taxaGrp <- subset(volbio\_all,

select = c(samp\_ev, exp, rep, mag, Group, type, grp\_typ, grp\_sz, esd, counts\_per\_ml,

bio\_per\_vol\_pgc\_ml)) %>%

mutate(totalMnCPM=counts\_per\_ml,totalMnBPM= bio\_per\_vol\_pgc\_ml)

taxaGrp$szesd <- paste(taxaGrp$grp\_sz, taxaGrp$esd)

taxaGrp <- filter(taxaGrp, Group %in% c("diatom", "cyanobacteria", "ciliate",

"flagellate", "unidentified")) %>%

filter(type %in% c("cone", "other", "round", "aphanizomenon", "dolichospermum", "centric", "pennate", "pennate fragillaria","pennate pleurosigma","colonial","cryptomonas"))

taxaGrp <- filter(taxaGrp, !grp\_typ %in% c("ciliate other", "unidentified other", "cyanobacteria other","diatom other"))

taxaGrp <- subset(taxaGrp,totalMnCPM !=0)

taxaGrp <- subset(taxaGrp, select = c(samp\_ev, exp, rep, mag, Group, type, grp\_typ, esd, szesd,

totalMnCPM, totalMnBPM))

taxaGrp$totalMnCPM<- formattable(taxaGrp$totalMnCPM, format="f",digits=2)

taxaGrp$totalMnBPM<- formattable(taxaGrp$totalMnBPM, format="f",digits=2)

write\_xlsx(taxaGrp, "data/TopTen/MainGroups/taxaGrp.xlsx")

save(taxaGrp, file = "data/TopTen/MainGroups/taxaGrp.Rdata")

### Add up the counts per ml for each distinct cilate

## size/esd name but keep the esd and biomass columns

taxaGrplumpC <- aggregate(totalMnCPM ~ szesd +esd,

data = taxaGrp, FUN = sum, na.rm =TRUE)

taxaGrplumpB <- aggregate(totalMnBPM ~ szesd + esd,data = taxaGrp, FUN = sum, na.rm =TRUE)

taxaGrplump <- merge(taxaGrplumpC, taxaGrplumpB, by="szesd")

taxaGrplump<- subset(taxaGrplump,

select = c(szesd, esd.x, totalMnCPM, totalMnBPM))

colnames(taxaGrplump)[1] = "Group"

colnames(taxaGrplump)[2] = "esd"

#colnames(taxaGrplump)[3] = "totalMnCPM"

#colnames(taxaGrplump)[4] = "totalMnBPM"