



Fall 2021 Data Science Intern Challenge

Please find the code using this [link](#)

- Answer 1.A
 - The calculation wasn't accurate enough since the data that was collected was raw data, and the operations were performed on that data without performing necessary cleaning operations.
 - The average was achieved since the calculated mean constituted of outliers.
- Answer 1.B
 - The metric used to get rid of all the outliers was IQR. Since the data was fairly small and didn't require any of the complex solutions.
- Answer 1.C
 - The final value that I got was an AOV of \$293.72

SQL Queries & Solutions

- How many orders were shipped by Speedy Express in total?
 - Statement –

```
SELECT COUNT(ShipperID)
FROM orders o
WHERE o.ShipperID in (select s.ShipperID from Shippers s where
ShipperName = 'Speedy Express')
```
 - Value – 54 orders

SQL Queries & Solutions

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

- Statement –

```
SELECT e.LastName, COUNT(o.EmployeeID) as C  
FROM  
    "Employees" AS e  
JOIN "Orders" AS o on o.EmployeeID = e.EmployeeID  
GROUP BY o.EmployeeID  
Order BY C DESC
```

- Value – Peacock

SQL Queries & Solutions

- What product was ordered the most by customers in Germany?
- Statement –

```
SELECT
```

```
o.OrderID, SUM(d.Quantity), p.ProductName
```

```
FROM
```

```
"Orders" AS o
```

```
JOIN "Customers" AS c on o.CustomerID = c.CustomerID
```

```
JOIN "OrderDetails" AS d on o.OrderID = d.OrderID
```

```
JOIN "Products" AS p on p.ProductID = d.ProductID
```

```
where c.Country='Germany'
```

```
GROUP BY o.OrderID
```

```
ORDER by SUM(d.Quantity) DESC
```

Value – Chang