

Linear regression - exercises

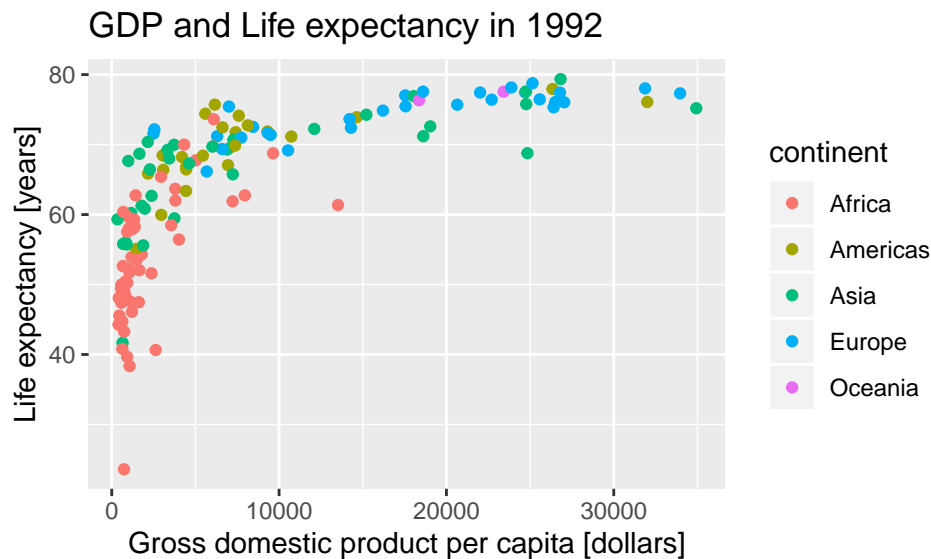
E. Pastucha

April 2020

Here are 4 exercises to solidify your knowledge about linear regression.

1. Investigating populations

The gapminder dataset contains information about expected life length, population size, gross domestic product for many countries in the timespan from 1952 to 2007. In this question you will have to investigate certain relations seen in the data from the year 1992.



Make a linear model of the expected life length in Europe, using gdp per capita as the input variable. Evaluate the model against linear regression conditions.

2. Cats body and heart weights

In this section the *catsM* dataset from the *boot* package is examined. The dataset contains measures of body and heart weight of a set of male cats.

Fit a linear model to the data and describe the fitted model. Are the assumptions for the linear model satisfied?

3. Average monthly temperatures from 1901 to 2015

The data in the file *temperatures.csv* is the basis for this exercise.

Make a linear fit to the average temperature in Denmark from 1975 to 2015. Check the assumptions (conditions) of the fit. Predict the average temperature for the year 2030.

4. Biogas usage over time

The data in the file *data/biogas_usage.csv* is the basis for this exercise.

- a) Fit a linear model to the biogas consumption. Use daynumber as the predictor.
- b) Predict the biogas consumption for day number 3650
- c) Locate the day with the highest consumption of biogas each year