

```

library(writexl)
library(readxl)
library(dplyr)
library(stringr)
library(dtplyr)
library(magrittr)
library(sjmisc)
library(striprtf)
library(splitstackshape)
library(pdftools)
library(openxlsx)
`%notin%` <- Negate(`%in%`)

#File names and locations
#These files change for each class

#Output file
outfile<-'P:/CPSC/PHBA 1213.xlsx'
#Chem and analog list
listfile<-'P:/CPSC/PHBA.xlsx'
listsheet1<-'CID CAS'
#Toxval
tfile1<-'P:/CPSC/toxval/comptox batch search_phba_11.09.2022.xlsx'
tsheet1<-'Toxval Details'
#OECD
ofile1<-'P:/CPSC/oecd/oecd qsar phba new analogs_11.09.2022.xlsx'
osheet1<-'Sheet1'
ofile2<-'P:/CPSC/oecd/oecd qsar toolbox phba.xlsx'
osheet2<-'Sheet1'
#WebICE
wfile1<-'P:/CPSC/WebIce/ice - phb Long Format data.xlsx'
wsheet1<-'Export'
#Pubchem bioassay & CTD
pfolder<-"P:/CPSC/Pubchem/PHBA"
#Location of IUCLID files
ifolder<-"P:/CPSC/IUCLID/PHBA"

#The rest of the files are constant
#Nothing needs to be changed from here on
#Template Excel
tempfile<-'P:/CPSC/All code revised 12_22/Template.xlsx'
#Comptox
cfile1<-'P:/CPSC/comptox/Vegosen and Martin 2020 - full supp.xlsx'
csheet1<-'Table S4 Score Records'
#PG
pgfile1<-'P:/CPSC/PG/pg_with_years_checked.xlsx'
pgsheet1<-'Sheet1'

#Webice ADME
wfile1<-'P:/CPSC/WebIce/adme_parameters.xlsx'
washeet1<-'Formatted_ADME_2'
#Webice Cancer
wcfile1<-'P:/CPSC/WebIce/cancer.xlsx'
wcsheet1<-'ForICE'

```

```

#Webice DART
wdfil1<-'P:/CPSC/WebIce/dart.xlsx'
wdsheet1<-'Chemicals'
#Webice Eye
wefil1<-'P:/CPSC/WebIce/eye_irritation.xlsx'
wesheet1<-'Chemicals'
#Webice Oral
wofil1<-'P:/CPSC/WebIce/Acute_Oral_Toxicity.xlsx'
wosheet1<-'Chemicals'
#Webice Skin Sensitization
wsnfil1<-'P:/CPSC/WebIce/Skin_Sensitization.xlsx'
wsnsheet1<-'Chemicals'
#Webice Skin Irritation
wskfil1<-'P:/CPSC/WebIce/Skin_Irritation.xlsx'
wksheet1<-'Chemicals'

#create skeleton dataframe
list<-read_excel(listfile,sheet=listsheet1)
attach(list)
list1<-cbind(list[,c('CID','CAS RN...2','INCHIKEY','SMILES')])
colnames(list1)=c("CID","CAS","Inchikey","SMILES")
list2<-cbind(list[,c('CID','CAS RN...3','INCHIKEY','SMILES')])
list2<-filter(list2,!is.na(list2[,2]))
colnames(list2)=c("CID","CAS","Inchikey","SMILES")
list3<-cbind(list[,c('CID','CAS RN...4','INCHIKEY','SMILES')])
list3<-filter(list3,!is.na(list3[,2]))
colnames(list3)=c("CID","CAS","Inchikey","SMILES")
tlist<-rbind(list1,list2,list3)
tlist<-tlist[order(tlist$CID,tlist$CAS,na.last=FALSE), ]
tlist[5:66]<-0
ncolt<-66
colnames(tlist)=c("CID","CAS","Inchikey","SMILES","animal","a_acute","a_syst",
"a_neuro","a_carc","a_muta","a_repro",

"a_irr","a_sens","a_endo","human","h_acute","h_syst","h_neuro","h_carc","h_muta",
"h_repro","h_irr",

"h_sens","h_endo","tox","t_habs","t_aabs","t_hmet","t_amet","t_vitro","t_pbpk",
"t_adme",

"exp","e_conn","e_noconn","qsar","qs_acute","qs_syst","qs_neuro","qs_carc",
"qs_muta","qs_repro",

"qs_irr","qs_sens","qs_endo","qual","ql_acute","ql_syst","ql_neuro","ql_carc",
"ql_muta","ql_repro",

"ql_irr","ql_sens","ql_endo","quan","qn_acute","qn_syst","qn_neuro","qn_carc",
"qn_repro","qn_sens",
"qn_endo","risk","r_nocanc","r_canc")

wb <- loadWorkbook(tempfile)
saveWorkbook(wb,outfile,overwrite = T)

```