

# Winter 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.
- b. What metric would you report for this dataset?
- c. What is its value?

**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 count(*) from (Select * from OrderDetails inner join Orders on
OrderDetails.OrderID = Orders.OrderID) as tab1 inner join Shippers on tab1.ShipperID =
Shippers.ShipperID and ShipperName like "Speedy Express"
```

149

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

```
select LastName,sum(Quantity) as Total_Orders from (Select * from OrderDetails inner
join Orders on OrderDetails.OrderID = Orders.OrderID) as table1 inner join Employees
on table1.EmployeeID = Employees.EmployeeID group by table1.EmployeeID order by
Total_Orders desc limit 1;
```

3232

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

```
select sum(Quantity) as Total_Orders,ProductName from (select * from (Select * from
OrderDetails inner join Orders on OrderDetails.OrderID = Orders.OrderID) as table1 inner join
Products on Products.ProductID = table1.ProductID) as table2 inner join Customers on
Customers.CustomerID = table2.CustomerID where Country like "Germany" Group by
ProductID order by Total_Orders desc limit 1;
```

Boston Crab Meat