

### Exercise 6

Suppose for a categorical variable with levels A, B, C, D and E you observe the data below:

E D A D C D B E C E D B B C B A B C C C A B C A A C A E E A

- (a) Enter the data into R (as factor variable) and compute the absolute as well as the relative frequency distribution (e.g. with the R function `table()`).
- (b) Illustrate the obtained frequencies using barplots and piecharts.
- (c) Also try out the function `qplot()` in the package `ggplot2`.

### Exercise 7

Check the dataset `HairEyeColor` and visualize the distribution of hair color with barplots for each eye color separately. Start by looking at women and men jointly, and in a second step investigate the situation for both genders separately. What conclusions can you draw?

### Exercise 8

In what follows, we will use the dataset `agriculture` from the R package `cluster`. For the below stated tasks also employ plot functions from the packages `lattice` and/or `ggplot2` if appropriate.

- (a) Check the helping page to get some more information on the data.
- (b) Visualize the relationship of `x` and `y` with a scatter plot.
- (c) Add another variable/column `region` to the data frame saying whether the respective country is rather located in northern or southern Europe (you can decide whether a country is southern or northern based on your individual point of view).
- (d) Draw boxplots of `x` and `y` for each `region` separately. Add appropriate titles to the plots. What conclusions can you draw?
- (e) Draw scatter plots of `y` vs. `x` for each `region` separately. What conclusions can you draw?

### Exercise 9

Delete the variable `region` from the dataset `agriculture` you defined in the previous Exercise and replace the variable names `x` and `y` by more informative ones. Check the help pages for information on the functions `plot()` and `text()`, and draw scatterplots of (former) `y` vs. `x` with (abbreviated) country names instead of simple points.