

Deadline 3 November 2020

For this assignment you need to use data frame wage collected in the March 1988 by the US Census Bureau. Here, wage is the wage in dollars per week, education and experience are measured in years, and ethnicity is a factor with levels Caucasian (cauc) and African-American (afam). There are three further factors, smsa, region, and parttime, indicating residence in a standard metropolitan statistical area (SMSA), the region within the United States of America, and whether the individual works part-time. Note that experience has been calculated as: age - education - 6. This quantity may become negative, which explains the 438 observations with this property in the wage data.

The aim of the analysis is to model the response variable wage as a function of the variables described above.

Using these data, write a report addressing the following points:

1. Justification of the chosen regression model specification. [5 marks]
2. Using the final model, provide a summary (e.g., using tables and figures) of the empirical findings as well as interpretation of the estimated model parameters. [5 marks]
3. Provide recommendations and limitations of your analysis. [5 marks]
4. What did you learn from the analysis? What is the answer, if any, to the questions you set out to address? Which recommendations can you give? How can the analysis be improved? [5 marks]

The remaining 5 marks will be for presentation.

The report must be in pdf format. It (excluding the title page) must not be longer than 6 pages (including graphs, tables, etc.) using font size 12pt with one and a half line spacing and at least 2.5 cm margin.