CDW Tally Analysis: D07

Cody Flagg, Courtney Meier

8th January 2017

```r
## Windows paths
# inpath <- "C:/Users/cflagg/Documents/GitHub/neonPlantSampling/cdw_tally_analysis/D07/"
# outpath <- inpath

##  macOS inputs
inpath <- "~/Documents/workDocuments/gitRepositories/neonPlantSampling/cdw_tally_analysis/fulcrumData_2
domain <- "D07"
outpath <- paste("~/Documents/workDocuments/gitRepositories/neonPlantSampling/cdw_tally_analysis/", dom


## import libraries
library(plyr)
library(dplyr)
library(ggplot2)

## import data
parent_cdw <- read.csv(paste(inpath,"tos_coarse_downed_wood_tally_prod.csv",sep=""), header=TRUE)
child_cdw <- read.csv(paste(inpath,"tos_coarse_downed_wood_tally_prod_per_plot_azimuth_log.csv",sep="")

## filter out parent fields that are unneeded
parent_cdw_fil <- dplyr::select(parent_cdw,
                        fulcrum_id, domainid, siteid, plotid_parent, tallydate, volumefactor, parti
## filter out child fields that are unneeded
child_cdw_fil <- dplyr::select(child_cdw,
                        fulcrum_parent_id, lidsazimuth, logdiameter, taxonid, decayclass, logid_inge

## join parent to child via parent::fulcrum_id to child::fulcrum_parent_id
cdw <- merge(x = parent_cdw_fil, y = child_cdw_fil, by.x = "fulcrum_id", by.y = "fulcrum_parent_id", al

## filter to the selected domain only
cdw <- dplyr::filter(cdw, domainid == domain)

## create diameter class factor
cdw$diameterClass <- ifelse(cdw$logdiameter >= 10, '>=10cm',
                        ifelse(cdw$logdiameter <5, "2-5cm", "5-10cm"))

## simplify decayclass to numeric value wrapped inside sapply(decayClassNum, "[[", 1) e.g. 'return firs
cdw$decayClassNum <- sapply(stringr::str_split(cdw$decayclass, pattern = " "),"[[", 1)


## write file
write.csv(cdw, file = paste(outpath, paste0("merged_",domain,"_cdw_rawdata.csv", sep=""), sep = "/"))
```

siteid	taxonid	decayClassNum	diameterClass	counts	totalLogs	relativeAbundance	cumulativeAbundance
GRSM	PINUS	4	>=10cm	30	388	7.73	7.73
GRSM	TSCA	3	>=10cm	20	388	5.15	12.88
GRSM	2PLANT-H	4	>=10cm	18	388	4.64	17.52

siteid	taxonid	decayClassNum	diameterClass	counts	totalLogs	relativeAbundance	cumulativeAbundance
GRSM	2PLANT-H	5	>=10cm	18	388	4.64	22.16
GRSM	PIRU	3	>=10cm	16	388	4.12	26.28
GRSM	TSCA	4	>=10cm	16	388	4.12	30.40
GRSM	TSCA	2	>=10cm	15	388	3.87	34.27
GRSM	PINUS	5	>=10cm	14	388	3.61	37.88
GRSM	TSCA	5	>=10cm	14	388	3.61	41.49
GRSM	LITU	3	>=10cm	12	388	3.09	44.58
GRSM	2PLANT-H	3	>=10cm	11	388	2.84	47.42
GRSM	LITU	4	>=10cm	11	388	2.84	50.26
GRSM	LITU	2	>=10cm	10	388	2.58	52.84
GRSM	PINUS	3	>=10cm	9	388	2.32	55.16
GRSM	ACRUR	3	>=10cm	8	388	2.06	57.22
GRSM	ACRUR	4	>=10cm	8	388	2.06	59.28
GRSM	2PLANT-H		5-10cm	8	388	2.06	61.34
GRSM	ACRUR	2	>=10cm	6	388	1.55	62.89
GRSM	PIRU	4	>=10cm	6	388	1.55	64.44
GRSM	PIPU5	4	>=10cm	5	388	1.29	65.73
GRSM	QUERC	4	>=10cm	5	388	1.29	67.02
GRSM	PIRU	5	>=10cm	5	388	1.29	68.31
GRSM	QUCO2		5-10cm	5	388	1.29	69.60
GRSM	QUMO4	2	>=10cm	4	388	1.03	70.63
GRSM	QUERC	3	>=10cm	4	388	1.03	71.66
GRSM	ABFR	4	>=10cm	4	388	1.03	72.69
GRSM	2PLANT-H	2	>=10cm	3	388	0.77	73.46
GRSM	PINUS	2	>=10cm	3	388	0.77	74.23
GRSM	PIRI	2	>=10cm	3	388	0.77	75.00
GRSM	PIVI2	2	>=10cm	3	388	0.77	75.77
GRSM	ROPS	2	>=10cm	3	388	0.77	76.54
GRSM	PIVI2	3	>=10cm	3	388	0.77	77.31
GRSM	QURU	3	>=10cm	3	388	0.77	78.08
GRSM	QUMO4	4	>=10cm	3	388	0.77	78.85
GRSM	2PLANT-S	5	>=10cm	3	388	0.77	79.62
GRSM	ACRUR	5	>=10cm	3	388	0.77	80.39
GRSM	TSCA		5-10cm	3	388	0.77	81.16
GRSM	FRPE	2	>=10cm	2	388	0.52	81.68
GRSM	PIPU5	2	>=10cm	2	388	0.52	82.20
GRSM	PIRU	2	>=10cm	2	388	0.52	82.72
GRSM	QUERC	2	>=10cm	2	388	0.52	83.24
GRSM	ACER	3	>=10cm	2	388	0.52	83.76
GRSM	ROPS	3	>=10cm	2	388	0.52	84.28
GRSM	CACA18	4	>=10cm	2	388	0.52	84.80
GRSM	HATE3	4	>=10cm	2	388	0.52	85.32
GRSM	OXAR	4	>=10cm	2	388	0.52	85.84
GRSM	PIPU5	5	>=10cm	2	388	0.52	86.36
GRSM	QUERC	5	>=10cm	2	388	0.52	86.88
GRSM	COFL2		5-10cm	2	388	0.52	87.40
GRSM	LITU		5-10cm	2	388	0.52	87.92
GRSM	QUMO4		5-10cm	2	388	0.52	88.44
GRSM	RHMA4		5-10cm	2	388	0.52	88.96
GRSM	VITIS		5-10cm	2	388	0.52	89.48
GRSM	HATE3	1	>=10cm	1	388	0.26	89.74
GRSM	LITU	1	>=10cm	1	388	0.26	90.00

siteid	taxonid	decayClassNum	diameterClass	counts	totalLogs	relativeAbundance	cumulativeAbundance
GRSM	PIRU	1	>=10cm	1	388	0.26	90.26
GRSM	QURU	1	>=10cm	1	388	0.26	90.52
GRSM	ROPS	1	>=10cm	1	388	0.26	90.78
GRSM	TSCA	1	>=10cm	1	388	0.26	91.04
GRSM	ACNE2	2	>=10cm	1	388	0.26	91.30
GRSM	ACRU	2	>=10cm	1	388	0.26	91.56
GRSM	BEAL2	2	>=10cm	1	388	0.26	91.82
GRSM	CACA18	2	>=10cm	1	388	0.26	92.08
GRSM	HATE3	2	>=10cm	1	388	0.26	92.34
GRSM	JUNI	2	>=10cm	1	388	0.26	92.60
GRSM	QUVE	2	>=10cm	1	388	0.26	92.86
GRSM	2PLANT-S	3	>=10cm	1	388	0.26	93.12
GRSM	ABFR	3	>=10cm	1	388	0.26	93.38
GRSM	COFL2	3	>=10cm	1	388	0.26	93.64
GRSM	HATE3	3	>=10cm	1	388	0.26	93.90
GRSM	JUNI	3	>=10cm	1	388	0.26	94.16
GRSM	PIPU5	3	>=10cm	1	388	0.26	94.42
GRSM	QUCO2	3	>=10cm	1	388	0.26	94.68
GRSM	QUMO4	3	>=10cm	1	388	0.26	94.94
GRSM	SASSA	3	>=10cm	1	388	0.26	95.20
GRSM	BEAL2	4	>=10cm	1	388	0.26	95.46
GRSM	BELE	4	>=10cm	1	388	0.26	95.72
GRSM	FRPE	4	>=10cm	1	388	0.26	95.98
GRSM	MATR	4	>=10cm	1	388	0.26	96.24
GRSM	NYSY	4	>=10cm	1	388	0.26	96.50
GRSM	PITA	4	>=10cm	1	388	0.26	96.76
GRSM	PIVI2	4	>=10cm	1	388	0.26	97.02
GRSM	QUCO2	4	>=10cm	1	388	0.26	97.28
GRSM	PLOC	5	>=10cm	1	388	0.26	97.54
GRSM	KALA		2-5cm	1	388	0.26	97.80
GRSM	QUCO2		2-5cm	1	388	0.26	98.06
GRSM	TSCA		2-5cm	1	388	0.26	98.32
GRSM	ABFR		5-10cm	1	388	0.26	98.58
GRSM	ACER		5-10cm	1	388	0.26	98.84
GRSM	ACRUR		5-10cm	1	388	0.26	99.10
GRSM	PINUS		5-10cm	1	388	0.26	99.36
GRSM	PIRU		5-10cm	1	388	0.26	99.62
GRSM	QUERC		5-10cm	1	388	0.26	99.88
GRSM	ROPS		5-10cm	1	388	0.26	100.14

Warning: Using size for a discrete variable is not advised.

Rank Abundance of decayClass x sizeCategory x taxonID (DST)
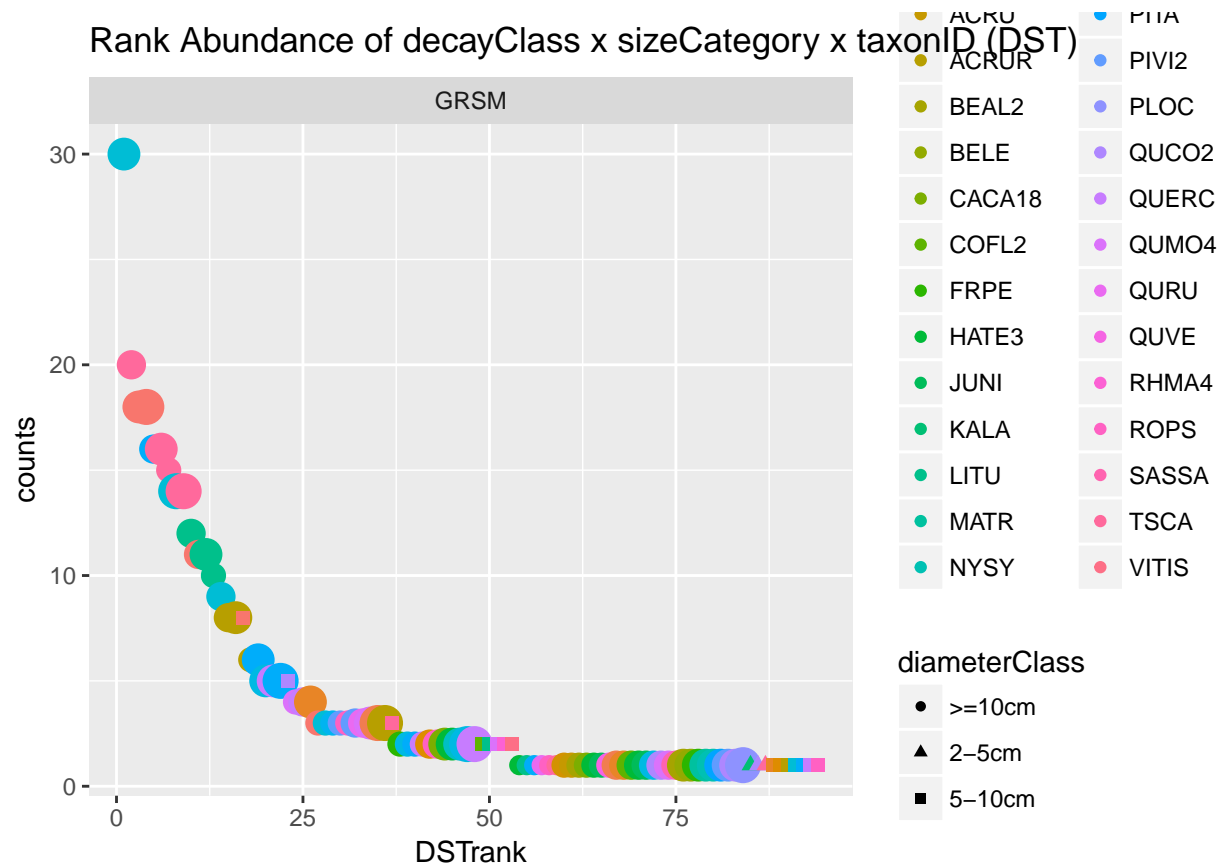
Code

```
## Windows paths
# inpath <- "C:/Users/cflagg/Documents/GitHub/neonPlantSampling/cdw_tally_analysis/D07/"
# outpath <- inpath

##  macOS inputs
inpath <- "~/Documents/workDocuments/gitRepositories/neonPlantSampling/cdw_tally_analysis/fulcrumData_2(
domain <- "D07"
outpath <- paste("~/Documents/workDocuments/gitRepositories/neonPlantSampling/cdw_tally_analysis/", doma


## import libraries
library(plyr)
library(dplyr)
library(ggplot2)

## import data
parent_cdw <- read.csv(paste(inpath,"tos_coarse_downed_wood_tally_prod.csv",sep=""), header=TRUE)
child_cdw <- read.csv(paste(inpath,"tos_coarse_downed_wood_tally_prod_per_plot_azimuth_log.csv",sep="")

## filter out parent fields that are unneeded
parent_cdw_fil <- dplyr::select(parent_cdw,
                        fulcrum_id, domainid, siteid, plotid_parent, tallydate, volumefactor, partic
## filter out child fields that are unneeded
```

```
child_cdw_fil <- dplyr::select(child_cdw,
                              fulcrum_parent_id, lidsazimuth, logdiameter, taxonid, decayclass, logid_inge

## join parent to child via parent::fulcrum_id to child::fulcrum_parent_id
cdw <- merge(x = parent_cdw_fil, y = child_cdw_fil, by.x = "fulcrum_id", by.y = "fulcrum_parent_id", all

## filter to the selected domain only
cdw <- dplyr::filter(cdw, domainid == domain)

## create diameter class factor
cdw$diameterClass <- ifelse(cdw$logdiameter >= 10, '>=10cm',
                           ifelse(cdw$logdiameter <5, "2-5cm", "5-10cm"))

## simplify decayclass to numeric value wrapped inside sapply(decayClassNum, "[[", 1) e.g. 'return first
cdw$decayClassNum <- sapply(stringr::str_split(cdw$decayclass, pattern = " "),"[[", 1)


## write file
write.csv(cdw, file = paste(outpath, paste0("merged_",domain,"_cdw_rawdata.csv", sep=""), sep = "/"))

#names(cdw)
## table 1 -- count of taxonIDs, report top 10?
## order ascending output by totalTaxa length() will count length of a vector

## Filter out 'blank' transects where TTP = NO
cdw_fil <- dplyr::filter(cdw, targettaxapresent == "Y")
t1 <- ddply(cdw_fil, ~siteid+taxonid+decayClassNum+diameterClass, summarize, counts = length(taxonid))

# an alternative to ddply call, returns similar result but output column name is defaulted
# plyr::count(cdw, c("siteid","taxonid", "decayClassNum", "diameterClass"))

## Sort rows by siteID then by descending number of counts
t1 <- dplyr::arrange(t1, siteid, desc(counts), diameterClass, decayClassNum, taxonid)

## Calculate relative abundances
t1$totalLogs <- nrow(cdw_fil)
t1$relativeAbundance <- round((t1$counts/t1$totalLogs)*100,2)
t1$cumulativeAbundance <- round(cumsum(t1$relativeAbundance),2)

## determine raw ranking by count of DST combo within siteid
t1 <- plyr::ddply(t1, ~siteid, mutate, DSTrank = as.numeric(paste(1:length(counts))), DSTcombo = paste0
t2 <- ddply(cdw, ~siteid+taxonid, summarize, counts = length(taxonid))

##  Determine required sample size per DST combo, and cumulative sample size
t1$sampledDiskNum <- ifelse(t1$diameterClass == ">=10cm", 10,5)
t1$cumulativeDiskNum <- cumsum(t1$sampledDiskNum)

## species abundance/total abundance
ggplot(data = t1, aes(x = DSTrank, y = counts, color = taxonid, shape = diameterClass, size = decayClas
```