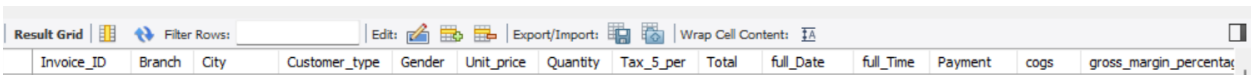
```
create database walmart_analysis;
use walmart_analysis;
```


```
create table walmart(
Invoice_ID      varchar(30) not null primary key,
Branch varchar(5) not null,
City varchar(30) not null,
Customer_type varchar(30) not null,
Gender varchar(10) not null,
Unit_price      decimal(10,2) not null,
Quantity int not null,
Tax_5_per float(6,4) not null,
Total decimal(10,2) not null,
full_Date date not null,
full_Time time not null,
Payment varchar(20) not null,
cogs decimal(10,2) not null,
gross_margin_percentage   decimal(10,9),
gross_income float(11,9),
Rating float(2,1) not null,
Sub_category varchar(150)
);
```




```
LOAD DATA INFILE 'walmart.csv'
INTO TABLE walmart
```

CHARACTER SET utf8mb4

FIELDS TERMINATED BY ','

ENCLOSED BY '"'

IGNORE 1 LINES;

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

Invoice_ID	Branch	City	Customer_type	Gender	Unit_price	Quantity	Tax_5_per	Total	full_Date	full_Time	Payment	cogs	gross_margin_percenta
101-17-6199	A	Yangon	Normal	Male	45.79	7	16.0265	336.56	2019-03-13	19:44:00	Credit card	320.53	4.761904762
101-81-4070	C	Naypyitaw	Member	Female	62.82	2	6.2820	131.92	2019-01-17	12:36:00	Ewallet	125.64	4.761904762
102-06-2002	C	Naypyitaw	Member	Male	25.25	5	6.3125	132.56	2019-03-20	17:52:00	Cash	126.25	4.761904762
102-77-2261	C	Naypyitaw	Member	Male	65.31	7	22.8585	480.03	2019-03-05	18:02:00	Credit card	457.17	4.761904762
105-10-6182	A	Yangon	Member	Male	21.48	2	2.1480	45.11	2019-02-27	12:22:00	Ewallet	42.96	4.761904762
105-31-1824	A	Yangon	Member	Male	69.52	7	24.3320	510.97	2019-02-01	15:10:00	Credit card	486.64	4.761904762

FEATURE ENGINEERING

1.Add a new column named `time_of_day` to give insight of sales in the Morning, Afternoon and Evening. This will help answer the question on which part of the day most sales are made.

alter table walmart

add column time_of_day varchar(30);

update walmart

set time_of_day=

CASE

when `full_Time` between "00:00:00" And "12:00:00" then "Morning"

when `full_Time` between "12:01:00" And "16:00:00" then "Afternoon"

else "Evening"

END;

time_of_day
Evening
Afternoon
Evening
Evening
Afternoon
Afternoon
Morning
Afternoon

2. Add a new column named `day_name` that contains the extracted days of the week on which the given transaction took place (Mon, Tue, Wed, Thur, Fri). This will help answer the question on which week of the day each branch is busiest.

```
alter table walmart

add column day_name varchar(30);


update walmart

set day_name=dayname(full_Date);
```

day_name
Wednesday
Thursday
Wednesday
Tuesday
Wednesday
Friday
Wednesday

3. Add a new column named `month_name` that contains the extracted months of the year on which the given transaction took place (Jan, Feb, Mar). Help determine which month of the year has the most sales and profit.

```
Alter Table walmart

add column month_name varchar(30);


update walmart

set month_name=

CASE

When month(full_Date) = 1 then 'January'

When month(full_Date) = 2 then 'Feburary'

When month(full_Date) = 3 then 'March'

When month(full_Date) =4 then 'April'

When month(full_Date) =5 then 'May'

When month(full_Date) =6 then 'June'

When month(full_Date) =7 then 'July'
```

When month(full_Date) =8 then 'August'

When month(full_Date) =9 then 'September'

When month(full_Date) =10 then 'October'

When month(full_Date) =11 then 'November'

Else 'December'

End;

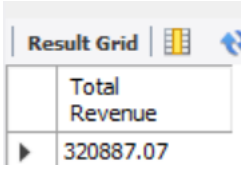
month_name
March
January
March
March
Feburary
Feburary
March

Business Questions To Answer

 # A-GENERIC QUESTION

1-Find The Total Sum Of Revenue in this data.

select round(sum(Total),2) as 'Total Revenue' from Walmart;

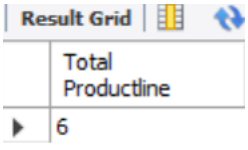


Result Grid

	Total Revenue
▶	320887.07

2. Find The Total Product Lines In this Data.

select count(distinct(Sub_category)) as 'Total Productline' from walmart;

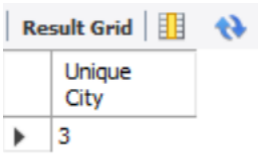


Result Grid

	Total Productline
▶	6

3-How many unique cities does the data have?

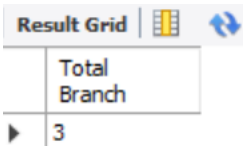
select count(distinct(City))as 'Unique City' from walmart;



	Unique City
▶	3

4-Find The Total Branches Available In This data.

select count(distinct(Branch)) as 'Total Branch' from walmart;



	Total Branch
▶	3

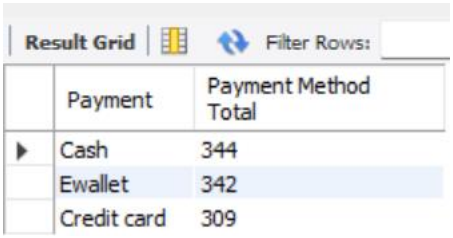
#B-PRODUCT QUESTION

1-What is the most common payment method?

select Payment,count(Payment) as 'Payment Method Total' from walmart

group by Payment

order by `Payment Method Total` desc;



	Payment	Payment Method Total
▶	Cash	344
	Ewallet	342
	Credit card	309

2-What is the most selling product line?

select Sub_category,count(Sub_Category) as 'Most Selling Productline' from walmart

group by Sub_category

order by `Most Selling Productline` desc;

Result Grid | Filter Rows:

Sub_category	Most Selling Productline
Fashion accessories	178
Food and beverages	174
Electronic accessories	169
Sports and travel	163
Home and lifestyle	160
Health and beauty	151

3-What is the most common product line by gender?

```
select Sub_Category,Gender,count(Sub_Category)as 'Gender Common Product Line' from walmart
group by Sub_category,Gender
order by `Gender Common Product Line` desc;
```

Result Grid | Filter Rows: | Export:

Sub_Category	Gender	Gender Common Product Line
Sports and travel	Female	86
Sports and travel	Male	77
Home and lifestyle	Female	79
Home and lifestyle	Male	81
Health and beauty	Female	63
Health and beauty	Male	88
Food and beverages	Female	90
Food and beverages	Male	84
Fashion accessories	Female	96
Fashion accessories	Male	82
Electronic accessories	Female	83
Electronic accessories	Male	86

4-What product line had the largest revenue?

```
select Sub_Category,round(sum(Total),2) as 'Product Line With Revenue' from walmart
group by Sub_category
order by `Product Line With Revenue` desc;
```

Result Grid | Filter Rows:

Sub_Category	Product Line With Revenue
Food and beverages	56144.96
Fashion accessories	54306.03
Sports and travel	53936.30
Home and lifestyle	53861.96
Electronic accessories	53783.34
Health and beauty	48854.48

5-What product line had the largest VAT?

```
select Sub_category,round(max(Tax_5_Per),2)as 'Product Line With Max Tax' from walmart
group by Sub_category
order by `Product Line With Max Tax` desc;
```

Result Grid | Filter Rows:

Sub_category	Product Line With Max Tax
Fashion accessories	49.65
Food and beverages	49.26
Home and lifestyle	48.75
Sports and travel	47.72
Health and beauty	45.25
Electronic accessories	44.88

6-What is the total revenue by month?

```
select month_name,round(sum(Total),2) as 'Month Revenue' from walmart
group by month_name
order by `Month Revenue` desc;
```

Result Grid | Filter Rows:

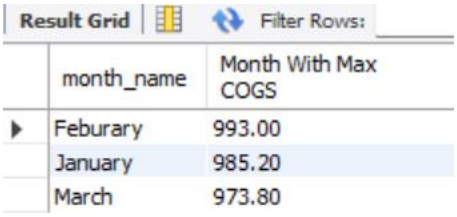
month_name	Month Revenue
January	116292.11
March	108867.38
Feburary	95727.58

7-What month had the largest COGS?

```
select month_name,max(cogs)as 'Month With Max COGS' from walmart
```

group by month_name

order by `Month With Max COGS` desc;



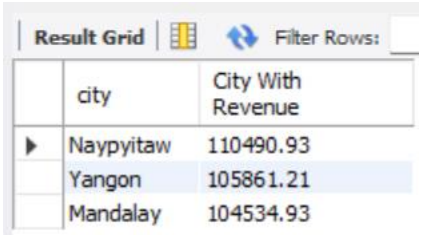
month_name	Month With Max COGS
Feburary	993.00
January	985.20
March	973.80

8-What is the city with the largest revenue?

select city,round(sum(Total),2) as 'City With Revenue' from walmart

group by city

order by `City With Revenue` desc;



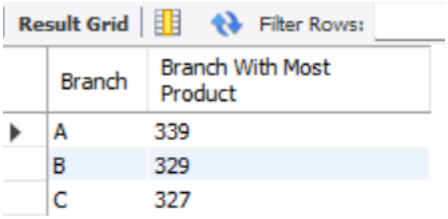
city	City With Revenue
Naypyitaw	110490.93
Yangon	105861.21
Mandalay	104534.93

9-Which branch sold more products than average product sold?

select Branch,round(count(Quantity),2)as 'Branch With Most Product' from walmart

group by Branch

order by `Branch With Most Product` desc;



Branch	Branch With Most Product
A	339
B	329
C	327

10-What is the average rating of each product line?

select Sub_category,round(avg(Rating),2) as 'Average Rating' from walmart

group by Sub_category

order by `Average Rating` desc;

Result Grid | Filter Rows:

Sub_category	Average Rating
Food and beverages	7.11
Fashion accessories	7.03
Health and beauty	6.98
Electronic accessories	6.91
Sports and travel	6.86
Home and lifestyle	6.84

#C-SALES QUESTION

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

```
select day_name,time_of_day,count(Quantity)as 'Sales Made In Quantity' from walmart
group by time_of_day,day_name
order by `Sales Made In Quantity` desc;
```

Result Grid | Filter Rows: | Exp

day_name	time_of_day	Sales Made In Quantity
Saturday	Evening	81
Tuesday	Evening	69
Wednesday	Afternoon	61
Wednesday	Evening	58
Friday	Afternoon	58
Sunday	Evening	58
Monday	Evening	56
Thursday	Evening	56
Saturday	Afternoon	55
Tuesday	Afternoon	53
Sunday	Afternoon	52
Friday	Evening	51
Thursday	Afternoon	49
Monday	Afternoon	48
Tuesday	Morning	36
Thursday	Morning	33

```
select time_of_day,count(Quantity)as 'Sales Made In Quantity' from walmart
where day_name='Monday'
group by time_of_day
order by `Sales Made In Quantity` desc;
```

2-Which of the customer types brings the most revenue?

```
select Customer_type,round(sum(Total),2) as 'Cust With Most Revenue' from walmart
group by Customer_type
order by `Cust With Most Revenue` desc;
```

Result Grid | Filter Rows:

Customer_type	Cust With Most Revenue
Member	163625.47
Normal	157261.60

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

```
select City,round(avg(Tax_5_per),2) as 'City With Max Tax' from walmart
group by City
order by `City With Max Tax` desc;
```

Result Grid | Filter Rows:

City	City With Max Tax
Naypyitaw	16.09
Mandalay	15.13
Yangon	14.87

4-Which customer type pays the most in VAT?

```
select Customer_type,round(avg(Tax_5_per),2) as 'Customer Type With Most Tax' from walmart
group by Customer_type
order by `Customer Type With Most Tax` desc;
```

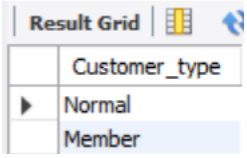
Result Grid | Filter Rows:

Customer_type	Customer Type With Most Tax
Member	15.61
Normal	15.1

#D-CUSTOMER QUESTION

1-How many unique customer types does the data have?

select distinct(Customer_type) from walmart;

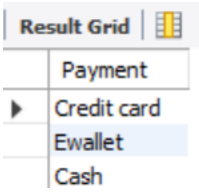


Result Grid

Customer_type
Normal
Member

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

select distinct(Payment) from walmart;



Result Grid

Payment
Credit card
Ewallet
Cash

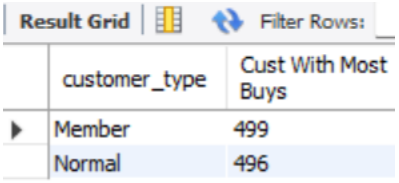
3-What is the most common customer type?

4-Which customer type buys the most?

select customer_type,count(Total)as 'Cust With Most Buys' from walmart

group by `customer_type`

order by `Cust With Most Buys` desc;



Result Grid Filter Rows:

customer_type	Cust With Most Buys
Member	499
Normal	496

5-What is the gender of most of the customers?

select gender,count(gender) as 'Cust Gender Count' from walmart

group by gender

order by `Cust Gender Count` desc;

Result Grid | Filter Rows:

	gender	Cust Gender Count
▶	Male	498
	Female	497

6-What is the gender distribution per branch?

```
select Branch,gender,count(gender) as 'Cust Gender Count' from walmart
group by Branch,gender
order by `Branch`;
```

Result Grid | Filter Rows:

	Branch	gender	Cust Gender Count
▶	A	Female	160
	A	Male	179
	B	Female	160
	B	Male	169
	C	Female	177
	C	Male	150

7-Which time of the day do customers give most ratings?

```
select time_of_day,round(avg(Rating),2)as 'Day Time Rating' from walmart
group by time_of_day
order by `Day Time Rating` desc;
```

Result Grid | Filter Rows:

	time_of_day	Day Time Rating
▶	Afternoon	7.02
	Morning	6.94
	Evening	6.91

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

```
select time_of_day,Branch,round(avg(Rating),2)as 'Day Time Rating' from walmart
group by time_of_day,Branch
order by `Branch`;
```

time_of_day	Branch	Day Time Rating
Evening	A	6.87
Afternoon	A	7.19
Morning	A	7.01
Afternoon	B	6.81
Morning	B	6.84
Evening	B	6.75
Afternoon	C	7.07
Evening	C	7.1
Morning	C	6.97

9-Which day fo the week has the best avg ratings?

```
select day_name,round(avg(Rating),1)as 'Avg_Rating' from walmart
group by day_name
order by `Avg_Rating` desc;
```

Result Grid | Filter Ro

day_name	Avg_Rating
Friday	7.1
Monday	7.1
Tuesday	7
Sunday	7
Thursday	6.9
Saturday	6.9
Wednesday	6.8

10-Which day of the week has the best average ratings per branch?

```
select day_name,Branch,round(avg(Rating),2)as 'Avg_Rating' from walmart
group by day_name,Branch
order by day_name;
```

Result Grid | Filter Rows:

day_name	Branch	Avg_Rating
Friday	A	7.31
Friday	B	6.69
Friday	C	7.21
Monday	A	7.1
Monday	B	7.27
Monday	C	7.04
Saturday	A	6.75
Saturday	B	6.74
Saturday	C	7.23
Sunday	A	7.08
Sunday	B	6.8
Sunday	C	7.03
Thursday	A	6.96
Thursday	B	6.75
Thursday	C	6.95
Tuesday	A	7.06
Tuesday	B	7