Extreme Events Exploratory Analysis

Erin Larson

11/13/2019

Importing data.

```r
EEdata<-read.csv("EEData.csv")
```

```r
EEdata_min<-EEdata[, 1:47]
```

Manipulating data to only include rows that have Biological response info.

```r
EEdata_Bio<-subset(EEdata_min, BiologicalResponse_OrganismalGroup!="")
EEdata_Phys<-subset(EEdata_min, PhysicalResponseCategory!="")
EEdata_Chem<-subset(EEdata_min, ChemicalResponse_Category!="")
EEdata_BioPhys<-subset(EEdata_min, BiologicalResponse_OrganismalGroup!="" & PhysicalResponseCategory!="
EEdata_BioChem<-subset(EEdata_min, BiologicalResponse_OrganismalGroup!="" & ChemicalResponse_Category!=
EEdata_PhysChem<-subset(EEdata_min, PhysicalResponseCategory!="" & ChemicalResponse_Category!="")
EEdata_BioPhysChem<-subset(EEdata_min, BiologicalResponse_OrganismalGroup!="" & PhysicalResponseCategory
```

Counting number of studies that had biological, chemical, and physical responses.

```r
EE_responsetype<-data.frame(ResponseType=c("All Studies", "Biological", "Physical", "Chemical", "Biolog
EE_responsetype$Count[1]<-nrow(distinct(EEdata_min, UniqueAccession, .keep_all=TRUE))
EE_responsetype$Count[2]<-nrow(distinct(EEdata_Bio, UniqueAccession, .keep_all=TRUE))
EE_responsetype$Count[3]<-nrow(distinct(EEdata_Phys, UniqueAccession, .keep_all=TRUE))
EE_responsetype$Count[4]<-nrow(distinct(EEdata_Chem, UniqueAccession, .keep_all=TRUE))
EE_responsetype$Count[5]<-nrow(distinct(EEdata_BioPhys, UniqueAccession, .keep_all=TRUE))
EE_responsetype$Count[6]<-nrow(distinct(EEdata_BioChem, UniqueAccession, .keep_all=TRUE))
EE_responsetype$Count[7]<-nrow(distinct(EEdata_PhysChem, UniqueAccession, .keep_all=TRUE))
EE_responsetype$Count[8]<-nrow(distinct(EEdata_BioPhysChem, UniqueAccession, .keep_all=TRUE))
```
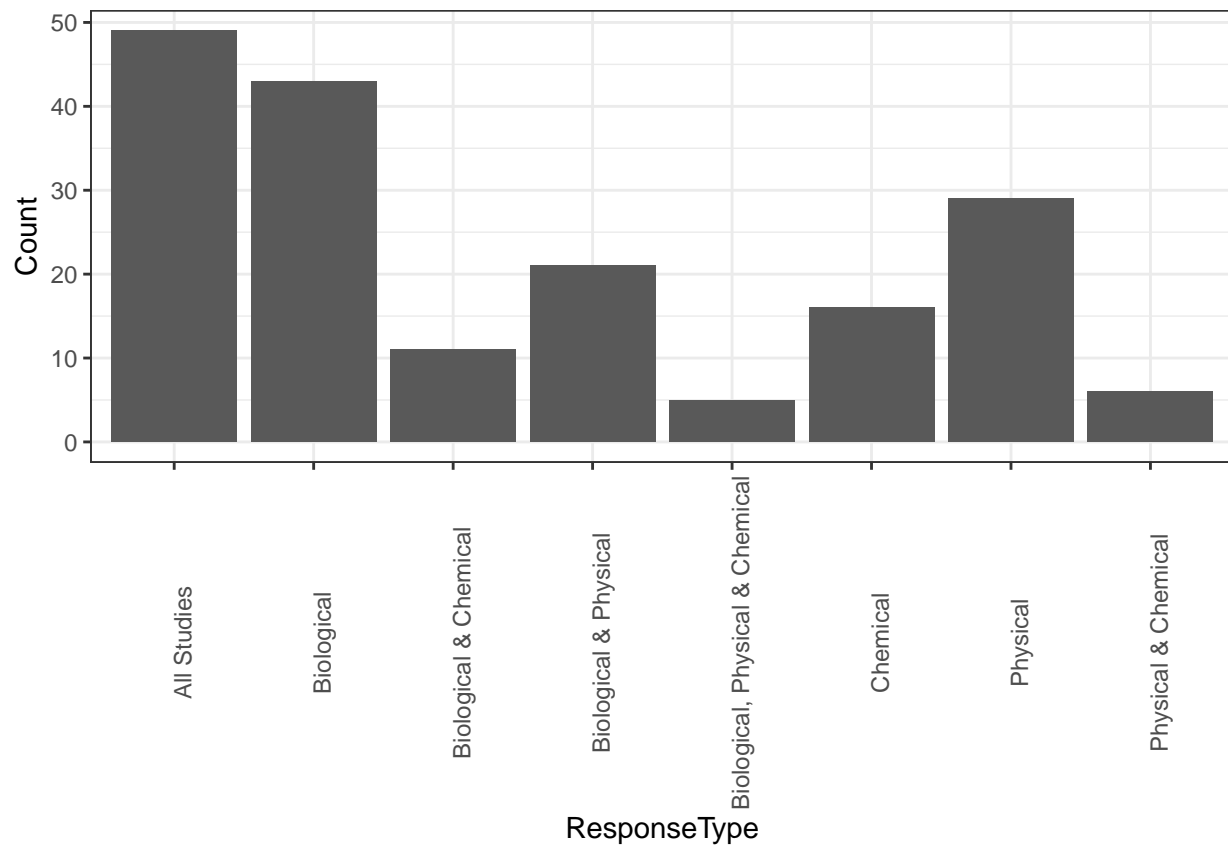
Making a plot of the number of studies that have looked at each type of response variable (biological, physical, chemical)

```r
responseplot<-ggplot(EE_responsetype, aes(x=ResponseType, y=Count))+geom_bar(stat="identity")+theme_bw(
```
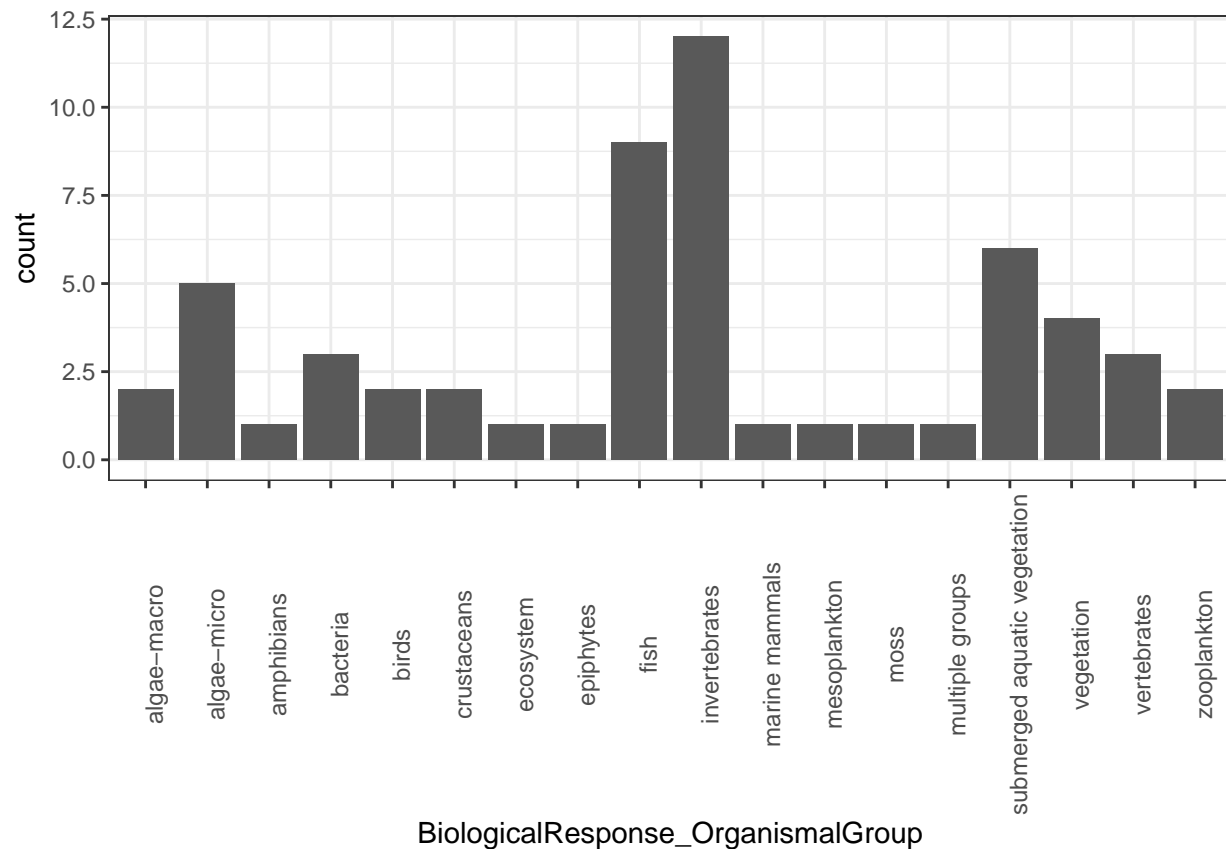
```r
responseplot
```

```
EEdatasum <- EEdata_Bio %>%
    group_by(UniqueAccession, BiologicalResponse_OrganismalGroup) %>%
    summarize()

EEBioGroup <- EEdatasum %>%
    group_by(BiologicalResponse_OrganismalGroup) %>%
    summarize(count=n())

biogroupplot<-ggplot(EEBioGroup, aes(x=BiologicalResponse_OrganismalGroup, y=count))+geom_bar(stat="iden

biogroupplot
```

```
EEtrophicsum <- EEdata_Bio %>%
    group_by(UniqueAccession, BiologicalResponse_TrophicLevel) %>%
    summarize()
```

```
## Warning: Factor `BiologicalResponse_TrophicLevel` contains implicit NA,
## consider using `forcats::fct_explicit_na`
```

```
EETrophic <- EEtrophicsum %>%
    group_by(BiologicalResponse_TrophicLevel) %>%
    summarize(count=n())
```

```
## Warning: Factor `BiologicalResponse_TrophicLevel` contains implicit NA,
## consider using `forcats::fct_explicit_na`
```

```
trophicplot<-ggplot(EETrophic, aes(x=BiologicalResponse_TrophicLevel, y=count))+geom_bar(stat="identity"

trophicplot
```
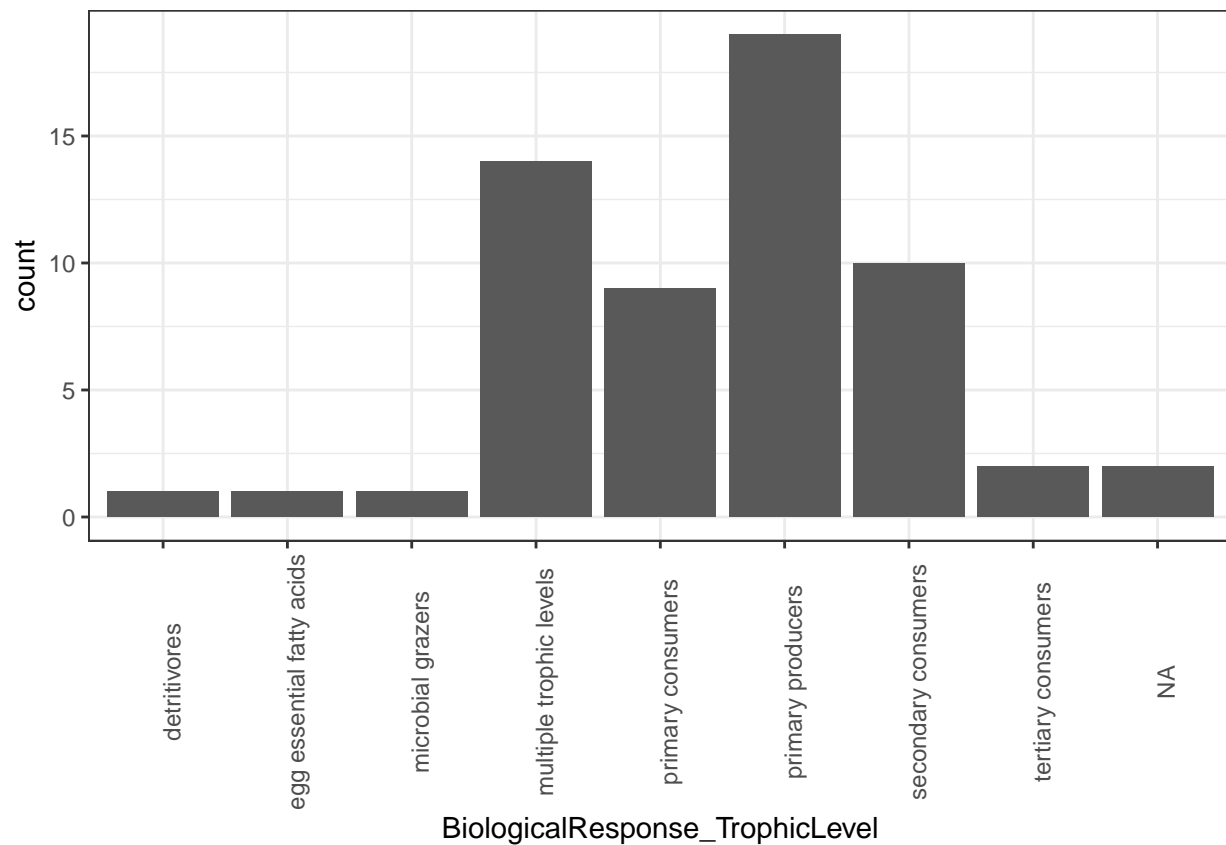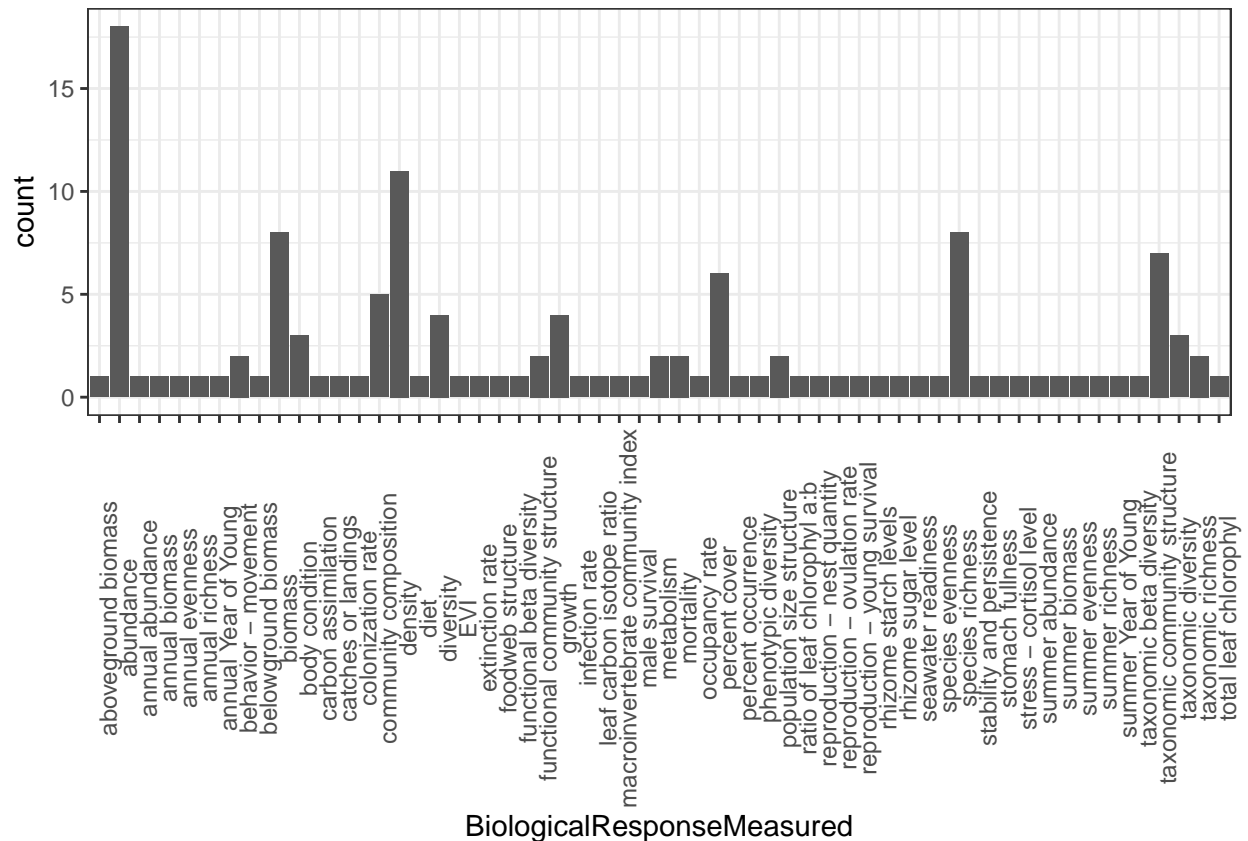
This creates an overwhelming plot...

```r
EEBioResponse <- EEdata_Bio %>%
    group_by(UniqueAccession, BiologicalResponseMeasured) %>%
    summarize()

EEBioMeasure <- EEBioResponse %>%
    group_by(BiologicalResponseMeasured) %>%
    summarize(count=n())

bioresponseplot<-ggplot(EEBioMeasure, aes(x=BiologicalResponseMeasured, y=count))+geom_bar(stat="identi

bioresponseplot
```

```
EEBioResponse$BroadMeasureType <- factor(rep(NA,length(EEBioResponse$BiologicalResponseMeasured)),
                          levels=c("biomass", "abundance/density", "movement", "community stru

EEBioResponse$BroadMeasureType[EEBioResponse$BiologicalResponseMeasured %in% c("annual biomass","abovegr

EEBioResponse$BroadMeasureType[EEBioResponse$BiologicalResponseMeasured %in% c("abundance","annual abun

EEBioResponse$BroadMeasureType[EEBioResponse$BiologicalResponseMeasured %in% c("behavior - movement", "

EEBioResponse$BroadMeasureType[EEBioResponse$BiologicalResponseMeasured %in% c("annual evenness", "annu

EEBioResponse$BroadMeasureType[EEBioResponse$BiologicalResponseMeasured %in% c("carbon assimilation", "

EEBioResponse$BroadMeasureType[EEBioResponse$BiologicalResponseMeasured %in% c( "male survival", "repro

EEBioResponse$BroadMeasureType[EEBioResponse$BiologicalResponseMeasured %in% c("body condition", "diet"

EEBioResponse$BroadMeasureType[EEBioResponse$BiologicalResponseMeasured %in% c("extinction rate")] <- "

##getting unique studies so not double-counting the broad measure type
EEBioMeasureUnique<-distinct_at(EEBioResponse, vars(UniqueAccession, BroadMeasureType))

broadbiomeasureplot<-ggplot(EEBioMeasureUnique, aes(x=BroadMeasureType))+geom_bar()+theme_bw()+theme(ax
broadbiomeasureplot
```