Log Transforming the Data

Xingyao Chen

6/21/2017

```
mydat_small=read.csv('pollinator_visitation_fullData.csv')


mydat_small$Experiment.Week=as.factor(mydat_small$Experiment.Week)
mydat_small$Plant.Number=as.factor(mydat_small$Plant.Number)
mydat_small$Pair=as.factor(mydat_small$Pair)
```

Make some adjustments before Log10-ing

```
mydat_small=mydat_small[,-1]
boo=c()
for (i in 1:ncol(mydat_small)){
  boo[i]=is.numeric(mydat_small[,i])
}

#Cannot take the log of 0, so added a small value to each column that contains 0's
mydat_small_val_log=mydat_small
mydat_small_val_log$Honeybees=mydat_small_val_log$Honeybees+1
mydat_small_val_log$Speed=mydat_small_val_log$Speed+0.1

#log10 transformation happenning
for(i in 1:ncol(mydat_small_val_log)){
  if(boo[i])
    mydat_small_val_log[,i]=log10(mydat_small_val_log[,i])
}
#peek
summary(mydat_small_val_log)
```

```
##         Date    Experiment.Week         Location  Plant.Number
##  5/17/2017:12   1:12            Pair 2 Large: 7   1:54
##  5/23/2017:12   2:24            Pair 2 Small: 7   2:30
##  5/25/2017:12   3:24            Pair 4 Large: 7
##  5/30/2017:12   4:24            Pair 4 Small: 7
##  6/2/2017 :12                   Pair 5 Large: 7
##  6/6/2017 :12                   Pair 5 Small: 7
##  6/8/2017 :12                   (Other)     :42
##       Pair       Size                   Start.Time                End.Time
##  2      :14   Large:42   2017-05-25 10:33:00: 3   2017-05-25 10:48:00: 3
##  4      :14   Small:42   2017-05-25 10:55:00: 2   2017-05-25 11:10:00: 2
##  5      :14              2017-05-30 10:12:00: 2   2017-05-30 10:27:00: 2
##  6      :14              2017-05-30 10:33:00: 2   2017-05-30 10:48:00: 2
##  1      :10              2017-06-06 11:44:00: 2   2017-06-06 11:59:00: 2
##  3      :10              2017-06-08 10:11:00: 2   (Other)            :72
##  (Other): 8              (Other)            :71   NA's               : 1
##  Avg.open.flowers.per.inflorescence Total.inflorescenses Total.Flowers
```

```
##  Min.   :0.716                     Min.   :0.6021       Min.   :1.591
##  1st Qu.:1.394                     1st Qu.:0.9886       1st Qu.:2.428
##  Median :1.498                     Median :1.1761       Median :2.656
##  Mean   :1.501                     Mean   :1.2249       Mean   :2.726
##  3rd Qu.:1.630                     3rd Qu.:1.4771       3rd Qu.:3.087
##  Max.   :2.027                     Max.   :1.9138       Max.   :3.896
##
##    Honeybees          Volume           Sugar_conc       Sugar_content
##  Min.   :0.0000   Min.   :-0.71627   Min.   :-0.47190   Min.   :-1.0869
##  1st Qu.:0.5708   1st Qu.:-0.19146   1st Qu.:-0.35018   1st Qu.:-0.3927
##  Median :1.1139   Median : 0.09611   Median :-0.28247   Median :-0.2401
##  Mean   :0.9720   Mean   : 0.01588   Mean   :-0.27797   Mean   :-0.2747
##  3rd Qu.:1.4472   3rd Qu.: 0.18799   3rd Qu.:-0.22349   3rd Qu.:-0.1523
##  Max.   :1.7559   Max.   : 0.54562   Max.   :-0.04194   Max.   : 0.1775
##
##    Volume_var      Sugar_conc_var    Sugar_content_var   Temperature
##  Min.   :-2.6178   Min.   :-3.4579   Min.   :-2.41525   Min.   :1.734
##  1st Qu.:-1.1264   1st Qu.:-2.1166   1st Qu.:-1.47495   1st Qu.:1.789
##  Median :-0.4814   Median :-1.7340   Median :-1.31852   Median :1.826
##  Mean   :-0.6730   Mean   :-1.8229   Mean   :-1.27494   Mean   :1.831
##  3rd Qu.:-0.1518   3rd Qu.:-1.4018   3rd Qu.:-0.90989   3rd Qu.:1.873
##  Max.   : 0.2757   Max.   :-0.9406   Max.   : 0.05588   Max.   :1.963
##  NA's   :3         NA's   :3         NA's   :3
##     Humidity       Wind        Speed              Gust       Pressure
##  66%    : 7   West   :26   Min.   :-1.0000   6mph   :11   29.91in:12
##  77%    : 7   WNW    :17   1st Qu.: 0.2553   9.4mph :11   30.04in:12
##  64%    : 6   SSW    :15   Median : 0.5441   11.9mph:10   29.96in:10
##  72%    : 5   SW     :12   Mean   : 0.3641   7.7mph : 9   29.94in: 9
##  76%    : 5   NW     : 6   3rd Qu.: 0.6616   6.8mph : 8   29.97in: 8
##  89%    : 5   NNW    : 3   Max.   : 0.8921   10.2mph: 7   29.98in: 8
##  (Other):49   (Other): 5                     (Other):28   (Other):25
##      Solar          Clustered          East           hotDay
##  Min.   :-0.2041   Mode :logical   Mode :logical   Mode :logical
##  1st Qu.: 2.4243   FALSE:38        FALSE:56        FALSE:46
##  Median : 2.6265   TRUE :46        TRUE :28        TRUE :38
##  Mean   : 2.2829
##  3rd Qu.: 2.9280
##  Max.   : 3.0115
##
##    visits
##  Mode :logical
##  FALSE:42
##  TRUE :42
##
##
##
##
```

```
data_dt=split(mydat_small_val_log, mydat_small$Date)
###median as theshold value
theshold=c()
wholedf=data.frame()
for( i in 1:length(data_dt)){
  theshold[i]=median(data_dt[[i]]$Honeybees)
```

```
  data_dt[[i]]$visits=(data_dt[[i]]$Honeybees>=theshold[i])
  wholedf=rbind(wholedf, data_dt[[i]])
}
mydat_small_val_log$visits=wholedf$visits
head(mydat_small_val_log)
```

```
##        Date Experiment.Week     Location Plant.Number Pair  Size
## 1 5/17/2017               1 Pair 1 Large            1    1 Large
## 2 5/17/2017               1 Pair 1 Small            1    1 Small
## 3 5/17/2017               1 Pair 2 Large            1    2 Large
## 4 5/17/2017               1 Pair 2 Small            1    2 Small
## 5 5/17/2017               1 Pair 3 Large            1    3 Large
## 6 5/17/2017               1 Pair 3 Small            1    3 Small
##            Start.Time            End.Time
## 1 2017-05-17 11:08:00 2017-05-17 11:23:00
## 2 2017-05-17 11:32:00 2017-05-17 11:47:00
## 3 2017-05-17 11:39:00 2017-05-17 11:49:00
## 4 2017-05-17 11:17:00 2017-05-17 11:27:00
## 5 2017-05-17 11:36:00 2017-05-17 11:51:00
## 6 2017-05-17 10:59:00 2017-05-17 11:14:00
##   Avg.open.flowers.per.inflorescence Total.inflorescenses Total.Flowers
## 1                           1.827369             1.913814      3.741183
## 2                           1.660865             1.518514      3.179379
## 3                           1.755875             1.278754      3.034628
## 4                           1.732394             1.000000      2.732394
## 5                           1.799341             1.518514      3.317854
## 6                           1.526339             1.041393      2.567732
##   Honeybees     Volume Sugar_conc Sugar_content  Volume_var Sugar_conc_var
## 1 1.1461280 0.18326984 -0.4088061   -0.20156517 -0.33525522      -2.101924
## 2 0.3010300 0.30776338 -0.3601468   -0.08461666 -0.23311488      -1.886816
## 3 0.7781513 0.15381486 -0.1476765   -0.01049456 -0.14413081      -1.619066
## 4 0.0000000 0.32990612 -0.2911848    0.03135796 -0.40521049      -2.184032
## 5 0.6989700 0.08813609 -0.2502637   -0.18037382  0.07291591      -2.135489
## 6 0.0000000 0.25072488 -0.2146702    0.05993007  0.27570717      -1.913640
##   Sugar_content_var Temperature Humidity Wind Speed   Gust Pressure
## 1       -0.90989453    1.749736      75%  NNW    -1 5.1mph  29.98in
## 2       -2.19291310    1.752048      74%  SSW    -1 3.4mph  29.98in
## 3       -0.68387831    1.748963      74%  NNW     0 6.8mph  29.98in
## 4       -1.00059914    1.749736      75%  NNW    -1 5.1mph  29.98in
## 5       -0.51122616    1.752048      74%  SSW    -1 3.4mph  29.98in
## 6        0.05588432    1.748188      76% West    -1 5.1mph  29.98in
##      Solar Clustered  East hotDay visits
## 1 2.076913     FALSE  TRUE  FALSE   TRUE
## 2 2.090346     FALSE  TRUE  FALSE  FALSE
## 3 2.028876      TRUE  TRUE  FALSE   TRUE
## 4 2.076913      TRUE  TRUE  FALSE  FALSE
## 5 2.090346     FALSE FALSE  FALSE   TRUE
## 6 2.163236     FALSE FALSE  FALSE  FALSE
```

```
write.csv(mydat_small_val_log,'pollinator_visitation_fullData_logTrans.csv')
```