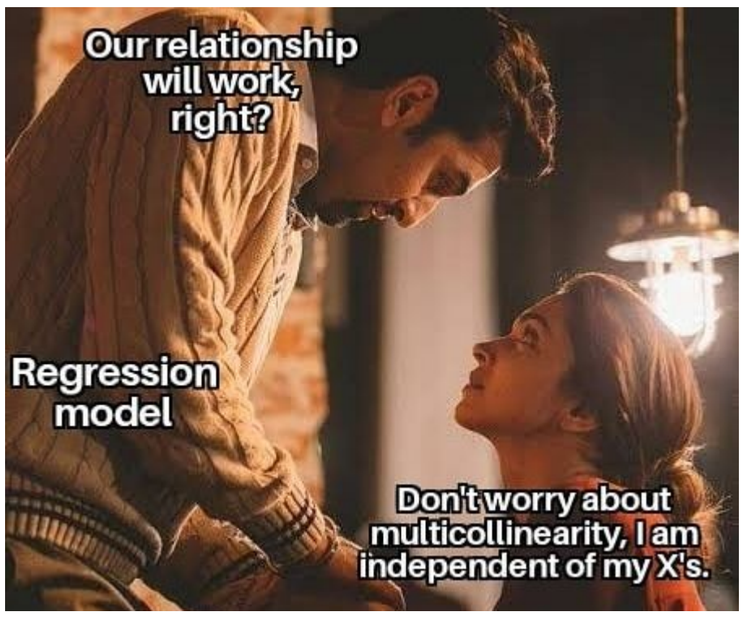
Welcome to Week 6!

Welcome to Week 6!

Now that we have been introduced to the concepts of correlation and causation (in week 5), we can look at multiple linear regression (MLR). In fact, MLR is one of the most commonly used tests in public health this is because it allows us to have multiple independent variables (predictors) for just one dependent variable.

As this test increases in complexity over standard Pearson's correlation and simple linear regression, we do not tend to try to calculate these by hand (Excel) and rather just use SPSS to do the calculations.

One big point to make about multiple linear regression is that there are a number of assumptions (like any other test) for it. In an ideal world, it would be nice to satisfy all of these assumptions but in practice, this is not always possible. One of the assumptions that holds a lot of importance are the correlations between predictor variables. This is called 'Multicollinearity'. When the correlation coefficients between predictor variables is high, the issue with this is that it greatly affects the data and results produced may not be reliable. It is essential to check the correlations (using Pearson's in SPSS) between the predictor variables (also known as "Xs") to see that they are not too high (above 0.8 as a rough measure) i.e. not highly correlated or are independent. Perhaps the image below might help to explain it (For those who enjoy a good love story!).



When doing a multiple regression, there are a number of approaches. These will be covered in the lecture but be sure to give them all a go in SPSS, particularly 'Enter', 'Backward' and 'Forward' options.