Question 1

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

after analysis plotting the amount order column to identify average value for all customers I found out there is 141 unusual customers that have had ordered with amount of money not less than 7000$ from 5000 orders these outliers had significant effect on the data leads to 3145.13

which is not the true average

IQR would be the middle range of sneakers data which will calculate the outlier fence using percentages in the analysis data is right skewed because of 141 orders and its only 2.5% out of the 5000 orders those customers will be removed by Interquartile Range(IQR)

b. median matrix will be used instead of mean which is highly sensitive for outliers

c- Average Value Order is 280.0

**Question 2:**

1. **How many orders were shipped by Speedy Express in total?**

Select sh.shipperName, count(ord.orderId) as TotalOrderedBySpeedyExpress

from shippers sh

left join

orders ord on sh.shipperID = ord.shipperID

where sh.shipperName ='speedy express'

group by sh.shipperName

;

**Answer:**

**54**

1. **What is the last name of the employee with the most orders?**

Select em.LastName, count(ord.OrderID) as Number\_of\_orders

from Employees em

left join

Orders ord on em.EmployeeID = ord.EmployeeID

group by em.LastName

order by Number\_of\_orders desc

limit 1

;

**Answer:**

**Peacock Number\_of\_orders = 40**

1. **What product was ordered the most by customers in Germany?**

Select prod.ProductName, cus.Country, sum(ordt.Quantity) as Total\_of\_Products

from Products prod

join

OrderDetails ordt on prod.ProductID = ordt.ProductID

join

Orders ords on ords.OrderID = ordt.OrderID

join

Customers cus on ords.CustomerID = cus.CustomerID

where cus.Country ='Germany'

group by prod.ProductName

order by Total\_of\_Products desc

limit 1

;

**Answer:**

**Boston Crab Meat Total ordered = 160**