

CDW Tally Analysis: D03

Cody Flagg, Courtney Meier

7th January 2017

```
## 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 <- "D03"
outpath <- paste("~/Documents/workDocuments/gitRepositories/neonPlantSampling/cdw_tally_analysis/", domain)

## 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_ingest

## 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"))

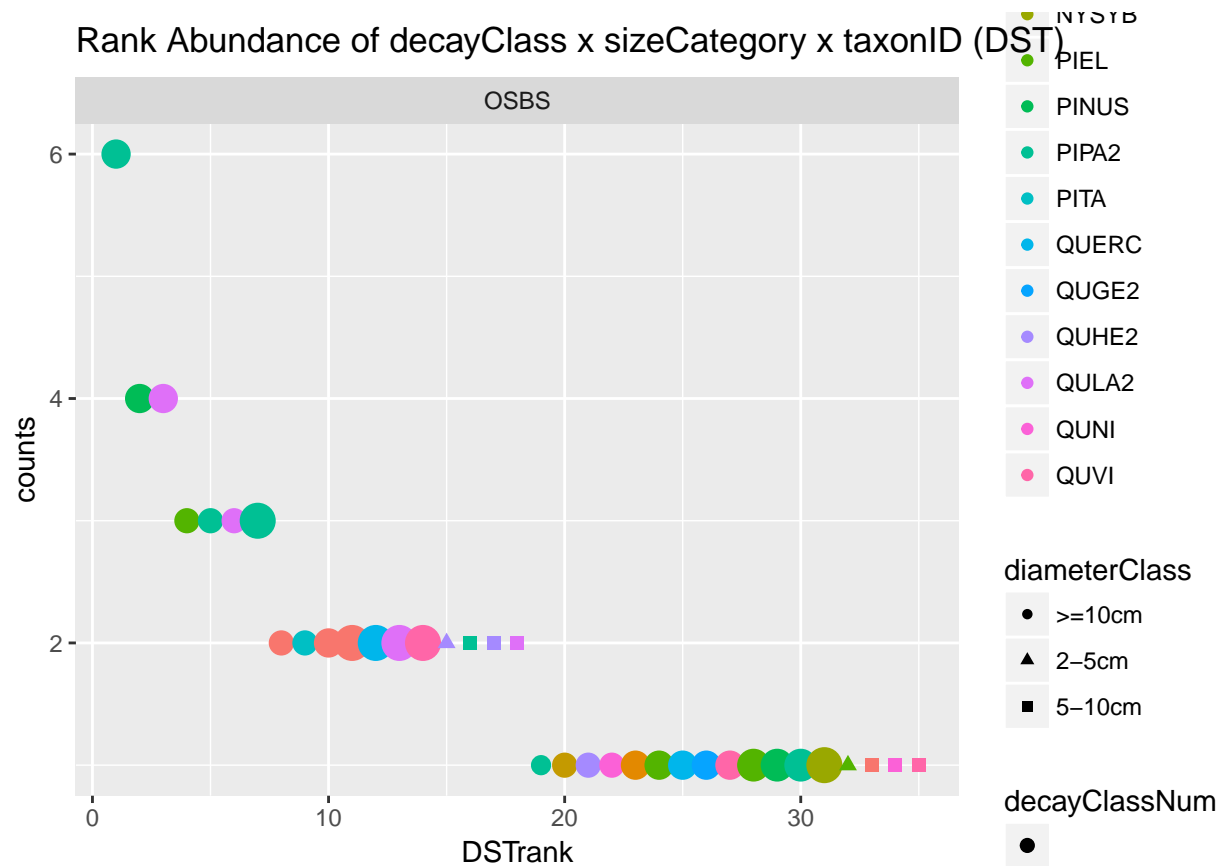
## simplify decayclass to numeric value wrapped inside sapply(decayClassNum, "[", 1) e.g. 'return first
cdw$decayClassNum <- sapply(stringr::str_split(cdw$decayclass, pattern = " "), "[", 1)

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

siteid	taxonid	decayClassNum	diameterClass	counts	totalLogs	relativeAbundance	cumulativeAbundance
OSBS	PIPA2	3	>=10cm	6	65	9.23	9.23
OSBS	PINUS	3	>=10cm	4	65	6.15	15.38
OSBS	QULA2	3	>=10cm	4	65	6.15	21.53

siteid	taxonid	decayClassNum	diameterClass	counts	totalLogs	relativeAbundance	cumulativeAbundance
OSBS	PIEL	2	>=10cm	3	65	4.62	26.15
OSBS	PIPA2	2	>=10cm	3	65	4.62	30.77
OSBS	QULA2	2	>=10cm	3	65	4.62	35.39
OSBS	PIPA2	5	>=10cm	3	65	4.62	40.01
OSBS	2PLANT-H	2	>=10cm	2	65	3.08	43.09
OSBS	PITA	2	>=10cm	2	65	3.08	46.17
OSBS	2PLANT-H	3	>=10cm	2	65	3.08	49.25
OSBS	2PLANT-H	5	>=10cm	2	65	3.08	52.33
OSBS	QUERC	5	>=10cm	2	65	3.08	55.41
OSBS	QULA2	5	>=10cm	2	65	3.08	58.49
OSBS	QUVI	5	>=10cm	2	65	3.08	61.57
OSBS	QUHE2		2-5cm	2	65	3.08	64.65
OSBS	PIPA2		5-10cm	2	65	3.08	67.73
OSBS	QUHE2		5-10cm	2	65	3.08	70.81
OSBS	QULA2		5-10cm	2	65	3.08	73.89
OSBS	PIPA2	1	>=10cm	1	65	1.54	75.43
OSBS	NYSSA	2	>=10cm	1	65	1.54	76.97
OSBS	QUHE2	2	>=10cm	1	65	1.54	78.51
OSBS	QUNI	2	>=10cm	1	65	1.54	80.05
OSBS	ACRU	3	>=10cm	1	65	1.54	81.59
OSBS	PIEL	3	>=10cm	1	65	1.54	83.13
OSBS	QUERC	3	>=10cm	1	65	1.54	84.67
OSBS	QUGE2	3	>=10cm	1	65	1.54	86.21
OSBS	QUVI	3	>=10cm	1	65	1.54	87.75
OSBS	PIEL	4	>=10cm	1	65	1.54	89.29
OSBS	PINUS	4	>=10cm	1	65	1.54	90.83
OSBS	PIPA2	4	>=10cm	1	65	1.54	92.37
OSBS	NYSYB	5	>=10cm	1	65	1.54	93.91
OSBS	PIEL		2-5cm	1	65	1.54	95.45
OSBS	2PLANT-H		5-10cm	1	65	1.54	96.99
OSBS	QUNI		5-10cm	1	65	1.54	98.53
OSBS	QUVI		5-10cm	1	65	1.54	100.07

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



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 <- "D03"
outpath <- paste("~/Documents/workDocuments/gitRepositories/neonPlantSampling/cdw_tally_analysis/", domain)

## 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, volume, partid)

## 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 = decayClass

```