from Walmart_Sales
group by
      day_name
order by
      avg_rating desc
```

day_name	avg_rating
Monday	7.15
Friday	7.08
Sunday	7.01
Tuesday	7
Saturday	6.9
Thursday	6.89
Wednesday	6.81

10. which day of the week has the best average ratings per branch?

```
 with cte as
(select Branch, day_name,
     ROUND( AVG(rating),2) as avg_rating,
```

```
      RANK() over (partition by branch order by ROUND(
AVG(rating),2)desc) as rn
from Walmart_Sales
group by
      branch, day_name)

select Branch, day_name, avg_rating
from cte
where rn = 1
order by
     Branch, avg_rating desc
```

Branch	day_name	avg_rating
A	Friday	7.31
B	Monday	7.34
C	Friday	7.28