Practice I

a)

A dataset from mouse experiments at 18 weeks is available in the file mice_data.csv in the dataset folder. Let's explore the dataset to see what it contains.

1. Open a new script file in R studio, comment it and save it.

2. Have look at the csv file in R studio's file explorer. What do you need to check in order to be able to read in the file correctly ?

3. Read the file into R, assign its content to object "mice_data". Examine the object.

4. How many observations and variables does the dataset have ?

5. What is the structure of the dataset? What are the names and classes of the variables ?

b)

1. Which variables appear to be categorical ? Convert them to factors.

2. Get the summary statistics of "mice_data"

3. Use the function table() to compute the number of observations in different mouse groups.

3.1. How many mice are included of each genotype (WT, KO) ?

3.2. How many mice are included per diet (HFD, CHOW) ?

3.3. Make a 2x2 table by genotype and diet crossed.

c)

1. Subsets

1.1. Isolate the observations for the mice on high fat diet (HFD) using subset().

1.2. Compute the average weights of the subset.

1.3. Do the same for the mice on regular chow diet (CHOW).

1.4. Export the data of each subgroup to a csv file.

2. Look at the results from the two previous exercises. What does this initial exploration of the data suggest about mouse weights ?

3. Optional: Compute the means and standard deviations for WT and KO mouse weights using tapply(). Then do the same for CHOW and HFD groups.