R homework excercises

Ex 1:

Using the NHANES data provide in the R package “NHANES”, generate two tables similar to Table 1 and Table 2 of the publication “The Cardiometabolic Risk Profile of Young Adults With Diabetes in the U.S.” (Diabetes Care, 2019). Calculate frequencies, means, standard deviations, but also medians, quartiles , and min and max. Use the package “tableone” for this purpose and export/save the resulting tables in a word/html/etc. file which can then be used for further editing.

Ex 2:

Same task as in Ex 1, but this time use the package “table1”.

Ex 3:

“Sleepdata” contains data of 50 individuals who reported on their habitual bed time and wake time on free days and work days (columns self\_BT\_Work, self\_WT\_Work, self\_BT\_Free, self\_WT\_Free), as well as actigraphy-measured bed and wake time for 8 consecutive days (columns 2 to 17; information on work or free day not known).

* Calculate the self-reported sleep duration, and sleep midpoint on work days and on free days.
* Calculate the mean (over all eight days) actigraphy-measured bed time, wake time, sleep duration, and sleep midpoint.
* Display the results in a table (use “table1” or “tableone” package).

Furthermore, restructure the dataframe from “wide format” (i.e. one line per participant) into “long format” (i.e. one line per day of actigraphy, i.e. 8 lines per participant), and plot bed time and wake time of one participant over the eight days using ggplot2.

### Covid-19 data from Austria is publicly available on <https://www.data.gv.at/covid-19/> (we need in particular the datasets “COVID-19: Zeitliche Darstellung von Daten zu Covid19-Fällen je Bundesland“, “[COVID-19: Daten Covid19-Fällen je Altergruppe](https://www.data.gv.at/katalog/dataset/3765ed62-0f9d-49ad-83b0-1405ed833108)”, and [COVID-19: Daten zur Auslastung in Spitälern und Testergebnissen](https://www.data.gv.at/katalog/dataset/846448a5-a26e-4297-ac08-ad7040af20f1)).

### Ex 4:

### Using the Covid-19 data, produce (using ggplot2) a graph as similar as possible to the following graph from the COVID Austria dashboard (https://covid19-dashboard.ages.at/):



Produce also a similar graph, but with cases per week (instead of per day), and save it as a pdf file.

Ex 5:

Plot (using ggplot2) the number of cases per 100000 persons (to make states of different size better comparable) per state (Tirol, Salzburg, etc.). Save it as a pdf file. The results should like something like this:



Ex 6:

Using the Covid-19 data, calculate (using ggplot2) the proportion of positive tests (i.e. positives/totals) per week and graph it (as a line chart). Save it as a pdf file.

Ex 7:

Using the Covid-19 data, produce (using ggplot2) plots as similar as possible to the following graph from the COVID Austria dashboard (<https://covid19-dashboard.ages.at/>). Save it as a pdf file.



Ex 8:

Simulate the draw of 20 independent random numbers X and Y from a standard normal distribution (R command (hint: R command “rnorm”). Repeat this simulation 100 times.

The resulting dataset should look like this (the dataset should have 80 rows):

Experiment Draw X Y

1 1 xx yy

1 2 xx yy

…

1 20 xx yy

2 1 xx yy

…

2 20 xx yy

…

4 20 xx yy

Draw a scatterplot (use ggplot2) for each experiment, and combine them in one plot with 4 panels. Save it as a pdf file.