

# Logistic Regression and Modeling

*Xingyao Chen*

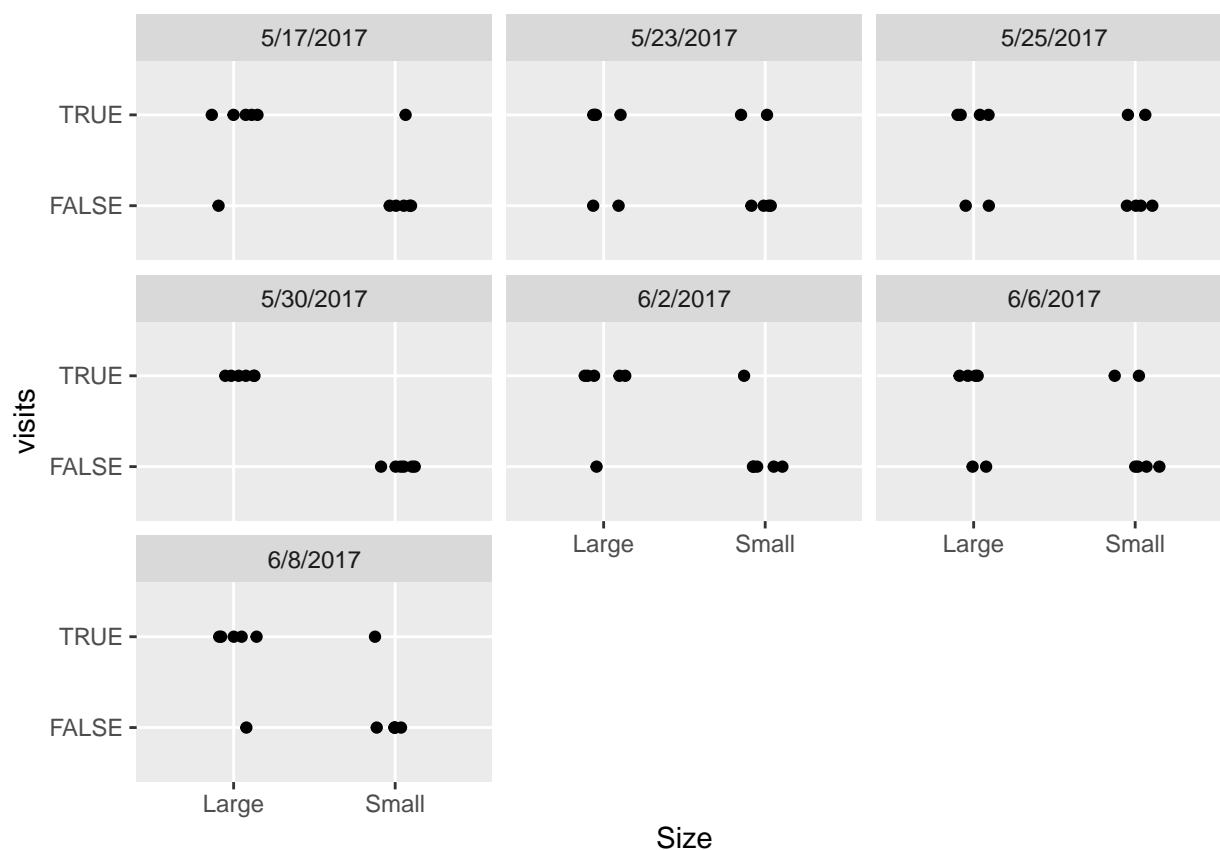
*6/21/2017*

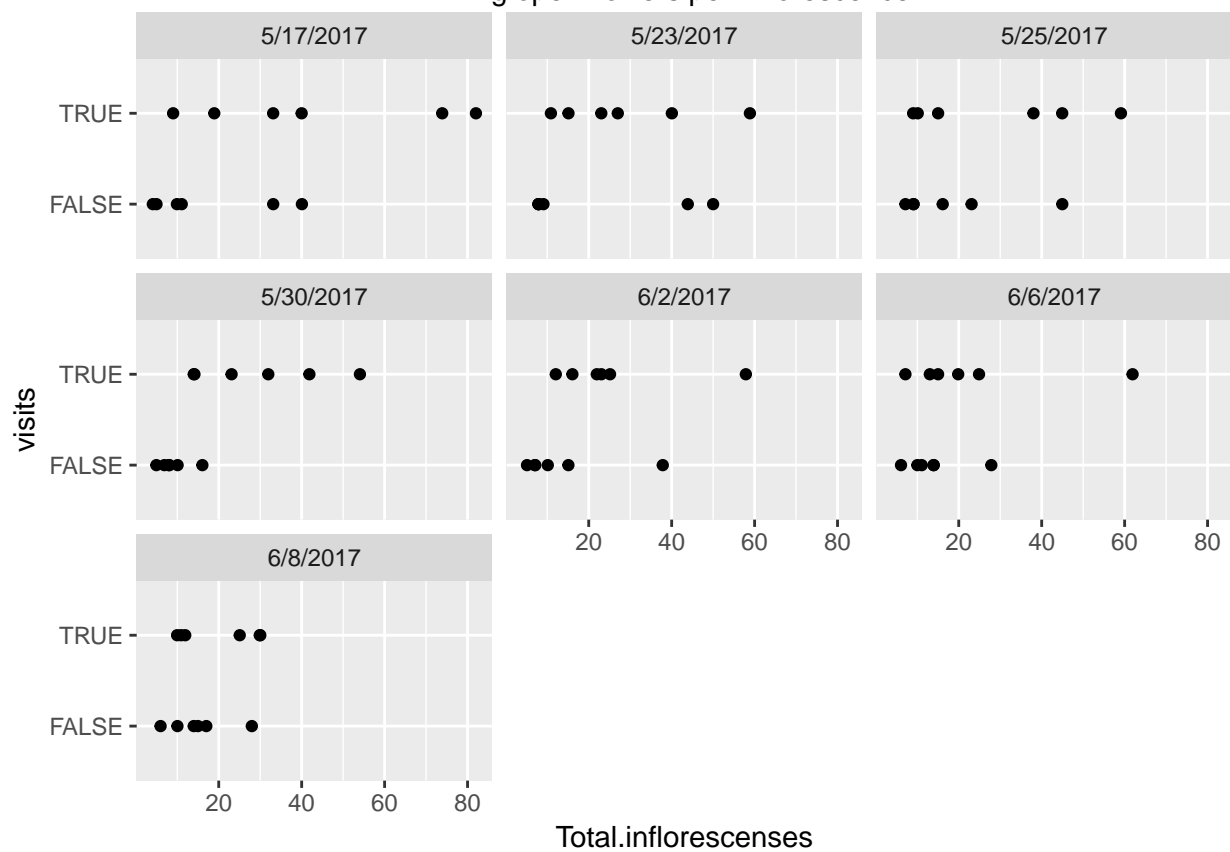
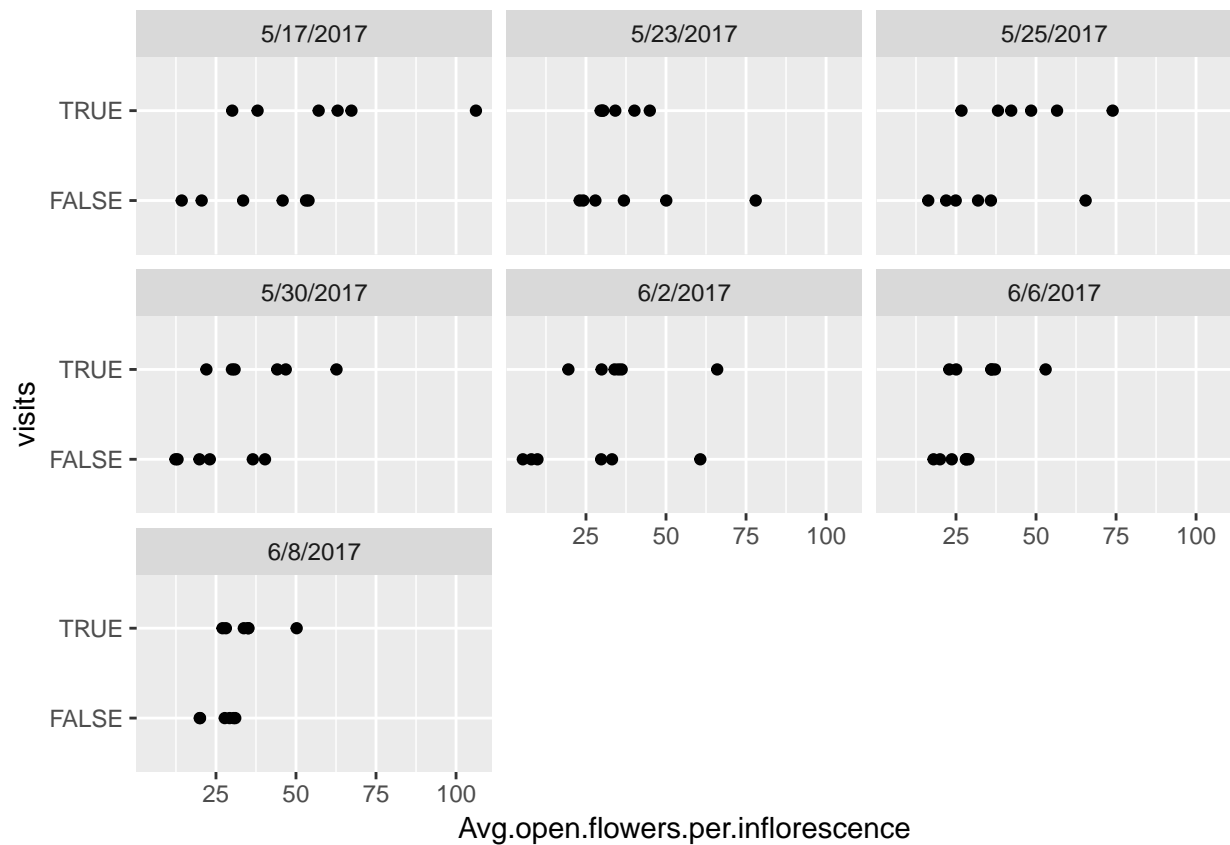
## Logistic Regressions with Non-Transformed Data

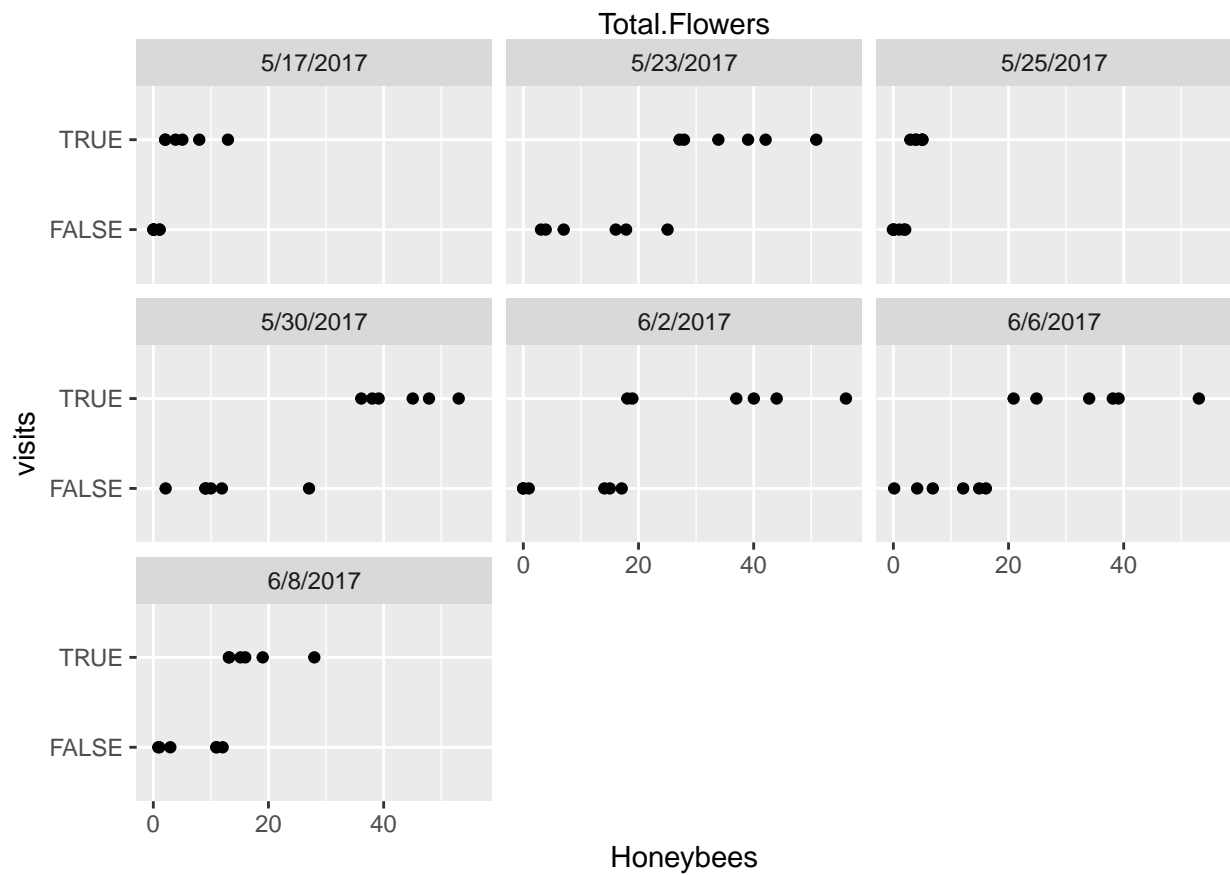
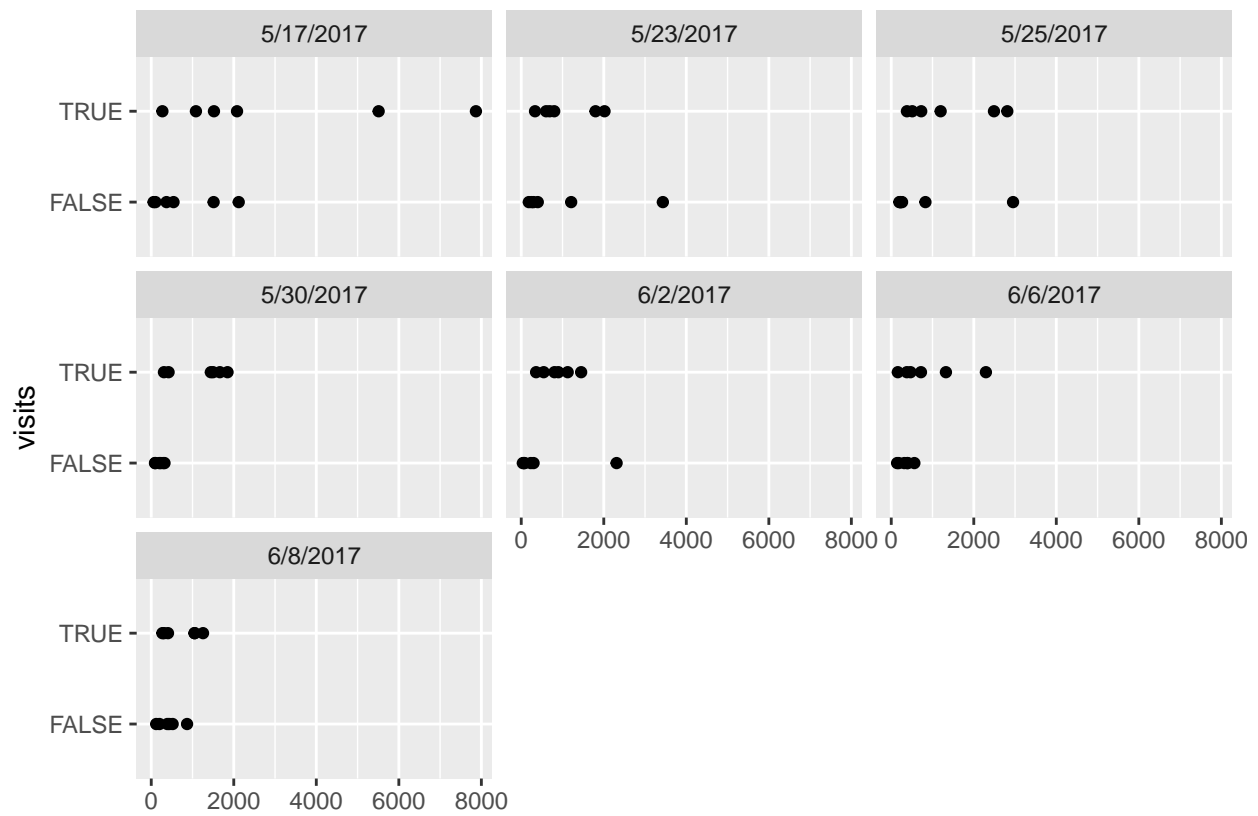
```
data=read.csv('pollinator_visitation_fullData.csv')
data$Experiment.Week=as.factor(data$Experiment.Week)
data$Plant.Number=as.factor(data$Plant.Number)
data$Pair=as.factor(data$Pair)

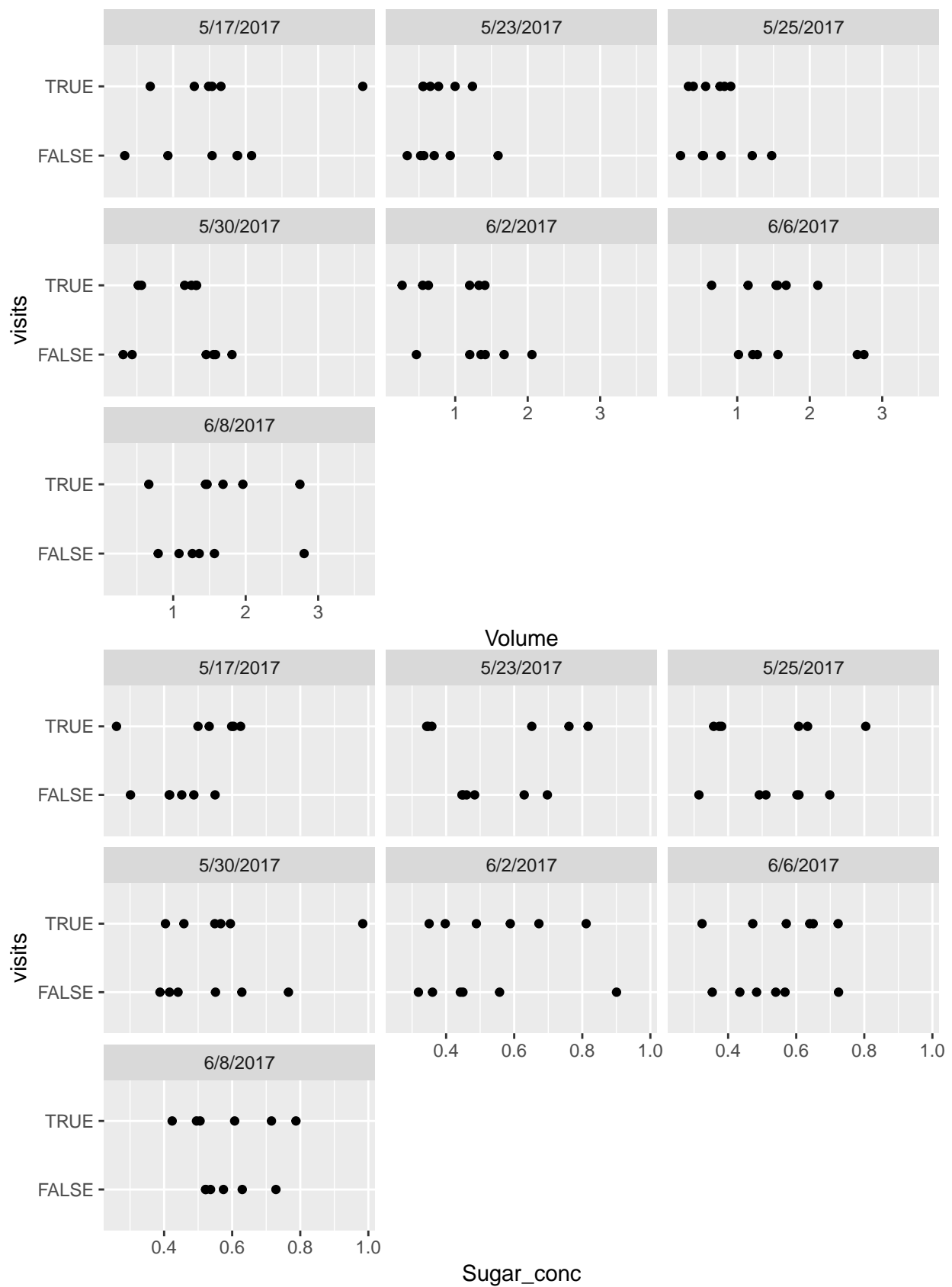
col=c('Size','Avg.open.flowers.per.inflorescence','Total.inflorescences','Total.Flowers','Honeybees', 'Honeybees')

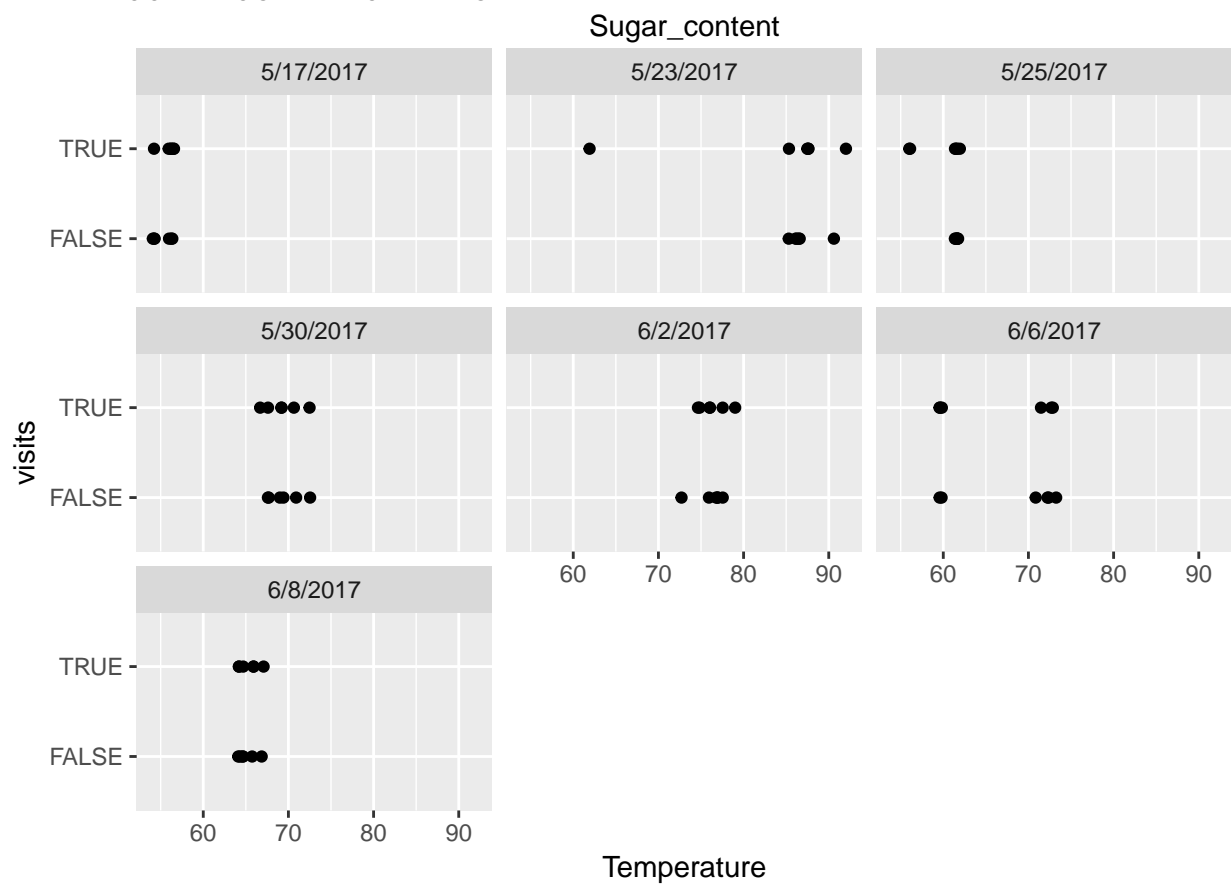
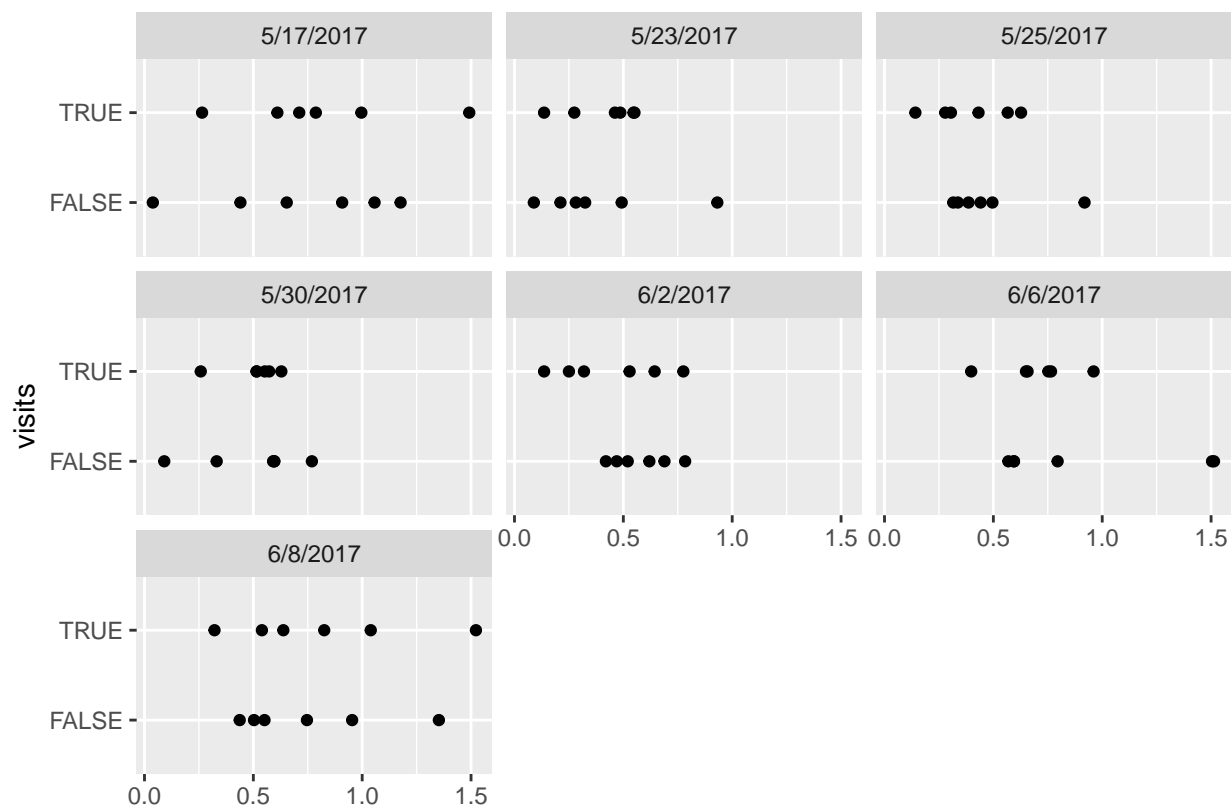
for (i in col){
  vars=data[,c('Date',i, 'visits')]
  p=ggplot(data=vars, aes(x=vars[,2], y=visits))+
    geom_jitter(height=0, width=0.15)+
    labs(x=i)+
    facet_wrap(~Date)
  plot(p)
}
```

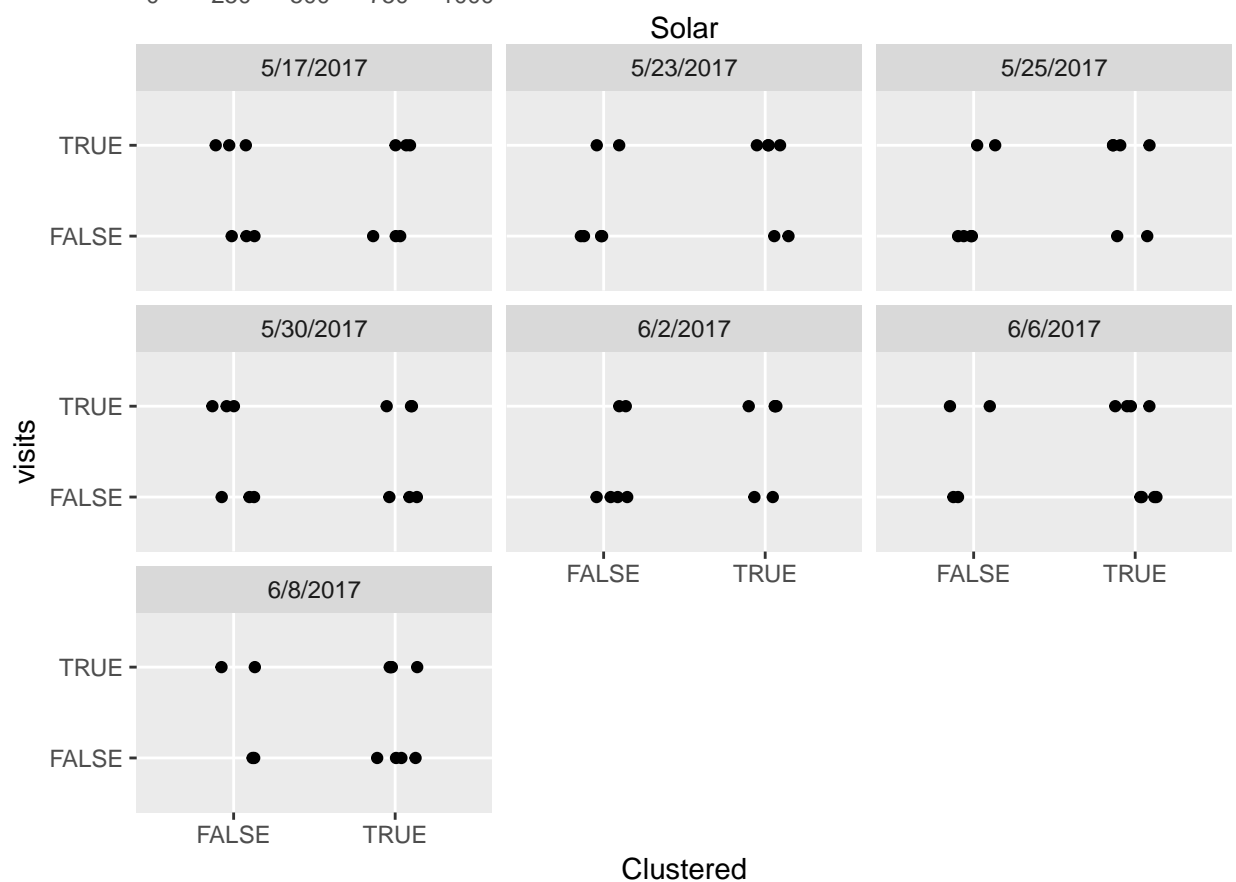
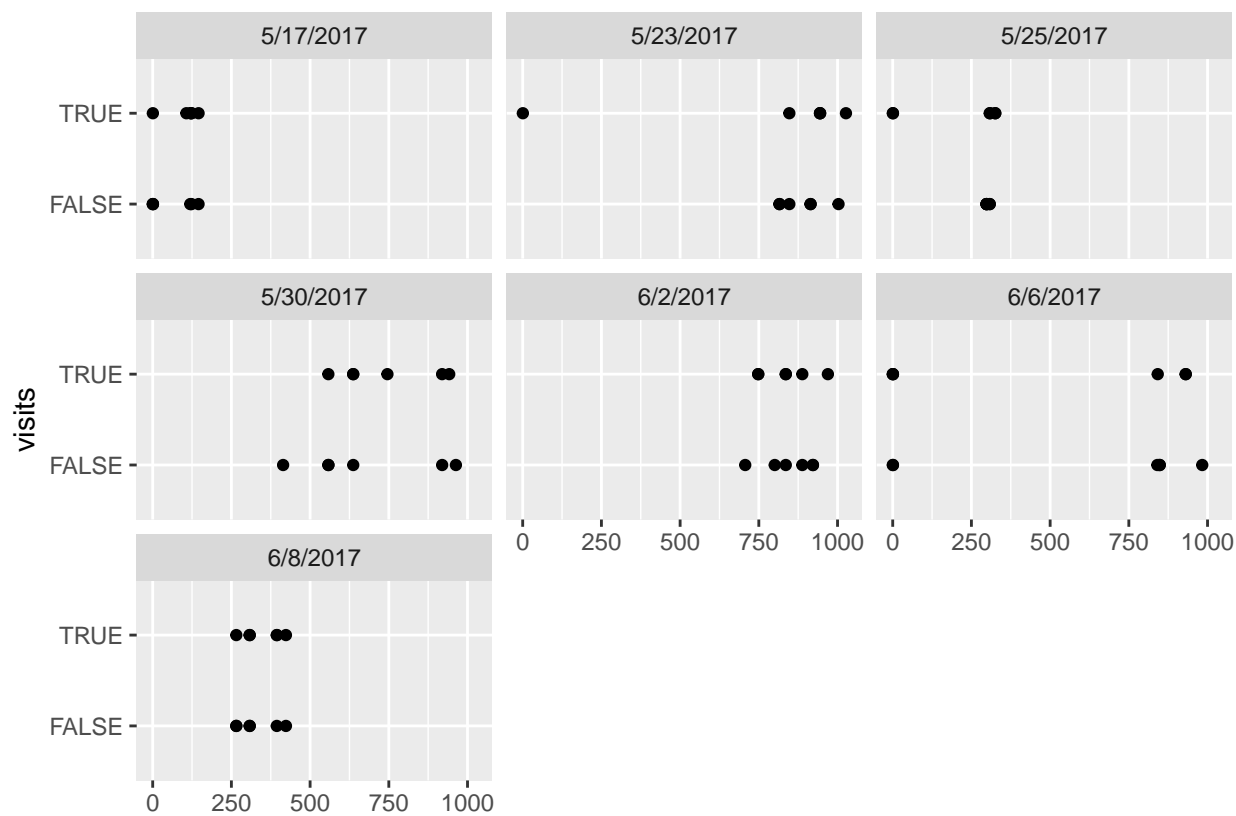










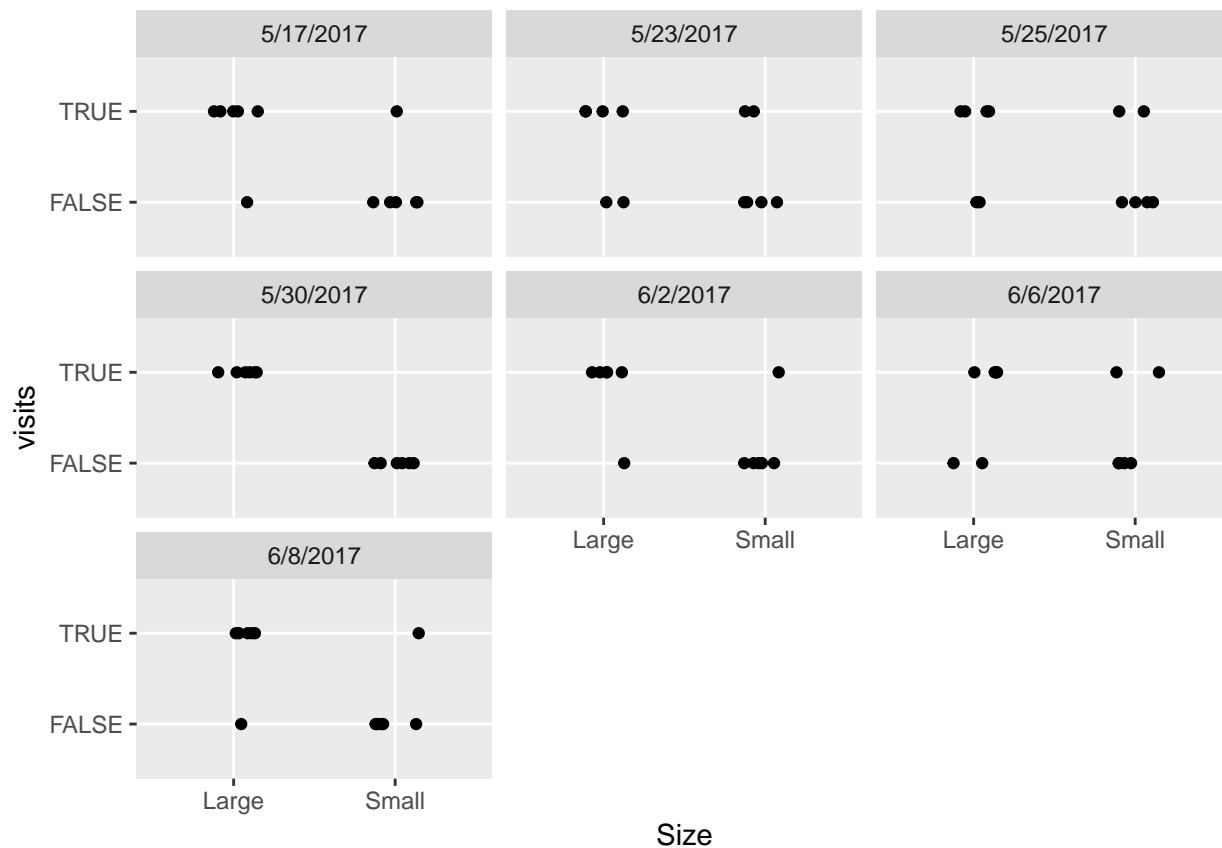


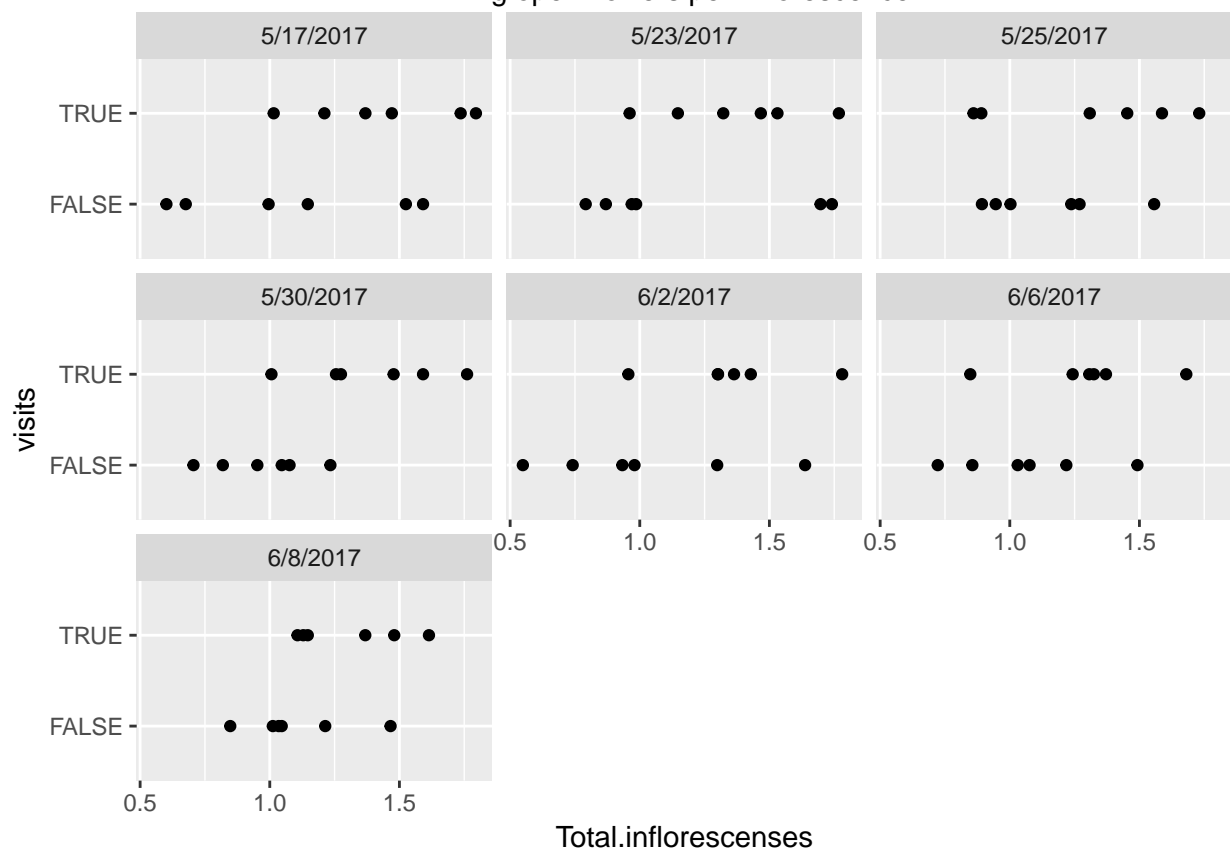
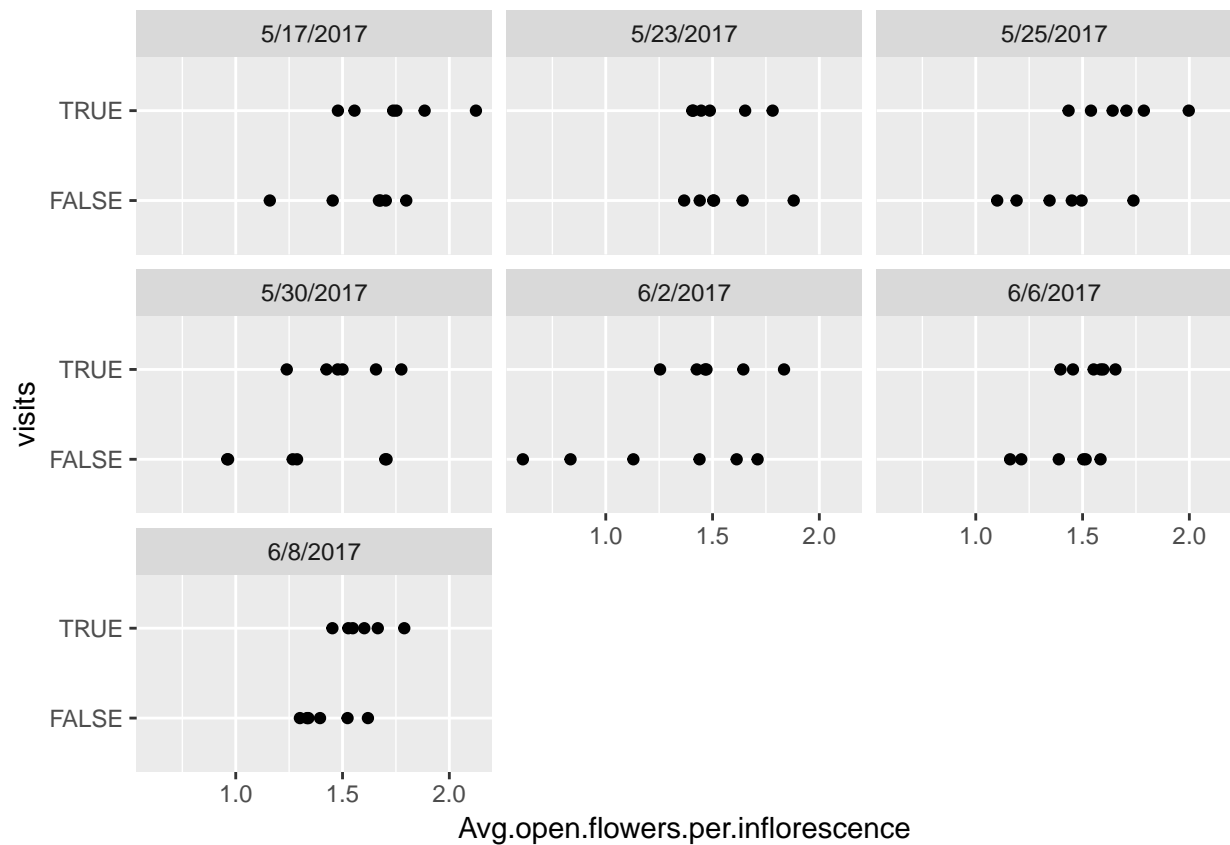
## Logistic Regression with Log-Transformed Data

```
data2=read.csv('pollinator_visitation_fullData_logTrans.csv')
data2$Experiment.Week=as.factor(data2$Experiment.Week)
data2$Plant.Number=as.factor(data2$Plant.Number)
data2$Pair=as.factor(data2$Pair)

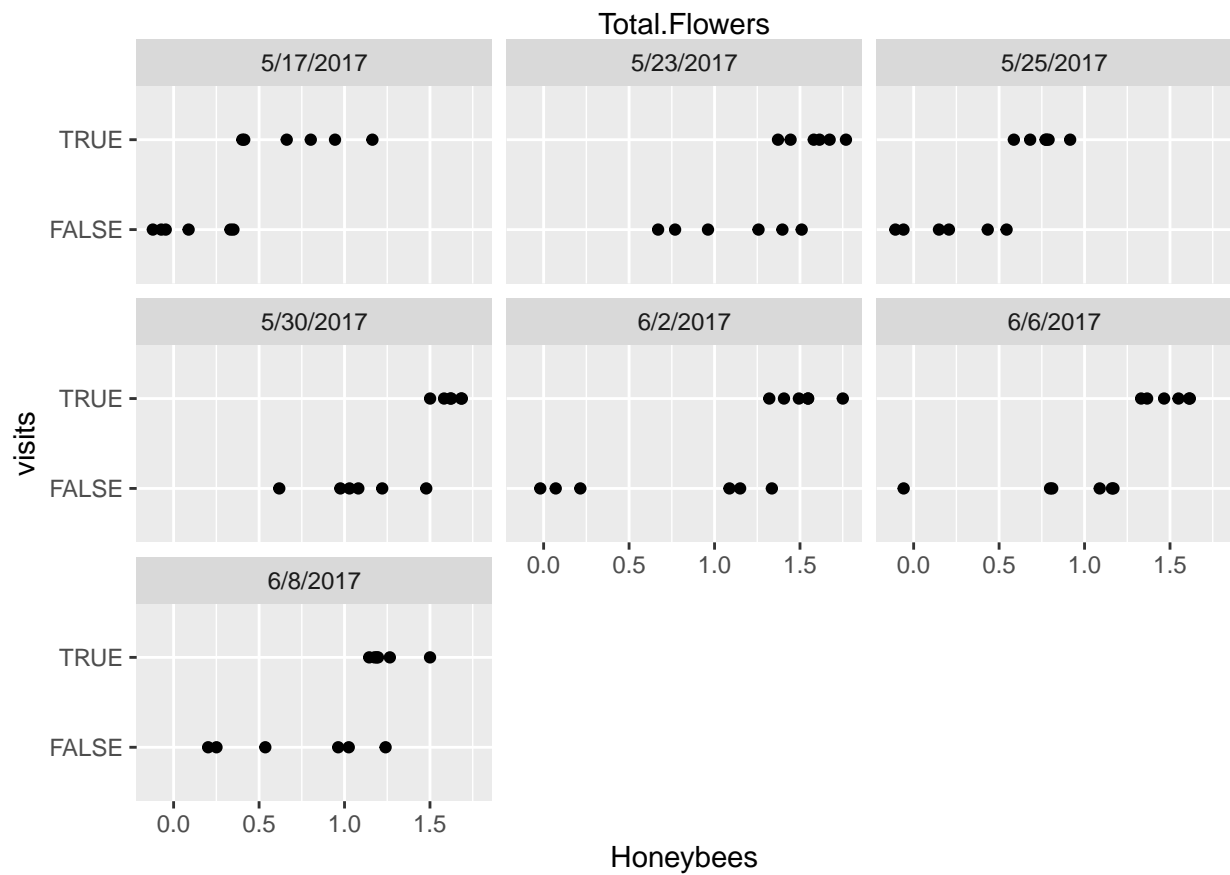
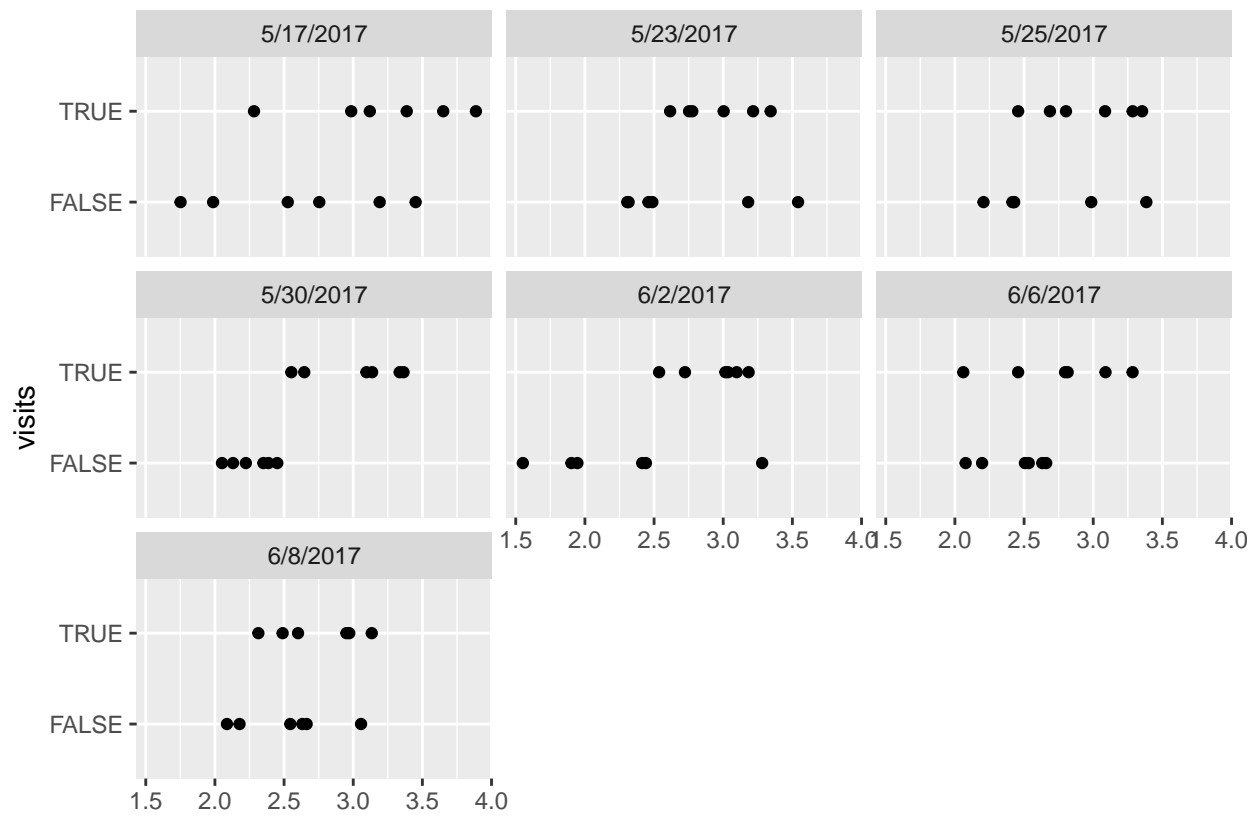
col2=c('Size','Avg.open.flowers.per.inflorescence','Total.inflorescences','Total.Flowers','Honeybees' ,

for (j in col2){
  vars=data2[,c('Date',j, 'visits')]
  p2=ggplot(data=vars, aes(x=vars[,2], y=visits))+
    geom_jitter(height=0, width=0.15)+
    labs(x=j)+
    facet_wrap(~Date)
  plot(p2)
}
```

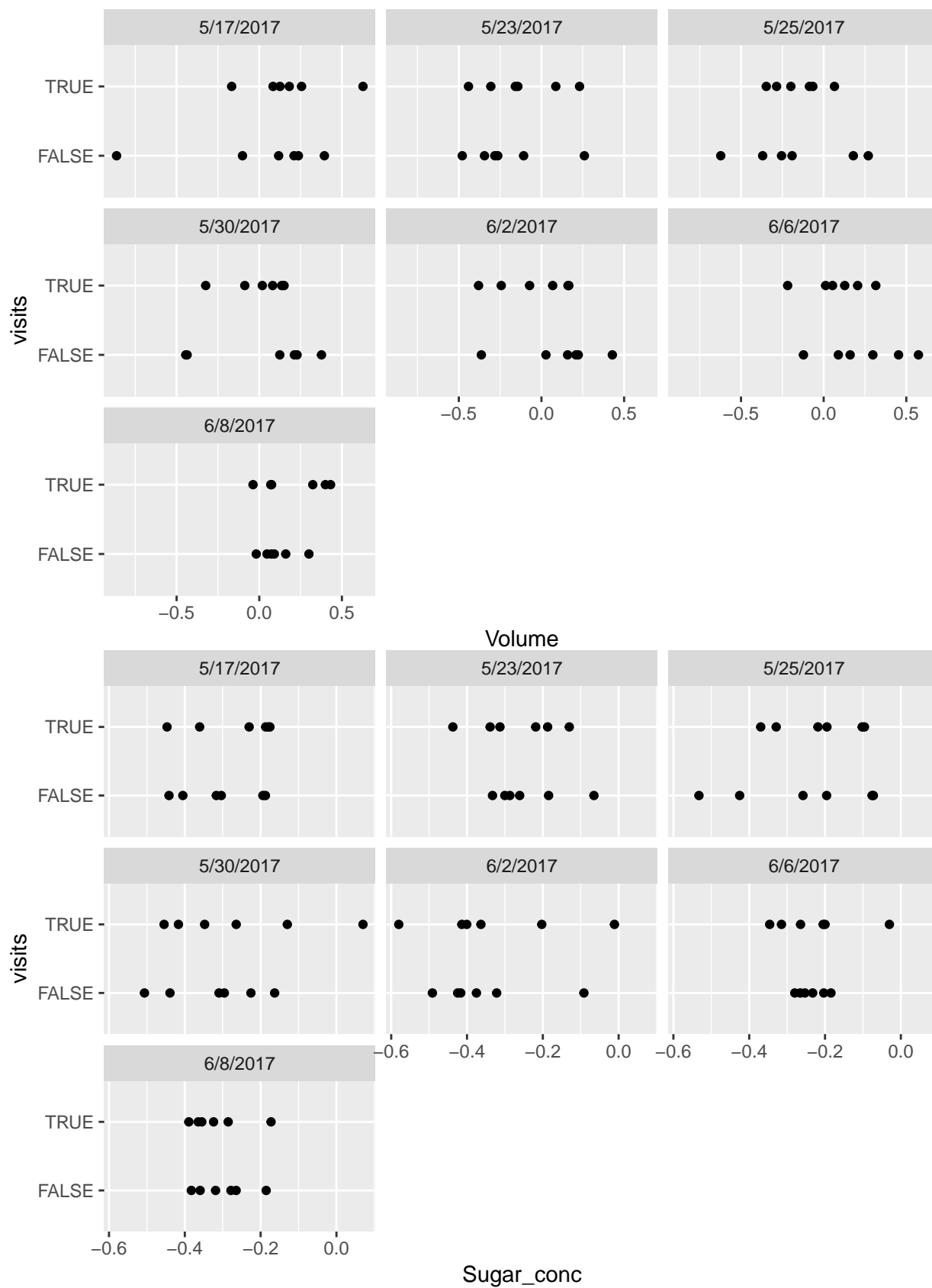


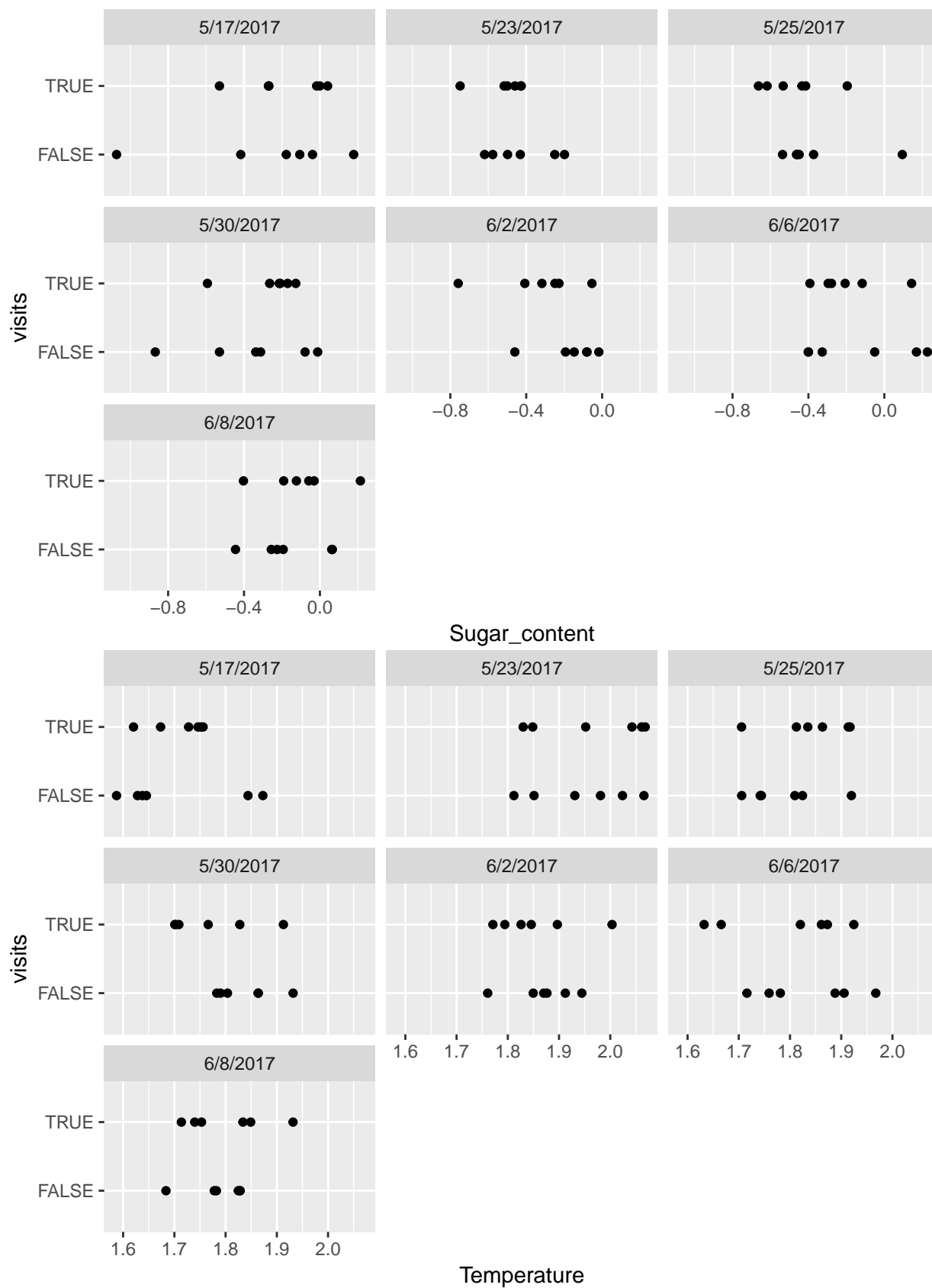


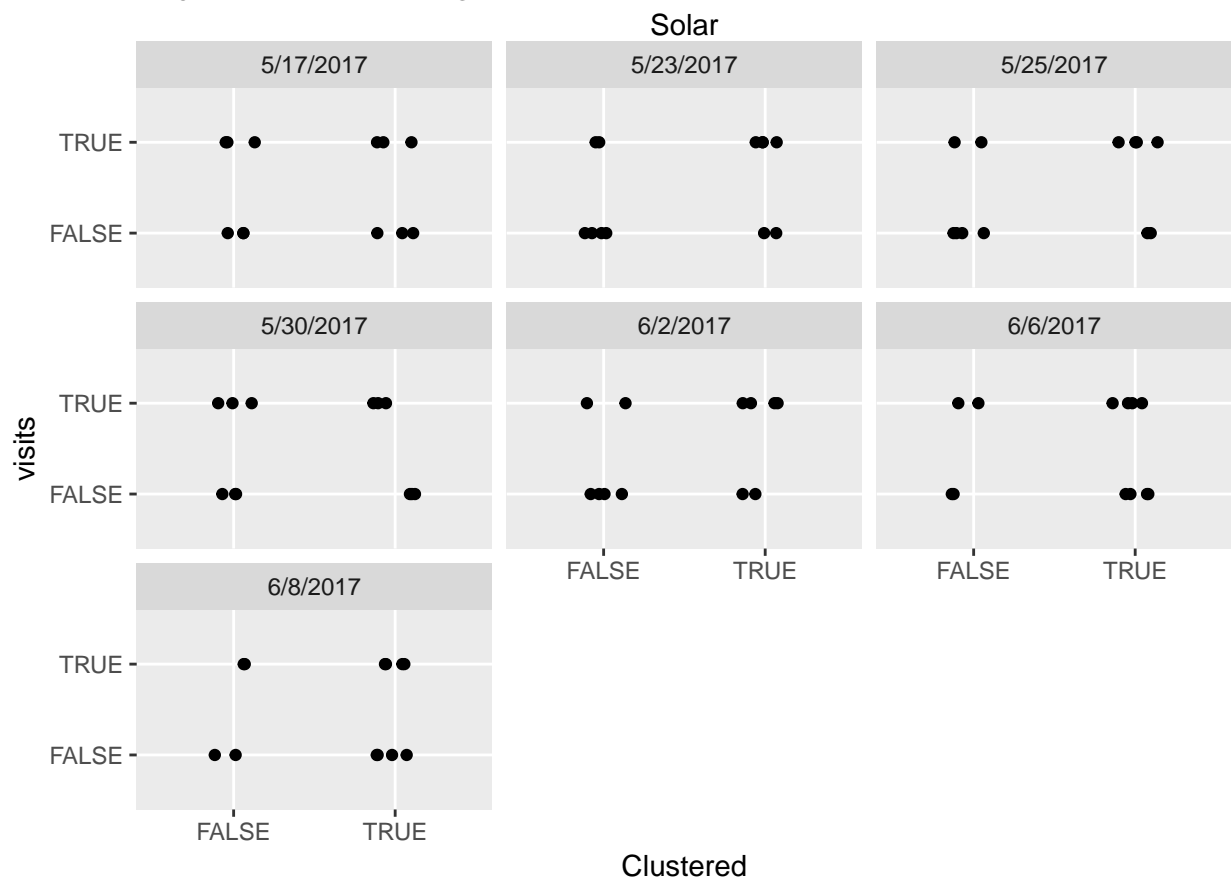
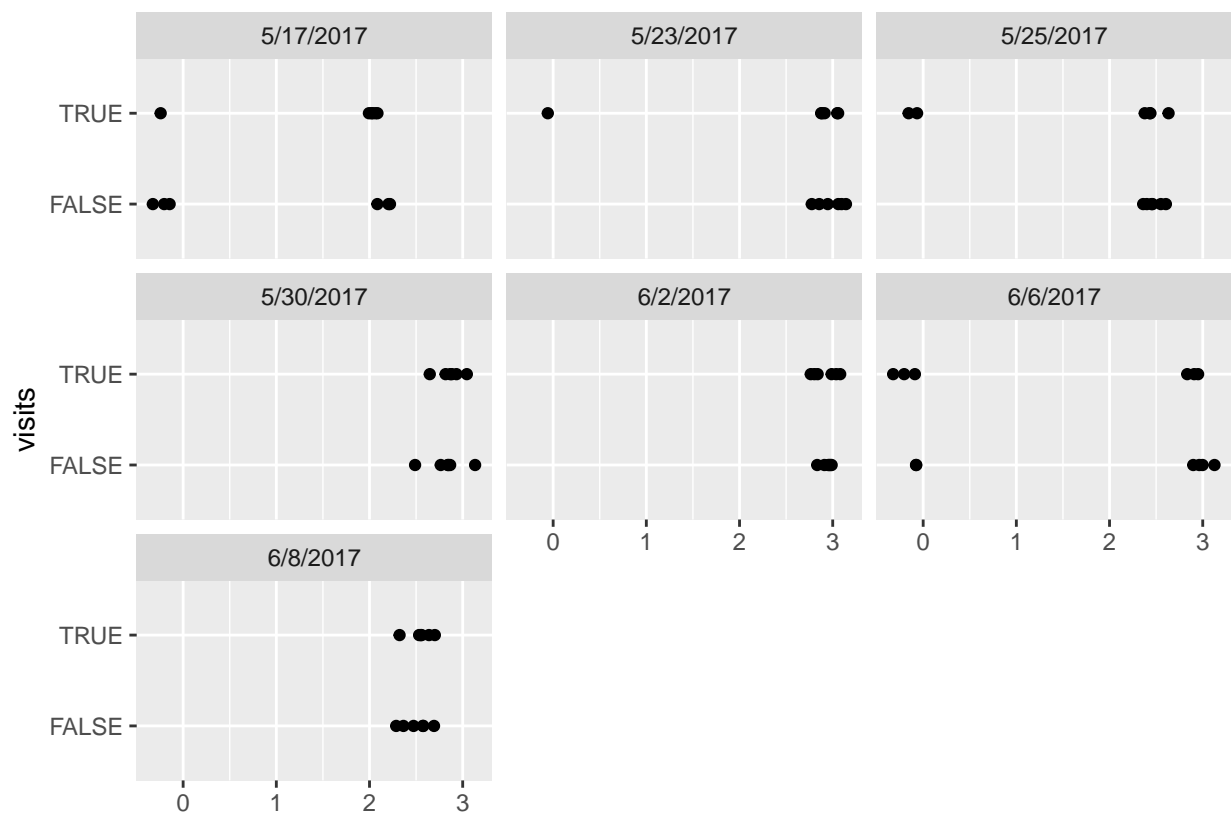




Honeybees







## General Linear Modeling With Mixed Effects

### Non-transformed first

```
mylogit1 <- glmm(visits ~ Avg.open.flowers.per.inflorescence+
                  Total.inflorescences
                  , random = list( ~ 0 + factor(Pair), ~0+factor(East)),
                  varcomps.names = c( "Location", 'East'),data = data,
                  family.glmm = binomial.glmm, m = 10^4)

summary(mylogit1)

##
## Call:
## glmm(fixed = visits ~ Avg.open.flowers.per.inflorescence + Total.inflorescences,
##      random = list(~0 + factor(Pair), ~0 + factor(East)), varcomps.names = c("Location",
##      "East"), data = data, family.glmm = binomial.glmm, m = 10^4)
##
##
## Link is: "logit (log odds)"
##
## Fixed Effects:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -2.60773    0.76851  -3.393 0.000691
## Avg.open.flowers.per.inflorescence  0.04025    0.02171   1.853 0.063818
## Total.inflorescences    0.06547    0.02358   2.777 0.005493
##
## (Intercept)          ***
## Avg.open.flowers.per.inflorescence .
## Total.inflorescences      **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Variance Components for Random Effects (P-values are one-tailed):
##              Estimate Std. Error z value Pr(>|z|)/2
## Location 1.8599330  1.0381685   1.792   0.0366 *
## East    0.0007843  0.0015071   0.520   0.3014
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

mylogit2 <- glmm(visits ~ Size+Clustered
                  , random = list( ~ 0 + factor(Pair), ~0+factor(East)),
                  varcomps.names = c( "Location", 'East'),data = data,
                  family.glmm = binomial.glmm, m = 10^4)

summary(mylogit2)

##
## Call:
## glmm(fixed = visits ~ Size + Clustered, random = list(~0 + factor(Pair),
##      ~0 + factor(East)), varcomps.names = c("Location", "East"),
##      data = data, family.glmm = binomial.glmm, m = 10^4)
```

```
##
##
## Link is: "logit (log odds)"
##
## Fixed Effects:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.0662    0.7813   1.365   0.172
## SizeSmall     -3.6055    0.7235  -4.984 6.24e-07 ***
## ClusteredTRUE   0.7344    1.0055   0.730   0.465
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Variance Components for Random Effects (P-values are one-tailed):
##           Estimate Std. Error z value Pr(>|z|)/2
## Location  2.1045185   1.5189477   1.386   0.0829 .
## East      0.0006555   0.0021500   0.305   0.3802
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Log-transformed second

```
mylogit1_log <- glmm(visits ~ Avg.open.flowers.per.inflorescence+
  Total.inflorescences
  , random = list( ~ 0 + factor(Pair), ~0+factor(East)),
  varcomps.names = c( "Location", 'East'),data = data2,
  family.glmm = binomial.glmm, m = 10^4)

summary(mylogit1_log)

##
## Call:
## glmm(fixed = visits ~ Avg.open.flowers.per.inflorescence + Total.inflorescences,
##      random = list(~0 + factor(Pair), ~0 + factor(East)), varcomps.names = c("Location",
##      "East"), data = data2, family.glmm = binomial.glmm, m = 10^4)
##
##
## Link is: "logit (log odds)"
##
## Fixed Effects:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -9.112     2.876  -3.168 0.00154 **
## Avg.open.flowers.per.inflorescence    2.967     2.003   1.481 0.13857
## Total.inflorescences    3.534     1.386   2.549 0.01080 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Variance Components for Random Effects (P-values are one-tailed):
##           Estimate Std. Error z value Pr(>|z|)/2
## Location  1.340658   1.366281   0.981   0.163
## East      0.001132   0.001979   0.572   0.284
```

```
mylogit2_log <- glmm(visits ~ Size+Clustered
, random = list( ~ 0 + factor(Pair), ~0+factor(East)),
varcomps.names = c( "Location", 'East'),data = data2,
family.glmm = binomial.glmm, m = 10^4)
```

```
summary(mylogit2_log)
```

```
##
## Call:
## glmm(fixed = visits ~ Size + Clustered, random = list(~0 + factor(Pair),
## ~0 + factor(East)), varcomps.names = c("Location", "East"),
## data = data2, family.glmm = binomial.glmm, m = 10^4)
##
##
## Link is: "logit (log odds)"
##
## Fixed Effects:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.3979    1.0622   1.316 0.188136
## SizeSmall    -3.9187    1.0101  -3.879 0.000105 ***
## ClusteredTRUE  0.5792    1.3125   0.441 0.658969
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Variance Components for Random Effects (P-values are one-tailed):
##           Estimate Std. Error z value Pr(>|z|)/2
## Location   3.07505    3.14259   0.979    0.164
## East        0.00106    0.00382   0.277    0.391
```