```
##########################################
### Author: Dr. Wei-Wen Hsu  #############
###        University of Cincinnati  ####
###        hsuwe@uc.edu           #####
### Latest modification: 11/17/2023  #####
##########################################

### This R code is for Blood serum data  ###



rm(list=ls(all=TRUE))


#install.packages("basictabler")
#install.packages("survey")
#install.packages("jtools")
#install.packages("remotes")
#install.packages("writexl")
#remotes::install_github("carlganz/svrepmisc")

#After the packages are downloaded, they need to be loaded. This needs to be done at the beginning of each R session.


library("haven")
library("survey")
library("jtools")
library("remotes")
library("svrepmisc")
library("readxl")
```

```
library("dplyr")

library(tidyverse)

library("writexl")

library(basictabler)

library(openxlsx)


# Litstream Data

#("C:\\Users\\hsuwe\\OneDrive - University of Cincinnati\\Desktop\\F\\Dr. Lynne - biomonitoring task\\Documents for calculat\\Task 16 PHOP Full extraction results from Litstream_10.25.23 cleanup.xlsx", sheet="")

path_litstream = "D:\\UC-OneDrive\\OneDrive - University of Cincinnati\\Desktop\\F\\Dr. Lynne - biomonitoring task\\Documents for calculations\\Task 16 PHOP Litstream data_MB data check_v3_MB2.xlsx";

#path_litstream = "C:\\Users\\hsuwe\\OneDrive - University of Cincinnati\\Desktop\\F\\Dr. Lynne - biomonitoring task\\Documents for calculations\\Task 16 PHOP Litstream data_MB data check_v3_MB2.xlsx";


## Location for the results

path_results ="D:\\UC-OneDrive\\OneDrive - University of Cincinnati\\Desktop\\F\\Dr. Lynne - biomonitoring task\\Results";

#path_results ="C:\\Users\\hsuwe\\OneDrive - University of Cincinnati\\Desktop\\F\\Dr. Lynne - biomonitoring task\\Results";



##

litsream_blood=read_excel(paste0(path_litstream), sheet='Blood-serum-plasma data')

litsream_blood_ulimits=read_excel(paste0(path_litstream), sheet="Blood-serum-plasma limits")


#View(litsream_blood)

#View(litsream_blood_ulimits)


names(litsream_blood)
```

```
names(litsream_blood_ulimits)



### Remove some letters from the variable names

names(litsream_blood_ulimits) = gsub(pattern =
"flexData:|initialQuestions:|matrix:|object_13:|object_17:|studyPopulations:|object_20:|Biomarker:|b
lood_limits:", replacement = "", x = names(litsream_blood_ulimits))

names(litsream_blood) = gsub(pattern =
"flexData:|initialQuestions:|matrix:|object_13:|object_17:|studyPopulations:|object_20:|Biomarker:",
replacement = "", x = names(litsream_blood))



#names(litsream_blood) = gsub(pattern = "IDX_matrix\\|object_13\\|$", replacement = "", x =
names(litsream_urine_HBM))



### preprocessing data

litsream_blood0 = litsream_blood
%>%mutate(statistics=ifelse(concentration=="Other",concentrationoth,concentration))

names(litsream_blood0)

litsream_blood1 = litsream_blood0 %>%

                select(studyId, studyLitstreamId, batch, year,location,sampyear,

                    samptype,dust, extract, blood_samptype, blood_normal,

                    blood_concunits, populationDescription, biomarker,

                    sampsize, percentfreq, statistics, value)


litsream_blood_ulimits0 = litsream_blood_ulimits%>%

                select(studyId,batch, biomarker, sampyear, location , dust ,blood_samptype,
blood_normal, blood_concunits, limit, value) %>%

                 filter(!is.na(value))

names(litsream_blood_ulimits0)
```

```
litsream_blood_ulimits0 = litsream_blood_ulimits0 %>% rename("blood_concunits_limits" =
"blood_concunits")


### obtain the list of studies

study_id_list_blood= names(table(litsream_blood1$studyId));

study_id_list_bloodL= names(table(litsream_blood_ulimits0 $studyId));



#########################################################

### Transform the data from long to wide      ######

#########################################################


### blood data


wide_blood_data = NULL;


for (study_id in study_id_list_blood){


litsream_blood2= litsream_blood1 %>% filter(studyId==study_id)


litsream_blood3 = litsream_blood2 %>% spread(statistics, value)


wide_blood_data = bind_rows(wide_blood_data,litsream_blood3);


cat("study ID=", study_id,"\n")


}
```

```
############################
### blood limits data   ####
############################

wide_bloodL_data =  NULL;
for (study_id in study_id_list_bloodL){

litsream_blood_ulimits1= litsream_blood_ulimits0 %>% filter(studyId==study_id)

litsream_blood_ulimits2 = litsream_blood_ulimits1 %>% spread(limit, value)

wide_bloodL_data = bind_rows(wide_bloodL_data,litsream_blood_ulimits2);

cat("study ID=", study_id,"\n")

}


#View(wide_blood_data)
names(wide_blood_data)
#View(wide_bloodL_data)
names(wide_bloodL_data)


##########################################
### merge Blood and blood Limits data  ####
##########################################

## select the needed variables for merging
```

```
wide_bloodL_data1 = wide_bloodL_data %>% select(studyId, biomarker, sampyear,
location,blood_concunits_limits,MDL, LOD, LOQ)



## merge and select the studies indicating "specific gravity"


blood_data =  wide_blood_data %>% left_join(wide_bloodL_data1,
by=c("studyId","biomarker","sampyear","location"))



#View(blood_data)


names(table(blood_data$studyId))  # the number of studies is 4.


table(blood_data$studyId, blood_data$biomarker) # a cross table to see how many studies for each
chemical



#####################################################

####### Save as a physical data in excel here  ####

#####################################################


write.xlsx(blood_data, file=paste0(path_results,"\\litstream_blood_wide.xlsx"),

      sheetName = "litstream_blood", colNames = TRUE, rowNames = F, append = FALSE)


#####################################################
```