

```

setwd(pfolder)
files <- list.files(pattern = "CTD")

all<-tlist

for(file in files){
all<-all[,1:ncolt]
a<-read_excel(file,sheet='A',col_types="text")
c<-read_excel(file,sheet='C',col_types="text")
d<-read_excel(file,sheet='D',col_types="text")
e<-read_excel(file,sheet='E',col_types="text")
f<-read_excel(file,sheet='F',col_types="text")

a2<-filter(a,!grepl( "[Object not found]", a$'# Input', fixed = TRUE) &
!grepl( "[No associated data found]", a$'# Input', fixed =
TRUE) &
!grepl( "Alligator", a$Organism, fixed = TRUE) &
!grepl( "Aves", a$Organism, fixed = TRUE) &
!grepl( "Azumapecten", a$Organism, fixed = TRUE) &
!grepl( "Callinectes", a$Organism, fixed = TRUE) &
!grepl( "Carassius", a$Organism, fixed = TRUE) &
!grepl( "Catostomus", a$Organism, fixed = TRUE) &
!grepl( "Chrysemys", a$Organism, fixed = TRUE) &
!grepl( "Ctenopharyngodon", a$Organism, fixed = TRUE) &
!grepl( "Daphnia", a$Organism, fixed = TRUE) &
!grepl( "Diplodus", a$Organism, fixed = TRUE) &
!grepl( "Eisenia", a$Organism, fixed = TRUE) &
!grepl( "Falco", a$Organism, fixed = TRUE) &
!grepl( "Fundulus", a$Organism, fixed = TRUE) &
!grepl( "Gadus", a$Organism, fixed = TRUE) &
!grepl( "Gallus", a$Organism, fixed = TRUE) &
!grepl( "Gammarus", a$Organism, fixed = TRUE) &
!grepl( "Gobiocypris", a$Organism, fixed = TRUE) &
!grepl( "Gymnogyps", a$Organism, fixed = TRUE) &
!grepl( "Ictalurus", a$Organism, fixed = TRUE) &
!grepl( "Larus", a$Organism, fixed = TRUE) &
!grepl( "Limecola", a$Organism, fixed = TRUE) &
!grepl( "Microgadus", a$Organism, fixed = TRUE) &
!grepl( "Monopylephorus", a$Organism, fixed = TRUE) &
!grepl( "Mytilus", a$Organism, fixed = TRUE) &
!grepl( "Oncorhynchus", a$Organism, fixed = TRUE) &
!grepl( "Oryzias", a$Organism, fixed = TRUE) &
!grepl( "Paracyclopina", a$Organism, fixed = TRUE) &
!grepl( "Phalacrocorax", a$Organism, fixed = TRUE) &
!grepl( "Pimephales", a$Organism, fixed = TRUE) &
!grepl( "Platichthys", a$Organism, fixed = TRUE) &
!grepl( "Psammechinus", a$Organism, fixed = TRUE) &
!grepl( "Pseudacris", a$Organism, fixed = TRUE) &
!grepl( "Pusa", a$Organism, fixed = TRUE) &
!grepl( "Salmo", a$Organism, fixed = TRUE) &
!grepl( "Solen", a$Organism, fixed = TRUE) &
!grepl( "Sparus", a$Organism, fixed = TRUE) &
!grepl( "Tigriopus", a$Organism, fixed = TRUE) &

```

```

!grepl( "Xenopus", a$Organism, fixed = TRUE))
c2<-filter(c,!grepl( "[Object not found]", c$'# Input', fixed = TRUE) &
!grepl( "[No associated data found]", c$'# Input', fixed =
TRUE) &
!grepl( "Alligator", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Aves", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Azumapecten", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Callinectes", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Carassius", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Catostomus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Chrysemys", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Ctenopharyngodon", c$'OrganismName^ID', fixed =
TRUE) &
!grepl( "Daphnia", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Diplodus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Eisenia", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Falco", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Fundulus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Gadus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Gallus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Gammarus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Gobiocypris", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Gymnogyps", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Ictalurus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Larus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Limecola", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Microgadus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Monopylephorus", c$'OrganismName^ID', fixed = TRUE)
&
!grepl( "Mytilus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Oncorhynchus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Oryzias", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Paracyclops", c$'OrganismName^ID', fixed = TRUE)
&
!grepl( "Phalacrocorax", c$'OrganismName^ID', fixed = TRUE)
&
!grepl( "Pimephales", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Platichthys", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Psammechinus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Pseudacris", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Pusa", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Salmo", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Solen", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Sparus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Tigriopus", c$'OrganismName^ID', fixed = TRUE) &
!grepl( "Xenopus", c$'OrganismName^ID', fixed = TRUE))
d2<-filter(d,!grepl( "[Object not found]", d$'# Input', fixed = TRUE) &
!grepl( "[No associated data found]", d$'# Input', fixed =
TRUE))
e2<-filter(e,!grepl( "[Object not found]", e$'# Input', fixed = TRUE) &
!grepl( "[No associated data found]", e$'# Input', fixed =
TRUE))
f2<-filter(f,!grepl( "[Object not found]", f$'# Input', fixed = TRUE) &

```

```

!grepl( "[No associated data found]", f$'# Input', fixed =
TRUE))

names(a2)[1] <- "CAS"
names(c2)[1] <- "CAS"
names(d2)[1] <- "CAS"
names(e2)[1] <- "CAS"
names(f2)[1] <- "CAS"

na<- a2 %>% group_by(CAS) %>%
  summarize(na=n())
nc<- c2 %>% group_by(CAS) %>%
  summarize(nc=n())
nd<- d2 %>% group_by(CAS) %>%
  summarize(nd=n())
ne<- e2 %>% group_by(CAS) %>%
  summarize(ne=n())
nf<- f2 %>% group_by(CAS) %>%
  summarize(nf=n())

all<-left_join(all, na, by= 'CAS')
all<-left_join(all, nc, by= 'CAS')
all<-left_join(all, nd, by= 'CAS')
all<-left_join(all, ne, by= 'CAS')
all<-left_join(all, nf, by= 'CAS')
all[is.na(all)] <- 0

for (i in 1:dim(all)[1]) {
  all$exp[i]<-
max(all$exp[i],sum(all$na[i],all$nc[i],all$nd[i],all$ne[i],all$nf[i],na.rm=TRUE))
  all$e_conn[i]<-
max(all$e_conn[i],sum(all$nc[i],all$nd[i],all$ne[i],all$nf[i],na.rm=TRUE))
)
  all$e_noconn[i]<-max(all$e_noconn[i],sum(all$na[i],na.rm=TRUE))}}

wb <- loadWorkbook(outfile)
writeData(wb, sheet = "CTD", all[1:ncol], colNames = F,startRow=3)
saveWorkbook(wb,outfile,overwrite = T)

```