

Assignment 3: Regression

This is an assignment on correlation, regression, multiple regression, non-linear regression and the use of dummy variables. It requires from you to:

- First, think of an interesting question that can be addressed by regression analysis. Try to find a short background on the problem.
- Find the appropriate data for your problem
 - You may use the data or at least the source from the previous assignments (optional, you can use new data)
 - The data should include at least two metric (interval or ratio) variables and at least one nominal variable to be converted into dummy
 - Careful about the number of observations (at least variables used times ten)!
- Plot the relationship between two variables in an appropriate chart
- Estimate the simple bivariate linear regression between the pictured variables
- Estimate a multiple regression
 - Code the nominal variable into dummies appropriately
 - Comment on the change in R^2 and regression coefficient compared to bivariate linear regression
 - Use one of the model building approaches (e.g. non-linear, interaction, forward or backward selection...)

For example (DO NOT USE THIS EXAMPLE!) take 27 EU members, FDI as % of GDP as a dependent variable, GDP per capita as an explanatory and an indicator of whether the country is a member of the euro area or not. First plot the relationship. Say the relationship is increasing at a decreasing rate. Compute the linear regression between FDI and GDP. Then estimate the logarithmic regression (suitable for this type of relationship). Code the »euro area« appropriately. Compute the regression on FDI using GDP and »euro area« as indicators.

What will be marked for this assignment:

- Appropriate data (15%)
- Presentation of the relationship (25%)
- Simple bivariate linear regression (30%)
- Multiple regression (30%)

The length limit is four pages.