

# Summer 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.

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

See code in my github @

[https://github.com/kathysimon99/shopify\\_intern\\_challenge/blob/main/shopify\\_intern\\_challenge.i  
pynb](https://github.com/kathysimon99/shopify_intern_challenge/blob/main/shopify_intern_challenge.ipynb)

**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.

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

QUERY	ANSWER
<b>SELECT Shippers.ShipperName, COUNT(Shippers.ShipperID) AS TotalOrders FROM Orders INNER JOIN Shippers ON Shippers.ShipperID=Orders.ShipperID</b>	<b>54</b>

<b>WHERE Shippers.ShipperName='Speedy Express';</b>	
---	--

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

QUERY	ANSWER
<b>SELECT Employees.LastName, COUNT(Employees.EmployeeID) AS TotalOrders  FROM Employees  INNER JOIN Orders ON Orders.EmployeeID = Employees.EmployeeID  GROUP BY Employees.EmployeeID  ORDER BY TotalOrders DESC;</b>	<b>Last Name:  Peacock   Number of Orders:  40</b>

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

QUERY	ANSWER
<b>SELECT OrderDetails.OrderID, OrderDetails.ProductID, Products.ProductName, Customers.Country, SUM(OrderDetails.Quantity) AS TotalOrder  FROM OrderDetails  INNER JOIN Products ON Products.ProductID = OrderDetails.ProductID  INNER JOIN Orders ON Orders.OrderID = OrderDetails.OrderID  INNER JOIN Customers ON Customers.CustomerID = Orders.CustomerID  WHERE Customers.Country = 'Germany'  GROUP BY Products.ProductName  ORDER BY TotalOrder DESC;</b>	<b>Boston Crab Meat;   Total Quantity:  160</b>