Winter 2022 Data Science Intern Challenge

Ankur Sanjay Limbashia

Question 1: [click here to access the required data set](#)

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

In your calculation, you have divided the total order amount by total number of entries.
But, In actual case we need to divide the sum of total order amount by the total items sold to get the correct average order value which is called as weighted average.

b.) What metric would you report for this dataset?

For this dataset, to find the average order value we need to find the weighted average which is (total sum of order value/ total number of orders).

c.) What is its value?

Average order value is 357.9215222

Question 2:

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

Query:
SELECT sum(od.quantity) AS "total quantity shipper by Speedy Express" FROM [OrderDetails] AS od
JOIN [orders] AS o ON od.orderid = o.orderid
JOIN [shippers] AS s ON s.shipperid = o.shipperid
WHERE s.shippername = "Speedy Express"

Answer: 3575

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

Query:
SELECT e.lastname FROM [Orders] AS o
JOIN orderdetails AS od ON o.orderid = od.orderid
JOIN employees AS e ON e.employeeid = o.employeeid
GROUP BY e.lastname
ORDER BY sum(od.quantity) desc
LIMIT 1

Answer: Peacock

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

Query:
SELECT p.productname FROM [Orders] AS o
JOIN orderdetails AS od ON o.orderid = od.orderid
JOIN customers AS c ON c.customerid = o.customerid
JOIN products AS p ON p.productid = od.productid
group by p.productname
HAVING c.country = "Germany"
ORDER BY sum(od.quantity) desc
LIMIT 1

Answer: Boston Crab Meat