Fall 2022 Data Science Intern Challenge
Submission for Bharath Punati (Resume attached in last page)

QUESTION 1:

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

Answer: Looking at the data, there are outliers which are very evident from checking the median, min and max value. Further investigating the data, the shops with ShopID = 42 and ShopID = 78 have very high number of orders and the value of sneakers from shopID = 78 is relatively very high with a value of $25725. I would assume that the there was a mistake while entering the price and it's entered in cents, converting into dollars and editing the table for the sneakers amount for shopID = 78 to be $257.25.

1) Changing the price of sneakers for ShopID = 78 could be one better way to find the mean.
2) For shopID = 42, the number of orders is extremely high relative to other shops and diving the revenue by number of orders could possibly be a better measure.

b) What metric would you report for this dataset?

Answer: Considering the value of item for ShopID = 78 is incorrectly entered, median would be a good measure since the median is less sensitive and affected by the outliers in the dataset and the AOV is around $300

c) What is its value?

Answer: Considering the median, AOV is about $284

QUESTION 2:

a) How many orders were shipped by Speedy Express in total?

```
SELECT Shippers.ShipperName, COUNT(Orders.OrderID) as NUM_ORDERS FROM Orders
LEFT JOIN Shippers ON Orders.ShipperID = Shippers.ShipperID
WHERE Shippers.ShipperName = 'Speedy Express'
GROUP BY ShipperName;
```

Answer : 54 Orders are shipped by Speedy Express

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

Answer : Peacock

```
Select Employees.LastName, Count(Orders.OrderID) As NumberOfOrders From Orders
Left Join Employees ON Orders.EmployeeID = Employees.EmployeeID
Group by LastName Order by NumberOfOrders Desc
Limit 1;
```

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

Answer : Steeleye Stout

```
SELECT ProductName
FROM Orders a JOIN Customers b ON a.CustomerID = b.CustomerID
JOIN OrderDetails c ON a.OrderID = c.OrderID
JOIN Products d ON d.ProductID = c.ProductID
WHERE Country='Germany'
ORDER BY Quantity DESC
LIMIT 1
```