

Weekly assignment 3

In this week's assignment the first task was to examine whether there is a linear or non-linear relationship between factors "project owner follower amount" and the "project pledged value". I started this task by first examining the quality of the data meaning checking if there are for example invalid or null values. With a quick review I did not come across any alarming issues, so I decided to continue to the regression.

I started the analysis by selecting linear regression analyze in SPSS. This is presented in figure 1.

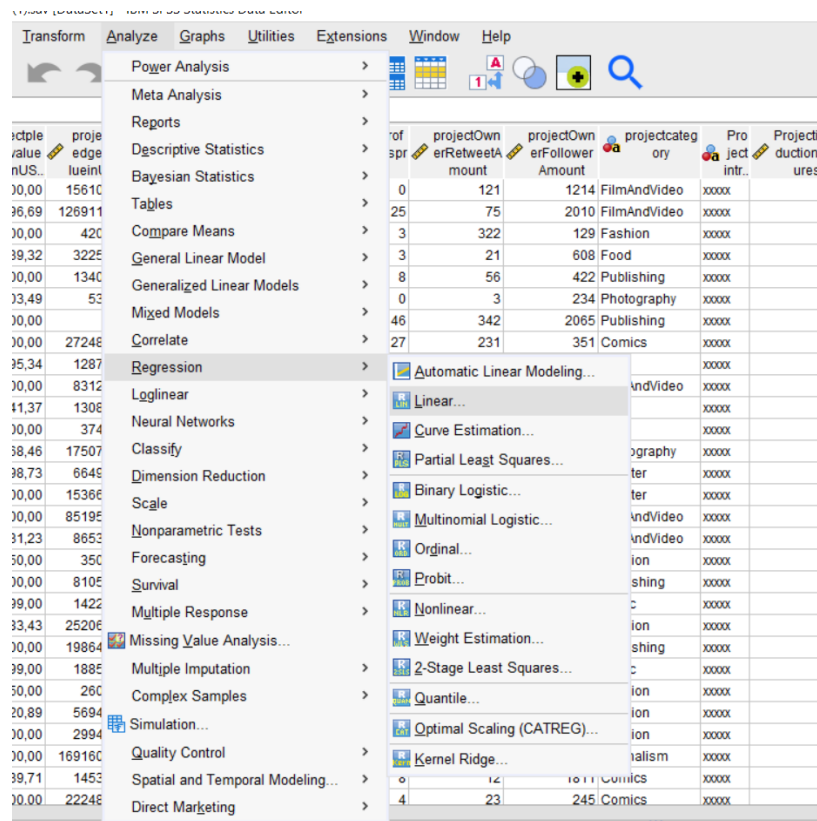


Figure 1 Start of the analysis

Next, I needed to figure out which one of the factors is dependent and which one independent variable. After a while I decided that "project owner follower amount" was going to be the dependent

variable and “project pledged value” the independent value. After this I set the statistics and plotting settings as presented in figures 2 and 3.

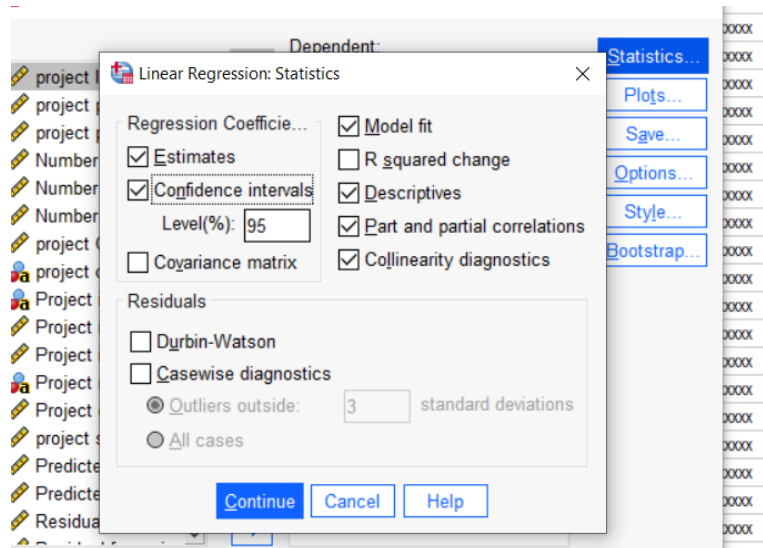


Figure 2 Statistics settings

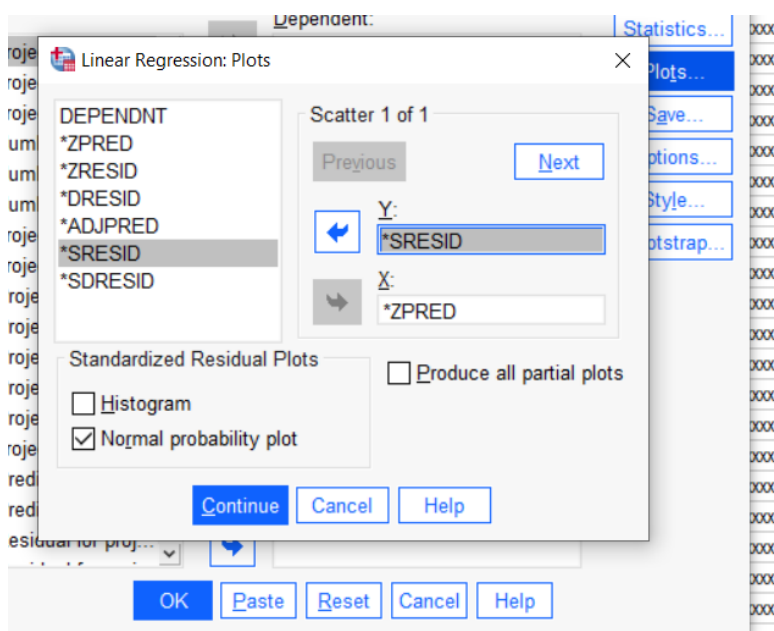


Figure 3 Plot settings

When these settings had been set, was it possible to run the analysis and get the results. The plots produced in the analysis are presented in figures 4 and 5 below.

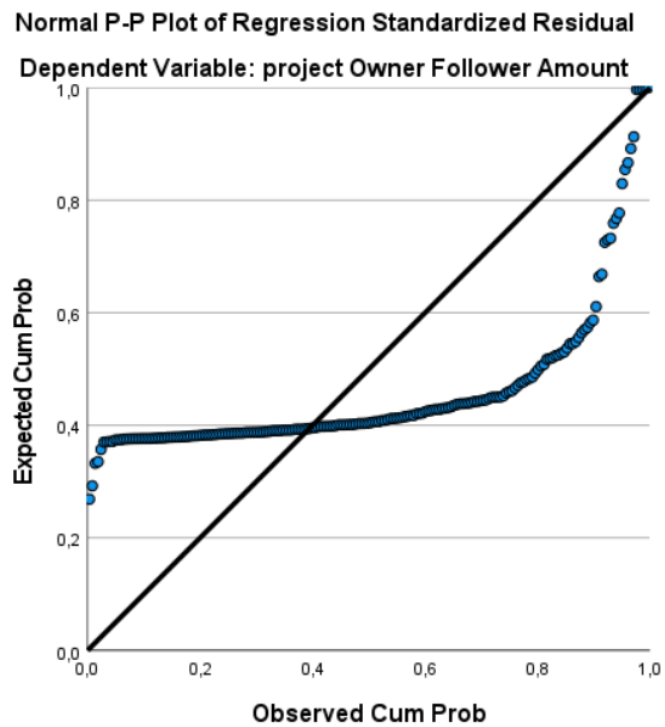


Figure 4 Plot of regression standardized residual

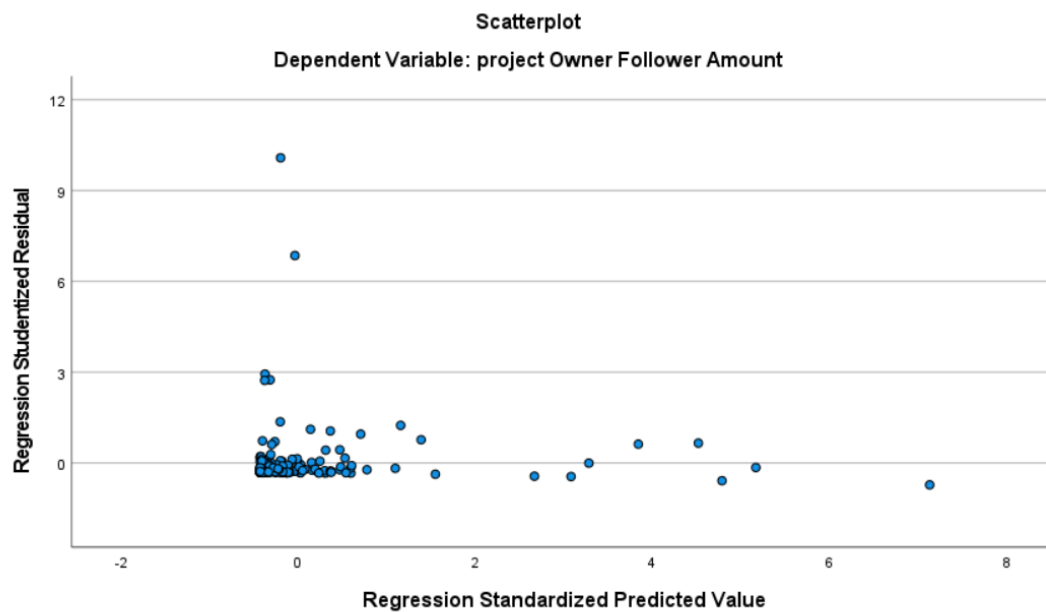


Figure 5 Scatterplot

Based on the plots can be determined that “project owner follower amount” and “project pleaded value” have a non-linear relationship because the data does not follow the linear model but instead

is closer to exponential model meaning that the pledged value increases more drastically when the project owner has a lot of followers versus when they have average amount of followers.

The second task in this week's assignment was to conduct a binary logistic analysis to investigate the effects of "project owner follower amount" and "the number of mentions of a project" on project success. This analysis was started by selecting binary logistic regression in SPSS as presented in figure 6.

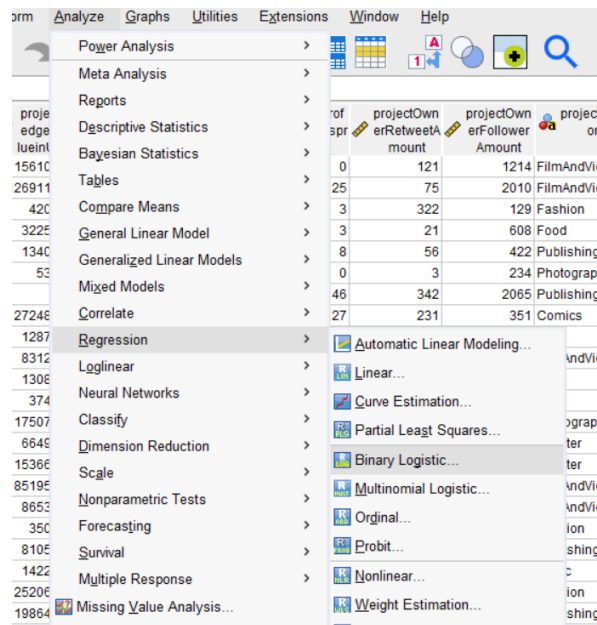


Figure 6 Start of the second analysis

In the analysis the dependent variable was project success and predictors were "project owner follower amount" and "the number of mentions of a project". Project success is selected as the dependent variable because according to the first assumption of binary logistic regression the dependent variable must be binary or a dichotomous scale.

The next step of the analysis was once again to set the settings. This time we set the options and save settings. These settings can be seen in figures 7 and 8.

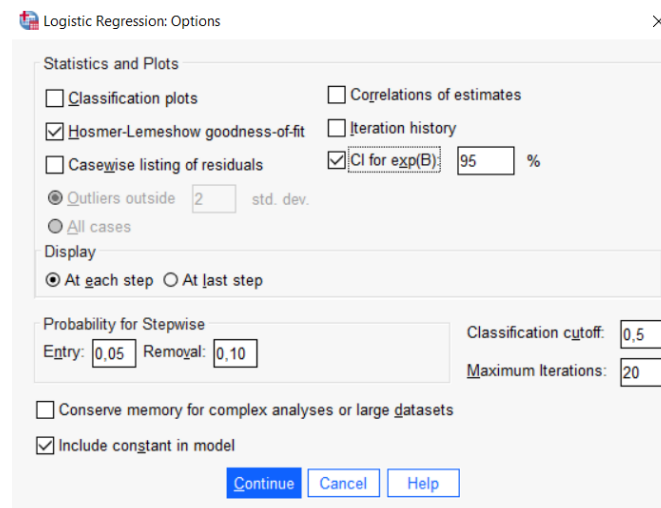


Figure 7 Options settings

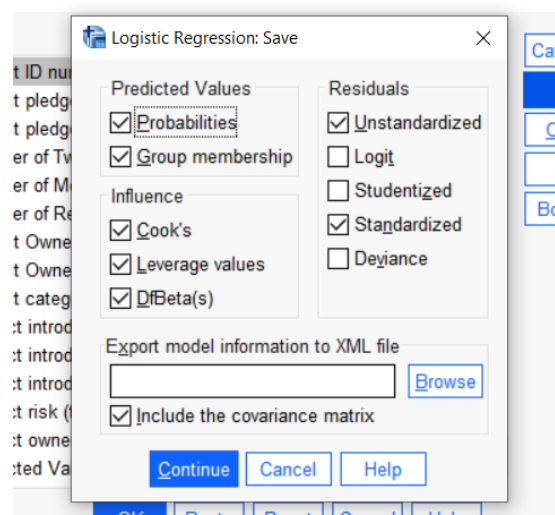


Figure 8 Save settings

After these settings were set was the analysis itself conducted. This produces a bunch of tables showing different aspects of the analysis. These tables are presented in figures 9, 10, 11 and 12.

Classification Table^a

| Observed | | Predicted | | Percentage Correct |
|--------------------|-----------------------------|-----------------------------|---------|--------------------|
| | | project success or not Fail | success | |
| Step 1 | project success or not Fail | 0 | 61 | ,0 |
| | success | 0 | 135 | 100,0 |
| Overall Percentage | | | | 68,9 |

^a The cutvalue is .500

Figure 9 Classification table

| Variables in the Equation | | | | | | | | | |
|---------------------------|---------------------------------|------|------|-------|----|------|--------|---------------------|-------|
| | | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I. for EXP(B) | |
| | | | | | | | | Lower | Upper |
| Step 1 ^a | project Owner Follower Amount | ,000 | ,000 | 2,750 | 1 | ,097 | 1,000 | 1,000 | 1,000 |
| | Number of Mentions of a project | ,000 | ,000 | ,425 | 1 | ,515 | 1,000 | 1,000 | 1,001 |
| | Constant | ,561 | ,188 | 8,890 | 1 | ,003 | 1,753 | | |

a. Variable(s) entered on step 1: project Owner Follower Amount, Number of Mentions of a project.

Figure 10 Variables in the Equation

| Omnibus Tests of Model Coefficients | | | | |
|-------------------------------------|-------|------------|----|------|
| | | Chi-square | df | Sig. |
| Step 1 | Step | 5,979 | 2 | ,050 |
| | Block | 5,979 | 2 | ,050 |
| | Model | 5,979 | 2 | ,050 |

| Model Summary | | | |
|---------------|----------------------|----------------------|---------------------|
| Step | -2 Log likelihood | Cox & Snell R Square | Nagelkerke R Square |
| 1 | 237,091 ^a | ,030 | ,042 |

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

| Hosmer and Lemeshow Test | | | |
|--------------------------|------------|----|------|
| Step | Chi-square | df | Sig. |
| 1 | 23,750 | 8 | ,003 |

Figure 11 Omnibus tests of model coefficients, model summary and Hosmer and Lemeshow test

| Contingency Table for Hosmer and Lemeshow Test | | | | | | |
|--|----|-------------------------------|----------|----------------------------------|----------|-------|
| | | project success or not = Fail | | project success or not = success | | Total |
| | | Observed | Expected | Observed | Expected | |
| Step 1 | 1 | 15 | 7,239 | 5 | 12,761 | 20 |
| | 2 | 7 | 7,168 | 13 | 12,832 | 20 |
| | 3 | 8 | 7,103 | 12 | 12,897 | 20 |
| | 4 | 3 | 7,016 | 17 | 12,984 | 20 |
| | 5 | 8 | 6,862 | 12 | 13,138 | 20 |
| | 6 | 8 | 6,644 | 12 | 13,356 | 20 |
| | 7 | 3 | 6,418 | 17 | 13,582 | 20 |
| | 8 | 2 | 5,845 | 18 | 14,155 | 20 |
| | 9 | 5 | 4,922 | 15 | 15,078 | 20 |
| | 10 | 2 | 1,783 | 14 | 14,217 | 16 |

Figure 12 Contingency table for Hosmer and Lemeshow test

From these tables would it be possible to answer the question whether the predictors effect the dependent variable. This finishes my binary logistics regression analysis.