Assignment 5, part 2

Plotting in R

1. Basic plotting

   (a) Points. Create the vectors
   
   \[ x \leftarrow \text{rnorm}(25) \]
   \[ y \leftarrow x \times \text{rnorm}(25, \text{mean}=2, \text{sd}=0.5) \]

   and draw a scatter plot of \( x \) versus \( y \). If you tried to draw a straight line through your scatter plot, would it fit the points? Do you expect that it should? Why or why not?

   (b) Lines and polygons. Plot the density function for a standard normal distribution (i.e., a normal distribution with \( \mu = 0 \) and \( \sigma = 1 \)) using the \texttt{dnorm} function. Draw a dashed, vertical line at the critical value corresponding to \( \alpha = 0.05 \) using either \texttt{lines} or \texttt{abline}. Fill in the area under the curve in the region \([c_\alpha, \infty)\) using \texttt{polygon}. Draw a text label near the shaded area reading “Area = 0.05” using \texttt{text} and \texttt{quote}. Draw a line connecting the label to the shaded area using \texttt{lines}.

2. Letter frequency

   (a) Download the letter frequency dataset from the course website and convert the data frame into a matrix using

   \[ \text{let.freq} \leftarrow \text{as.matrix(read.csv2("let-freq.csv", rownames=1))} \]

   (b) Plot four letter frequency histograms for English, Spanish, French and German in the same plot. Use appropriate plot titles, \( x \)-axis labels and \( y \)-axis labels for each of the four frequency plots.

   (c) Plot a heatmap for A-Z letter frequency for all 12 languages. Use appropriate \( x \)- and \( y \)-axis tick labels (i.e., use letters for row ticks and language names for column ticks). This may require suppressing the usual axes and plotting them yourself using \texttt{axis}. 