

Data Science Intern (Remote) - Winter 2022

Shopify Technical Challenge Submission

(<https://github.com/Ashwin3110/Shopify-Winter-2022-Intern-Challenge->)

Question 1: Given some sample data, write a program to answer the following: [click here to access the required data set](#)

On Shopify, we have exactly 100 sneaker shops, and each of these shops sells only one model of shoe. We want to do some analysis of the average order value (AOV). When we look at orders data over a 30 day window, we naively calculate an AOV of \$3145.13. Given that we know these shops are selling sneakers, a relatively affordable item, something seems wrong with our analysis.

- a. Think about what could be going wrong with our calculation. Think about a better way to evaluate this data.
- b. What metric would you report for this dataset?
- c. What is its value?

Solution(Please look at my github link for my analysis of the dataset)

- a) Proper analysis has not been done with the data. We could see that shop with id=78 Which sells the price of shoe with \$25725 Which is much higher than other shoes in other shop and it is not relatively affordable.
I guess there would be some error while entering the price of the shoe for the shop 78
- b) Same metric is used here (i.e) total order amount divided by total no of order ' But I have excluded the shoes sold by shop 78
- c) The value I got is :- 2717.37

Question 2: For this question you'll need to use SQL. [Follow this link](#) to access the data set required for the challenge. Please use queries to answer the following questions. Paste your queries along with your final numerical answers below.

- a. **How many orders were shipped by Speedy Express in total?**

```
SELECT COUNT(*) AS NumberOfOrders
FROM [Orders]
JOIN [Shippers]
    ON [Shippers].ShipperID = [Orders].ShipperID
WHERE [Shippers].ShipperName = 'Speedy Express'
```

ANSWER : 54

- b. **What is the last name of the employee with the most orders?**

```
SELECT [Employees].LastName
FROM [Orders]
JOIN [Employees]
    ON [Orders].EmployeeID = [Employees].EmployeeID
GROUP BY [Employees].LastName
ORDER BY NumberOfOrders DESC
LIMIT 1
```

ANSWER : Peacock

- c. **What product was ordered the most by customers in Germany?**

```
SELECT [Products].ProductName,
    SUM([OrderDetails].Quantity) AS "TotalOrdered"
FROM [Orders]
JOIN [Customers]
    ON [Customers].CustomerID = [Orders].CustomerID
JOIN [OrderDetails]
    ON [OrderDetails].OrderID = [Orders].OrderID
JOIN [Products]
    ON [Products].ProductID = [OrderDetails].ProductID
WHERE [Customers].Country = 'Germany'
GROUP BY [OrderDetails].ProductID
ORDER BY TotalOrdered DESC
LIMIT 1
```

ANSWER : Boston Crab Meat , 160