# 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](https://docs.google.com/spreadsheets/d/16i38oonuX1y1g7C_UAmiK9GkY7cS-64DfiDMNiR41LM/edit#gid=0)

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.

1. Think about what could be going wrong with our calculation. Think about a better way to evaluate this data.

***The calculation could be wrong because it might not be for just one item in the order. The customer may have ordered multiple items to get that large price tag.***

1. What metric would you report for this dataset?

***Since we know that they only sell one type of shoe, I would divide the total amount of orders\_values with the amount of total\_items sold to get a better idea of each shoe.***

1. What is its value?

***$357.92***

**Question 2:** For this question you’ll need to use SQL. [Follow this link](https://www.w3schools.com/SQL/TRYSQL.ASP?FILENAME=TRYSQL_SELECT_ALL) 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.

1. How many orders were shipped by Speedy Express in total?

SELECT COUNT(OrderID) AS "Amount of Orders Shipped by Speedy Express"

FROM Orders JOIN Shippers USING(ShipperID)

WHERE ShipperName="Speedy Express"

**Answer:** 54

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

SELECT LastName AS "Last Name"

FROM (SELECT Max(OrdersTotal),LastName

FROM (SELECT COUNT(OrderId) AS OrdersTotal, EmployeeID, LastName

FROM ORDERS JOIN Employees USING(EmployeeID)

GROUP BY EmployeeID))

**Answer:** Peacock

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

SELECT ProductName AS "Product Name"

FROM (SELECT Max(Total), ProductName

FROM (SELECT ProductID, SUM(Quantity) AS Total, ProductName

FROM (SELECT \* FROM Customers JOIN (SELECT OrderID, ProductID, Quantity, CustomerID

FROM Orders JOIN OrderDetails USING (OrderID))

USING (CustomerID)

WHERE (Country) LIKE "Germany")

JOIN Products USING (ProductID)

GROUP BY ProductID))

**Answer:** Boston Crab Meat