Inferences based on covariance and correlation values

1. What conclusion, if any, can be drawn from the values obtained in task 4 above?

The covariance for age and income is a positive value (139829.44) and the correlation for the same is positive with a value of 0.75.

Positive covariance implies that age and income vary positively upon a change in either of them. In simple words, if income increases, so does age.

From the greater correlation value, we can say that age and income are highly related to each other.

2. Suppose the income were measured in thousands instead of the actual value (for example, and income value of 17,456 now becomes 17.456), how would the results of task 4 above change (if at all)? Explain.

The value of covariance changes to 139.8294, which is significantly different from its original value. However, correlation remains same (0.75).

Explanation –

$$Cov_{xy} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{(n-1)} = \frac{\Sigma xy - n\overline{xy}}{(n-1)}$$

As shown in the formula above for computing covariance, let x=age and y=income. As mentioned in the question, if the income is measured in thousands, each of the current value of income gets divided by 1000. Hence, the new formula can be obtained by substituting y=y/1000 and $\bar{y}$=$\bar{y}$/1000. Thus, the covariance also gets divided by 1000.