Biostatistics & Epidemiological Data Analysis using R

2

Times and Dates in R

Stefan Konigorski

Health Intervention Analytics Group, HPI

October 26, 2022

HPI Hasso Plattner Institut
Digital Engineering · Universität Potsdam

Learning objective

- Get to know functions for importing date and time variables in R, and for working with them.

Overview

General

- Beside numeric, character, factor, and logical variable classes in R, dates are another class that you may sometimes encounter (e.g. in NoShow data in homework 1, exercise 7).
- There are different options how to work with dates and times in R, and different date/time classes and functions.

Overview

- The `as.Date()` function handles dates (without times).
- The function `chron()` in the package `chron` handles dates and times, but not time zones.
- The `POSIXct` (and `POSIXlt`) classes allow for dates and times with control for time zones.

Details and Suggestions

Details

- The `as.Date()` function creates objects of the class `Date`, which count the number of days since January 1970-01-01.
- The `POSIXct` classes count the number of seconds since the beginning of 1970 (in the UTC time zone).
- Exercise: check with `class()` and `as.numeric()` functions!

Suggestions

Use the simplest representation possible:

- Work with `as.Date()` if you only have dates but no times, ...
- Or: work with the format you are given when data is automatically available/imported in a given date/time format.

Import and Transform

Use

- When reading in the NoShow csv dataset with the `readr::read_csv()`, it automatically detects dates/times and transforms it to `POSIXct` vectors.
- The `as.Date()` and `as.POSIXct()` functions can be used to transform character (and POSIXct) vectors to `Date` or `POSIXct` classes.

Why bother?

- When date/time variables are in Date formats, you can use them easily in calculations, e.g. substract times.

Exercise 4

- See `R_2_exercises.Rmd`.

Questions?