



The red horizontal lines represent the mean for each variable - if a participant is below the mean on one variable, notice that they are also below the mean for the other variable - this suggests the two variables co-vary.

- For participants 2, 3 and 5, their scores on each variable are all below the respective means for each variable, for participants 1 and 4 their scores are all above the respective means for each variable.
- To formalise this, we can calculate the average combined differences.....

$$\text{cov}(x,y) = \frac{\sum (x_i - \bar{x}) (y_i - \bar{y})}{N - 1}$$