

Winter 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](#)

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?

Please refer to the Jupyter Notebook for answers to Question 1.

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 OrderDetails.OrderID, OrderDetails.Quantity, Orders.ShipperID
FROM OrderDetails
JOIN Orders
ON OrderDetails.OrderID= Orders.OrderID
GROUP BY Orders.ShipperID
ORDER BY Orders.ShipperID, OrderDetails.Quantity;
```

Speedy Express shipped 12 orders in total.

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

```
SELECT COUNT(*) AS num_orders, Orders.EmployeeID, Employees.LastName
FROM Orders
JOIN Employees
ON Orders.EmployeeID = Employees.EmployeeID
GROUP BY Orders.EmployeeID
ORDER BY num_orders DESC;
```

The last name of the employee with the most orders is 'Peacock'

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

```
SELECT Customers.Country, Products.ProductID, Products.ProductName,
OrderDetails.Quantity
FROM Customers
JOIN Orders
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 Customers.Country;
```

The product ordered most by customers in Germany was Boston Crab Meat