

AMAZON SALES ANALYSIS - S10112

- ARIF SHAIK

OBJECTIVE OF PROJECT:

The primary objective of this project is to analyze Amazon's sales data to gain insights into the various factors influencing sales across different branches.

OVERVIEW OF PROJECT:

The dataset comprises sales records from three branches in Myanmar: Naypyitaw, Yangon, and Mandalay, covering the first quarter of 2019. It contains a total of 1,000 rows and 17 columns.

APPROACH USED:

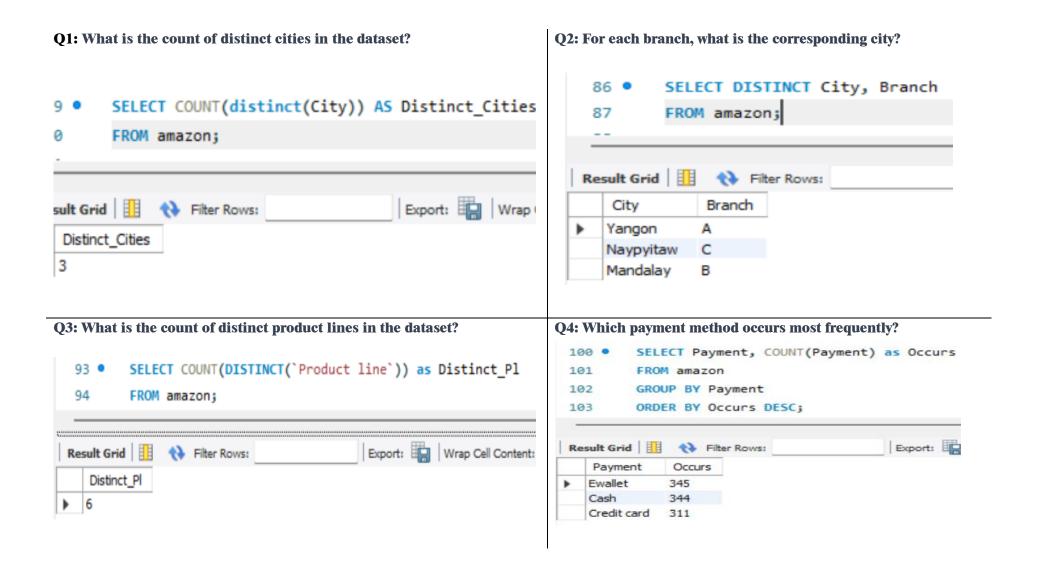
DATA WRANGLING:

This initial step involves inspecting the data to identify NULL and missing values. Appropriate data replacement methods are then applied to address these issues, ensuring the dataset is clean and ready for analysis.

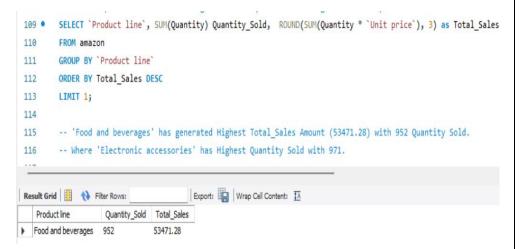
FEATURE ENGINEERING:

This process involves creating new columns derived from existing ones, enhancing the dataset's richness and enabling more insightful analysis.

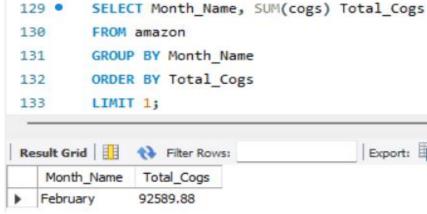
ANSWERING BUSINESS QUESTIONS



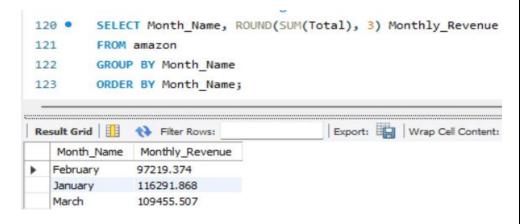
Q5: Which product line has the highest sales?



Q7: In which month did the cost of goods sold reach its peak?

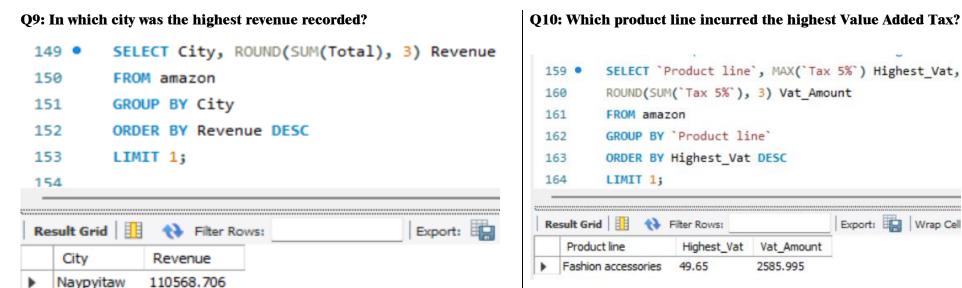


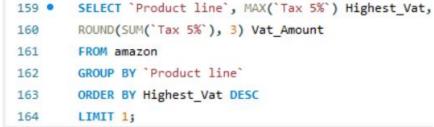
Q6: How much revenue is generated each month?



Q8: Which product line generated the highest revenue?

	Product line Food and beverages		Revenue				
Re	esult Gric	1 () F	ilter Rows:	1	Export:	Wrap	Cell Content
14	13	LIMIT 1;					
14	12	ORDER BY	Revenue DESC				
14	11	GROUP BY	Product line				
14	10	FROM amazo	on				
13	39 •	SELECT P	roduct line`,	ROUND (SUM	Total),	3) as	Revenue
		reconstruction grant		AND TO SERVICE		= 12	



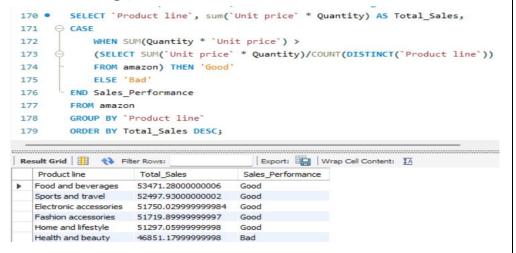


Highest Vat Vat Amount

2585.995

Export:

Q11: For each product line, add a column indicating "Good" if its sales are above average, otherwise "Bad."



Q12: Identify the branch that exceeded the average number of products sold.

Filter Rows:

49.65

Result Grid

Product line

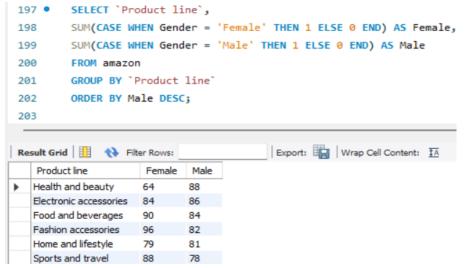
Fashion accessories

```
SELECT Branch, SUM(Quantity) as Total Units Sold
186
187
         FROM amazon
 188
         GROUP BY Branch
         HAVING Total Units Sold >
 189

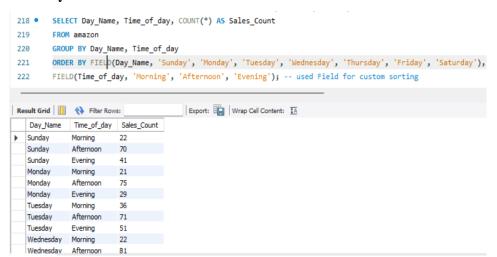
⊖ (SELECT SUM(Quantity) / COUNT(DISTINCT(Branch)) AS Avg Quantity

 190
         FROM amazon);
 191
                                       Export: Wrap Cell Content: TA
Total Units Sold
   Branch
▶ A
          1859
```

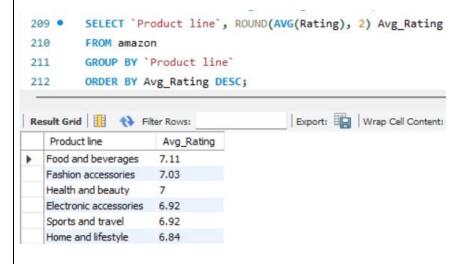
Q13: Which product line is most frequently associated with each gender?



Q15: Count the sales occurrences for each time of day on every weekday.

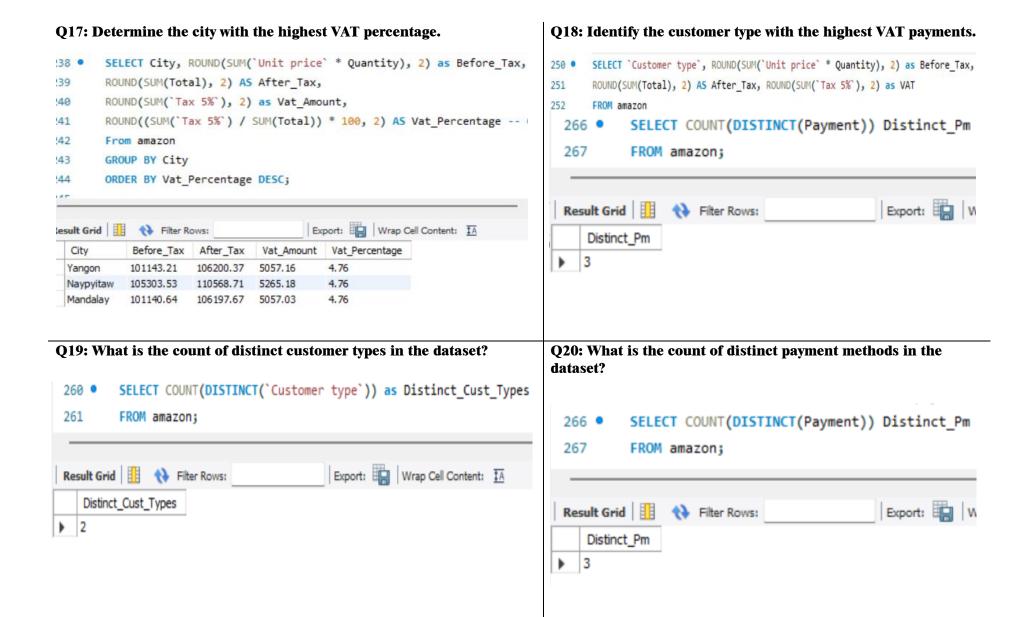


Q14: Calculate the average rating for each product line.

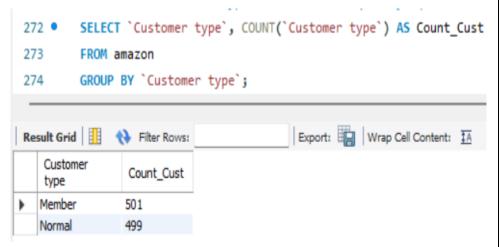


Q16: Identify the customer type contributing the highest revenue.

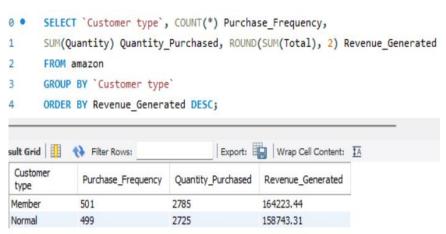
```
229
          select `Customer type`, ROUND(SUM(Total), 3) as Revenue
230
          FROM amazon
231
          GROUP BY 'Customer type'
232
          ORDER BY Revenue DESC;
 233
Result Grid
               Filter Rows:
                                             Export: Wrap Cell Content:
    Customer
                 Revenue
    type
Member
                 164223,444
   Normal
                 158743,305
```







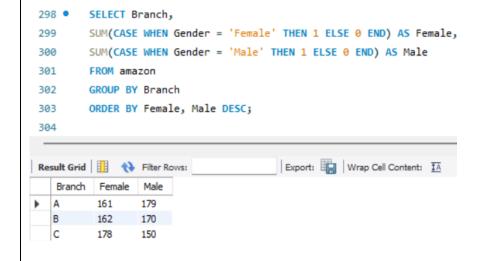
Q22: Identify the customer type with the highest purchase frequency.



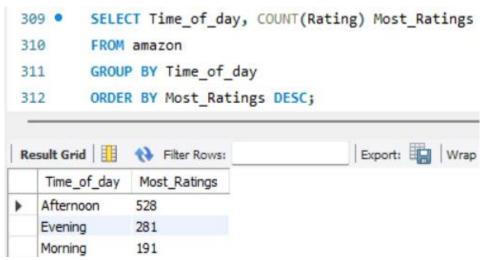
Q23: Determine the predominant gender among customers.



Q24: Examine the distribution of genders within each branch.



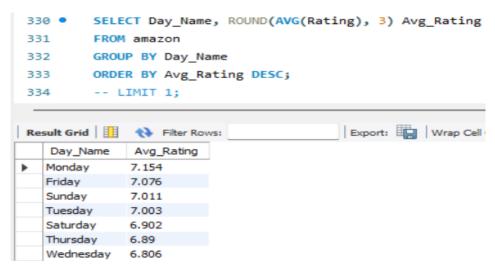
Q25: Identify the time of day when customers provide the most ratings.



Q26: Determine the time of day with the highest customer ratings for each branch.

```
SELECT Branch, Time_of_day, Most_Ratings
     319
           SELECT Branch, Time of day, COUNT(Rating) AS Most Ratings,
320
                  ROW NUMBER() OVER (PARTITION BY Branch ORDER BY COUNT(Rating) DESC) AS Rating
321
           FROM amazon
           GROUP BY Branch, Time_of_day
322
323
        ) AS ranked
        WHERE Rating = 1;
Export: Wrap Cell Content: 1A
  Branch
         Time of day Most Ratings
         Afternoon
                   162
  C
                   181
         Afternoon
```

Q27: Identify the day of the week with the highest average ratings.



Q28: Determine the day of the week with the highest average ratings for each branch.

```
SELECT Day_Name, Branch, Avg_Rating FROM
       (SELECT Branch, Day Name, ROUND(AVG(Rating), 3) AS Avg Rating,
42
       ROW NUMBER() OVER(PARTITION BY Branch ORDER BY AVG(Rating) DESC) AS High Rating
43
       FROM amazon
44
       GROUP BY Branch, Day Name
45
       ORDER BY Avg_Rating DESC) AS bb
       WHERE High_Rating = 1;
esult Grid
           Filter Rows:
                                        Export: Wrap Cell Content: IA
 Day_Name Branch Avg_Rating
 Monday
                  7.336
                  7.312
 Friday
Friday
                  7.279
```

INSIGHTS

PRODUCT ANALYSIS:

HIGHEST REVENUE PRODUCT LINE: Food and Beverages Generating Highest Revenue (\$56144.96)

LOWEST REVENUE PRODUCT LINE: Health and beauty Generating Lowest Revenue (\$49193.84)

HIGHEST SALES PRODUCT LINE: Electronic Accessories (Units Sold:971)

LOWEST SALES PRODUCT LINE: Health and Beauty (units Sold: 854)

SALES ANALYSIS:

MONTH WITH HIGHEST REVENUE: January has Generated Highest Revenue (\$ 116292.11)

MONTH WITH LOWEST REVENUE: February has Generated Lowest Revenue (\$ 97219.58)

CITY WITH HIGHEST REVENUE: Naypyitaw has Generated Highest Revenue (\$ 110568.86)

CITY WITH LOWEST REVENUE: Mandalay has Generated Lowest Revenue (\$ 106198.00)

PEAK SALES TIME OF DAY: Afternoon has recorded Highest Sales while Morning with Lowest Sales

PEAK SALES DAY OF WEEK: Saturday has recorded Highest Sales

CUSTOMER ANALYSIS:

PREDOMINANT GENDER: Females contributed more to Revenue, although difference between contribution is not much.

PREDOMINANT CUSTOMER TYPE: Member

HIGHEST REVENUE GENDER: Female (\$ 167883.26)

HIGHEST REVENUE CUSTOMER TYPE: Member (\$ 164223.81)

MOST POPULAR PRODUCT LINE (MALE): Health and Beauty

MOST POPULAR PRODUCT LINE (FEMALE): Fashion and Accessories

RECOMMENDATIONS:

- Health and Beauty products underperform across metrics and require a comprehensive plan to enhance the product line.
- Leverage the high sales and revenue in January by offering diverse customer options and implementing effective, targeted strategies.
- Prioritize new product launches or campaigns during the peak afternoon sales hours.
- Develop a plan to increase memberships, as member customers contribute more, potentially due to incentives or offers. This can reduce Acquisition Cost and drive higher revenue.

