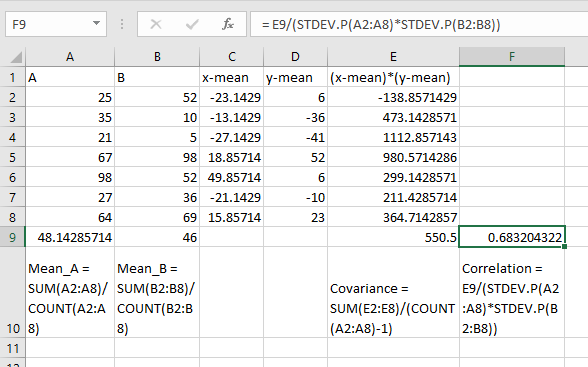
Edwisor: Calculate covariance and correlation between below two columns A and B

Uday: 

Edwisor: What are the different ways to deal with multi collinearity?

Uday:

When analyzing a dataset with multiple variables, collinearity helps in finding redundant variables ie collinearity indicates us the occurrence of highly correlated variables , so we can consider only one variable out of highly correlated variables in our analysis of dataset, which makes our analysis simpler.

If in case we are trying to predict an output using a dataset then we need to ignore the collinearity of the dataset.

Edwisor: What should be the correlation threshold value based on which we determine the highly collinear variables?

Uday: the value of correlation varies from -1 to 1. Negative value indicates negative correlation ie if one value increases, other value decreases. Positive value indicates positive correlation where if one variable increases other variable also increases. 0 indicates there is no correlation between variables.

The value of correlation should be closer to 1 to determine highly collinear variables.

Edwisor: What are the two different types of variable we used in ANOVA?

Uday:

1. Numerical variable

2. Categorical variable

Edwisor: What are the null and alternate hypothesis in chi-square test?

Uday:

H0: there is no relation between two variables or they are independent.

H1: there is some relation between two variables or they are dependent.