



Online Sales report



Divyansh Dinesh Pandey

2024

Introduction

To complete this project, I applied SQL querying skills, including writing queries with SELECT, GROUP BY, ORDER BY, and aggregate functions like SUM, AVG, and COUNT. Formatting outputs with functions like ROUND() ensured precision. Data analysis was key to extracting insights into customer behavior, sales trends, and product performance.

Tools like MySQL and MySQL Workbench helped manage the database, organize scripts, and execute queries efficiently. My analytical thinking enabled me to break tasks into smaller steps, identify patterns, and solve problems effectively. Combining technical skills and critical thinking, I transformed raw online shopping data into actionable insights.



Analysing Online Shopping Data Using CSV Import and MySQL

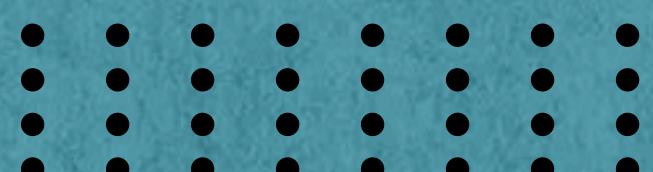
- Project Start:
 - Used a CSV file containing online shopping data (customer demographics, purchase details, product categories, payment methods).
 - Imported the CSV into MySQL using the import feature.
 - Ensured the data was structured with proper column names and data types for efficient querying.
- Tools Used:
 - Utilized MySQL Workbench for database visualization, schema verification, and query writing.

SQL Querying and Analysis:

- Applied SQL querying techniques like SELECT, GROUP BY, ORDER BY, and aggregate functions (SUM, AVG, COUNT).
- Used ROUND() function to format outputs for clarity.

Task Breakdown:

- Broke the project into smaller tasks to analyze customer behavior, sales performance, and product trends.
- Transformed raw CSV data into actionable insights, reflecting real business scenarios.



1.1 Customer Analysis

How many unique customers are in the dataset?



```
• SELECT COUNT(DISTINCT `Customer ID`) AS Unique_Customers  
FROM `project`.`shopping_trends`;
```

	Unique_Customers
▶	3900

This query counts the number of unique customers in the shopping_data table by counting distinct Customer ID values. The result is labeled as Unique_Customers.

1.2 Customer Analysis

What are the top 5 locations with the most customers?



- ```
SELECT Location, COUNT(`Customer ID`) AS Customer_Count
FROM project.shopping_trends
GROUP BY Location
ORDER BY Customer_Count DESC
LIMIT 5;
```

|   | Location   | Customer_Count |
|---|------------|----------------|
| ▶ | Montana    | 96             |
|   | Illinois   | 92             |
|   | Idaho      | 93             |
|   | California | 95             |
|   | Alabama    | 89             |

This query retrieves the top 5 locations with the highest number of customers from the shopping\_data table. It groups the data by Location, counts the number of customers in each location, orders the results in descending order by customer count, and limits the output to the top 5 locations.

# 1.3 Customer Analysis

How many orders has each customer placed?



```
• SELECT `Customer ID`, COUNT(*) AS Total_Orders
FROM project.shopping_trends
GROUP BY `Customer ID`
ORDER BY Total_Orders DESC;
```

|   | Customer ID | Total_Orders |
|---|-------------|--------------|
| 1 | 1           | 1            |
| 2 | 1           | 1            |
| 3 | 1           | 1            |
| 4 | 1           | 1            |
| 5 | 1           | 1            |
| 6 | 1           | 1            |
| 7 | 1           | 1            |

This query counts the total number of orders placed by each customer in the shopping\_data table. It groups the data by Customer ID, counts the orders for each customer using COUNT(\*), and orders the results in descending order by the total number of orders.

If the outcome is always 1, it indicates that each customer has placed only one order, as COUNT(\*) is counting each row (order) for the customer.

# 1.4 Customer

## Analysis

Who are the top 5 customers based on total spending?



- ```
SELECT `Customer ID`, SUM(`Purchase Amount (USD)`) AS Total_Spending
FROM project.shopping_trends
GROUP BY `Customer ID`
ORDER BY Total_Spending DESC
LIMIT 5;
```

Customer ID	Total_Spending
205	100
43	100
96	100
194	100
244	100

This query calculates the total spending of each customer by summing up their Purchase Amount (USD) in the shopping_data table. It groups the data by Customer ID, orders the results in descending order by total spending, and limits the output to the top 5 customers with the highest total spending.

1.5 Customer

Analysis

What is the average purchase amount per customer?

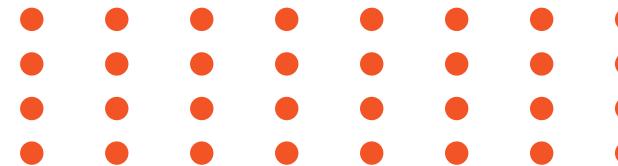


```
• SELECT `Customer ID`, ROUND(AVG(`Purchase Amount (USD)`), 2) AS Avg_Purchase_Amount  
FROM project.shopping_trends  
GROUP BY `Customer ID`  
LIMIT 5;
```

	Customer ID	Avg_Purchase_Amount
▶	1	53.00
	2	64.00
	3	73.00
	4	90.00
	5	49.00

This query calculates the average purchase amount for each customer by dividing the total Purchase Amount (USD) by the number of purchases. It groups the data by Customer ID and returns the average purchase amount for each customer in the shopping_data table.

2.1 Product Analysis



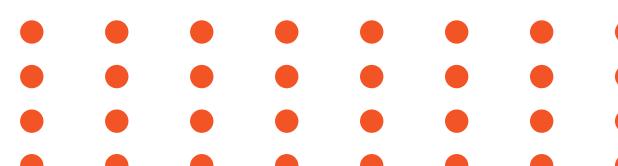
What are the top 3 best-selling categories based on total sales?



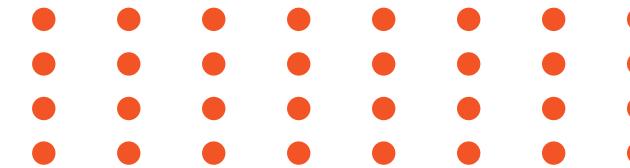
This query calculates the total sales for each product category by summing the Purchase Amount (USD) for each category in the shopping_data table. It groups the data by Category, orders the results in descending order by total sales, and limits the output to the top 3 categories with the highest total sales.

```
• SELECT Category, SUM(`Purchase Amount (USD)`) AS Total_Sales  
FROM project.shopping_trends  
GROUP BY Category  
ORDER BY Total_Sales DESC  
LIMIT 5;
```

	Category	Total_Sales
▶	Clothing	104264
	Accessories	74200
	Footwear	36093
	Outerwear	18524



2.2 Product Analysis



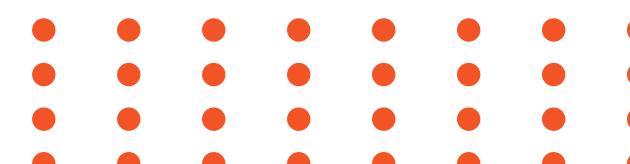
What is the average price of items in each category?



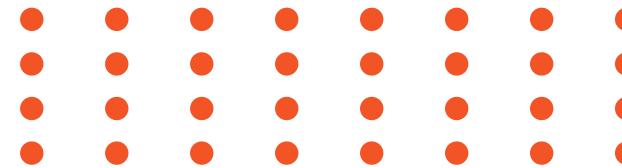
This query calculates the average purchase amount (price) for items in each product category by averaging the Purchase Amount (USD) values. It groups the data by Category and returns the average price for each category.

```
• SELECT Category, AVG(`Purchase Amount (USD)`) AS Avg_Item_Price  
FROM project.shopping_trends  
GROUP BY Category;
```

Result Grid		
	Category	Avg_Item_Price
▶	Clothing	60.0253
	Footwear	60.2554
	Outerwear	57.1728
	Accessories	59.8387



2.3 Product Analysis



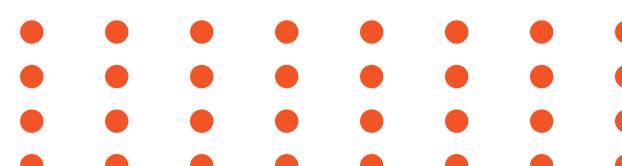
How many products were purchased by color?



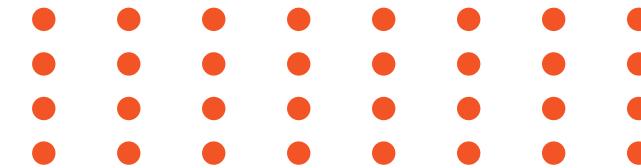
This query counts the total number of products (Item Purchased) for each color in the shopping_data table. It groups the data by Color and returns the total count of products for each color.

```
SELECT Color, COUNT(`Item Purchased`) AS Total_Products  
FROM project.shopping_trends  
GROUP BY Color;
```

	Color	Total_Products
▶	Gray	159
	Maroon	158
	Turquoise	145
	White	142
	Charcoal	153
	Silver	173
	Pink	153
	Burnle	151



2.4 Product Analysis



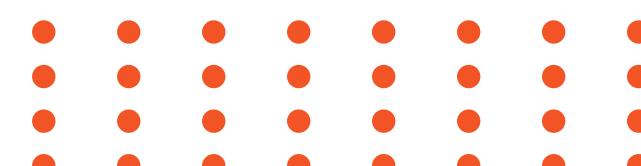
Which seasons generate the most purchases?



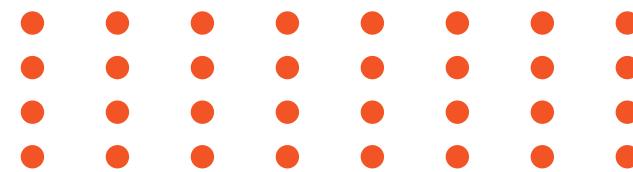
This query counts the total number of purchases made in each season by grouping the data by Season. It orders the results in descending order based on the total number of purchases, showing which seasons had the highest purchase activity.

```
• SELECT Season, COUNT(*) AS Total_Purchases  
  FROM project.shopping_trends  
  GROUP BY Season  
  ORDER BY Total_Purchases DESC;
```

	Season	Total_Purchases
▶	Spring	999
	Fall	975
	Winter	971
	Summer	955



3.1 Sales Performance



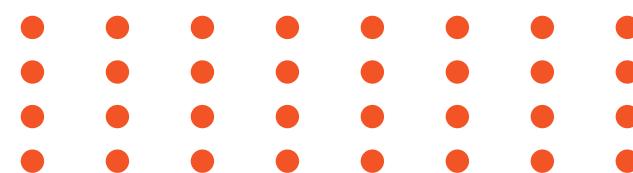
What is the total sales revenue?



This query calculates the total revenue by summing up all values in the Purchase Amount (USD) column from the shopping_data table. The result shows the overall revenue generated from all purchases.

- ```
SELECT SUM(`Purchase Amount (USD)`) AS Total_Revenue
FROM project.shopping_trends;
```

| Total_Revenue |
|---------------|
| 233081        |



## 3.2 Sales Performance

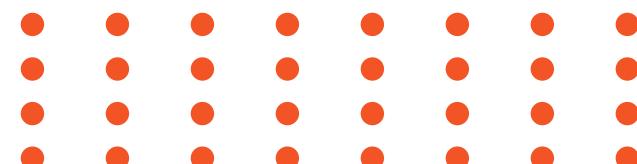
What is the average purchase amount across all purchases?



This query calculates the average purchase amount by taking the mean of all values in the Purchase Amount (USD) column from the shopping\_data table. The result shows the average amount spent per purchase.

- ```
SELECT AVG(`Purchase Amount (USD)`) AS Avg_Purchase_Amount  
FROM project.shopping_trends;
```

	Avg_Purchase_Amount
▶	59.7644



3.3 Sales Performance

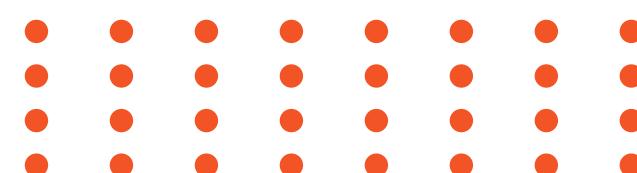
Which payment method is the most popular?



This query counts the number of purchases made using each payment method by grouping the data by Payment Method. It orders the results in descending order of purchase count, showing the most frequently used payment methods first.

- ```
SELECT `Payment Method`, COUNT(*) AS Payment_Count
FROM project.shopping_trends
GROUP BY `Payment Method`
ORDER BY Payment_Count DESC;
```

|   | Payment Method | Payment_Count |
|---|----------------|---------------|
| ▶ | Credit Card    | 696           |
|   | Venmo          | 653           |
|   | Cash           | 648           |
|   | PayPal         | 638           |
|   | Debit Card     | 633           |
|   | Bank Transfer  | 632           |



# 3.4 Sales Performance

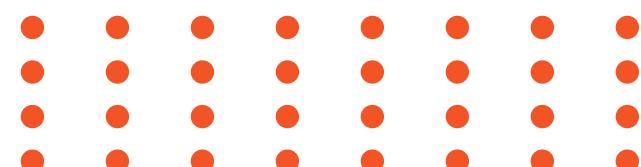
What is the total discount applied?



This query counts the total number of orders where a discount was applied by filtering rows with Discount Applied = 'Yes' in the shopping\_data table. The result shows the total discounted orders.

- ```
SELECT COUNT(*) AS Total_Discounted_Orders
FROM project.shopping_trends
WHERE `Discount Applied` = 'Yes';
```

Total_Discounted_Orders
1677



4.1 Order and Shipping Analysis

- ```
SELECT `Shipping Type`, COUNT(`Customer ID`) AS Customer_Count
FROM project.shopping_trends
GROUP BY `Shipping Type`;
```

|   | Shipping Type  | Customer_Count |
|---|----------------|----------------|
| ▶ | Express        | 646            |
|   | Free Shipping  | 675            |
|   | Next Day Air   | 648            |
|   | Standard       | 654            |
|   | 2-Day Shipping | 627            |
|   | Store Pickup   | 650            |

How many customers used each shipping type?



This query counts the number of customers (Customer ID) for each shipping type by grouping the data by Shipping Type. The result shows how many customers used each shipping method.

## 4.2 Order and Shipping Analysis

- SELECT ROUND(COUNT(\*) \* 100.0 / (SELECT COUNT(\*))  
FROM project.shopping\_trends))  
AS Promo\_Code\_Percentage  
FROM project.shopping\_trends  
WHERE `Promo Code Used` = 'Yes';

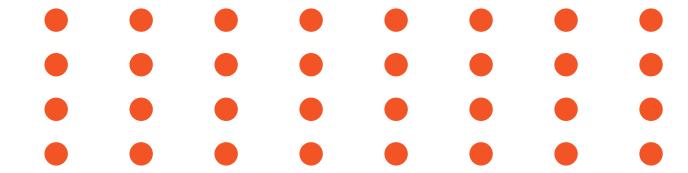
What percentage of orders applied promo codes?



This query calculates the percentage of orders where a promo code was used. It counts the rows where Promo Code Used = 'Yes', divides that count by the total number of rows in the shopping\_data table, and multiplies by 100 to get the percentage.

|   | Promo_Code_Percentage |
|---|-----------------------|
| ▶ | 43                    |

# 5.1 Customer Behavior



How many customers have a subscription status as 'Yes'?

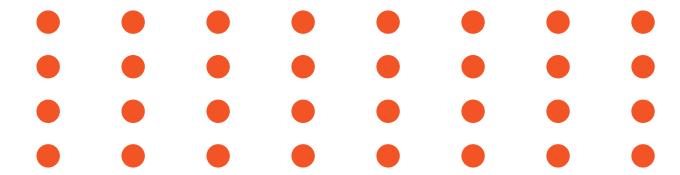


This query counts the number of unique customers who have a subscription (Subscription Status = 'Yes') in the shopping\_data table, returning the total number of subscribed customers.

- ```
SELECT COUNT(DISTINCT `Customer ID`) AS Subscribed_Customers
FROM project.shopping_trends
WHERE `Subscription Status` = 'Yes';
```

	Subscribed_Customers
▶	1053

5.2 Customer Behavior



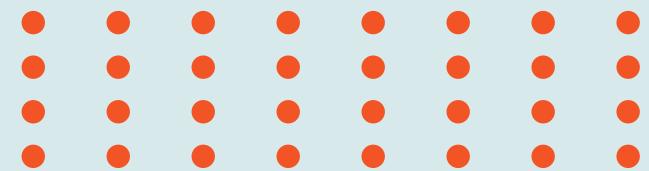
What is the average review rating for each product category?



- This query calculates the average review rating for each product category by grouping the data by Category and averaging the Review Rating for each group.

- ```
SELECT Category,round(AVG(`Review Rating`),2) AS Avg_Review_Rating
FROM project.shopping_trends
GROUP BY Category;
```

|   | Category    | Avg_Review_Rating |
|---|-------------|-------------------|
| ▶ | Clothing    | 3.72              |
|   | Footwear    | 3.79              |
|   | Outerwear   | 3.75              |
|   | Accessories | 3.77              |



## 6.1 Regional and Seasonal Insights

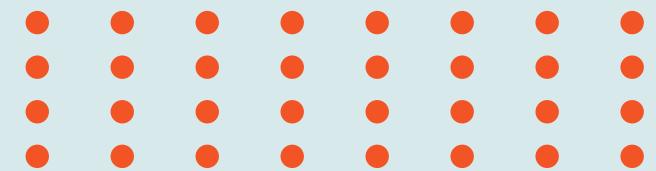
Which locations generate the most revenue?



This query calculates the total revenue for each location by summing the Purchase Amount (USD) for each Location. It then orders the results in descending order of total revenue, showing the locations with the highest revenue first.

- ```
SELECT Location, SUM(`Purchase Amount (USD)`) AS Total_Revenue
FROM project.shopping_trends
GROUP BY Location
ORDER BY Total_Revenue DESC
limit 8;
```

Location	Total_Revenue
Illinois	5617
California	5605
Idaho	5587
Nevada	5514
Alabama	5261
New York	5257
North Dakota	5220



6.2 Regional and Seasonal Insights

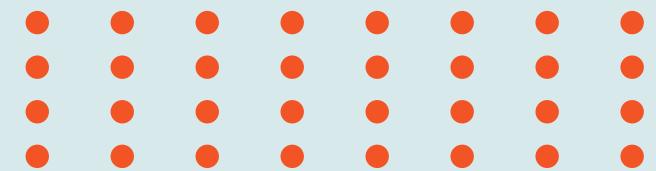
What are the most popular product colors in each season?



This query counts the number of products for each color in each season by grouping the data by Season and Color. It then orders the results by season and color count in descending order, showing the most popular colors within each season.

- ```
SELECT Season, Color, COUNT(*) AS Color_Count
FROM project.shopping_trends
GROUP BY Season, Color
ORDER BY Season, Color_Count DESC
LIMIT 7;
```

|   | Season | Color   | Color_Count |
|---|--------|---------|-------------|
| ▶ | Fall   | Yellow  | 50          |
|   | Fall   | Magenta | 50          |
|   | Fall   | Olive   | 47          |
|   | Fall   | Orange  | 45          |
|   | Fall   | Gray    | 44          |
|   | Fall   | Violet  | 44          |
|   | Fall   | Cyan    | 43          |



## 6.3 Regional and Seasonal Insights

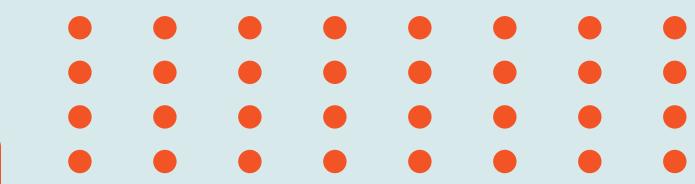
How many purchases were made in each season?



This query counts the total number of purchases made in each season by grouping the data by Season and ordering the results in descending order of total purchases, showing the seasons with the highest number of purchases first.

- ```
SELECT Season, COUNT(*) AS Total_Purchases
FROM project.shopping_trends
GROUP BY Season
ORDER BY Total_Purchases DESC;
```

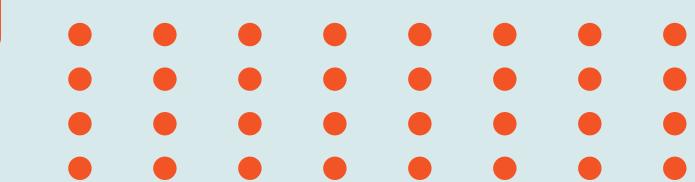
	Season	Total_Purchases
▶	Spring	999
	Fall	975
	Winter	971
	Summer	955



7. Return and Refund Analysis

How many items have a review rating below 3? (Assume low ratings mean dissatisfaction)

	Low_Rated_Items
▶	685



- ```
SELECT COUNT(*) AS Low_Rated_Items
FROM project.shopping_trends
WHERE `Review Rating` < 3;
```

This query counts the number of items with a review rating below 3 by filtering the shopping\_data table for rows where Review Rating is less than 3 and returning the total count of such items.

# Contact



+91 7208243180



divyanshpandey910@gmail.com



[linkedin.com/in/divyansh-pandey-3675132b9](https://www.linkedin.com/in/divyansh-pandey-3675132b9)





# Thank you!



*Divyansh Dinesh Pandey*

2024