# ANALYSIS INTERPRETATION QUESTIONS

1. What was the correlation ( R2 value ) between sales and income ?
2. What was the correlation ( R2 value ) between customer ratings and product return rate ?
3. What are the regression formulas to predict customer from customer sales ?
4. Which customer do you predict has the highest income
5. Which product will be advertised the most.

Conclusively, It can seen that niche customers were those with an income range of 70000 – 80000 dollars as they can be seen to be the most frequent customers .Interestingly, the niche customer were those from Maryland as per our heat map .Equally , it clear that the return rate of products seems low when the rating of products are high and this could seen from the strength of the correlation between those two quantities

AS per our question , it can seen from our visualization between sales and income that its correlation value reads at 0.78 which shows a strong level of correlation . Meaning that there is a high linking strength of impact between this quantities as an increase in the average last 6 month income increases with an increase in average income by state.

At a second glance , it can seen from our visualization between customer ratings and product return rate that its correlation value reads at 0.88 which shows a strong level of correlation . Meaning that there is a high linking strength of impact between these quantities as an increase in the customer rating lead to decrease in the return rate.

Regarding the regression formula for this analysis , the equation is Y = 0.01x – 722.17 , where the 0.01 is the gradian or the slope of the linear regression and the -722.17 is the intercept.

As per the customer with the highest income , this could be seen to be a customer from District Columbia.

Regarding which product to advertised most , It can seen from our visualization that leader back will be advertised the most .This is because its price standards from other products leading to a request in high level of advertisement .

Regarding proposals for further analysis, the design of the current study is subject to one limitation such as not all correlation and distribution has been studied in the dataset rendering it difficult for an absolute and appropriate further analysis to recommend specific product in specific area .In Future , such dataset should be analyzed completely by studying all correlation and distribution of the dataset so as to ensure an intuitive and understandable analysis.