

# Fall 2022 Data Science Intern Challenge

Please complete the following questions, and provide your thought process/work. You can attach your work in a text file, link, etc. on the application page. Please ensure answers are easily visible for reviewers!

**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.

**Answer:**

On inspecting the data, it was clear that \$3145.13 was the Mean order amount and not the Average Order Value.

In order to calculate AOV, we need an alternative metric.

I propose this metric, sum all orders amount and divide by the number of items in the order.

In this way, we are calculating the average price of sneakers selling by Shopify's partners.

- b. What metric would you report for this dataset?

**Answer:**

Calculate Sum of Order Amount.

Calculate Sum of Total Items

Divide them

This will be your AOV.

- c. What is its value?

**Answer:**

The Average Order Value (AOV) is: \$357.92

**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?

**Answer:**

```
SELECT s.ShipperName,COUNT(o.OrderID) as Counts
FROM Orders o
LEFT JOIN Shippers s
ON o.ShipperID = s.ShipperID
GROUP BY s.ShipperName;
```

**Total\_Items shipped by Speedy Express is 54**

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

**Answer:**

```
SELECT o.EmployeeID,COUNT(ORDERID) as Total_Orders,LastName
FROM Orders o
JOIN Employees e
ON o.EmployeeID = e.EmployeeID
GROUP BY o.EmployeeID
ORDER BY Total_Orders DESC
LIMIT 1;
```

**Lastname is Peacock**

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

**Answer:**

```
SELECT p.ProductID,p.ProductName,COUNT(p.ProductID) as Count_Products
FROM Products p
JOIN OrderDetails od
ON p.ProductID = od.ProductID
JOIN Orders o
ON od.OrderID = o.OrderID
JOIN Customers c
ON o.CustomerID = c.CustomerID
WHERE Country = 'Germany'
```

*GROUP BY p.ProductID,p.ProductName  
ORDER BY Count\_Products DESC;*

**The Product that was ordered the most was “Gorgonzola Telino”**