

August 8, 2022

The results below are generated from an R script.

```
#### Detection des anthroponymes ####
l_anthroponymes <- NULL
caps <- NULL
for (i in df_main$rente)
{
  v_anthroponymes <- ren_extract(i)
  k <- i
  for (j in v_anthroponymes) {
    k <- str_replace_all(k,j,str_to_upper(j))
  }
  caps <- c(caps,k)
  l_anthroponymes <- c(l_anthroponymes,v_anthroponymes)
}
l_anthroponymes <- unique(l_anthroponymes)
l_anthroponymes <- l_anthroponymes[!is.na(l_anthroponymes)]
df_main$rente <- caps
l_anthroponymes <- str_to_upper(l_anthroponymes)

#### calcul de distance ####
#m_distance <- myDamereauLevenstheinDist(l_anthroponymes)

#### clustering ####
#l_cluster <- myClustering(l_anthroponymes,clustering_lim=3,m_distance)
load("./export/l_cluster_corr.RData")

#### Remplacement des variants ####
for(i in 1:length(df_main$rente)){
  for (j in l_cluster) {
    for(k in j){
      if(!is.na(df_main$rente[i])){
        if (str_detect(df_main$rente[i],k)) {
          df_main$rente[i] <- str_replace_all(df_main$rente[i],k,j[1])
        }
      }
    }
  }
}
}
```

The R session information (including the OS info, R version and all packages used):

```
sessionInfo()

## R version 4.0.3 (2020-10-10)
```

```
## Platform: x86_64-apple-darwin17.0 (64-bit)
## Running under: macOS 12.3.1
##
## Matrix products: default
## LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib
##
## locale:
## [1] fr_BE.UTF-8/fr_BE.UTF-8/fr_BE.UTF-8/C/fr_BE.UTF-8/fr_BE.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices datasets  utils      methods    base
##
## other attached packages:
## [1] RColorBrewer_1.1-3 concaveman_1.1.0 ggforce_0.3.3 scales_1.2.0
## [5] ggrepel_0.9.1 readxl_1.3.1 tidygeocoder_1.0.5 ggraph_2.0.5.9000
## [9] ggmap_3.0.0 igraph_1.3.0 comparator_0.1.2 forcats_0.5.1
## [13] dplyr_1.0.9 purrr_0.3.4 readr_2.1.2 tidyr_1.2.0
## [17] tibble_3.1.8 ggplot2_3.3.6 tidyverse_1.3.1 stringr_1.4.0.9000
##
## loaded via a namespace (and not attached):
## [1] bitops_1.0-7 fs_1.5.2 lubridate_1.8.0 httr_1.4.2
## [5] tools_4.0.3 backports_1.4.1 utf8_1.2.2 R6_2.5.1
## [9] DBI_1.1.2 colorspace_2.0-3 withr_2.5.0 sp_1.5-0
## [13] tidyselect_1.1.2 gridExtra_2.3 curl_4.3.2 compiler_4.0.3
## [17] cli_3.3.0 rvest_1.0.2 xml2_1.3.3 proxy_0.4-26
## [21] digest_0.6.29 jpeg_0.1-9 pkgconfig_2.0.3 highr_0.9
## [25] dbplyr_2.1.1 rlang_1.0.4 rstudioapi_0.13 farver_2.1.1
## [29] generics_0.1.3 jsonlite_1.8.0 magrittr_2.0.3 Rcpp_1.0.9
## [33] munsell_0.5.0 fansi_1.0.3 viridis_0.6.2 lifecycle_1.0.1
## [37] stringi_1.7.6 MASS_7.3-53 plyr_1.8.7 grid_4.0.3
## [41] crayon_1.5.0 lattice_0.20-41 graphlayouts_0.8.0 haven_2.4.3
## [45] hms_1.1.1 knitr_1.37 pillar_1.8.0 rjson_0.2.21
## [49] reprex_2.0.1 glue_1.6.2 evaluate_0.15 renv_0.15.4
## [53] modelr_0.1.8 png_0.1-7 vctrs_0.4.1 tzdb_0.2.0
## [57] tweenr_1.0.2 RgoogleMaps_1.4.5.3 cellranger_1.1.0 gtable_0.3.0
## [61] polyclip_1.10-0 clue_0.3-60 assertthat_0.2.1 xfun_0.30
## [65] broom_0.7.12 tidygraph_1.2.1 viridisLite_0.4.0 tinytex_0.37
## [69] cluster_2.1.0 ellipsis_0.3.2

Sys.time()

## [1] "2022-08-08 07:41:38 CEST"
```