CDW Tally Analysis: D08

Cody Flagg, Courtney Meier

8th January 2017

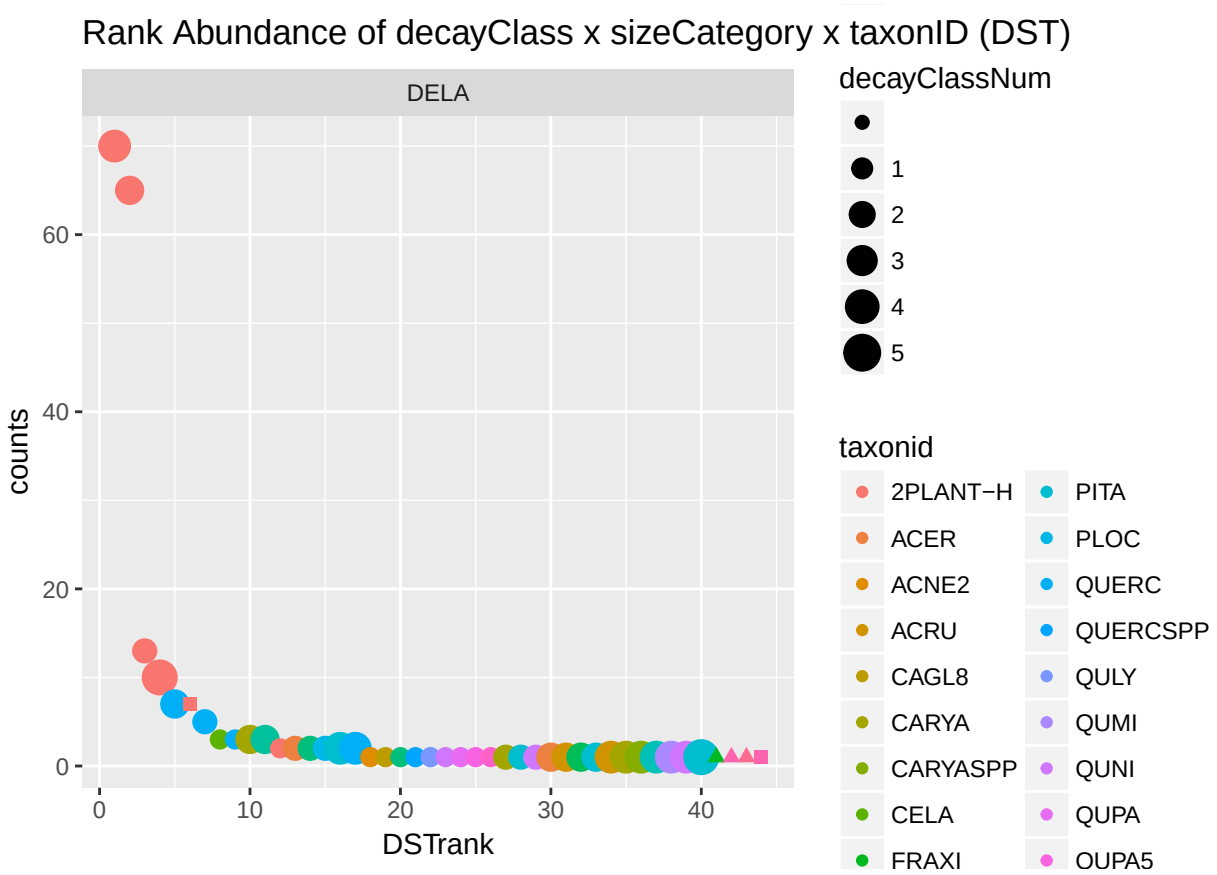
```
## Windows paths
# inpath <- "C:/Users/cflagg/Documents/GitHub/neonPlantSampling/cdw_tally_analysis/D07/"
# outpath <- inpath

##  macOS inputs
inpath <- "~/Documents/workDocuments/gitRepositories/neonPlantSampling/cdw_tally_analysis/fulcrumData_2
domain <- "D08"
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, 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
DELA	2PLANT-H	4	>=10cm	70	228	30.70	30.7
DELA	2PLANT-H	3	>=10cm	65	228	28.51	59.2
DELA	2PLANT-H	2	>=10cm	13	228	5.70	64.9

siteid	taxonid	decayClassNum	diameterClass	counts	totalLogs	relativeAbundance	cumulativeAbundance
DELA	2PLANT-H	5	>=10cm	10	228	4.39	69.30
DELA	QUERC	3	>=10cm	7	228	3.07	72.37
DELA	2PLANT-H		5-10cm	7	228	3.07	75.44
DELA	QUERC	2	>=10cm	5	228	2.19	77.63
DELA	CELA	1	>=10cm	3	228	1.32	78.95
DELA	QUERC	1	>=10cm	3	228	1.32	80.27
DELA	CARYA	3	>=10cm	3	228	1.32	81.59
DELA	PINUS	3	>=10cm	3	228	1.32	82.91
DELA	2PLANT-H	1	>=10cm	2	228	0.88	83.79
DELA	ACER	2	>=10cm	2	228	0.88	84.67
DELA	LIST2	2	>=10cm	2	228	0.88	85.55
DELA	PLOC	2	>=10cm	2	228	0.88	86.43
DELA	PITA	4	>=10cm	2	228	0.88	87.31
DELA	QUERC	4	>=10cm	2	228	0.88	88.19
DELA	ACNE2	1	>=10cm	1	228	0.44	88.63
DELA	CAGL8	1	>=10cm	1	228	0.44	89.07
DELA	LIST2	1	>=10cm	1	228	0.44	89.51
DELA	QUERCSPP	1	>=10cm	1	228	0.44	89.95
DELA	QULY	1	>=10cm	1	228	0.44	90.39
DELA	QUNI	1	>=10cm	1	228	0.44	90.83
DELA	QUPA	1	>=10cm	1	228	0.44	91.27
DELA	QUPA5	1	>=10cm	1	228	0.44	91.71
DELA	QUVE	1	>=10cm	1	228	0.44	92.15
DELA	CARYA	2	>=10cm	1	228	0.44	92.59
DELA	PITA	2	>=10cm	1	228	0.44	93.03
DELA	QUNI	2	>=10cm	1	228	0.44	93.47
DELA	ACER	3	>=10cm	1	228	0.44	93.91
DELA	ACRU	3	>=10cm	1	228	0.44	94.35
DELA	FRPE	3	>=10cm	1	228	0.44	94.79
DELA	PITA	3	>=10cm	1	228	0.44	95.23
DELA	ACRU	4	>=10cm	1	228	0.44	95.67
DELA	CARYA	4	>=10cm	1	228	0.44	96.11
DELA	CARYASPP	4	>=10cm	1	228	0.44	96.55
DELA	PIPA2	4	>=10cm	1	228	0.44	96.99
DELA	QUMI	4	>=10cm	1	228	0.44	97.43
DELA	QUNI	4	>=10cm	1	228	0.44	97.87
DELA	PITA	5	>=10cm	1	228	0.44	98.31
DELA	FRAXI		2-5cm	1	228	0.44	98.75
DELA	TORA2		2-5cm	1	228	0.44	99.19
DELA	VITIS		2-5cm	1	228	0.44	99.63
DELA	TORA2		5-10cm	1	228	0.44	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 <- "D08"
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, 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 = "/"))
```

```
#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
```