# Fall 2021 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 of AOV of $3145.13 is the average of total order\_amount. A better way to evaluate this data could be if we find out order\_amount/total\_items for each order\_id in data. AOV in this case is $387.74, which gives us amount per item.

1. What metric would you report for this dataset?

Metric that I report for this data would be total\_items order per day.

1. What is its value?

Value is 1465 total\_items per day

**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?

**Query:**

SELECT Orders.ShipperID, Shippers.ShipperName

FROM Orders

JOIN Shippers

ON Orders.ShipperID = Shippers.ShipperID

WHERE Orders.ShipperID = 1 AND Shippers.ShipperName = 'Speedy Express'

**Answer:** Number of Records: 54

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

**Query:**

SELECT Orders.OrderID, Orders.EmployeeID, Employees.LastName,

COUNT(\*)

FROM Orders

JOIN Employees

ON Orders.EmployeeID = Employees.EmployeeID

GROUP BY Orders.EmployeeID;

**Answer :** Peacock

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

**Query:**

SELECT Customers.Country, Customers.CustomerID, Orders.OrderID, OrderDetails.ProductID, OrderDetails.Quantity

FROM Customers

JOIN Orders

ON Orders.CustomerID = Customers.CustomerID

JOIN OrderDetails

ON Orders.OrderID = OrderDetails.OrderID

Where Customers.Country = 'Germany'

Order by OrderDetails.Quantity DESC

Product ID = 35

**Query:**

SELECT ProductName

FROM Products

WHERE ProductID = 35

**Answer :** Product Name is Steeleye Stout