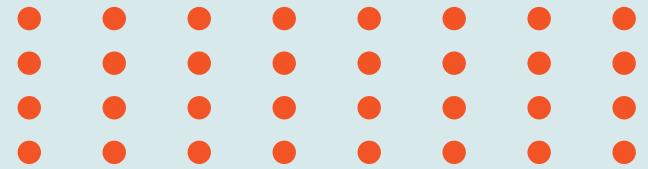




Hello I am Nishant
this is sales report

AMAZON SALES REPORT OF Q1





To gain insight into the sales data of Amazon to understand the different factors that affect sales of the different branches.

Introduction



Amazon

1.What is the count of distinct cities in the dataset?

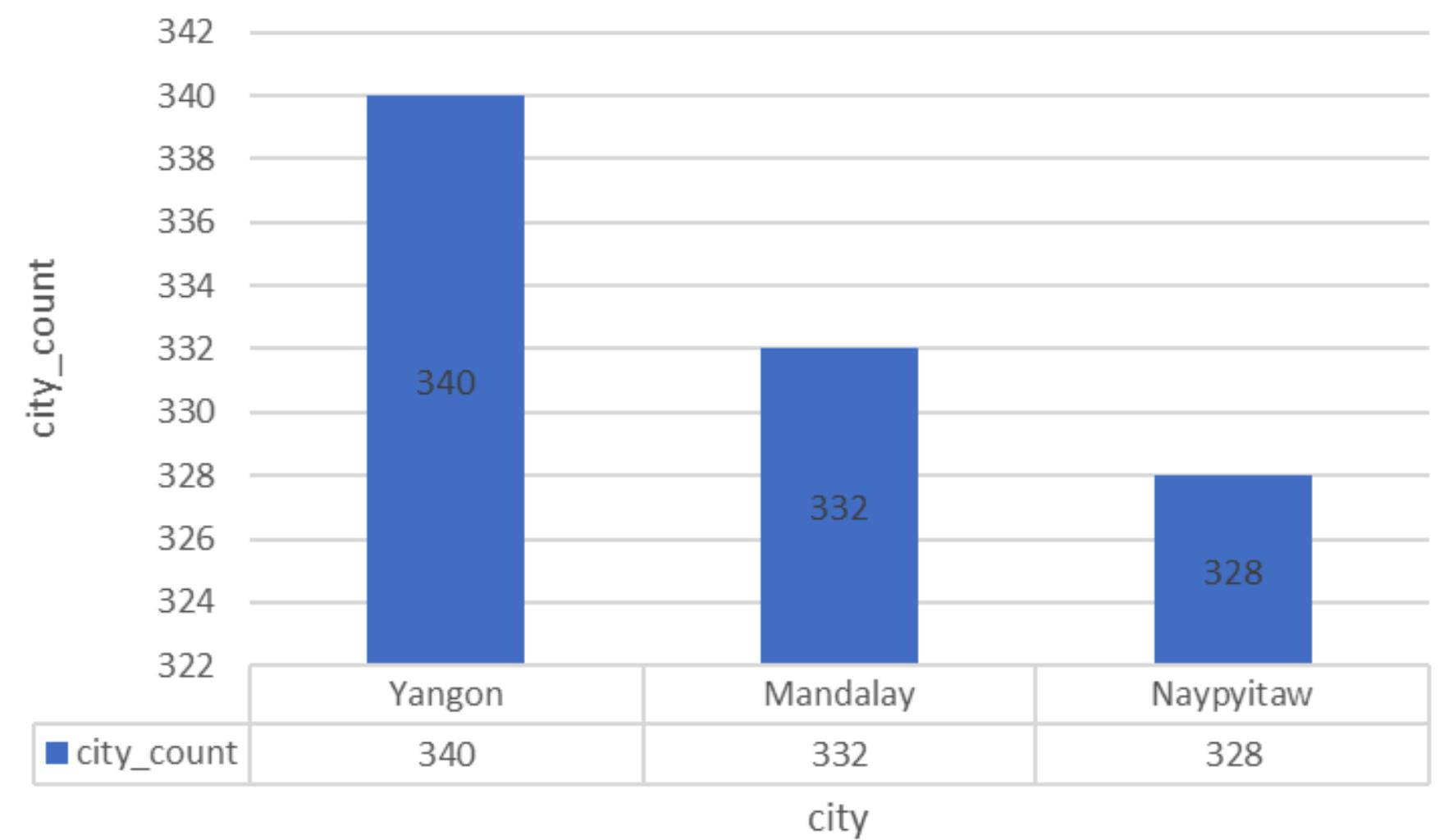
Query

```
SELECT DISTINCT  
    city, COUNT(city) AS city_count  
FROM  
    amazon  
GROUP BY city  
ORDER BY city_count DESC;
```

	city	city_count
▶	Yangon	340
	Mandalay	332
	Naypyitaw	328



Output



2. For each branch, what is the corresponding city?

Query

```
SELECT DISTINCT  
    branch, city  
FROM  
    amazon;
```



Output

	branch	city
▶	A	Yangon
	C	Naypyitaw
	B	Mandalay

3.What is the count of distinct product lines in the dataset?

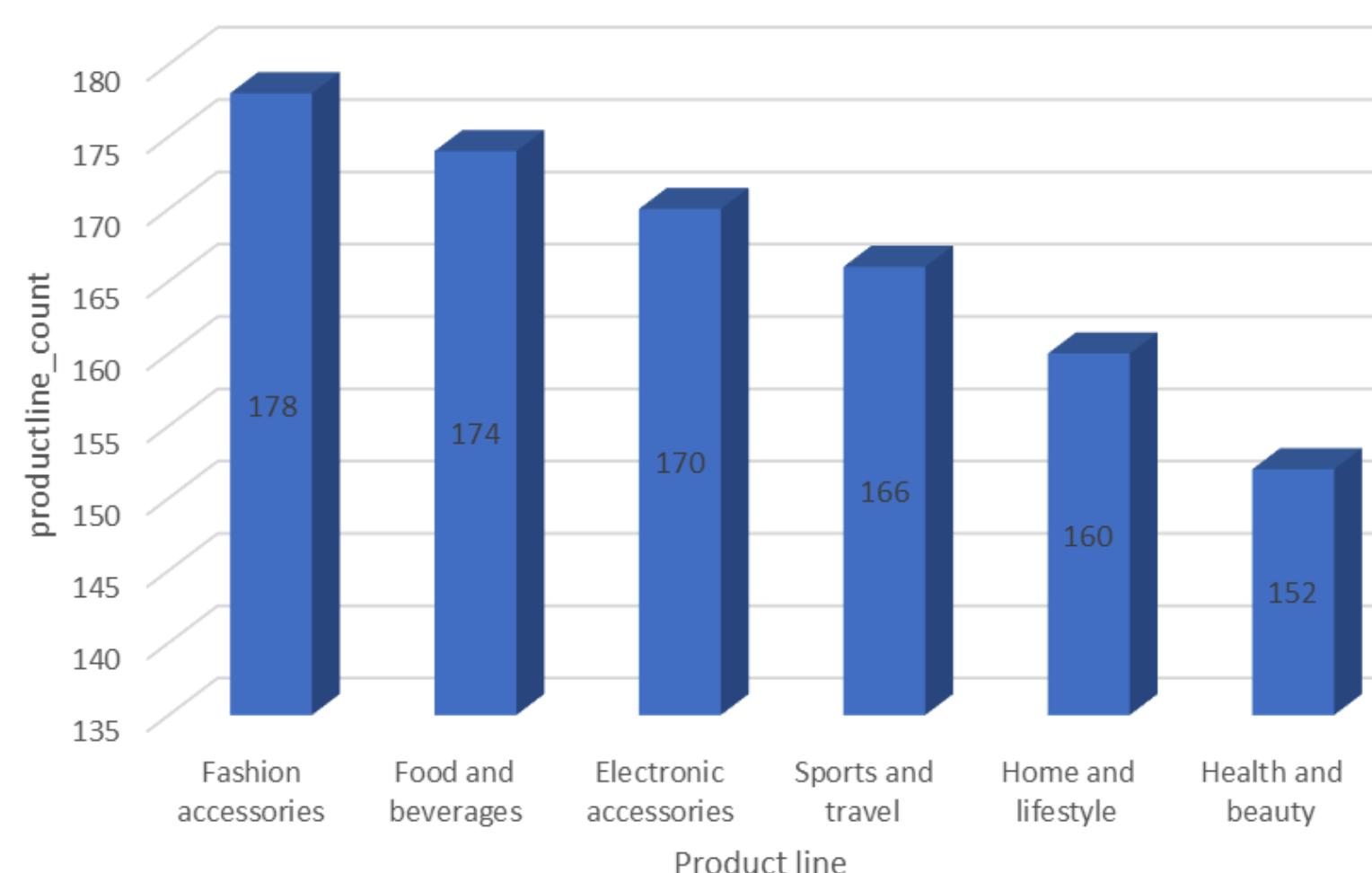
Query

```
SELECT DISTINCT
    `Product line`, COUNT(`Product line`) AS productline_count
FROM
    amazon
GROUP BY `Product line`
ORDER BY productline_count DESC;
```

Product line	productline_count
Fashion accessories	178
Food and beverages	174
Electronic accessories	170
Sports and travel	166
Home and lifestyle	160
Health and beauty	152



Output



4.Which payment method occurs most frequently?

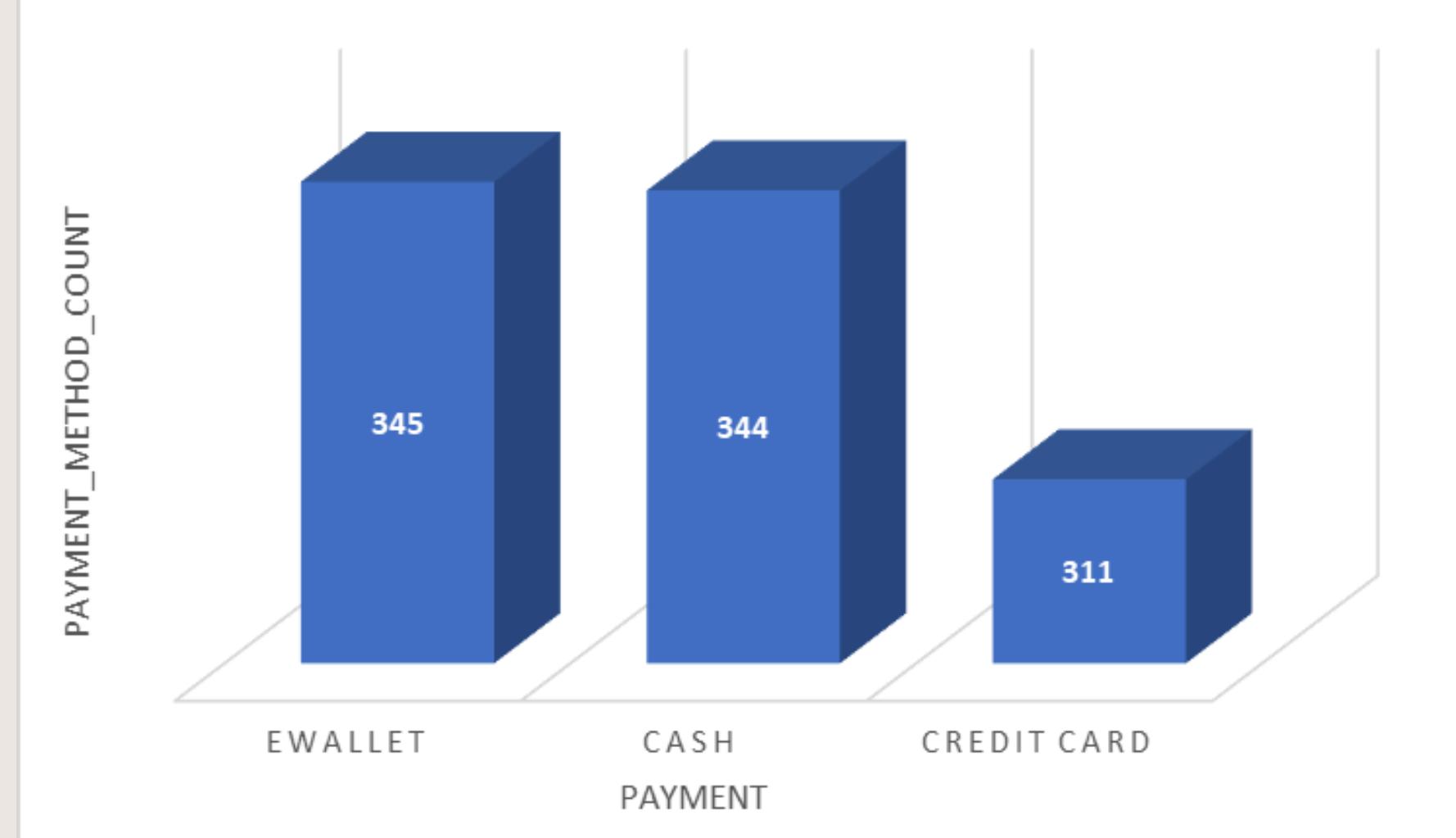
Query

```
SELECT  
    Payment, COUNT(Payment) AS payment_method_count  
FROM  
    amazon  
GROUP BY Payment  
ORDER BY payment_method_count DESC;
```

Payment	payment_method_count
Ewallet	345
Cash	344
Credit card	311



Output



5.Which product line has the highest sales?

Query

```
SELECT
    `product line`, SUM(total) AS total_sales
FROM
    amazon
GROUP BY `product line`
ORDER BY total_sales DESC
LIMIT 1;
```



Output

product line	total_sales
Food and beverages	56144.84400000005

6.How much revenue is generated each month?

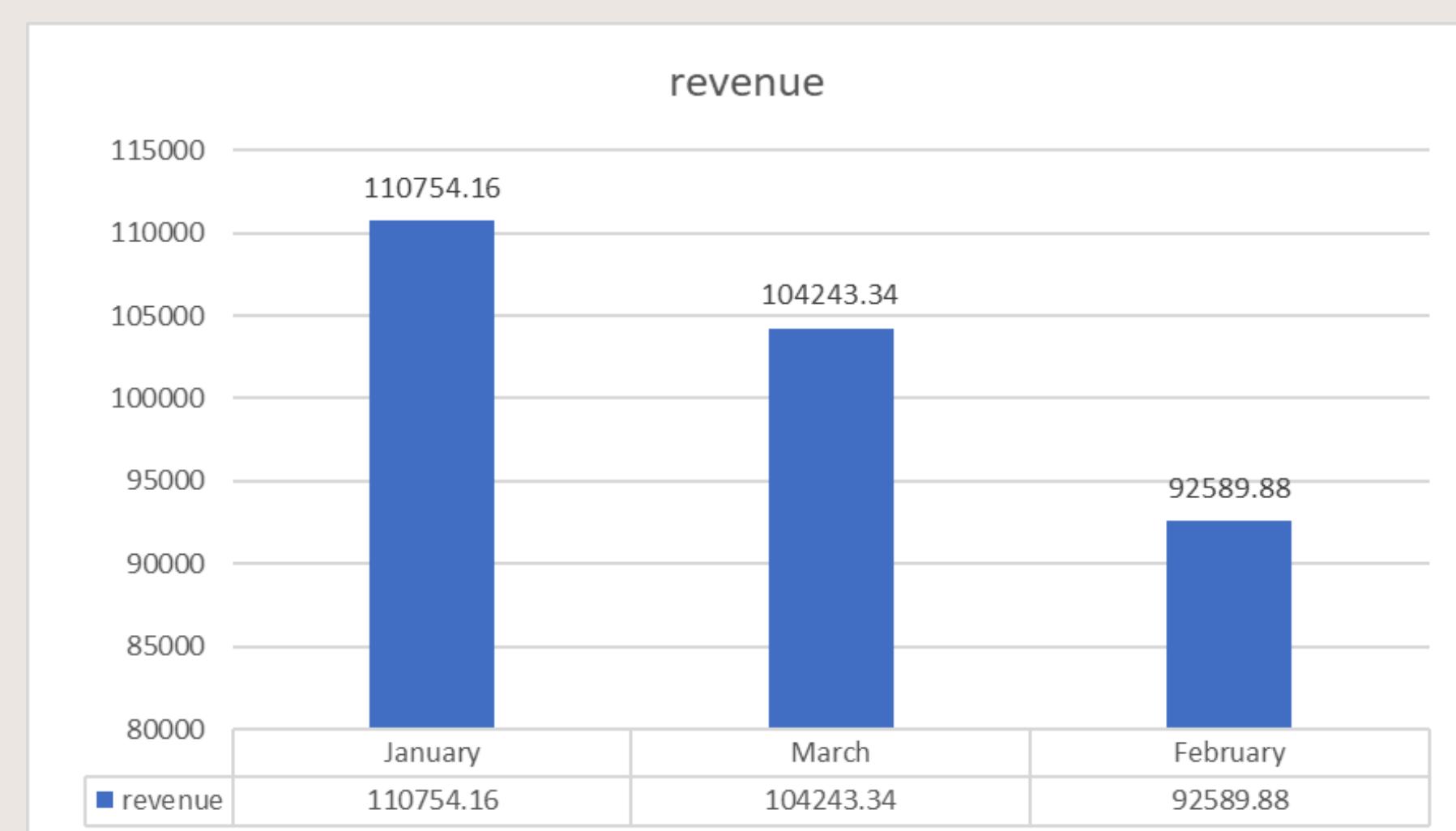
Query

```
SELECT  
    monthname, ROUND(SUM(quantity * `unit price`), 2) AS revenue  
FROM  
    amazon  
GROUP BY Monthname  
ORDER BY revenue DESC;
```

monthname	revenue
January	110754.16
March	104243.34
February	92589.88



Output



7.In which month did the cost of goods sold reach its peak?

Query

```
SELECT  
    monthname, SUM(cogs) AS total_cogs  
FROM  
    amazon  
GROUP BY monthname  
ORDER BY total_cogs DESC  
LIMIT 1;
```



Output

monthname	total_cogs
January	110754.1600000002

8.Which product line generated the highest revenue?

Query

```
SELECT
    `product line`,
    ROUND(SUM(quantity * `unit price`), 2) AS revenue
FROM
    amazon
GROUP BY `product line`
ORDER BY revenue DESC
LIMIT 1;
```



Output

product line	revenue
Food and beverages	53471.28

9.In which city was the highest revenue recorded?

Query

```
SELECT
    city, SUM(quantity * `unit price`) AS revenue
FROM
    amazon
GROUP BY city
ORDER BY revenue DESC
LIMIT 1;
```



Output

city	revenue
Naypyitaw	105303.53

10.Which product line incurred the highest Value Added Tax?

Query

```
SELECT  
    `Product line`, ROUND(SUM(`tax %`), 2) AS tax  
FROM  
    amazon  
GROUP BY `Product line`  
ORDER BY tax DESC  
LIMIT 1;
```



Output

Product line	tax
Food and beverages	2673.56

11. For each product line, add a column indicating "Good" if its sales are above average, otherwise "Bad."

Query

```
SELECT
    `Product line`,
    SUM(Total) AS total_sales,
    IF(SUM(Total) > (
        SELECT AVG(total_sales)
        FROM
            (SELECT
                `Product line`, SUM(Total) AS total_sales
                FROM
                    amazon
                    GROUP BY `Product line`) AS subquery),
        'good',
        'Bad') AS performance
FROM
    amazon
GROUP BY `Product line`
order by performance desc ;
```



Output

	Product line	total_sales	performance
▶	Electronic accessories	54337.531500000005	good
	Home and lifestyle	53861.91300000001	good
	Sports and travel	55122.82649999996	good
	Food and beverages	56144.84400000005	good
	Fashion accessories	54305.895	good
	Health and beauty	49193.739000000016	Bad

12. Identify the branch that exceeded the average number of products sold

Query

```
SELECT  
    branch, SUM(Quantity) AS total_quantity  
FROM  
    amazon  
GROUP BY branch  
HAVING SUM(Quantity) > (SELECT  
    AVG(total_quantity)  
    FROM  
        (SELECT  
            SUM(Quantity) AS total_quantity  
        FROM  
            amazon  
        GROUP BY branch) AS subquery);
```



Output

	branch	total_quantity
▶	A	1859

13. Which product line is most frequently associated with each gender

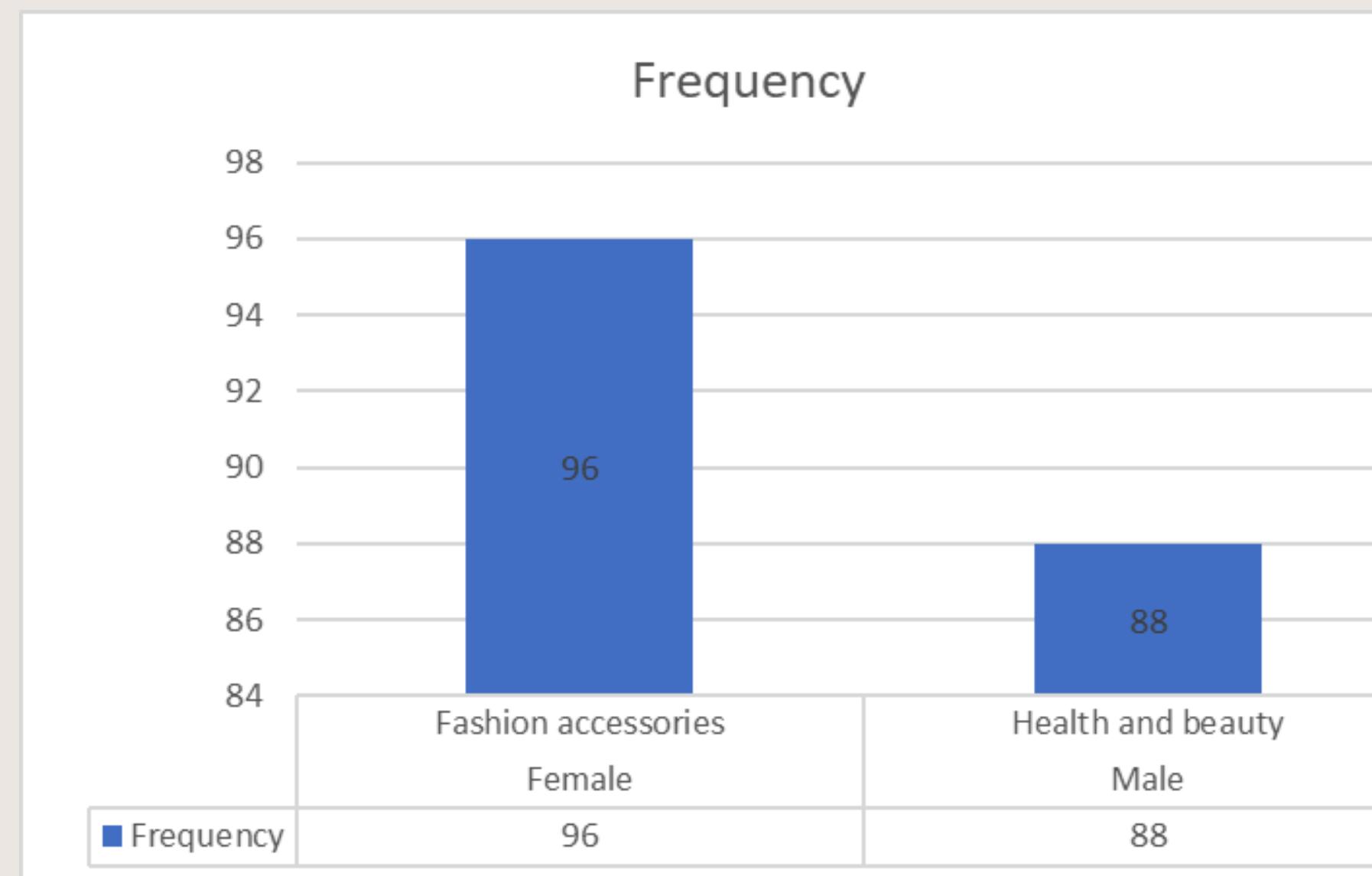
Query

```
WITH GenderProductFrequency AS (
    SELECT
        Gender,
        `Product line`,
        COUNT(*) AS Frequency,
        ROW_NUMBER() OVER (PARTITION BY Gender ORDER BY COUNT(*) DESC) AS Rank_
    FROM amazon
    GROUP BY Gender, `Product line`
)
SELECT Gender, `Product line`, Frequency
FROM GenderProductFrequency
WHERE Rank_ = 1;
```

	Gender	Product line	Frequency
▶	Female	Fashion accessories	96
	Male	Health and beauty	88



Output



14. Calculate the average rating for each product line

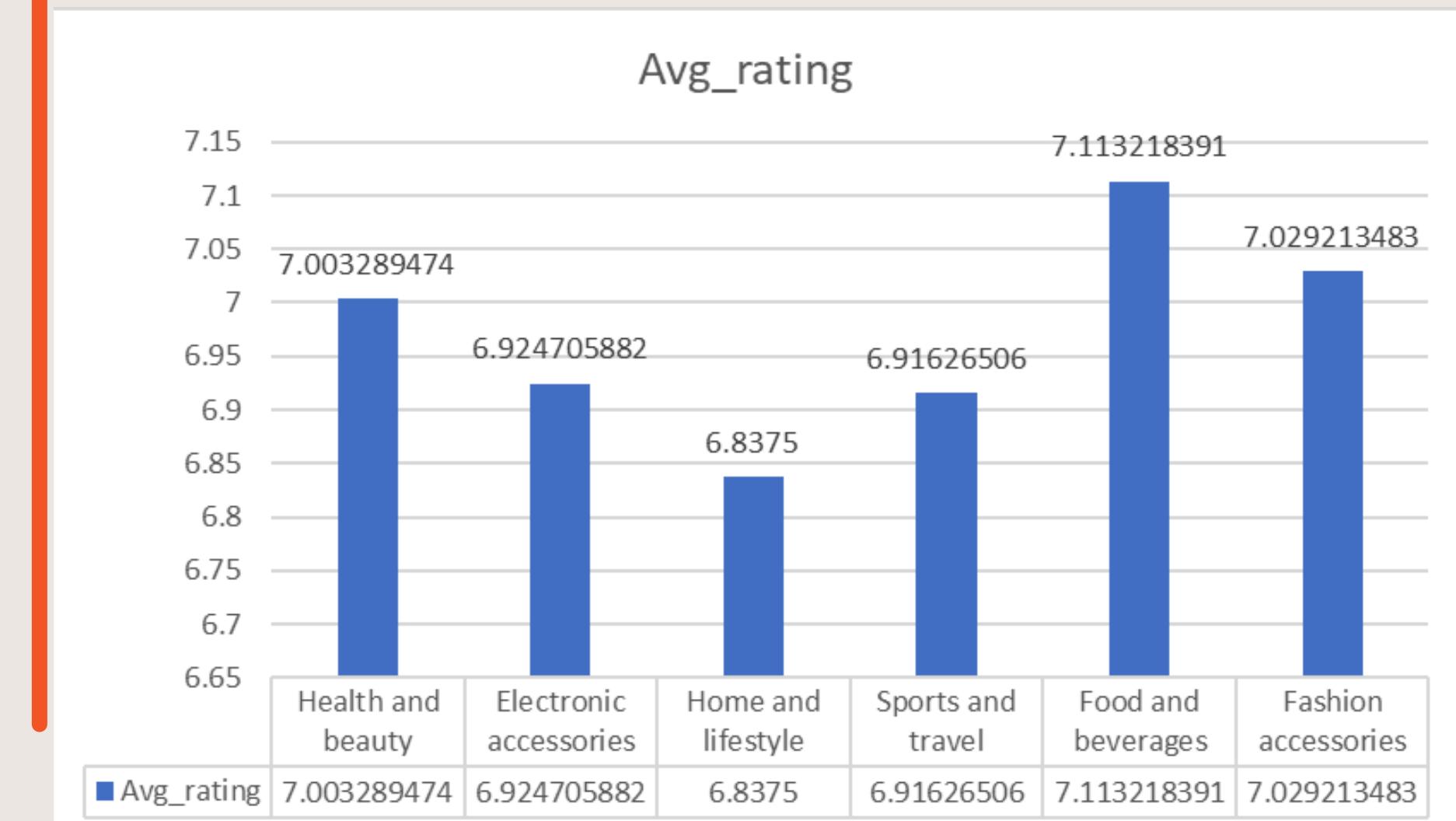
Query

```
select `Product line`, avg(rating) as Avg_rating from amazon  
group by `Product line`
```



Output

Product line	Avg_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



15. Count the sales occurrences for each time of day on every weekday.

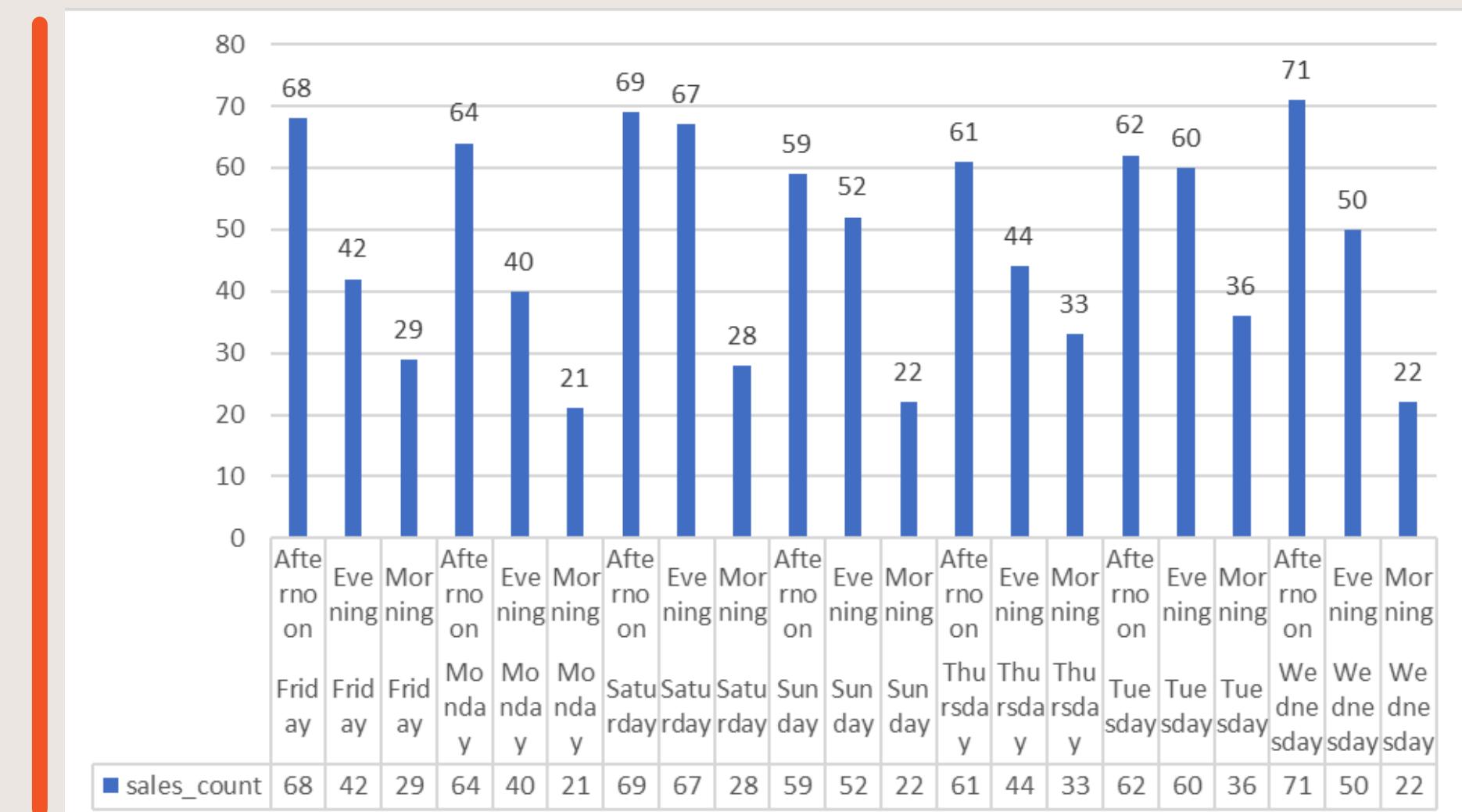
Query

```
SELECT  
    Dayname, Timeofday, COUNT(*) AS sales_count  
FROM  
    amazon  
GROUP BY Dayname , Timeofday  
ORDER BY Dayname , Timeofday;
```

	Dayname	Timeofday	sales_count
▶	Friday	Afternoon	68
	Friday	Evening	42
	Friday	Morning	29
	Monday	Afternoon	64
	Monday	Evening	40
	Monday	Morning	21
	Saturday	Afternoon	69
	Saturday	Evening	67
	Saturday	Morning	28
	Sunday	Afternoon	59
	Sunday	Evening	52
	Sunday	Morning	22
	Thursday	Afternoon	61
	Thursday	Evening	44
	Thursday	Morning	33
	Tuesday	Afternoon	62
	Tuesday	Evening	60
	Tuesday	Morning	36
	Wednesday	Afternoon	71
	Wednesday	Evening	50
	Wednesday	Morning	22



Output



16. Identify the customer type contributing the highest revenue.

Query

```
SELECT  
    `Customer type`, SUM(`Unit price` * Quantity) AS revenue  
FROM  
    amazon  
GROUP BY `Customer type`  
ORDER BY revenue DESC  
LIMIT 1;
```



Output

	Customer type	revenue
▶	Member	156403.2799999985

17. Determine the city with the highest VAT percentage

Query

```
SELECT
    City, SUM(`Tax 5%`) AS Total_VAT
FROM
    amazon
GROUP BY City
ORDER BY Total_VAT DESC
LIMIT 1;
```



Output

City	Total_VAT
Naypyitaw	5265.176500000002

18.Identify the customer type with the highest VAT payments.

Query

```
SELECT  
    `Customer type`, SUM(`Tax 5%`) AS VAT  
FROM  
    amazon  
GROUP BY `Customer type`  
ORDER BY vat DESC  
LIMIT 1;
```



Output

Customer type	VAT
Member	7820.164000000002

19.What is the count of distinct customer types in the dataset?

Query

```
SELECT  
    COUNT(DISTINCT `Customer type`) AS customer_types  
FROM  
    amazon
```



Output

	customer_types
→	2

20.What is the count of distinct payment methods in the dataset?

Query

```
SELECT  
    COUNT(DISTINCT payment) AS payment_methods  
FROM  
    amazon;
```



Output

	payment_methods
3	

21. Which customer type occurs most frequently?

Query

```
SELECT  
    `Customer type`, COUNT(*) AS frequency  
FROM  
    amazon  
GROUP BY `Customer type`  
ORDER BY frequency DESC  
LIMIT 1;
```



Output

Customer type	frequency
Member	501

22. Identify the customer type with the highest purchase frequency.

Query

```
SELECT  
    `Customer type`, COUNT(*) AS purchase  
FROM  
    amazon  
GROUP BY `Customer type`  
ORDER BY purchase DESC  
LIMIT 1;
```



Output

Customer type	purchase
Member	501

23. Determine the predominant gender among customers.

Query

```
SELECT  
    gender, COUNT(`Customer type`) AS customer  
FROM  
    amazon  
GROUP BY gender  
ORDER BY customer DESC  
LIMIT 1;
```



Output

	gender	customer
▶	Female	501

24.Examine the distribution of genders within each branch.

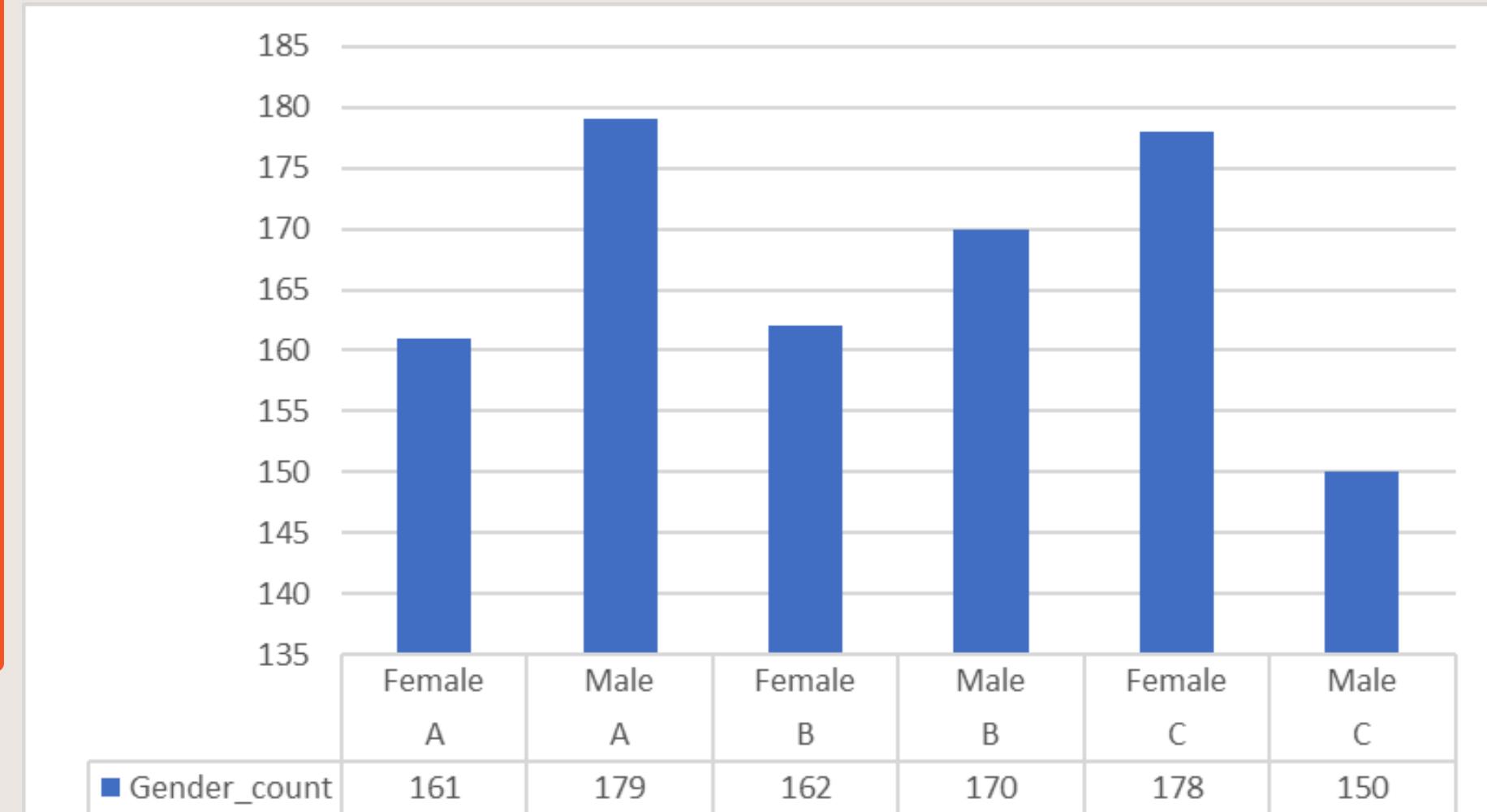
Query

```
• SELECT  
    branch, gender, COUNT(*) AS Gender_count  
FROM  
    amazon  
GROUP BY  
    branch , gender  
ORDER BY  
    branch , gender;
```

branch	gender	Gender_count
A	Female	161
A	Male	179
B	Female	162
B	Male	170
C	Female	178
C	Male	150



Output



25. Identify the time of day when customers provide the most ratings.

Query

```
SELECT  
    timeofday, COUNT(rating) AS rating  
FROM  
    amazon  
GROUP BY timeofday  
ORDER BY rating DESC  
LIMIT 1;
```



Output

	timeofday	rating
→	Afternoon	454

26. Determine the time of day with the highest customer ratings for each branch.

Query

```
with highrating as (
  select `Customer type`, timeofday ,Branch, sum(Rating) as ratings
  from amazon
  group by Branch,`Customer type`, timeofday
  order by ratings
)
select `Customer type`, timeofday ,Branch ,ratings
from highrating as hr
where ratings = (
  select max(ratings)
  from highrating
  where branch = hr.branch)
order by branch ;
```



Output

Customer type	timeofday	Branch	ratings
Member	Afternoon	A	566.8000000000003
Member	Afternoon	B	524.3
Member	Afternoon	C	574.5

27. Identify the day of the week with the highest average ratings.

Query

```
SELECT  
    Dayname, ROUND(AVG(rating), 2) AS avg_rating  
FROM  
    amazon  
GROUP BY Dayname  
ORDER BY avg_rating DESC  
LIMIT 1;
```



Output

Dayname	avg_rating
Monday	7.15

28. Determine the day of the week with the highest average ratings for each branch.

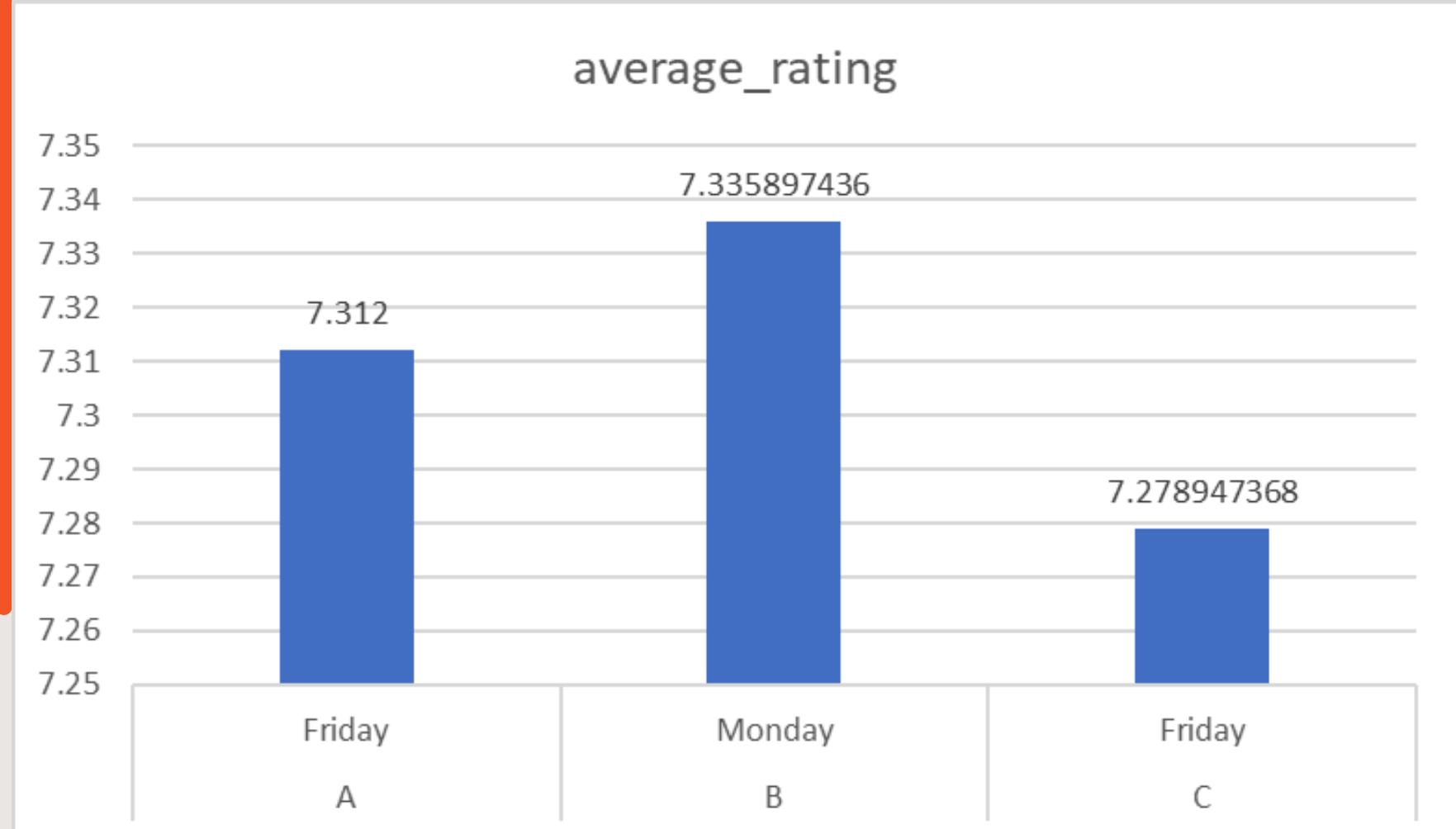
Query

```
WITH AvgRatings AS (
    SELECT
        Branch,
        Dayname,
        AVG(Rating) AS average_rating
    FROM
        amazon
    GROUP BY
        Branch, Dayname
)
SELECT
    Branch,
    Dayname,
    average_rating
FROM
    AvgRatings AS ar
WHERE
    average_rating = (
        SELECT MAX(average_rating)
        FROM AvgRatings
        WHERE Branch = ar.Branch
    )
ORDER BY
    Branch;
```

Branch	Dayname	average_rating
A	Friday	7.3119999999999985
B	Monday	7.335897435897434
C	Friday	7.278947368421051



Output



1. Product Analysis

- There are total six product line Food and beverages, Fashion accessories, Health and beauty, Electronic accessories, Sports and travel, Home and lifestyle
- Among this 6 product line Food and beverages, Fashion accessories, Health and beauty are the top 3 best performing. This product line has generated most number sales.
- Sports and travel, Home and lifestyle are the product line which has the most low rating among all the product line
- Electronic accessories, Sports and travel, Home and lifestyle this are the bottom 3 product . among this product Sports and travel, Home and lifestyle needs to improve the sales
- Food and beverages, Sports and travel are the most profitable product in product line ,
- Health and beauty ,Home and Lifestyle are the product in product line which have generate list amount of profit

1. Sales Analysis

- Their are sales data of three month January, March, February from this January has generate the highest sales and February has generated lowest sales
- In the weekly sales saturday is the has generated most amount of sales from which we can concluded people purchase more on holiday .
- Monday , Wednesday and Friday has generated list amount of sales
- Their are the payment methods cash, Ewallet and credit card among this the most use payment method is cash and the second is Ewallet credit card is the one which is been use the list times to pay
- Their are 3 branch A, B, C and from this C branch has generate most amount of sale and branch A is at second and the branch B is the branch which has generate list sale
- Their are 3 branch A, B, C and from this C branch has generate most amount of sale and branch A is at second and the branch B is the branch which has generate list sale

1. Customer Analysis

- In this data base there are two customer type member and normal
- The customer type which has paid most amount of VAT is normal
- The customer with Highest purchase frequency is Member
- Maximum Number of customer are female
- Afternoon is the day time when customer provide most numbers of rating .
- For each branch member customer type that provide most number of eating and at afternoon time
- For each branch member customer type that provide most number of eating and at afternoon time
- For branch A the highest average rating is on Friday and for branch B it is on Monday and for branch C it on Friday again



Thank You

