

# Project 2 – Leading cause of deaths among individuals aged 15-24

MTH 161 – Fall 2024

Tehya Bailey

2024-11-12

## Part 1: choose a dataset and propose a question

Due Nov. 15th

For this project, I have chosen a dataset that explores the leading causes of death for individuals aged 15-24 in the United States. This led to my research question: “What are the leading causes of death among individuals aged 15-24 in the United States, and how have these causes changed over time?”

```
death_data_2022 <- tibble(  
  Cause_of_Death = c("Accidents", "Suicide", "Homicide", "Cancer", "Drug Overdose", "Heart D.  
  Death_Count = c(3500, 2800, 1500, 1200, 1500, 300, 1000),  
  Gender = rep(c("Male", "Female"), times = c(4, 3)),  
  Age_Group = rep("15-24", 7),  
)  
  
print(data)
```

```
function (... , list = character(), package = NULL, lib.loc = NULL,  
  verbose = getOption("verbose"), envir = .GlobalEnv, overwrite = TRUE)  
{  
  fileExt <- function(x) {  
    db <- grepl("\\.[^.]+"\\.(gz|bz2|xz)$", x)  
    ans <- sub(".*\\.", "", x)  
    ans[db] <- sub(".*\\.[^.]+"\\.(gz|bz2|xz)$", "\\1\\2",  
      x[db])  
    ans  
  }  
}
```

```

}
my_read_table <- function(...) {
  lcc <- Sys.getlocale("LC_COLLATE")
  on.exit(Sys.setlocale("LC_COLLATE", lcc))
  Sys.setlocale("LC_COLLATE", "C")
  read.table(...)
}
stopifnot(is.character(list))
names <- c(as.character(substitute(list(...))[-1L]), list)
if (!is.null(package)) {
  if (!is.character(package))
    stop("'package' must be a character vector or NULL")
}
paths <- find.package(package, lib.loc, verbose = verbose)
if (is.null(lib.loc))
  paths <- c(path.package(package, TRUE), if (!length(package)) getwd(),
    paths)
paths <- unique(normalizePath(paths[file.exists(paths)]))
paths <- paths[dir.exists(file.path(paths, "data"))]
dataExts <- tools:::make_file_exts("data")
if (length(names) == 0L) {
  db <- matrix(character(), nrow = 0L, ncol = 4L)
  for (path in paths) {
    entries <- NULL
    packageName <- if (file_test("-f", file.path(path,
      "DESCRIPTION")))
      basename(path)
    else "."
    if (file_test("-f", INDEX <- file.path(path, "Meta",
      "data.rds"))) {
      entries <- readRDS(INDEX)
    }
    else {
      dataDir <- file.path(path, "data")
      entries <- tools::list_files_with_type(dataDir,
        "data")
      if (length(entries)) {
        entries <- unique(tools::file_path_sans_ext(basename(entries)))
        entries <- cbind(entries, "")
      }
    }
  }
  if (NROW(entries)) {
    if (is.matrix(entries) && ncol(entries) == 2L)

```

```

        db <- rbind(db, cbind(packageName, dirname(path),
                             entries))
      else warning(gettextf("data index for package %s is invalid and will be ignored",
                           sQuote(packageName)), domain = NA, call. = FALSE)
    }
  }
  colnames(db) <- c("Package", "LibPath", "Item", "Title")
  footer <- if (missing(package))
    paste0("Use ", sQuote(paste("data(package = ", ".packages(all.available = TRUE)))",
                                "\n", "to list the data sets in all *available* packages."))
  else NULL
  y <- list(title = "Data sets", header = NULL, results = db,
           footer = footer)
  class(y) <- "packageIQR"
  return(y)
}

paths <- file.path(paths, "data")
for (name in names) {
  found <- FALSE
  for (p in paths) {
    tmp_env <- if (overwrite)
      envir
    else new.env()
    if (file_test("-f", file.path(p, "Rdata.rds"))) {
      rds <- readRDS(file.path(p, "Rdata.rds"))
      if (name %in% names(rds)) {
        found <- TRUE
        if (verbose)
          message(sprintf("name=%s:\t found in Rdata.rds",
                          name), domain = NA)
        thispkg <- sub(".*(/[~/*])/data$", "\\1", p)
        thispkg <- sub("_.*$", "", thispkg)
        thispkg <- paste0("package:", thispkg)
        objs <- rds[[name]]
        lazyLoad(file.path(p, "Rdata"), envir = tmp_env,
                  filter = function(x) x %in% objs)
        break
      }
    }
    else if (verbose)
      message(sprintf("name=%s:\t NOT found in names() of Rdata.rds, i.e.,\n\t%s",
                      name, paste(names(rds), collapse = ",")),
              domain = NA)
  }
}

```

```

files <- list.files(p, full.names = TRUE)
files <- files[grep(name, files, fixed = TRUE)]
if (length(files) > 1L) {
  o <- match(fileExt(files), dataExts, nomatch = 100L)
  paths0 <- dirname(files)
  paths0 <- factor(paths0, levels = unique(paths0))
  files <- files[order(paths0, o)]
}
if (length(files)) {
  for (file in files) {
    if (verbose)
      message("name=", name, ":\t file= ...", .Platform$file.sep,
              basename(file), ":\t", appendLF = FALSE,
              domain = NA)
    ext <- fileExt(file)
    if (basename(file) != paste0(name, ".", ext))
      found <- FALSE
    else {
      found <- TRUE
      switch(ext, R = , r = {
        library("utils")
        sys.source(file, chdir = TRUE, envir = tmp_env)
      }, RData = , rdata = , rda = load(file, envir = tmp_env),
      TXT = , txt = , tab = , tab.gz = , tab.bz2 = ,
      tab.xz = , txt.gz = , txt.bz2 = , txt.xz = assign(name,
        my_read_table(file, header = TRUE, as.is = FALSE),
        envir = tmp_env), CSV = , csv = , csv.gz = ,
      csv.bz2 = , csv.xz = assign(name, my_read_table(file,
        header = TRUE, sep = ";", as.is = FALSE),
        envir = tmp_env), found <- FALSE)
    }
    if (found)
      break
  }
  if (verbose)
    message(if (!found)
      "*NOT* ", "found", domain = NA)
}
if (found)
  break
}
if (!found) {
  warning(gettextf("data set %s not found", sQuote(name)),

```

```

        domain = NA)
    }
    else if (!overwrite) {
      for (o in ls(envir = tmp_env, all.names = TRUE)) {
        if (exists(o, envir = envir, inherits = FALSE))
          warning(gettextf("an object named %s already exists and will not be overwr
            sQuote(o)))
        else assign(o, get(o, envir = tmp_env, inherits = FALSE),
          envir = envir)
      }
      rm(tmp_env)
    }
  }
invisible(names)
}
<bytecode: 0x5eb9c8504b88>
<environment: namespace:utils>

```

### *Key Observations/Variables*

*Cause\_of\_Death:* The medical condition or event that caused death (e.g., “Accidents,” “Suicide,” “Homicide,” etc.)

*Death\_Count:* The number of deaths attributed to each cause

*Age\_Group:* The age group for which the data is reported (e.g., “15-19”, “20-24”)

*Year:* The year in which the deaths occurred.

*Gender:* The gender of the individuals (e.g., “Male,” “Female”)

### *Link*

<https://wisqars.cdc.gov/lcd/?o=LCD&y1=2022&y2=2022&ct=10&cc=ALL&g=00&s=0&r=0&ry=2&e=0&ar>

I will investigate which causes of death are most prevalent in this age group and explore any significant trends or shifts over the years. By analyzing this data, it can highlight the change in health trends and flaws in public health prevention. Knowing these aspects can raise awareness for people ages 15-24 years physically and behaviorally.