Dynamic Reporting with knitr

Felix Hofmann, Samuel Pawel

9 November 2022

In this document we will give an example of dynamic reporting using a data set containing the names of newborn children in the city of Zurich since the year 1993. The data were downloaded from `https://data.stadt-zuerich.ch/dataset/bev_vornamen\_baby\_od3700`.

Import data and perform manipulations

We will now import the data and perform some manipulations, such as rename the variables into English and computing the most common name by year and sex.

```r
## load data
dat <- read.csv(file = "../data/namesZurich.csv")

## data manipulations
library(dplyr)
datCleaned <- dat %>%
    rename(year = StichtagDatJahr,
           name = Vorname,
           sex = SexLang,
           births = AnzGebuWir) %>%
    group_by(year, sex) %>%
    top_n(n = 1, wt = births)
```
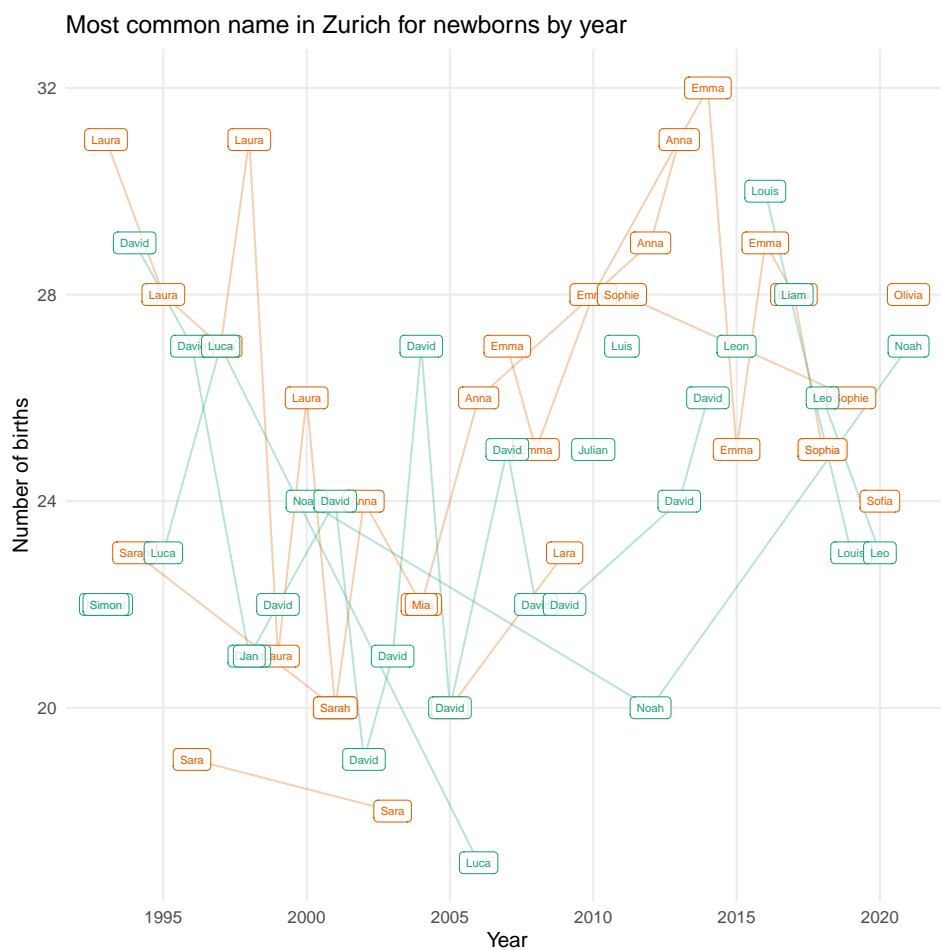
Data visualization

We will now visualize the most common name by year and sex.

```r
library(ggplot2)
ggplot(data = datCleaned, aes(x = year, y = births, color = sex)) +
    geom_line(aes(group = name), alpha = 0.3, show.legend = FALSE) +
    geom_label(aes(label = name), alpha = 0.8, show.legend = FALSE,
```

```
              size = 2) +
    scale_color_brewer(palette = "Dark2") +
    scale_x_continuous(breaks = seq(1995, 2020, 5)) +
    labs(x = "Year", y = "Number of births",
         title = "Most common name in Zurich for newborns by year") +
    theme_minimal() +
    theme(panel.grid.minor = element_blank())
```

Most common name in Zurich for newborns by year

Software versions and computational environment

It is good practice to save details on software versions and computational
environment when creating dynamic reports.

```
sessionInfo()

## R version 4.2.2 Patched (2022-11-10 r83330)
```

```
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 20.04.5 LTS
##
## Matrix products: default
## BLAS:   /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.9.0
## LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.9.0
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=de_CH.UTF-8        LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=de_CH.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=de_CH.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=de_CH.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] ggplot2_3.4.0 dplyr_1.0.10
##
## loaded via a namespace (and not attached):
##  [1] knitr_1.40      magrittr_2.0.3    munsell_0.5.0     tidyselect_1.2.0
##  [5] colorspace_2.0-3 R6_2.5.1         rlang_1.0.6       fansi_1.0.3
##  [9] stringr_1.4.1   highr_0.9         tools_4.2.2       grid_4.2.2
## [13] gtable_0.3.1    xfun_0.34        utf8_1.2.2        cli_3.4.1
## [17] DBI_1.1.3       withr_2.5.0      assertthat_0.2.1  tibble_3.1.8
## [21] lifecycle_1.0.3 farver_2.1.1     RColorBrewer_1.1-3 vctrs_0.5.0
## [25] glue_1.6.2      evaluate_0.18    labeling_0.4.2    stringi_1.7.8
## [29] compiler_4.2.2  pillar_1.8.1     scales_1.2.1      generics_0.1.3
## [33] pkgconfig_2.0.3
```