



Question 3

Not yet answered

Marked out of 15.00

Linear Models:

You are asked to analyse a subset of the `CPS1985` dataset (you could check `?CPS1985` after loading the R package `AER` - this is not necessary now!). The sample consists of wages of persons and relevant variables to explain wages. The following variables are to be considered:

- `wage` (hourly wage of the person, in US Dollar),
- `age` (age, in years),
- `education` (education, in years),
- `experience` (work experience, in years),
- `region` (region, factor variable with 2 values),
- `union` (union membership, factor variable with 2 values),
- `gender` (gender, factor variable with 2 values).

Your tasks (please do save all your R code to upload it with your solutions):

- Download the dataset `cps-9538.csv` (click to download). Note, that this is an English .csv file.
- Estimate the linear regression models `m1` - `m5` described in the following.
- Find the answers for a. - e. and enter them directly.
- Provide solutions to the additional questions f. - i., either by writing them into the text input field below or as comments within your R code.
- Finally, save your R code (into an R script or R Markdown file) and save it in the upload area at the end of this page.

Models to estimate

(Please use `wage` as given, no transformation.)

- Model `m1` : Explain `wage` by `experience`.
- Model `m2` : Explain `wage` by `experience` and `experience^2`.
- Model `m3` : Explain `wage` by `experience`, `education` and `union`.
- Model `m4` : Explain `wage` by `experience`, `experience^2`, `education` and `union`.
- Model `m5` : Explain `wage` by using all explanatory variables in the dataset.

Additional questions

- For models `m1` and `m2` do a scatterplot of the data and draw the estimated regression curves. Interpret the curves: How would you describe the effect of `experience` on `wage`?
- Compare `m1` and `m2` using an F test. Comment the result (also reflecting your interpretation before): which model would you choose?
- Explain using model `m3`: What is the effect of the different values of `union` on `wage`?
- Find the best linear regression model for explaining `wage` based on AIC.

Give the following answers right here. For c. give the numerical values with at least 2 decimal places.

- The sample size equals:
- How many persons from the south (`region`) in the sample are no union members?
- With `m3` do predict the expected wage for a person with 12 years of education and 15 years of experience who is not a union member :
- Which of the five models has the largest coefficient of determination? Answer by entering `m1`, `m2`, `m3`, `m4` or `m5`:
-

Marked out of 14.00

A screenshot of a web-based file manager interface. At the top left, there's a small icon of a document. To its right are three icons: a grid of dots, a list view icon, and a folder icon. Below these icons, the word "Dateien" is written in a teal color. The main area of the interface is a large rectangle with a dashed border. Inside this rectangle, centered at the bottom, is the text "You can drag and drop files here to add them." in a teal color.

Uploads of additional files on September 29 /
Additional comments ►