Shopify Application

Question 1:

There are two potential metrics I would use instead of the AOV. The first one is to utilize the **median value ($284.0)**. This is a good metric because I do not have context behind why certain values were inputted. For example the highest order value is $704000. This may have been due to an input error or it may be a correct inputted value; however, it is evident that this value is an outlier which heavily skews the data. Utilizing the median would reduce the effect of these high order value data points allowing us to see a good metric for how well stores are doing. The second metric I would potentially use is **removing these outlier values using the 1.5 * IQR method and then calculating the mean ($293.71)** afterwards. This would remove high outlier values nullifying their effects on the mean leaving a better average for how well stores are doing. One reason I may stray away from the second metric is that the data is right skewed after removing the outliers. Overall the safest method is to use the median value but it would be important to investigate the data and determine the context behind these numbers

Code Used for analysis: https://github.com/Alan-Phan/ShopifyCode/blob/main/Shopify.ipynb

Question 2: SQL Portion

How many orders were shipped by Speedy Express in total?

```
SELECT ShipperName, COUNT(*)
FROM [Shippers]
INNER JOIN Orders
ON Shippers.ShipperID == Orders.ShipperID
WHERE Shippers.ShipperID == 1
```

Speedy Express shipped 54 orders.

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

```
SELECT LastName, COUNT(*)
FROM Employees
INNER JOIN Orders
ON Orders.EmployeeID == Employees.EmployeeID
GROUP BY Employees.LastName
ORDER BY COUNT(*) DESC
LIMIT(1)
```

The Last name of the employee with the most orders is Peacock with 40 orders.

What product was ordered the most by customers in Germany?

```
SELECT p.ProductName, SUM(Quantity) as TotalQuant, p.Unit
FROM Customers as c
        INNER JOIN Orders as o
   ON c.CustomerID == o.CustomerID
   INNER JOIN OrderDetails as od
   ON od.OrderID == o.OrderID
   INNER JOIN Products as p
   ON p.ProductID == od.ProductID
WHERE Country == "Germany"
GROUP BY p.ProductID
ORDER BY TotalQuant DESC
LIMIT(1)
```

The product that was ordered the most by customers in Germany is Boston Crab Meat at a quantity of 160 units (24 - 4oz tins).