

MATH1712 Probability and Statistics II

Homework 2

<http://www1.maths.leeds.ac.uk/~voss/2017/MATH1712/>

Jochen Voss, J.Voss@leeds.ac.uk

2017/18, semester 2

*This homework sheet does not correspond to a tutorial class. Solutions must be submitted online before the deadline of **Monday, 12th February, 4pm.***

Submit your answers here: <https://goo.gl/forms/btNXWdda6VIZa07L2> (after 4th February)

The topics of this homework are loading data into R, and using R to compute simple summary statistics. Help is available on the back of homework sheet 1, and in the section “Handling Data in R” in the lecture notes. The exercises below refer to a number $\kappa_4 \in \{1, \dots, 10000\}$. This number represents the last four digits of your student ID, plus one. For example, if your student ID is 200123456, then $\kappa_4 = 3457$.

Exercise 4. Data extracted from the handwriting samples I collect in recent lectures are available at <https://www.seeuhn.de/maths/letters>. The data are split into two groups, a “training data set” and a “test data set”. Following the instructions on the web page, load the training data set into R.

This data set contains samples $1, \dots, n$, where $n = 22138$. Each sample represents a scanned image, together with a “label” (A to Z). Which letter does the sample κ_4 represent?

Exercise 5. The image encoded in the sample κ_4 consists of 54 rows and 32 columns of pixels. Each pixel has an intensity value, ranging from 0 to 255. What is the intensity of the pixel in row 10, column 10 for the sample κ_4 ? (You may need to carefully read the web page to answer this question.)

Exercise 6. Continuing from the previous question, what is the average pixel intensity for sample κ_4 ?

Exercise 7. Continuing from the previous question, what is the standard deviation of the pixel intensities for sample κ_4 ?

Exercise 8. Continuing from the previous question, how many pixels in sample κ_4 have an intensity strictly greater than 127?