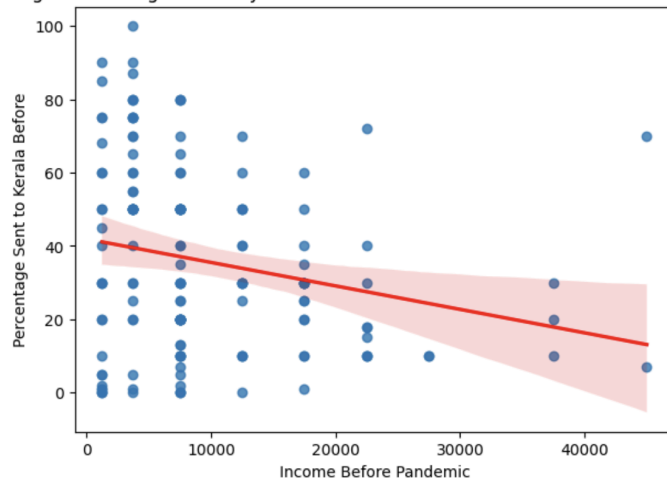


Misleading Data Piece

Slope: -0.0006403802068772716
R-squared: 0.05
P value: 0.01

Fig.4 Percentage of Money Sent to Kerala versus Income Before the Pandemic

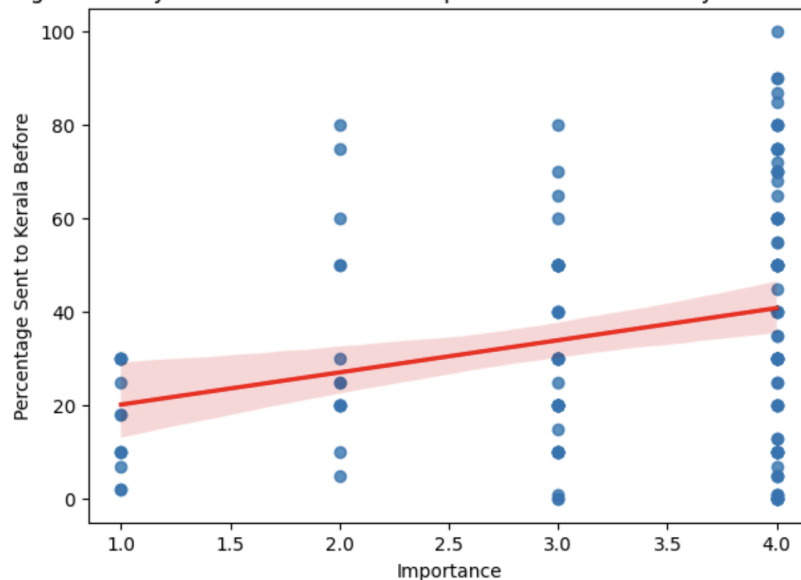


For each of these graphs we also calculated R^2 and the p-value.

1. The slope represents the slope of the regression line shown on the graph
2. The R^2 value tells you how well the regression line fits the data. It goes from 0 to 1– 0 is worse and 1 is better.
3. The p-value indicates whether the regression relationship indicated by the graph is strong or weak. If p-value < 0.05 , then the relationship is significant. Otherwise, the relationship is insignificant.

Slope: 6.882222985909951
R-squared: 0.07
P value: 0.00

Fig.6 Percentage of Money Sent to Kerala versus Importance to Send Money Home Before the Pandemic



Caption: Fig.4 and Fig.6

Explanation: These graphs are misleading, primarily because of the slope and regression line. When we see the slope of Figure 6, we are inclined to think that

Importance has a stronger impact on the percentage of money sent to Kerala as opposed to income. But this comparison isn't fair, because we are looking at unnormalized data.

When each type of data is in-between different ranges, regression analysis treats them differently based on their range of values. Since Income ranges from 1250 to 45000, a small change in income doesn't really lead to a large percentile change for percent sent to Kerala— a 1 AED change in income doesn't really mean much with such a large range. However, for Importance, since it ranges from 1 to 4, a small change in importance has a greater impact on percent sent to Kerala— a change in degree of importance by 1 unit is huge.

Because of this behavior, the two graphs are misleading because it implies importance has a bigger impact than income.