R Notebook

library(wisdotcrashdatabase)  
library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.4 v purrr 0.3.4  
## v tibble 3.1.2 v dplyr 1.0.6  
## v tidyr 1.1.3 v stringr 1.4.0  
## v readr 1.4.0 v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

# old and new cntycode  
file\_loc = "C:/CSV/csv\_from\_sas/fst/"

# crash <- import\_db\_data(file\_loc, "crash", years\_old = c("16"), years = c("17"), columns = c("RLTNTRWY", "CRSHLOC", "INTTYPE", "INTDIS", "CRSHTIME"))  
  
# all=crash %>% get\_crash\_times(combine\_with\_old = TRUE) %>%  
# select(CRSHTIME, CRSHTIME\_GROUP, newtime, newtime\_old, CRSHDATE, newtime\_both)# %>%  
# # filter(lubridate::year(CRSHDATE) == 2016, is.na(newtime))  
# # all=crash %>% county\_rename(combine\_with\_old = TRUE)  
# all$newtime\_both %>% unique()  
# all %>% filter(is.na(newtime\_both))

person <- import\_db\_data(file\_loc, "person", years\_old = c("16"), years = c("17"), columns = c("AGE"))  
  
get\_age\_groups(person) %>% select(AGE,age\_group\_5yr )

## AGE age\_group\_5yr  
## 1: 18 15-19  
## 2: 26 25-29  
## 3: 56 55-59  
## 4: 15 15-19  
## 5: NA <NA>  
## ---   
## 576013: NA <NA>  
## 576014: NA <NA>  
## 576015: NA <NA>  
## 576016: NA <NA>  
## 576017: NA <NA>

age\_group\_5yr <- function(person\_df, age\_column = "AGE") {  
 person\_df <- person\_df %>% dplyr::mutate(age\_group\_5yr = cut(  
 # add age\_group column, 5 year intervals  
 .data[[age\_column]],  
 right = FALSE,  
 c(0, 4, 9, 14, 19, 24, 29, 34, 39, 44, 49, 54, 59, 64, 69, 120),  
 labels = c(  
 "0-4",  
 "5-9",  
 "10-14",  
 "15-19",  
 "20-24",  
 "25-29",  
 "30-34",  
 "35-39",  
 "40-44",  
 "45-49",  
 "50-54",  
 "55-59",  
 "60-64",  
 "65-69",  
 "70+"  
 ),  
 include.lowest = T  
 ))  
 # Get levels of age\_group factor and add Unknown  
 # levels <- levels(.data$person\_df$age\_group\_5yr)  
 # levels[length(levels) + 1] <- "UNKNOWN"  
 # # refactor agegroup to include "None" as a factor level  
 # # and replace NA with "None"  
 # person\_df$age\_group\_5yr <- person\_df %>%  
 # factor(.data$age\_group\_5yr, levels = levels)  
 # .data$person\_df$age\_group\_5yr[is.na(.data$person\_df$age\_group\_5yr)] <-  
 # "UNKNOWN"  
 person\_df  
}

year1=as.character(seq(88, 99, 1))  
year2=formatC(seq(00, 16, 1), digits = 0, width = 2, format = "f", flag = "0")  
year1

## [1] "88" "89" "90" "91" "92" "93" "94" "95" "96" "97" "98" "99"

year2

## [1] "00" "01" "02" "03" "04" "05" "06" "07" "08" "09" "10" "11" "12" "13" "14"  
## [16] "15" "16"

get\_list\_of\_years <- function(start\_year = "94",  
 end\_year = "27") {  
 # between 1985 and 2030  
 if (start\_year > 85 & end\_year < 30) {  
 year1 = as.character(seq(start\_year, 99, 1))  
 year2 = formatC(  
 seq(00, end\_year, 1),  
 digits = 0,  
 width = 2,  
 format = "f",  
 flag = "0"  
 )  
 return(c(year1, year2))  
 # between 1985 and 1999  
 } else if (start\_year > 85 & end\_year <= 99) {  
 return(as.character(seq(start\_year, end\_year, 1)))  
 } # between 2000 and 2030  
 else if (start\_year >= 0 & end\_year < 30) {  
 return(formatC(seq(start\_year, end\_year, 1), digits = 0, width = 2, format = "f", flag = "0"))  
 }  
}  
  
get\_list\_of\_years("89", "89")

## [1] "89"