

Jan Hincz

Student number: 465718

Krzysztof Kuba

Student number: 456361

Tymoteusz Mętrak

Student number: 429741

Wojciech Szymczak

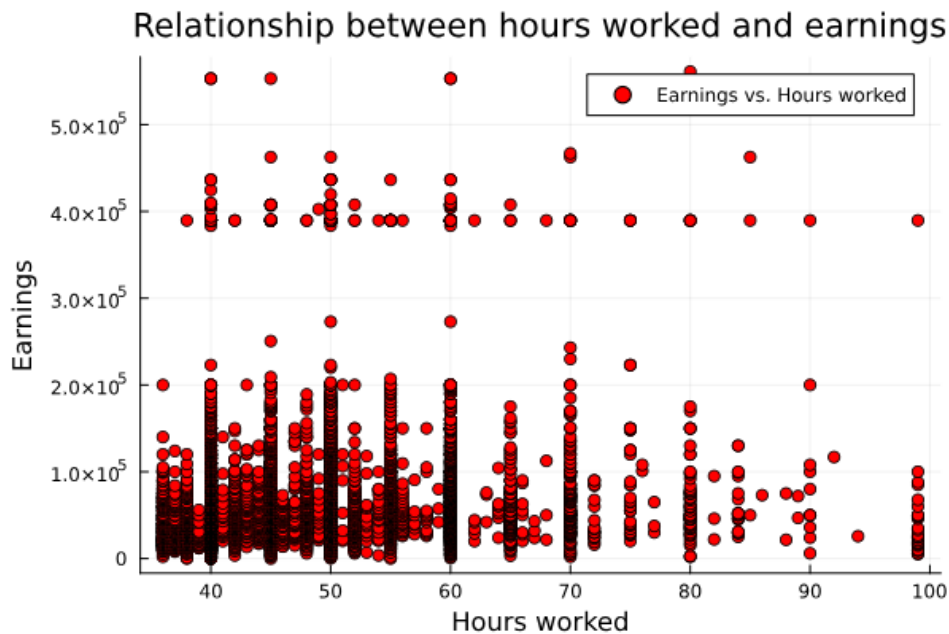
Student number: 424178

Quantitative Economics – Write-up for Problem Set 1

Warsaw, October 2024

Problem 7: Import the Data, Plot It, and Calculate Correlations

5. Analyze the results: discuss the findings

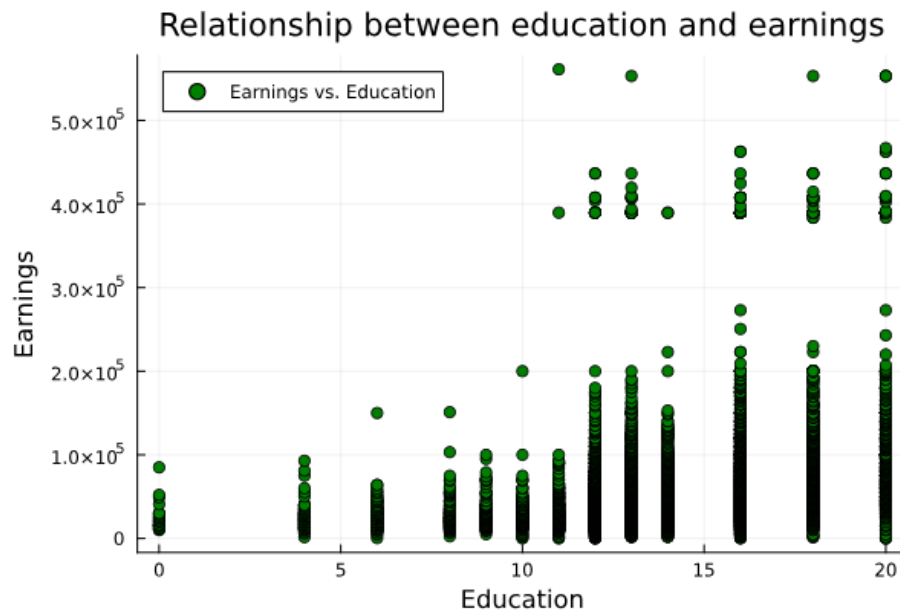


There is a small positive correlation between hours worked and earnings at ca. 0.25.

There seems to be a minute average increase of earnings with hours worked until the hours worked reach 70, then the earnings seem to slightly decrease. Numerically, the correlation is low also because e.g. people earning slightly below 4.0×10^5 seem to work various hours which are distributed roughly uniformly.

Additionally, we might suspect a feedback loop between x = hours worked and y = earnings, which would result in an inconsistent estimate of the effect of x on y in a hypothetical simple regression.

Ceteris paribus, people who work more should earn more as they exhibit more effort, but on the other hand leisure is a normal good for many people who would like to work less if they earned more.



The correlation between education and earnings is stronger and more straightforward – though still relatively mild at around 0.39.

It may be concluded that the more years of formal education, the more person earns. This is probably because additional education allows to gain incremental knowledge and skills valuable to some extent in the labour market. Moreover, additional passed years of education signal intrinsic ability.

All in all, this is probably not the best method to calculate the correlation, because education is a discrete variable. It may make more sense to compare means of earnings between groups of people with given years of education.

It is important to bear in mind that correlation is a linear measure of association, but both examined associations could be of non-linear variety. For both of the above relations more formal analysis (e.g. multivariate regression) is advised in order to draw more reliable conclusions