Data Summaries

Eric Huang

3/2/2019

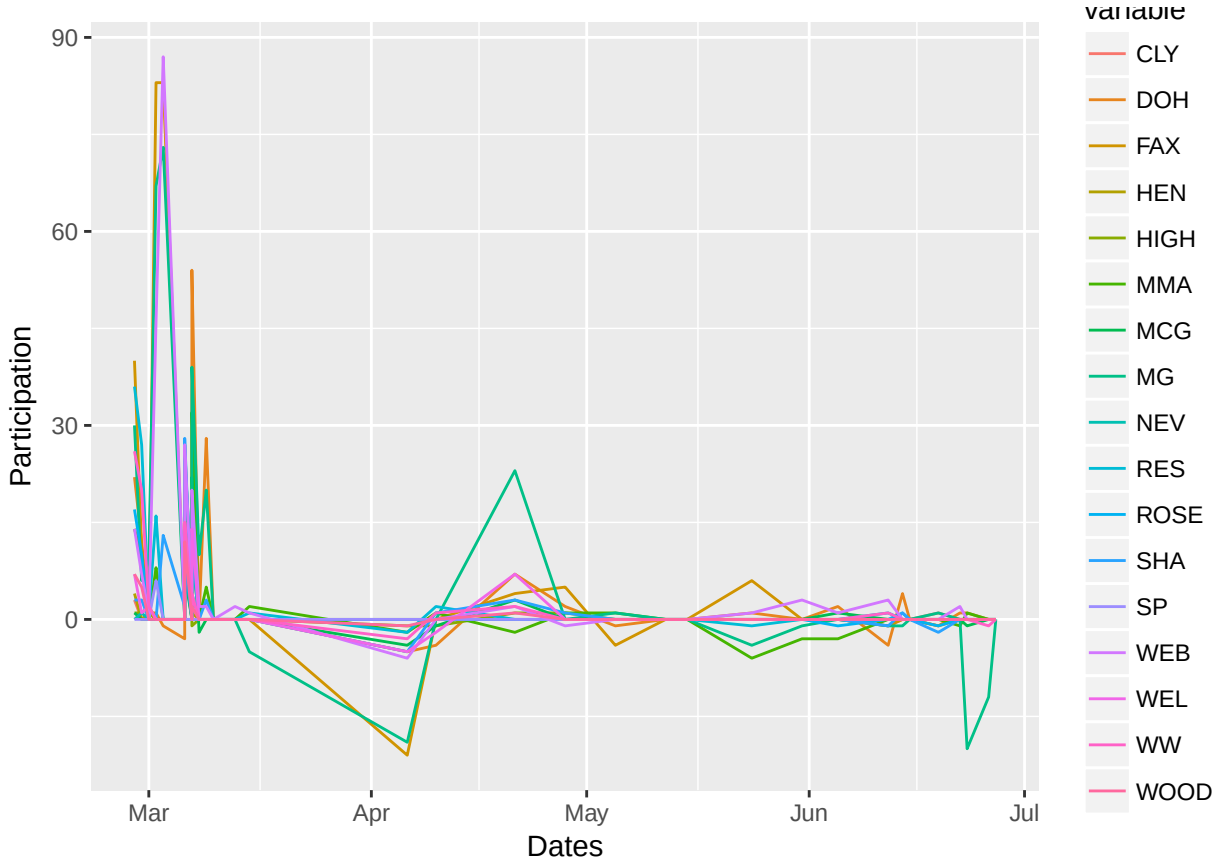
Data

```
rsNet1718 <- read.csv("~/Desktop/thesisDocuments/totalNet1718.csv")
vacancies1718 <- read.csv("~/Desktop/thesisDocuments/vacancies1718.csv")
occupancy1819 <- read.csv("~/Desktop/thesisDocuments/occupancy1819.csv")
surveyResults <- read.csv("~/Desktop/thesisDocuments/rsSurvey18.csv")
```

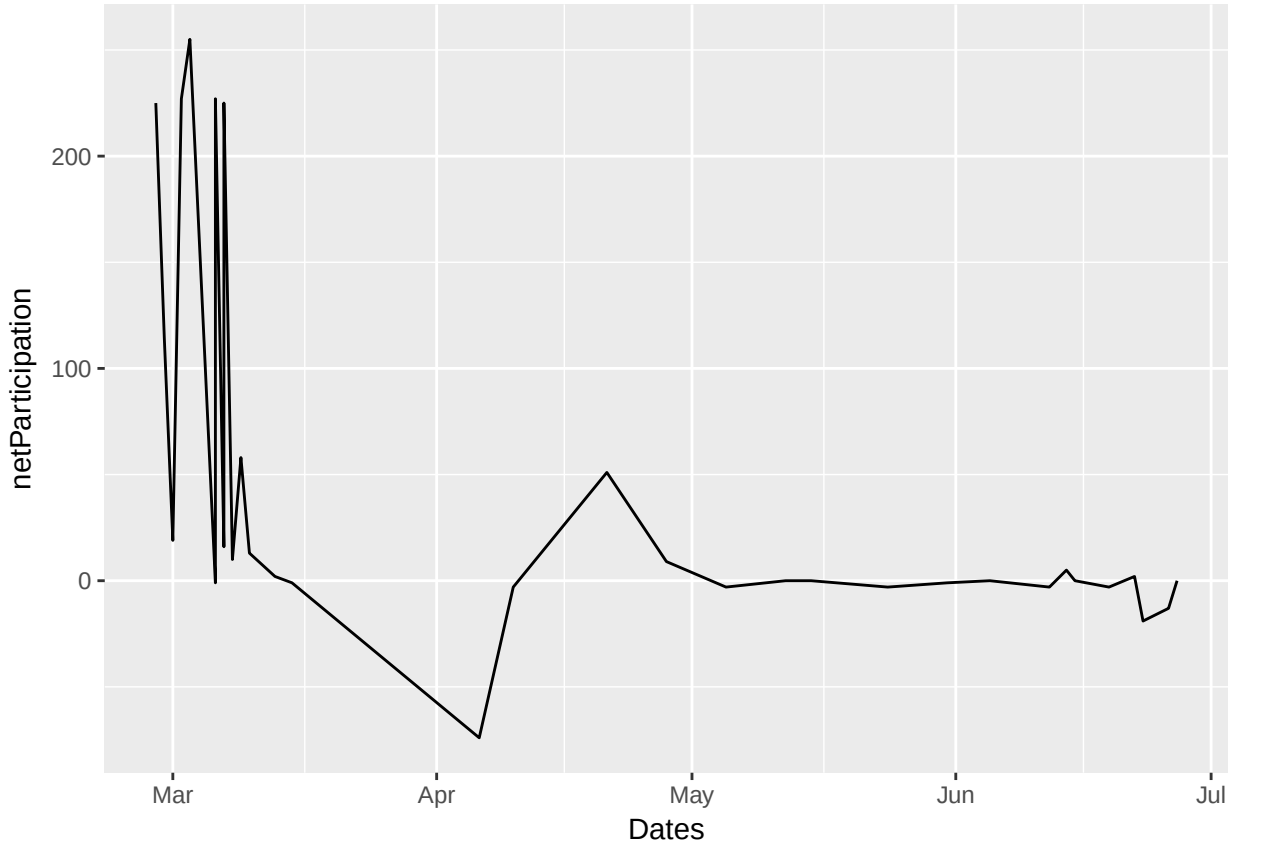
Exploratory Graphs

2017-18 Room Selection Participation

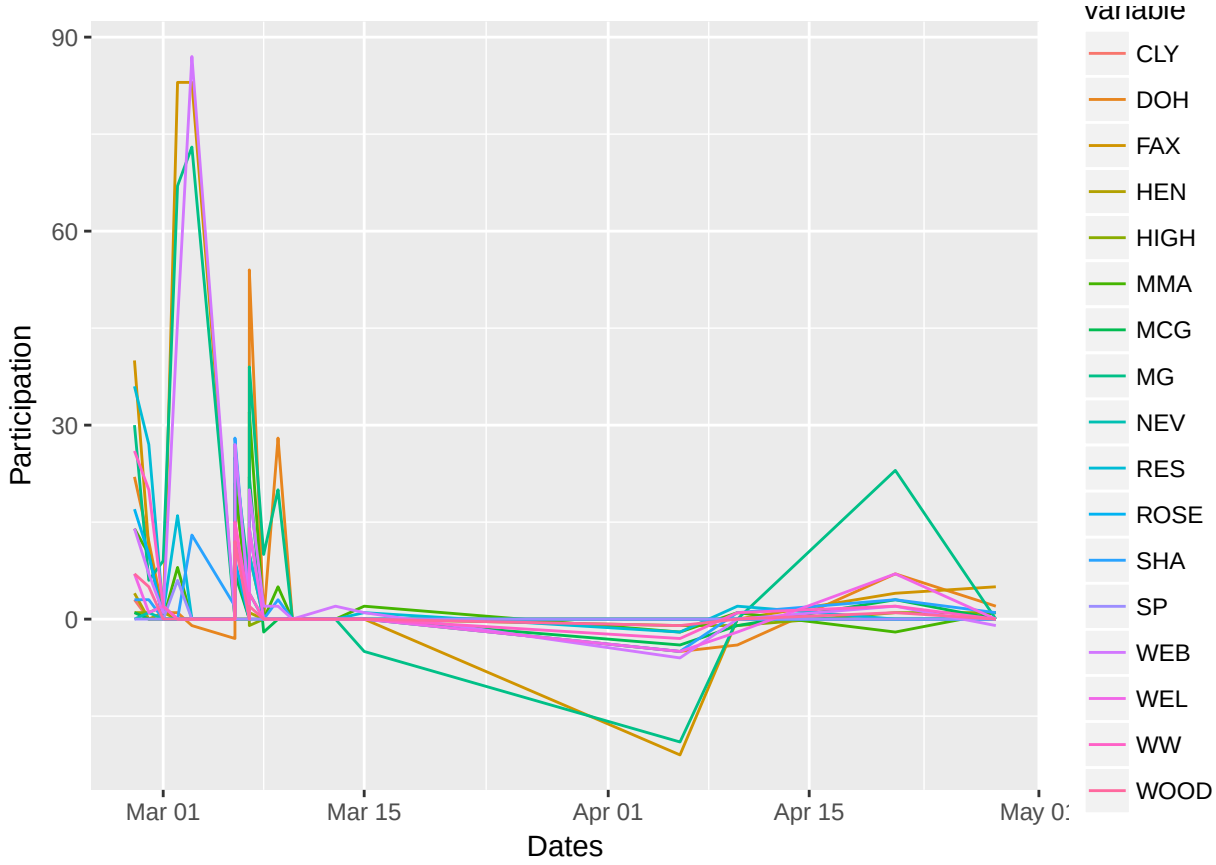
```
# rsNet1718 by Houses through June
ggplot(newRSNet1718ByHouses,aes(x = Dates, y = Participation, group = variable,
                                colour = variable)) +
  geom_line()
```



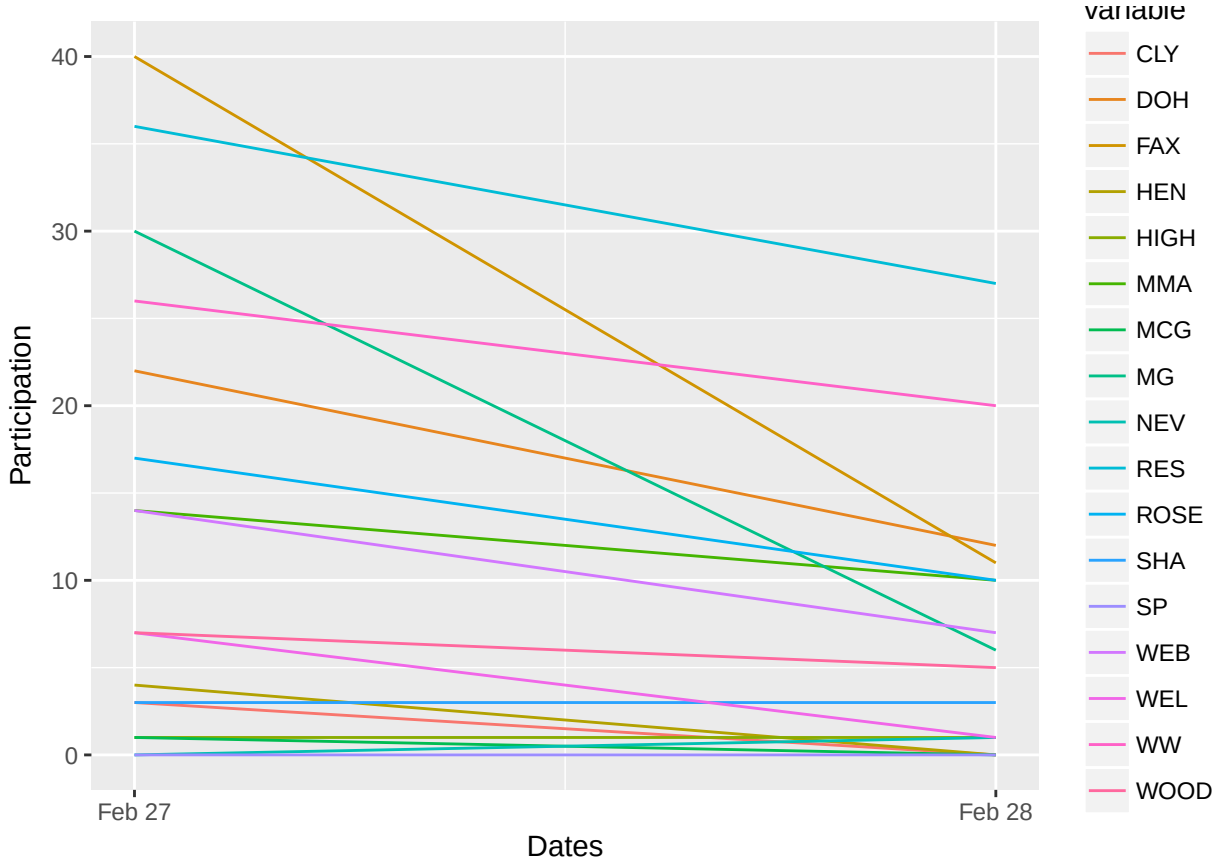

```
# rsNet1718 total participation through June
ggplot(rsNet1718Total,aes(x = Dates, y = netParticipation)) +
  geom_line()
```



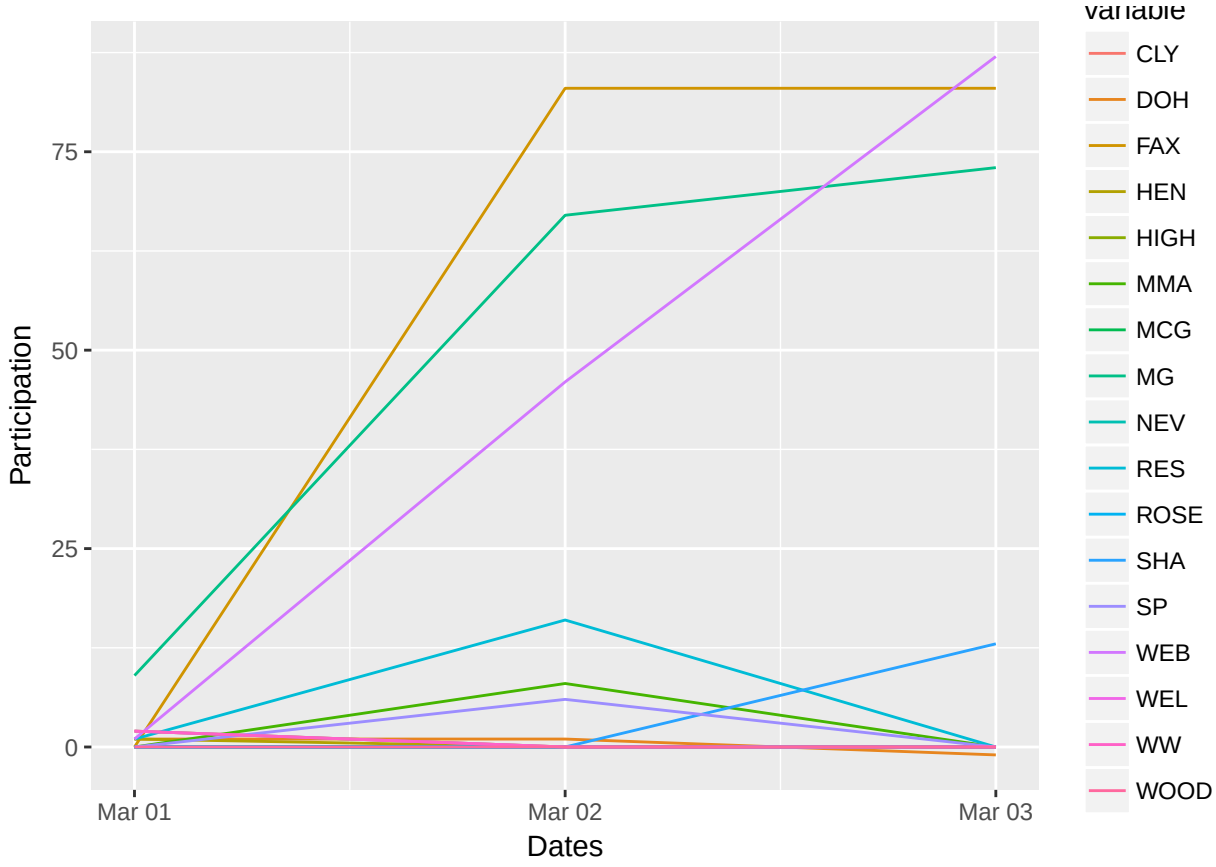

```
# rsNet1718 by Houses through April
ggplot(newRSNet1718ByHousesApr,aes(x = Dates, y = Participation, group = variable,
                                   colour = variable)) +
  geom_line() +
  labs(fill = "Houses")
```



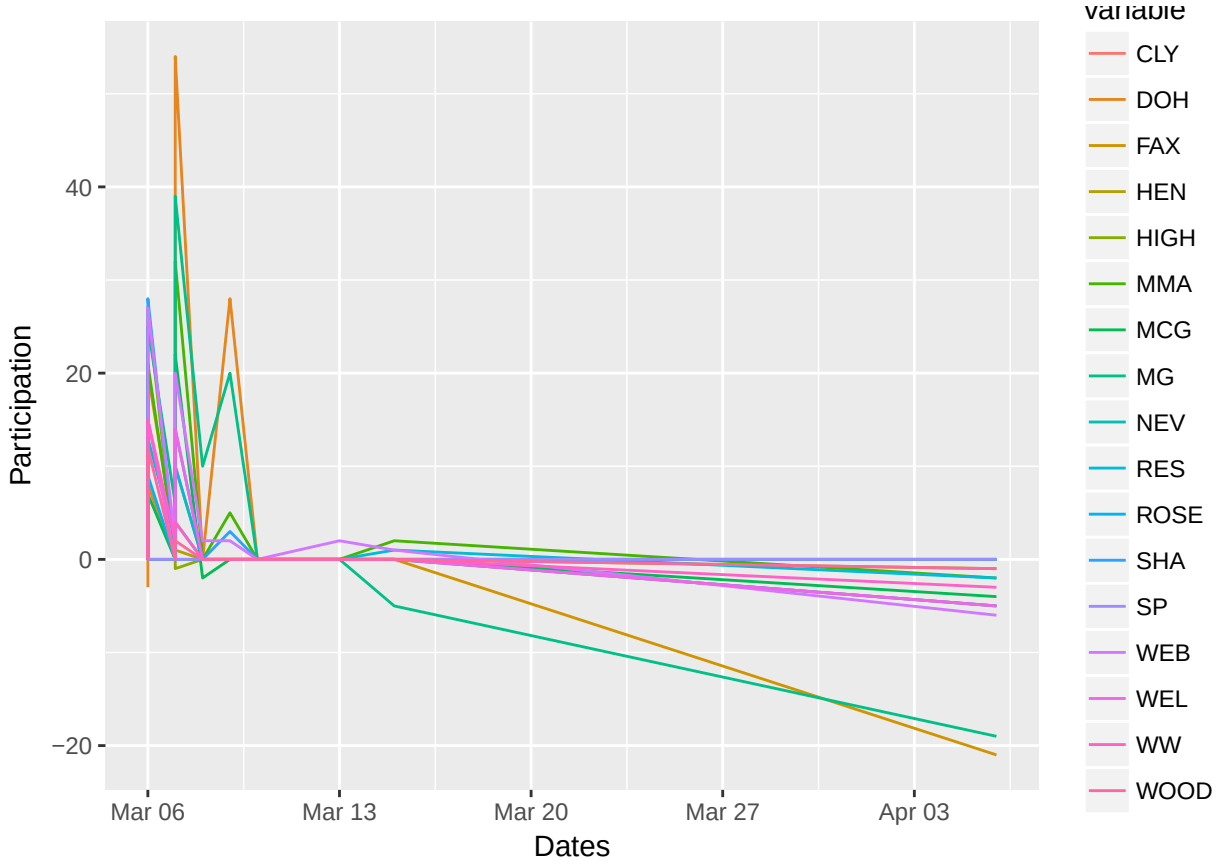

```
# rsNet1718 by Houses during Retention
ggplot(newRSNet1718ByHousesRetention,aes(x = Dates, y = Participation,
                                                    group = variable,
                                                    colour = variable)) +
                  geom_line() +
                  labs(fill = "Houses")
```



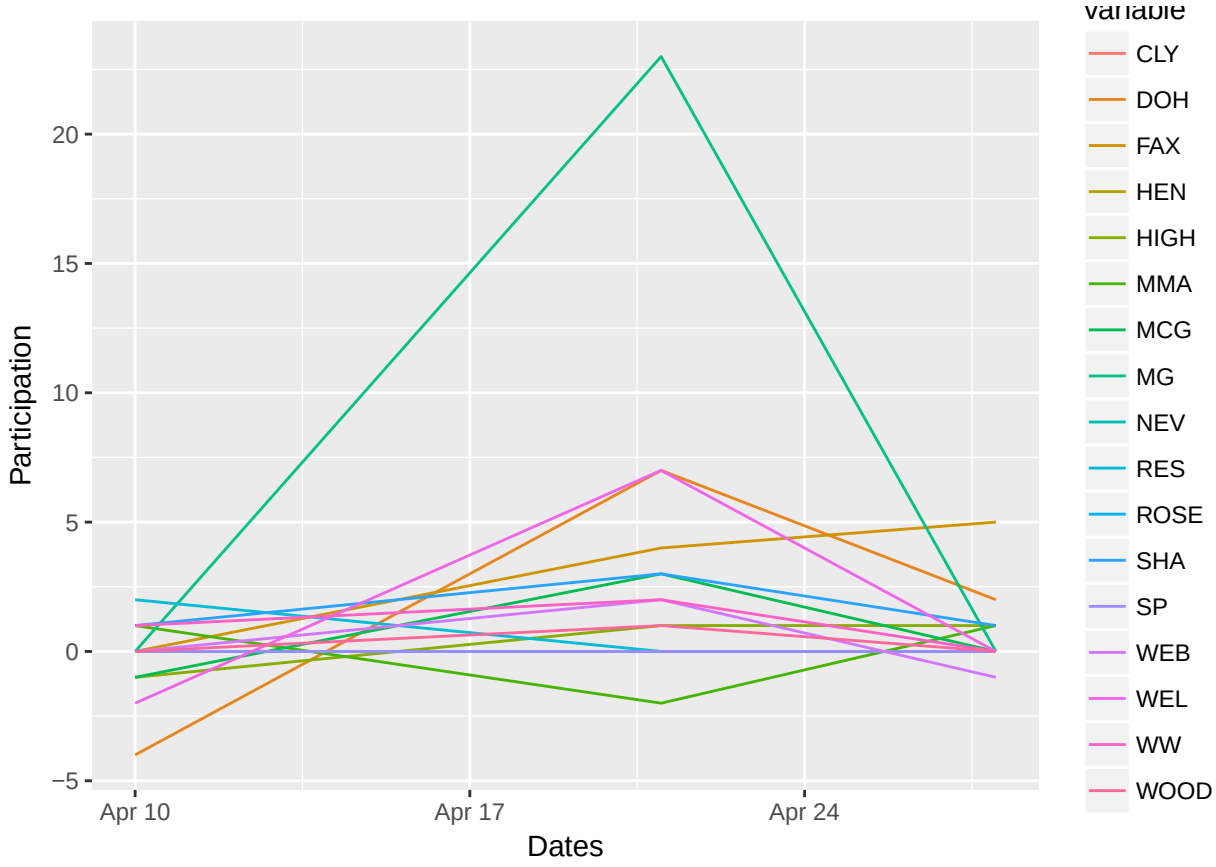

```
# rsNet1718 by Houses during Block Housing
ggplot(newRSNet1718ByHousesBlock,aes(x = Dates, y = Participation,
                                                group = variable,
                                                colour = variable)) +
                geom_line() +
                labs(fill = "Houses")
```



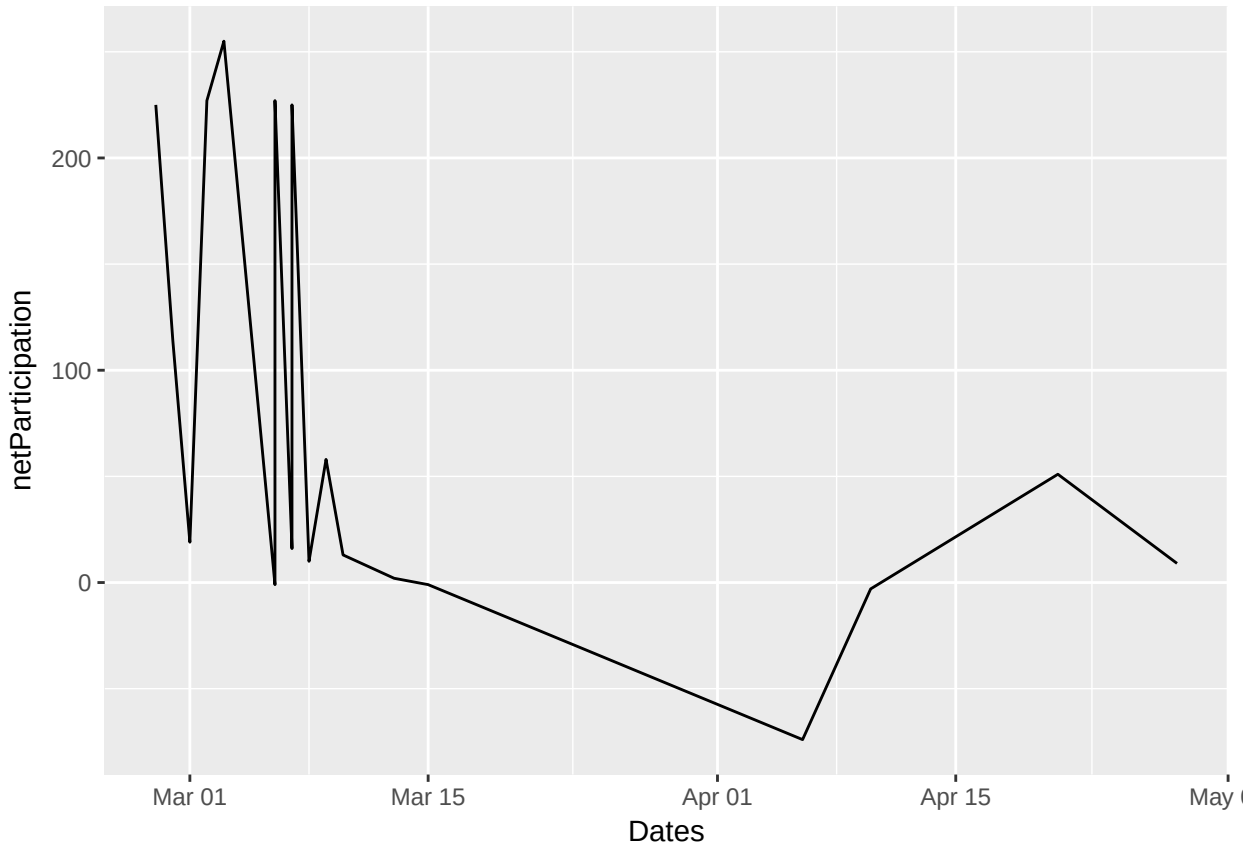

```
# rsNet1718 by Houses during General Selection
ggplot(newRSNet1718ByHousesGeneral,aes(x = Dates, y = Participation,
                                                  group = variable,
                                                  colour = variable)) +
                    geom_line() +
                    labs(fill = "Houses")
```



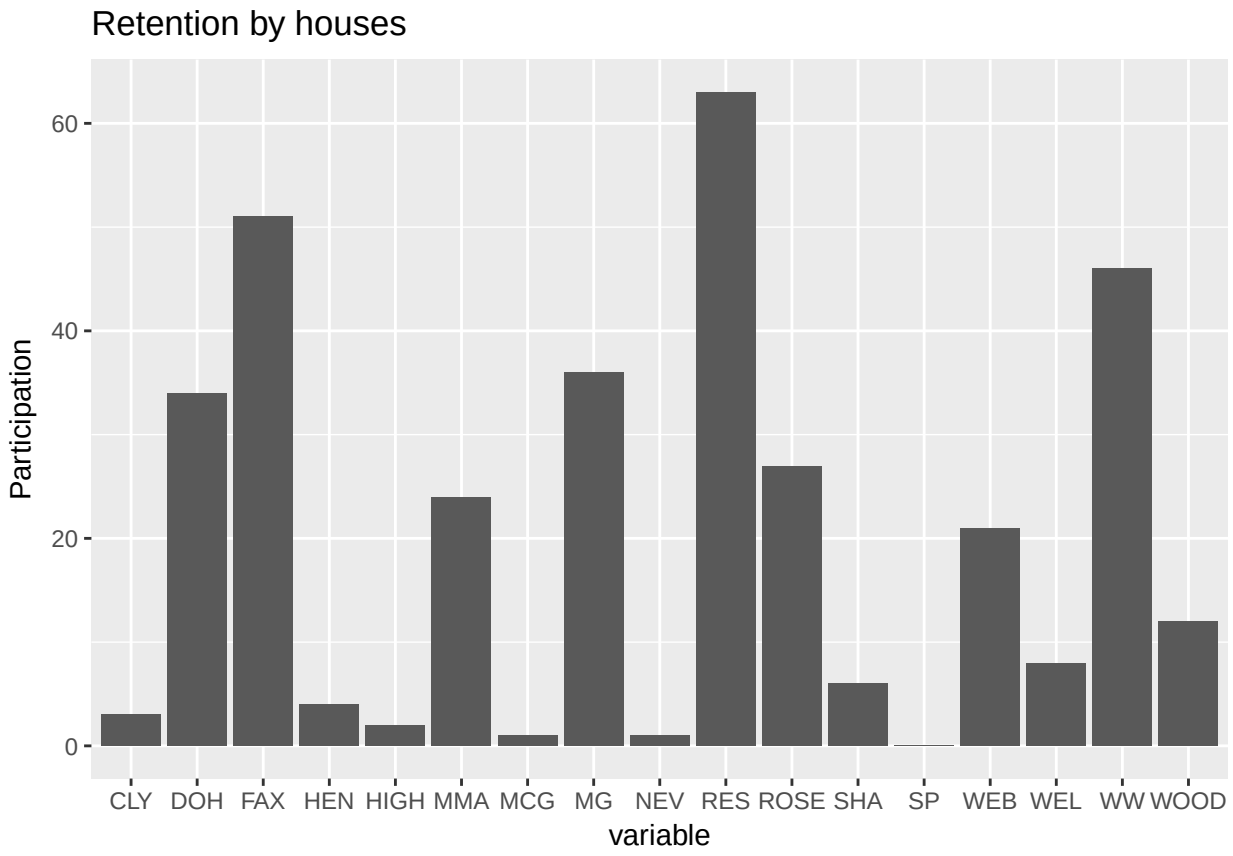

```
# rsNet1718 by Houses during Open Assignment
ggplot(newRSNet1718ByHousesOpen,aes(x = Dates, y = Participation,
                                    group = variable,
                                    colour = variable)) +
                geom_line() +
                labs(fill = "Houses")
```



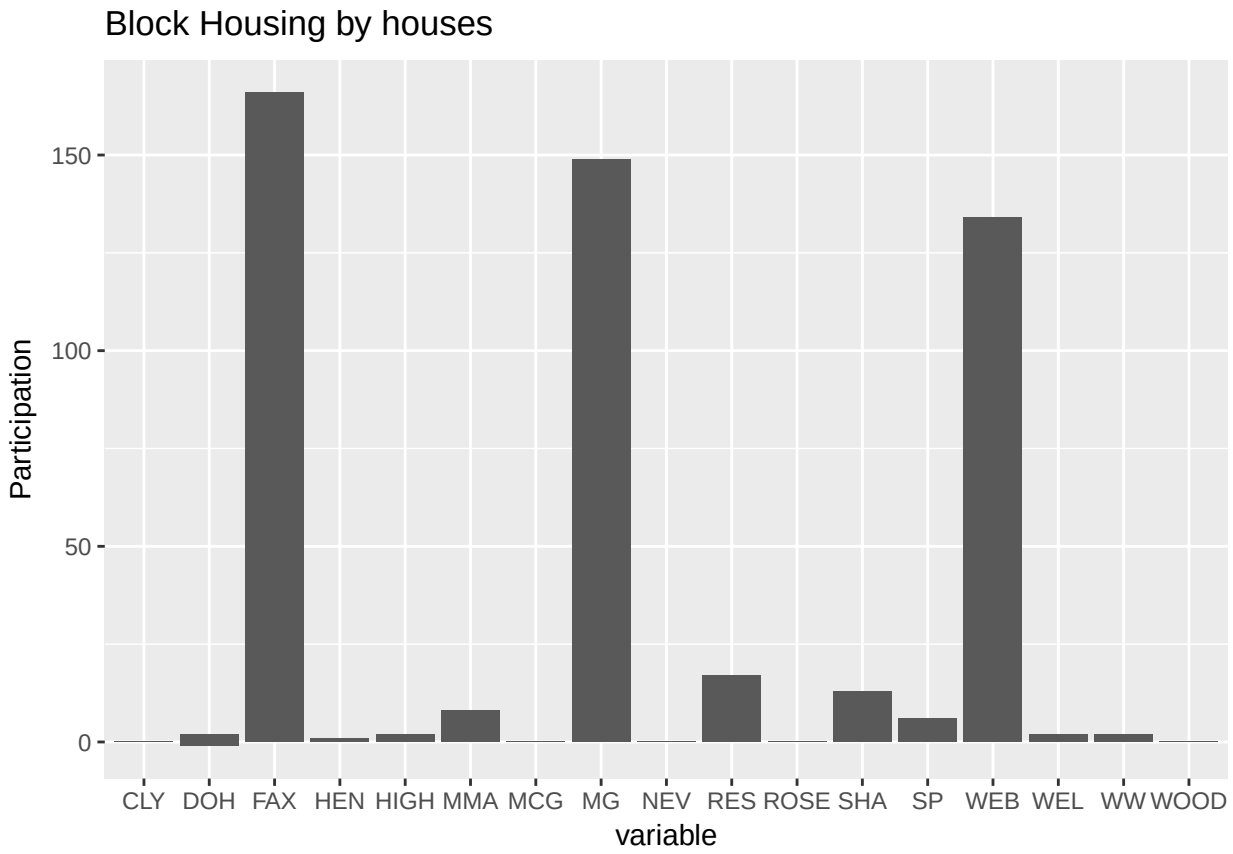

```
# rsNet1718 total participation through April
ggplot(rsNet1718TotalApr,aes(x = Dates, y = netParticipation)) +
  geom_line()
```




```
# bar charts of counts
ggplot(newRSNet1718ByHousesRetention,aes(x = variable, y = Participation)) +
                  geom_bar(stat = "identity") +
                  labs(title = "Retention by houses", fill = "Houses")
```

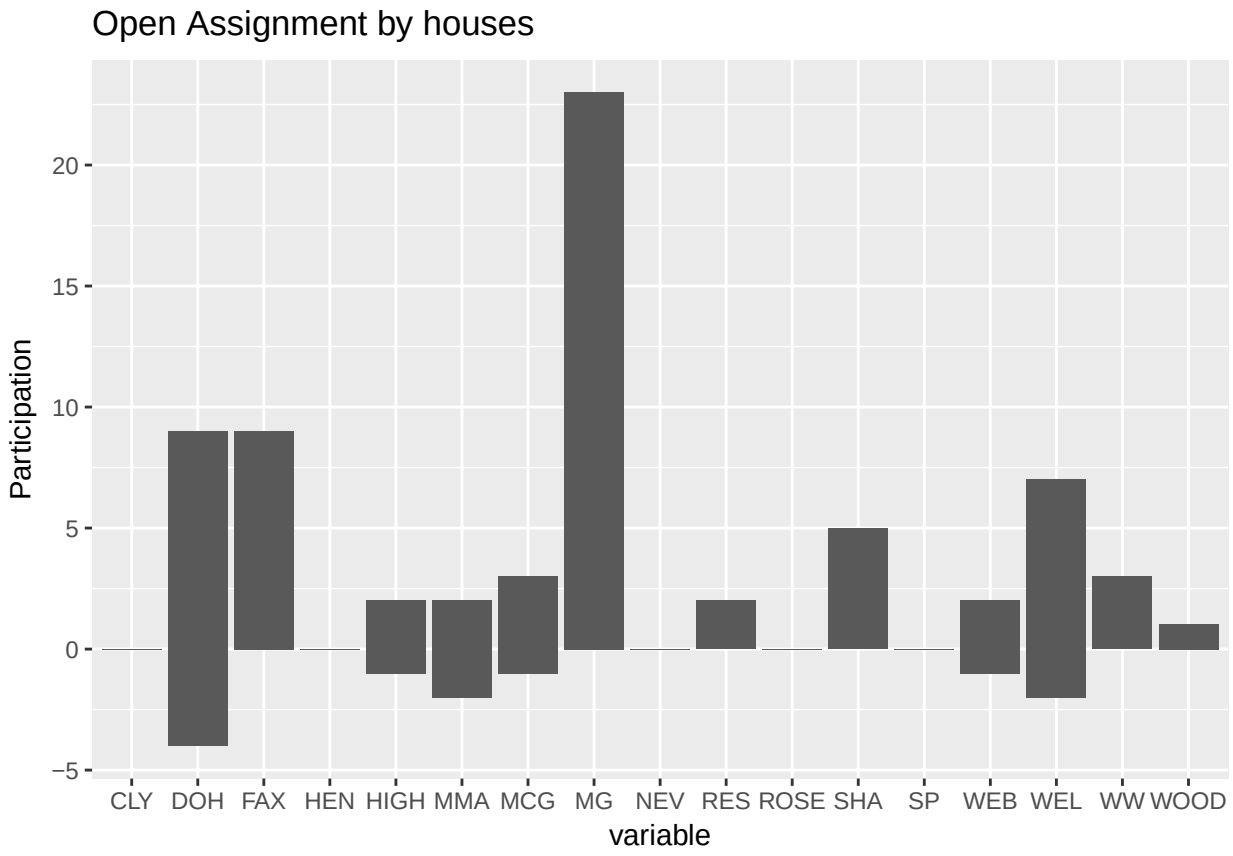



```
ggplot(newRSNet1718ByHousesBlock,aes(x = variable, y = Participation)) +
                  geom_bar(stat = "identity") +
                  labs(title = "Block Housing by houses", fill = "Houses")
```



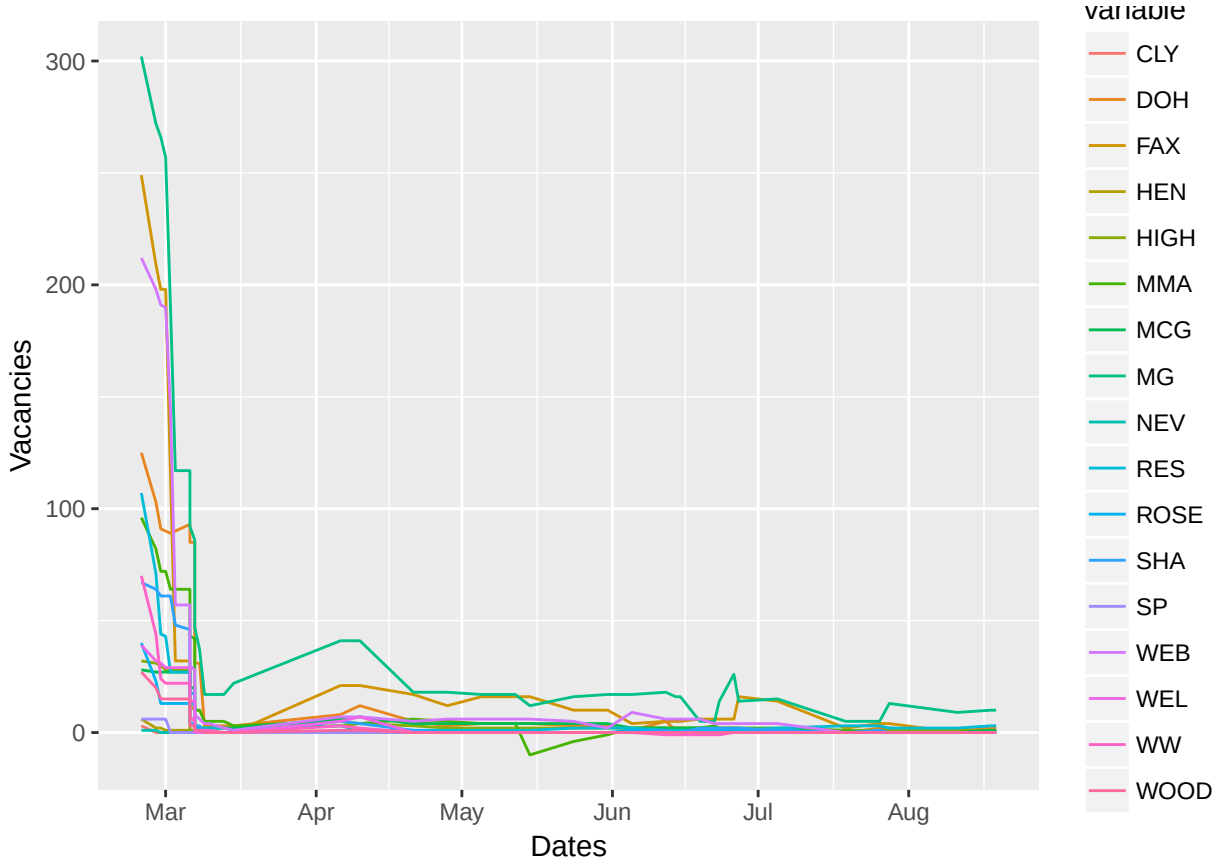

```
ggplot(newRSNet1718ByHousesGeneral,aes(x = variable, y = Participation)) +
                  geom_bar(stat = "identity") +
                  labs(title = "General Selection by houses", fill = "Houses")
```


```
ggplot(newRSNet1718ByHousesOpen,aes(x = variable, y = Participation)) +
                  geom_bar(stat = "identity") +
                  labs(title = "Open Assignment by houses", fill = "Houses")
```

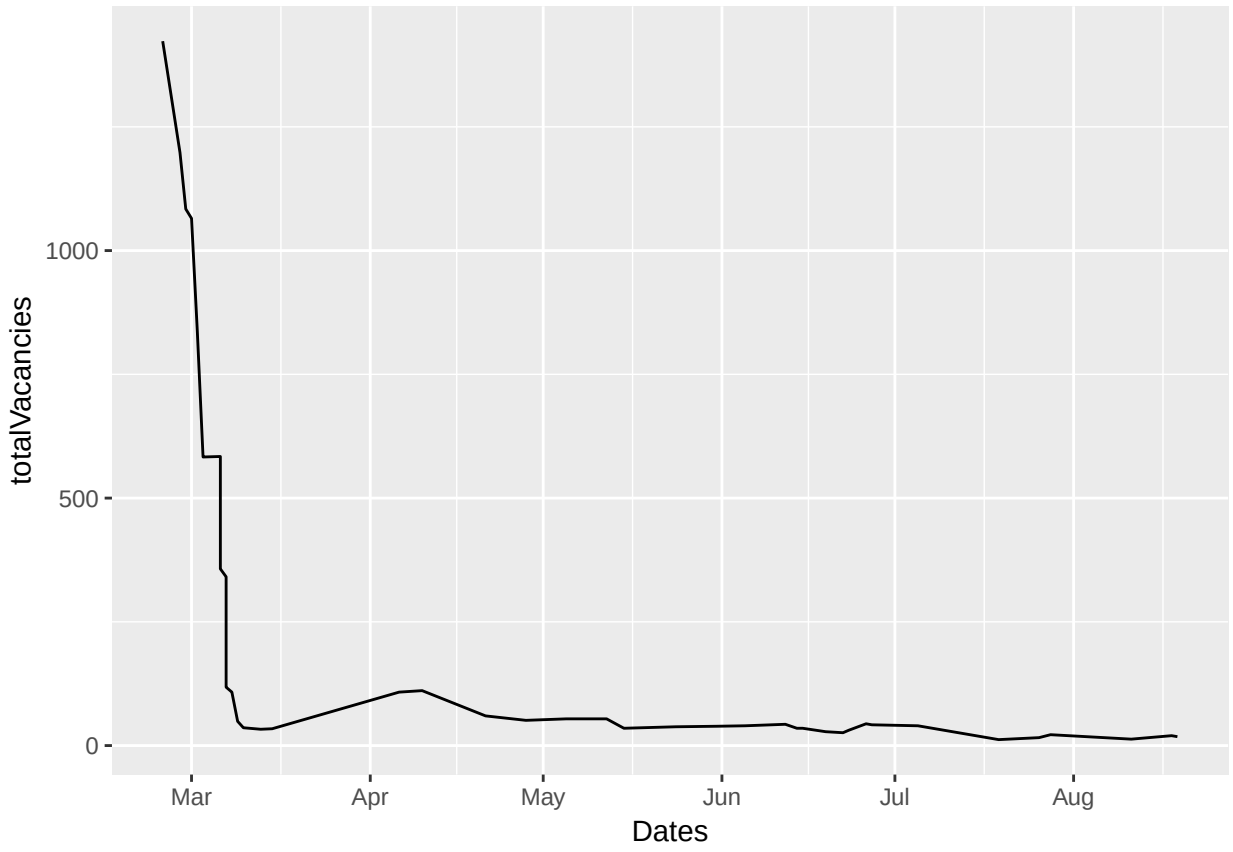


2017-18 Vacancies

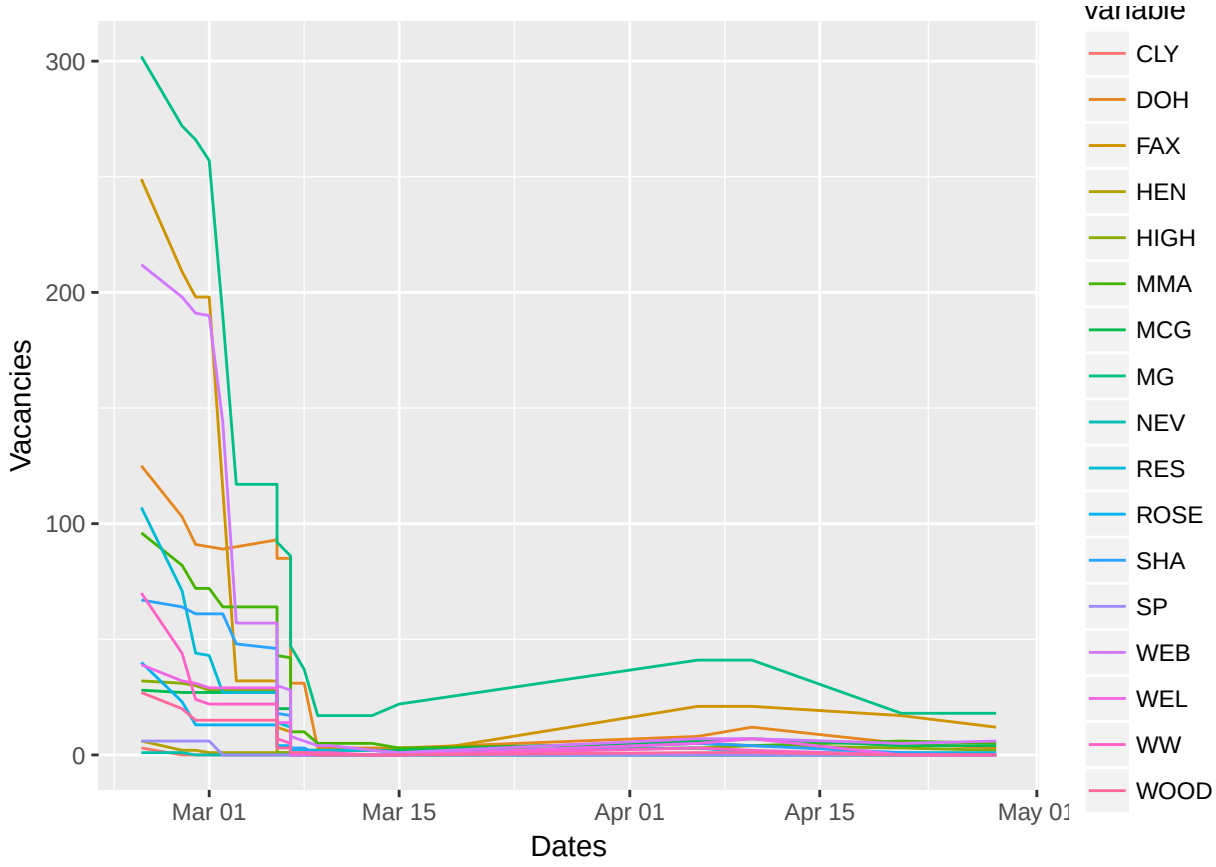
```
# vacancies1718 by Houses until start of Fall opening
ggplot(newVacancies1718ByHouses,aes(x = Dates, y = Vacancies, group = variable,
                                    colour = variable)) +
  geom_line()
```



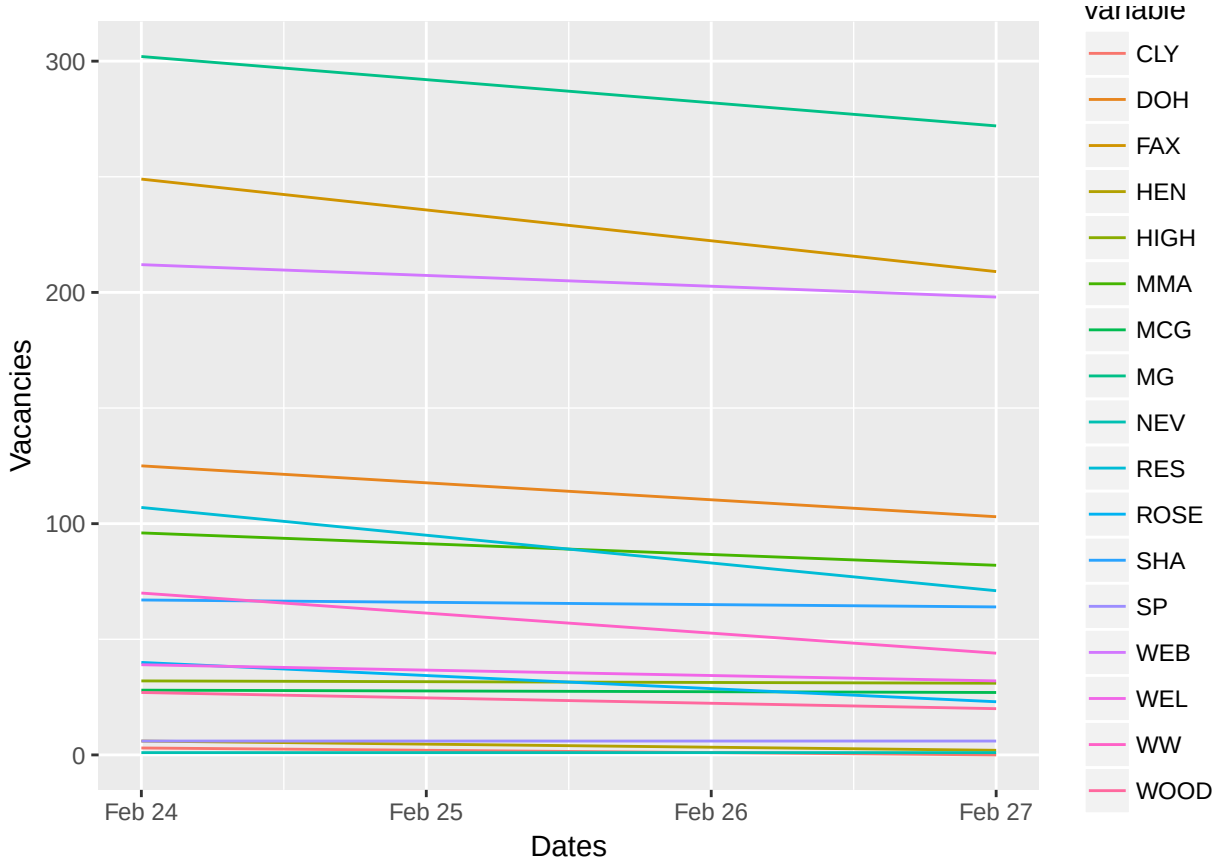

```
# vacancies1718 total participation until start of Fall opening
ggplot(vacancies1718Total,aes(x = Dates, y = totalVacancies)) +
  geom_line()
```



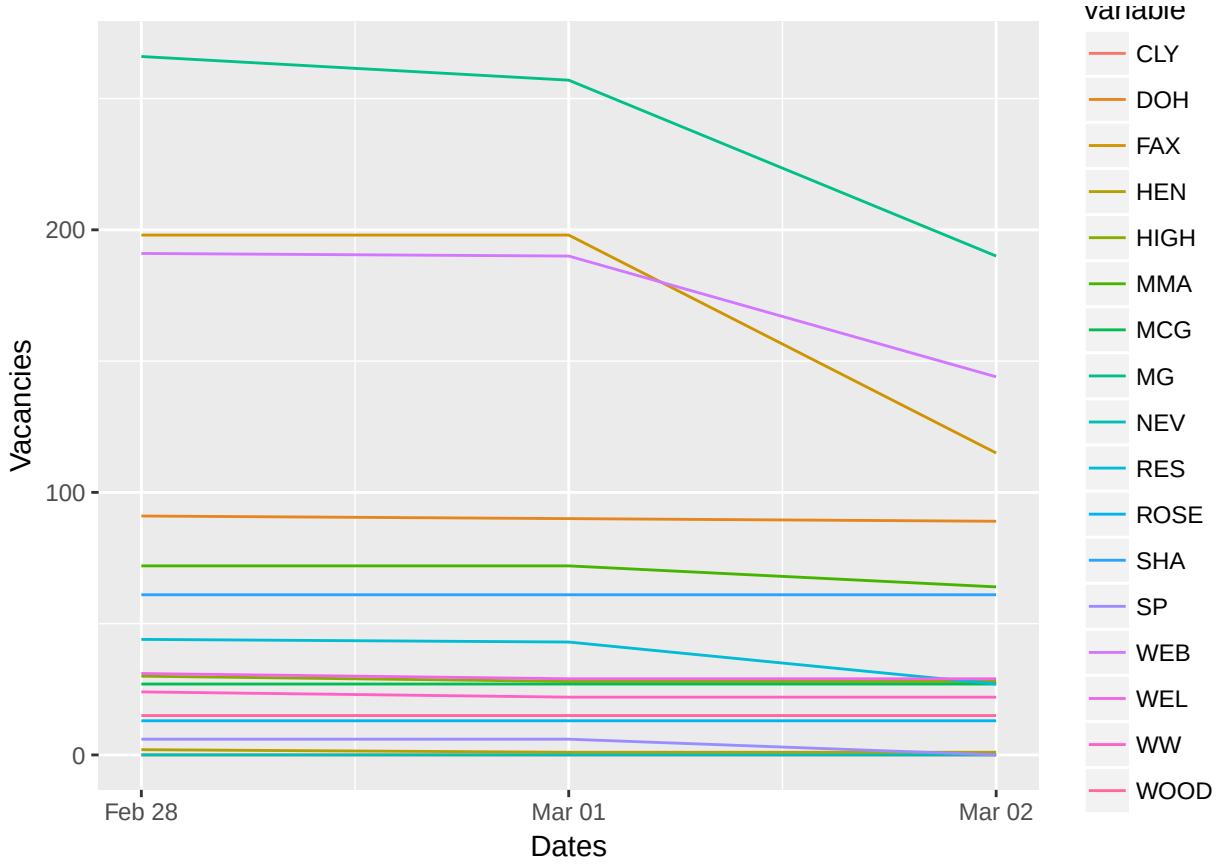

```
# vacancies1718 by Houses through April
ggplot(newVacancies1718ByHousesApr,aes(x = Dates, y = Vacancies, group = variable,
                                       colour = variable)) +
  geom_line()
```



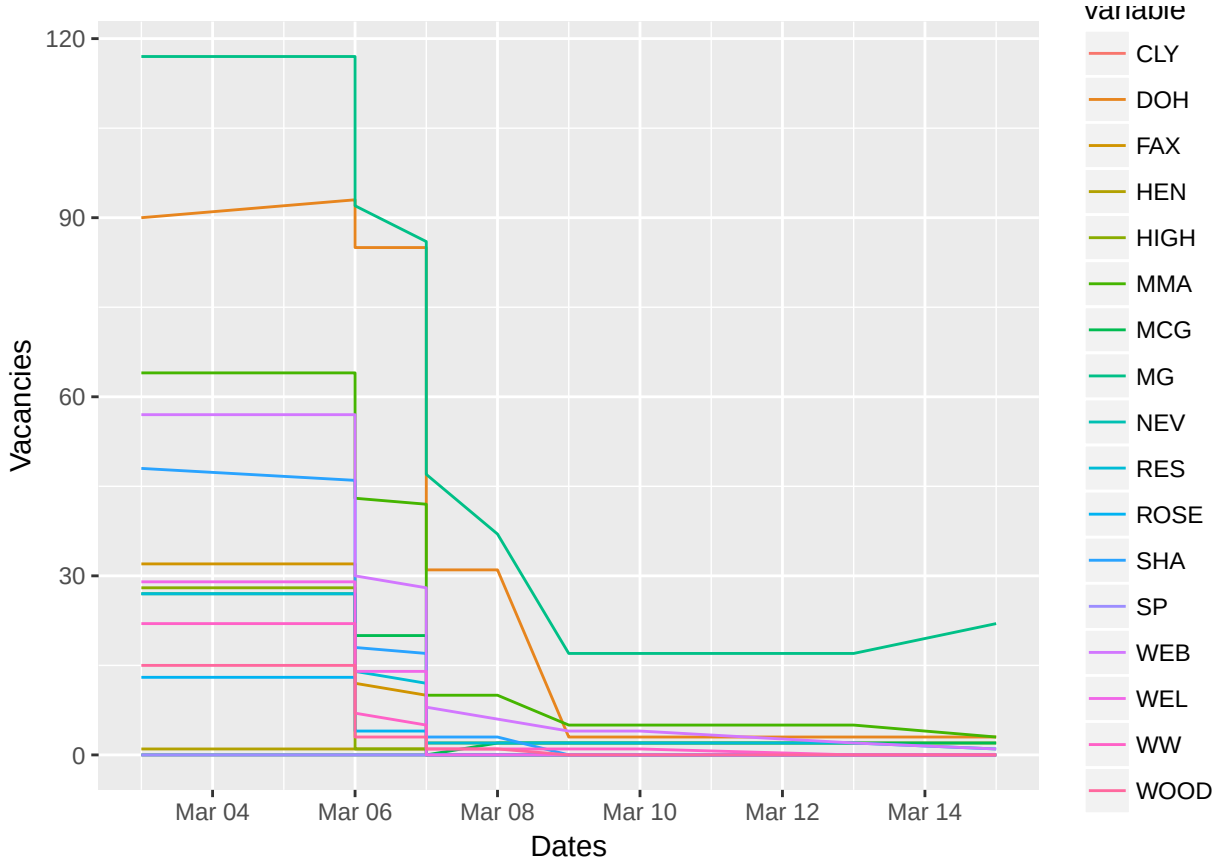

```
# vacancies1718 by Houses during Retention
ggplot(newVacancies1718ByHousesRetention,aes(x = Dates, y = Vacancies, group = variable,
                                             colour = variable)) +
  geom_line()
```



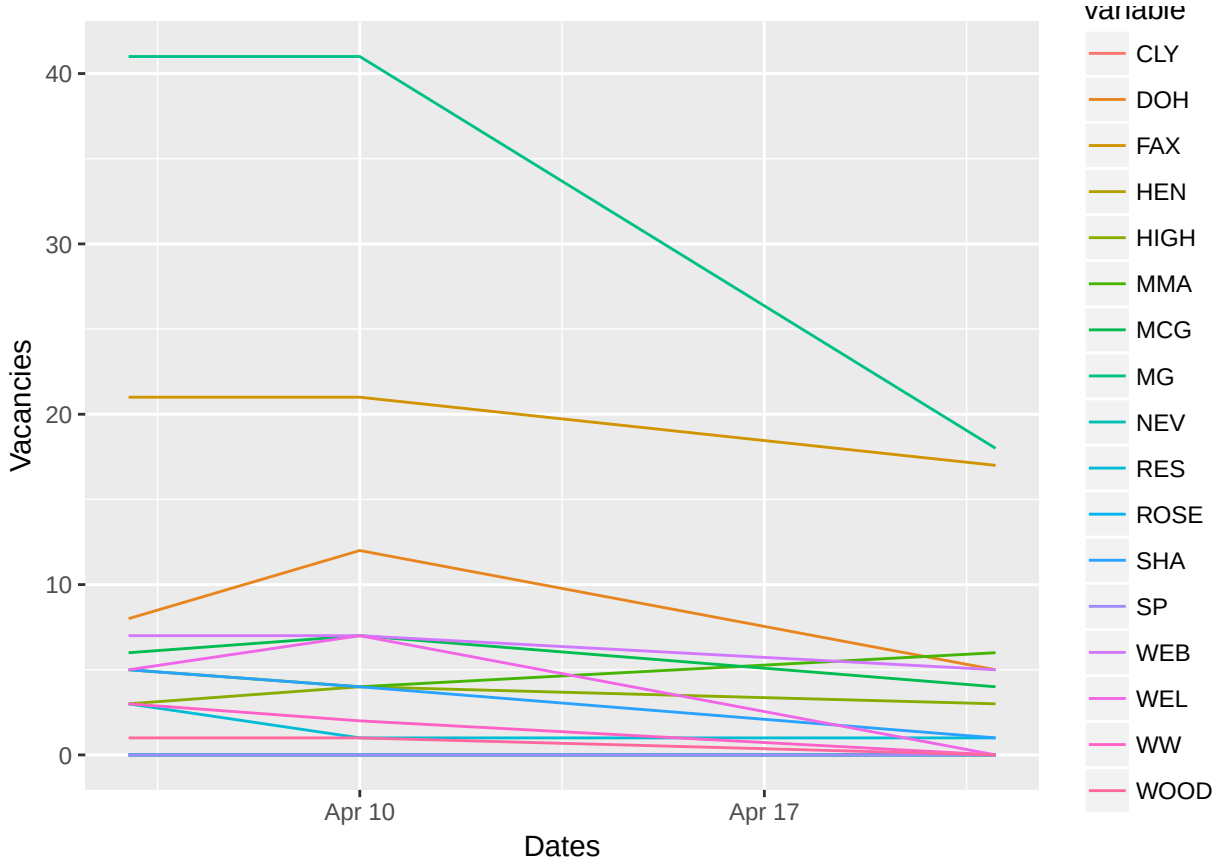

```
# vacancies1718 by Houses during Block Housing
ggplot(newVacancies1718ByHousesBlock,aes(x = Dates, y = Vacancies, group = variable,
                                         colour = variable)) +
  geom_line()
```



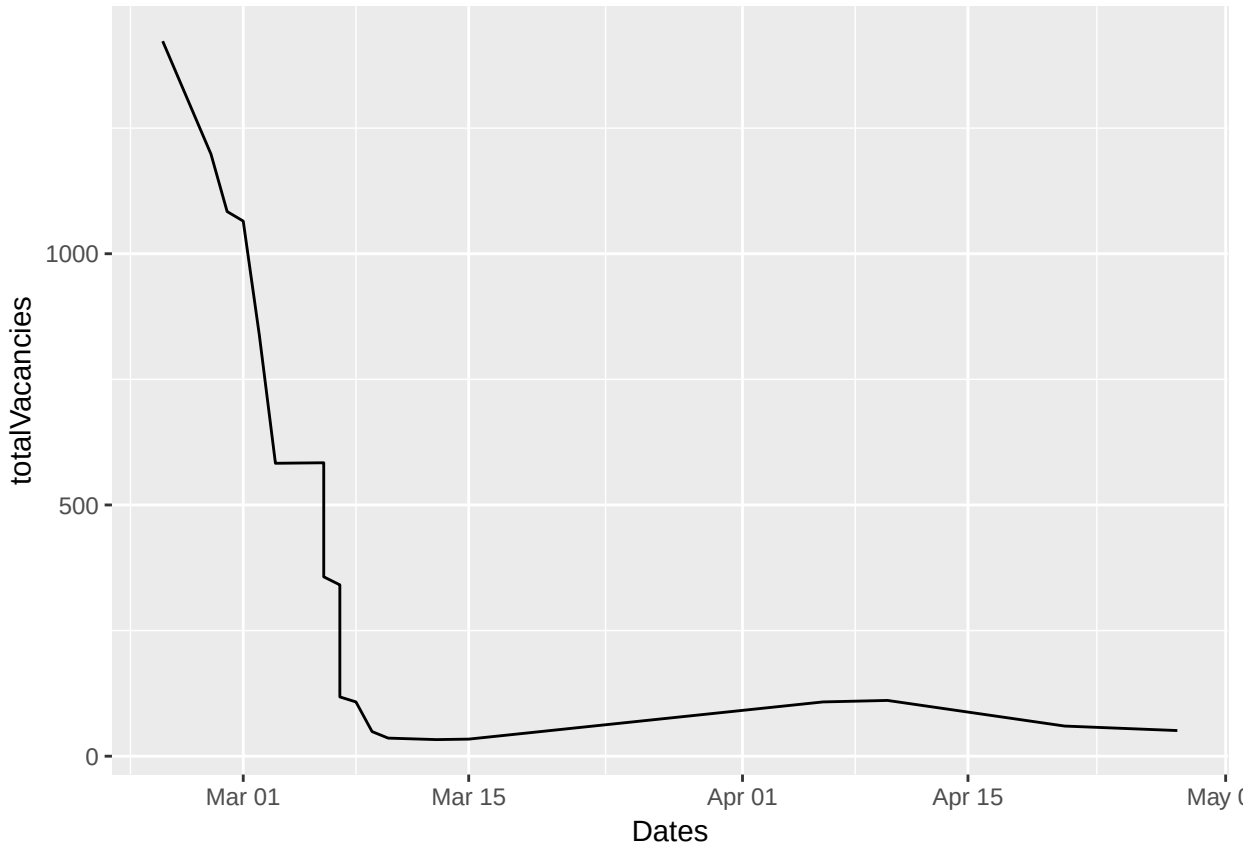

```
# vacancies1718 by Houses during General Selection
ggplot(newVacancies1718ByHousesGeneral,aes(x = Dates, y = Vacancies, group = variable,
                                           colour = variable)) +
  geom_line()
```



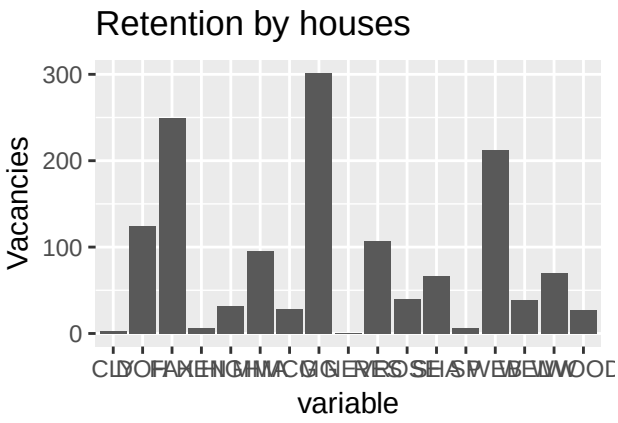

```
# vacancies1718 by Houses during Open Assignment
ggplot(newVacancies1718ByHousesOpen,aes(x = Dates, y = Vacancies, group = variable,
                                        colour = variable)) +
  geom_line()
```

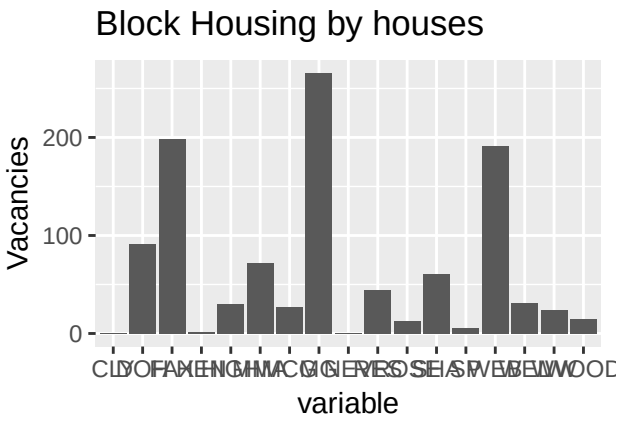


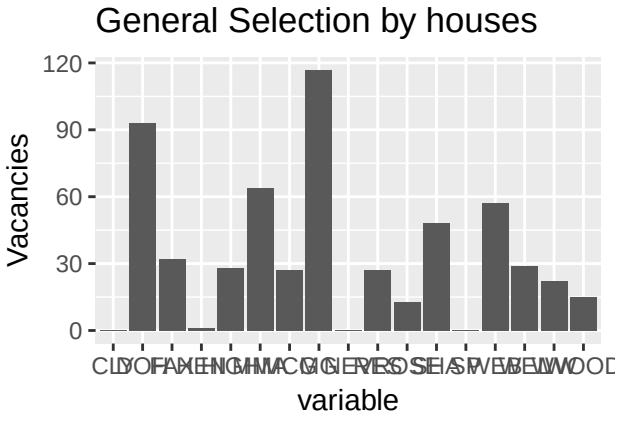

```
# vacancies1718 total participation through April
ggplot(vacancies1718TotalApr,aes(x = Dates, y = totalVacancies)) +
  geom_line()
```

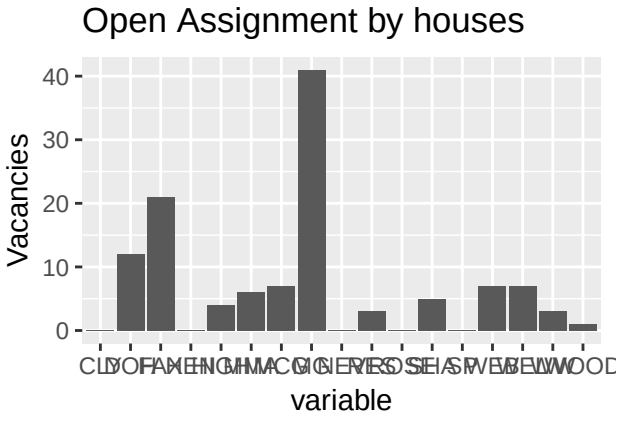



```
# bar chart of counts
r <- ggplot(newVacancies1718ByHousesRetention,aes(x = variable, y = Vacancies)) +
                  geom_bar(stat = "identity", position = "identity") +
                  labs(title = "Retention by houses", fill = "Houses")

b <- ggplot(newVacancies1718ByHousesBlock,aes(x = variable, y = Vacancies)) +
                  geom_bar(stat = "identity", position = "identity") +
                  labs(title = "Block Housing by houses", fill = "Houses")

g <- ggplot(newVacancies1718ByHousesGeneral,aes(x = variable, y = Vacancies)) +
                  geom_bar(stat = "identity", position = "identity") +
                  labs(title = "General Selection by houses", fill = "Houses")

o <- ggplot(newVacancies1718ByHousesOpen,aes(x = variable, y = Vacancies)) +
                  geom_bar(stat = "identity", position = "identity") +
                  labs(title = "Open Assignment by houses", fill = "Houses")

grid.arrange(r,b,g,o,ncol=2)
```



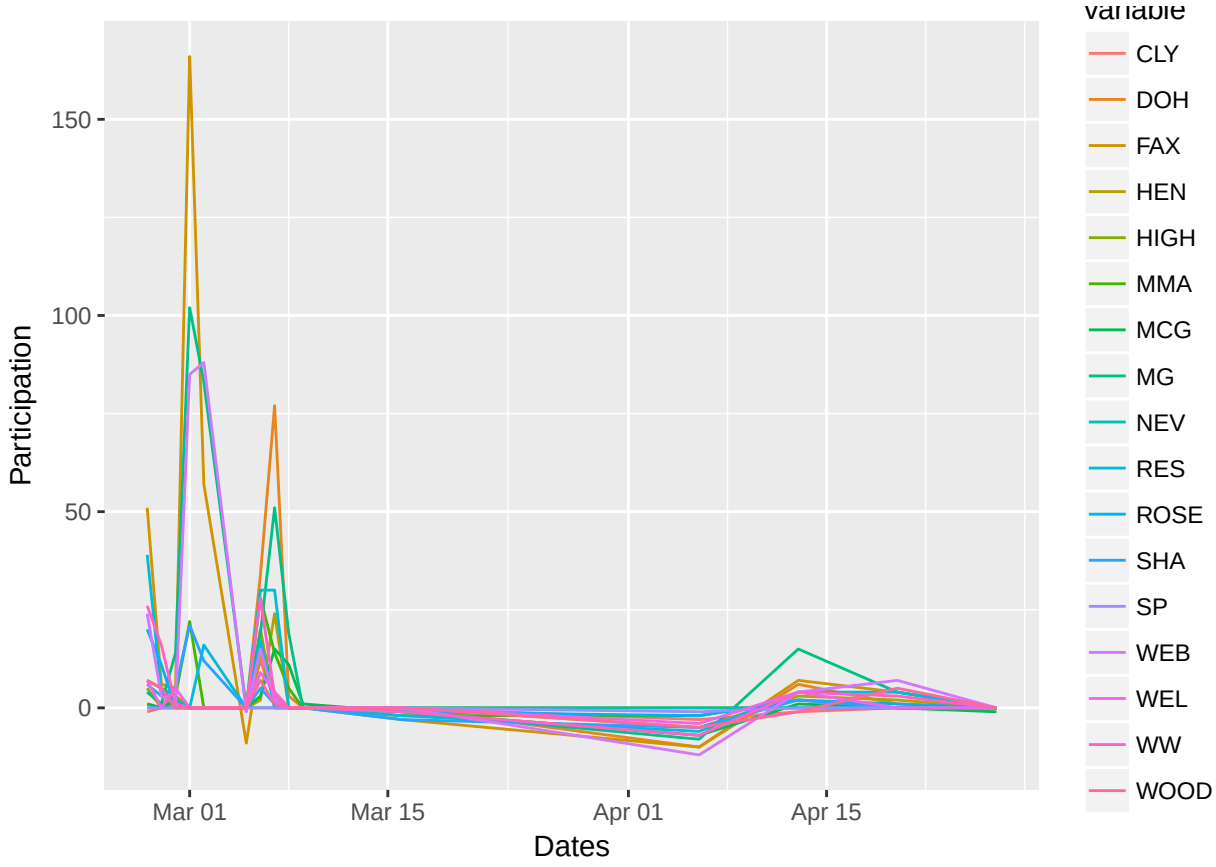




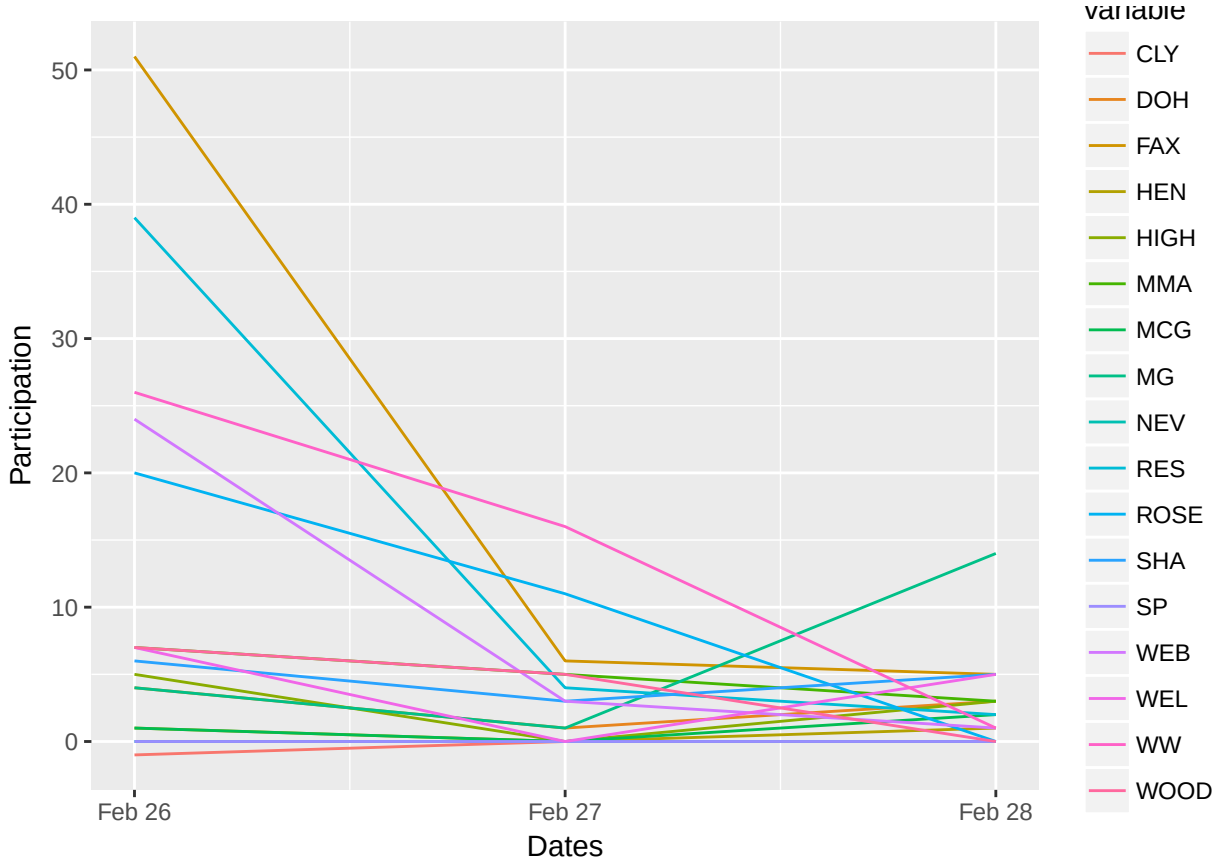


2018-19 Room Selection

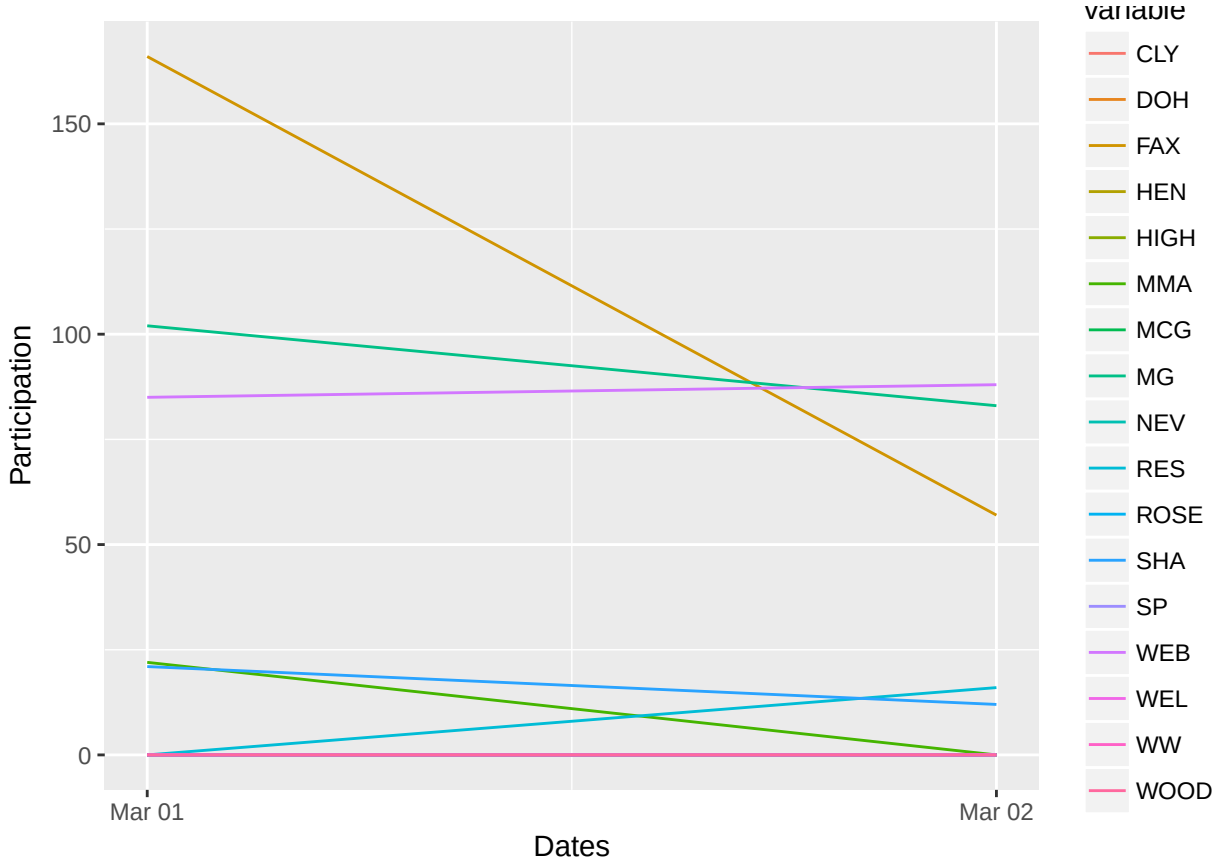
```
# occupancy1819 by Houses
ggplot(newOccupancyByHouses,aes(x = Dates, y = Participation, group = variable,
                                colour = variable)) +
  geom_line()
```



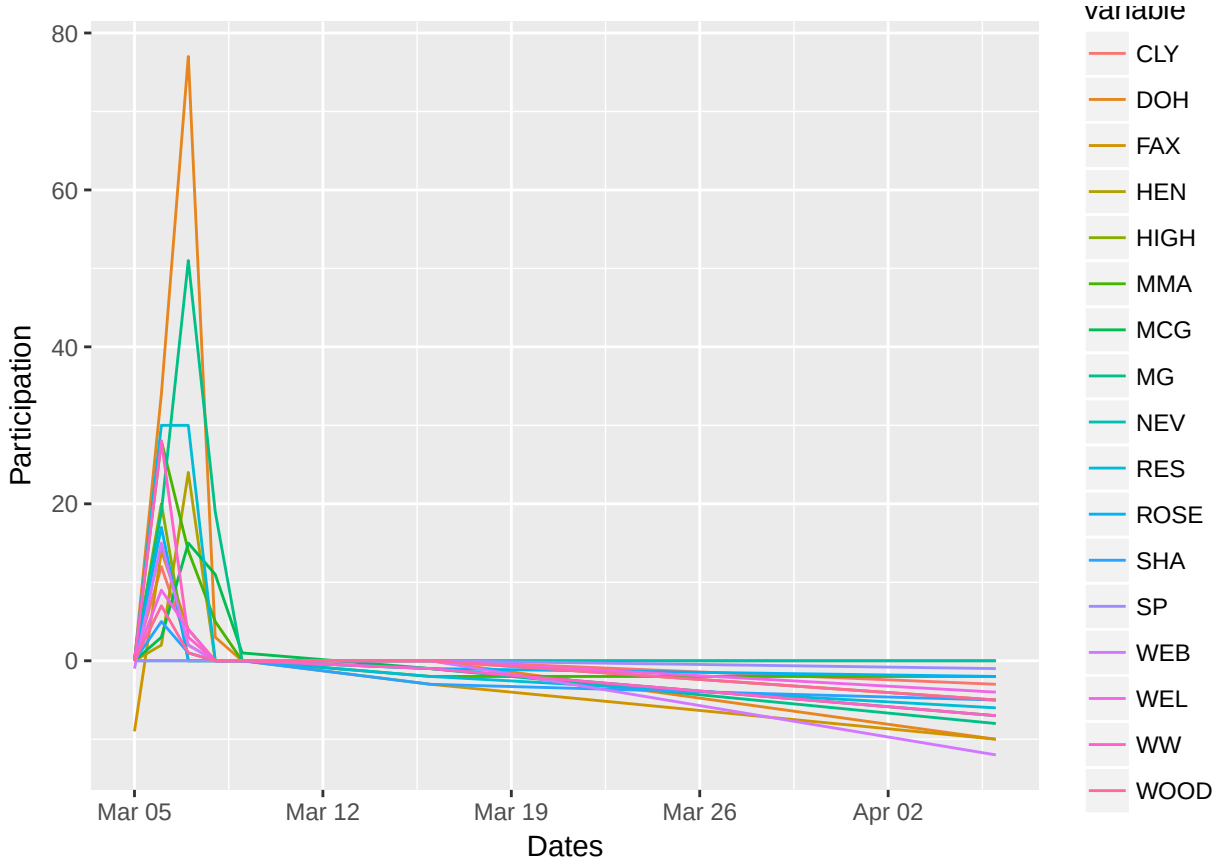

```
# occupancy1819 by Houses during Retention
ggplot(newOccupancyByHousesRetention,aes(x = Dates, y = Participation, group = variable,
                                         colour = variable)) +
  geom_line()
```



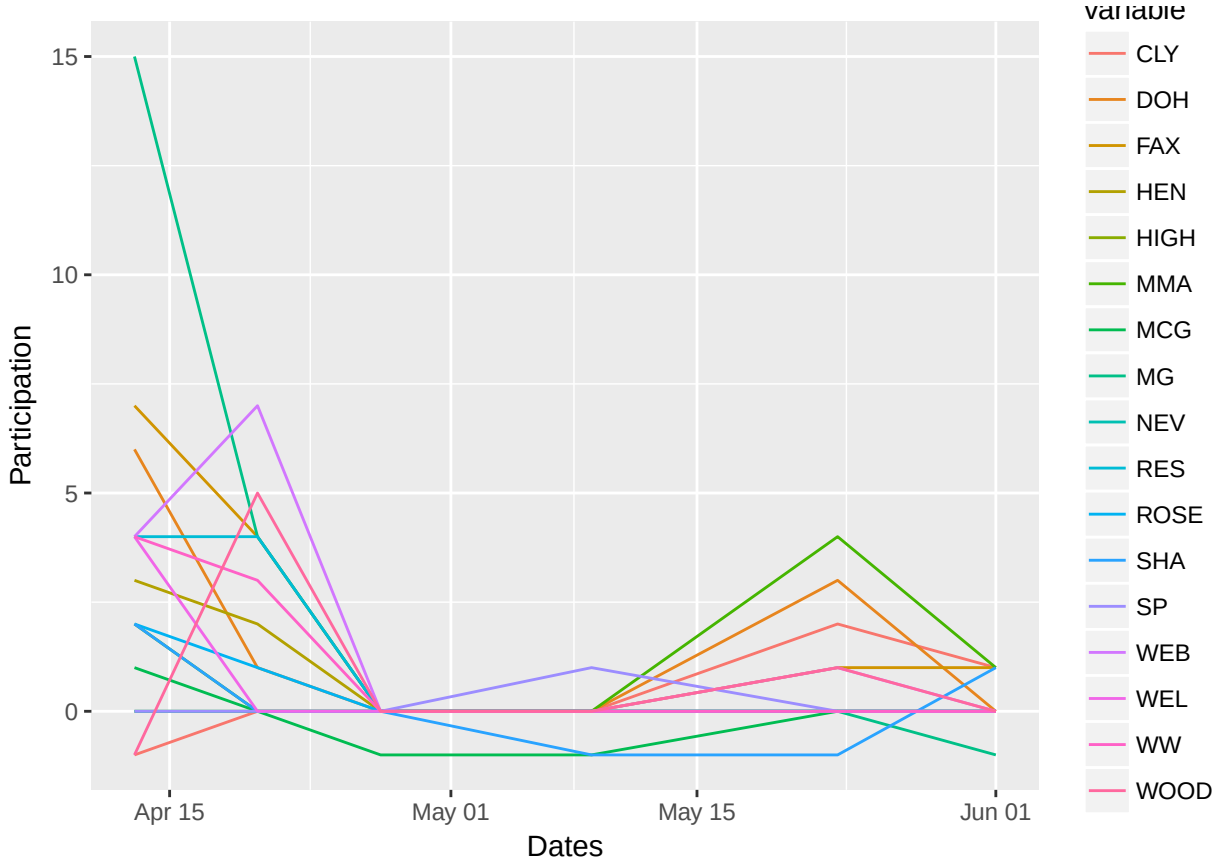

```
# occupancy1819 by Houses during Block
ggplot(newOccupancyByHousesBlock,aes(x = Dates, y = Participation, group = variable,
                                     colour = variable)) +
  geom_line()
```



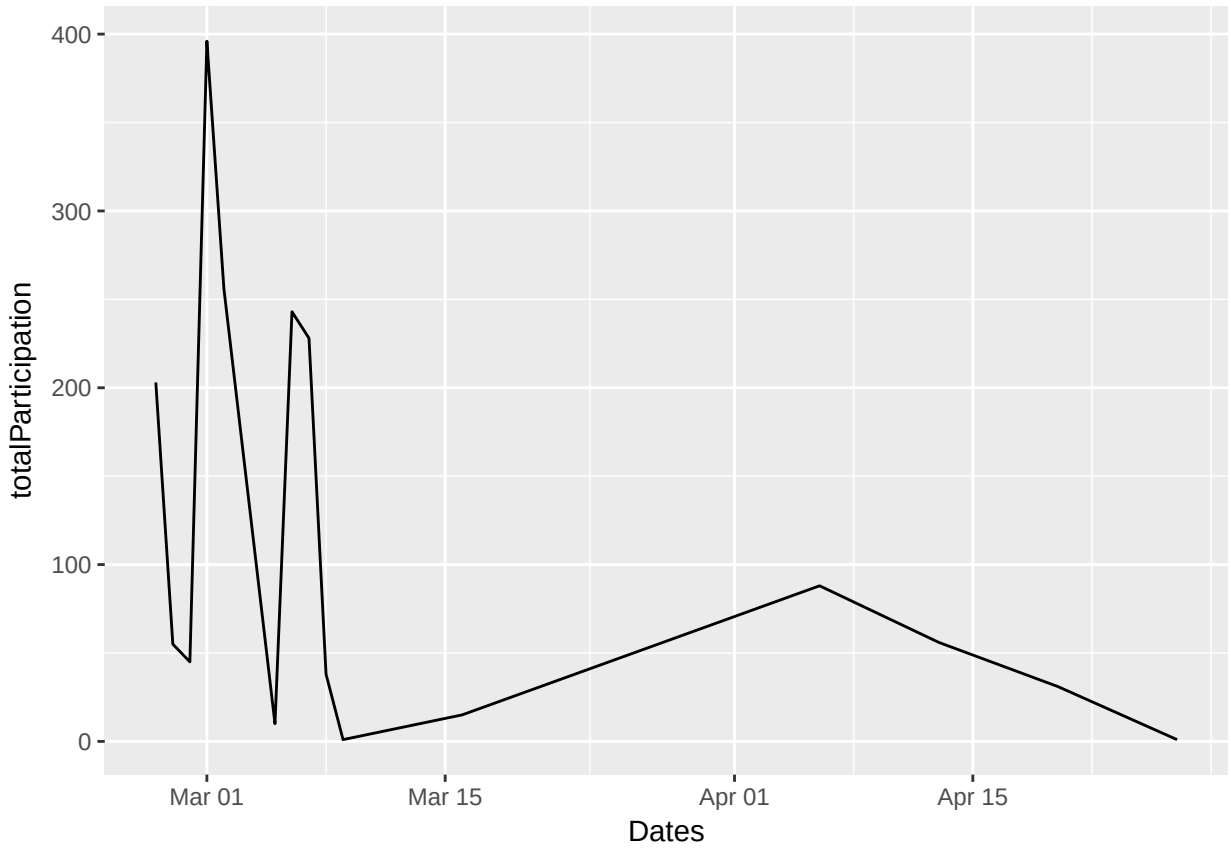

```
# occupancy1819 by Houses during General Selection
ggplot(newOccupancyByHousesGeneral,aes(x = Dates, y = Participation, group = variable,
                                       colour = variable)) +
  geom_line()
```




```
# occupancy1819 by Houses during Open Assignment
ggplot(newOccupancyByHousesOpen,aes(x = Dates, y = Participation, group = variable,
                                    colour = variable)) +
  geom_line()
```




```
# occupancy1819 total participation
ggplot(occupancy1819Total,aes(x = Dates, y = totalParticipation)) +
  geom_line()
```



2017-18 Room Selection

```
# # phases by houses
# ggplot(data = newPhasesByHouses, aes(x = newPhasesByHouses$Group,
#                                      fill = newPhasesByHouses$variable,
#                                      na.rm = TRUE)) +
#       geom_bar(aes(y = newPhasesByHouses$Participation), stat = "identity") +
#       labs(x = "Phase", y = "Count", fill = "House")
#
# # phases by neighborhood
# ggplot(data = newPhasesByNeighborhood, aes(x = newPhasesByNeighborhood$Group,
#                                      fill = newPhasesByNeighborhood$variable,
#                                      na.rm = TRUE)) +
#       geom_bar(aes(y = newPhasesByNeighborhood$Participation), stat = "identity") +
#       labs(x = "Phase", y = "Count", fill = "House")
#
# # phases by building/room types
# ggplot(data = newPhasesByBuilding, aes(x = newPhasesByBuilding$Group,
#                                      fill = newPhasesByBuilding$variable,
#                                      na.rm = TRUE)) +
#       geom_bar(aes(y = newPhasesByBuilding$Participation), stat = "identity") +
#       labs(x = "Phase", y = "Count", fill = "House")
```

2017-18 Vacancies

```
##############################################################################
```

2018-19 Room Selection

```
##############################################################################
```

By Neighborhood

- Hill: HEN, MMA, MCG, RES, ROSE, SP, WEL, WW, WOOD
- Morewood: DOH, MG
- Oakland: CLY, FAX, HIGH, NEV, SHA, WEB

2017-18 Room Selection

```
##     Hill Morewood  Oakland
##      409      406      544
```

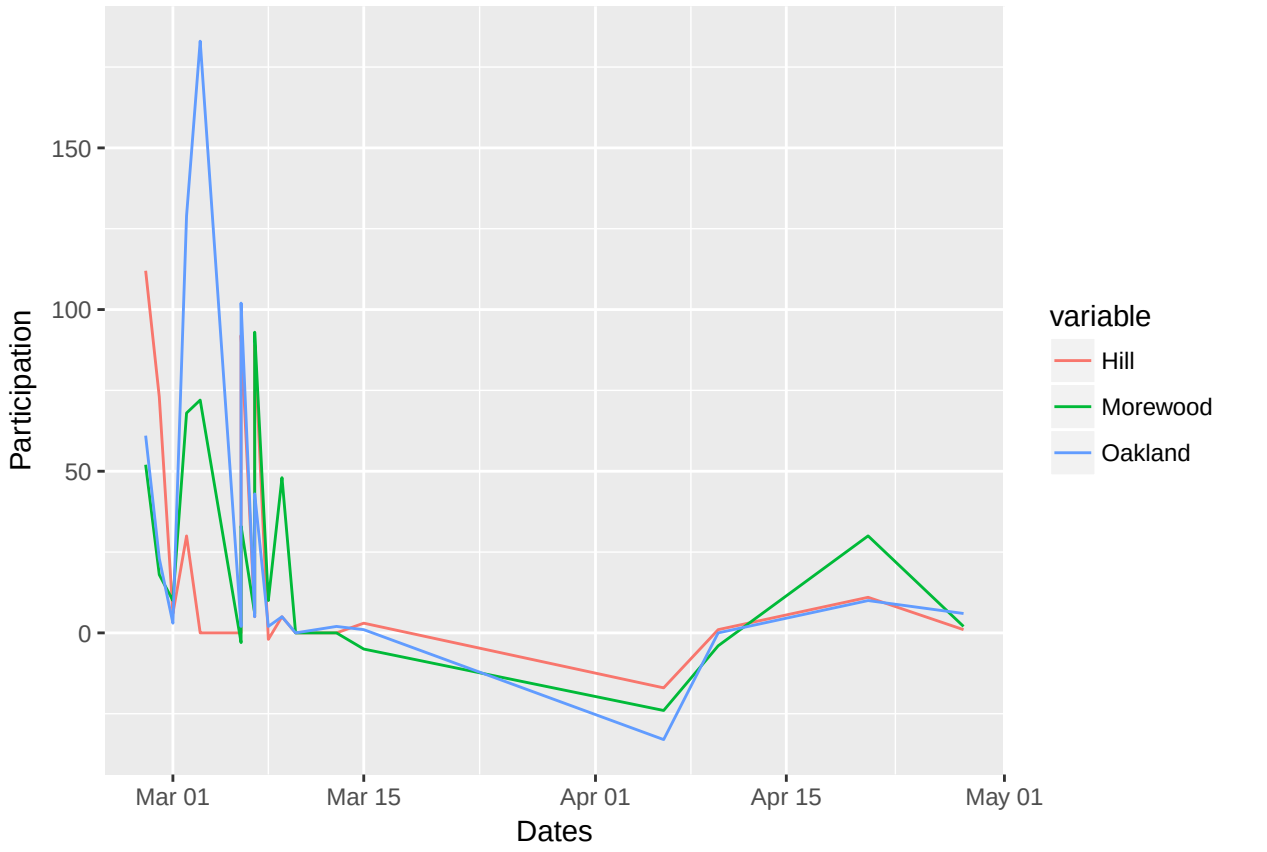
```
##     Hill Morewood  Oakland
##      185       70       84
```

```
##     Hill Morewood  Oakland
##       36      150      315
```

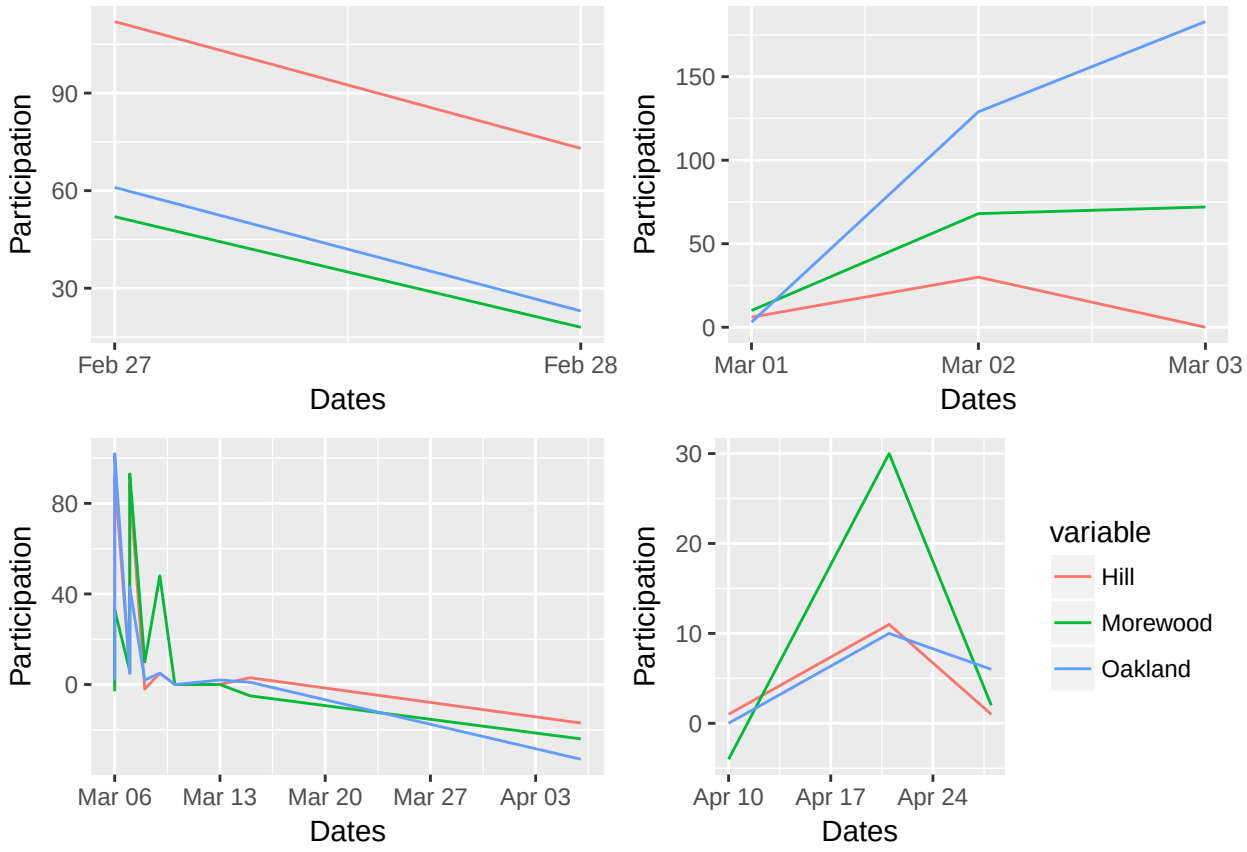
```
##     Hill Morewood  Oakland
##      175      158      129
```

```
##     Hill Morewood  Oakland
##       13       28       16
```

```
# rsNet1718 by Neighborhoods
ggplot(newRSNet1718AprNeighborhoods,aes(x = Dates, y = Participation,
                                        group = variable, colour = variable)) +
  geom_line()
```

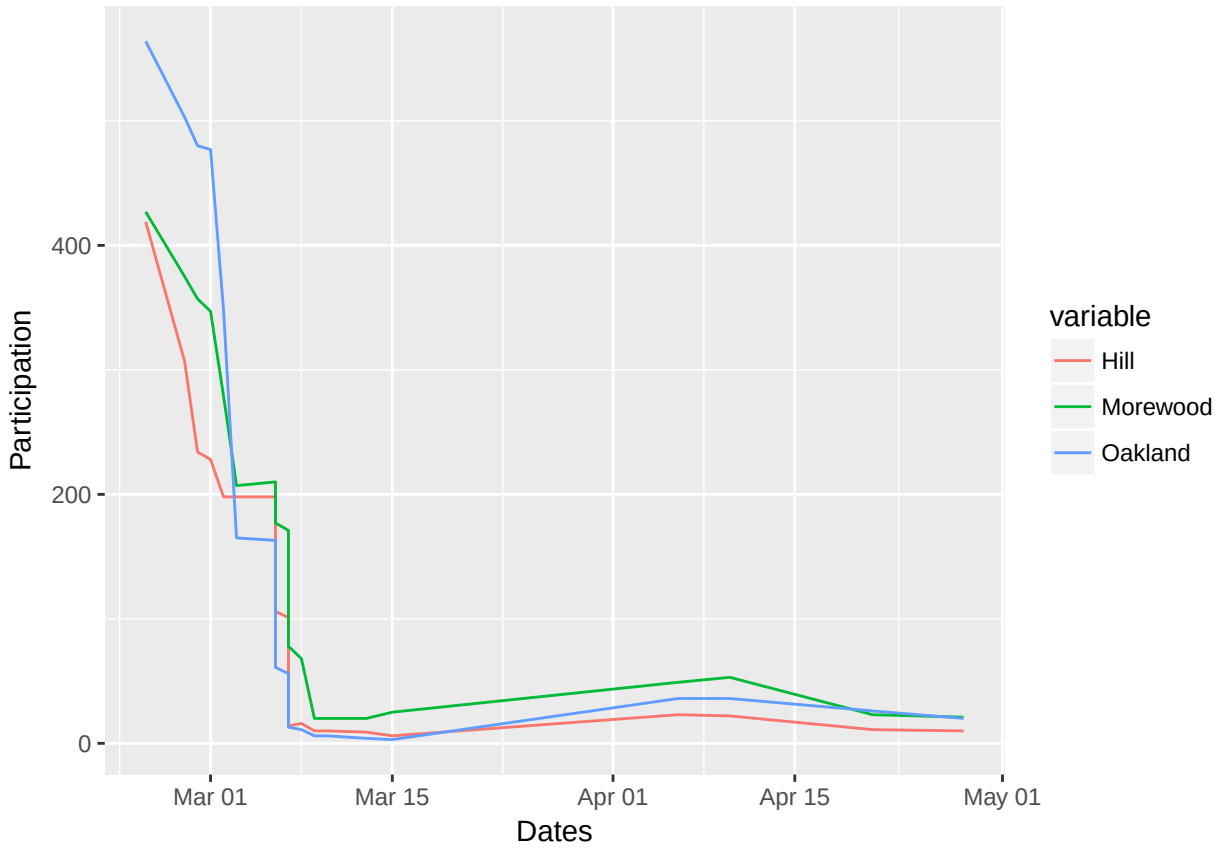



```
# rsNet1718 by Neighborhoods during Retention
retain1 <- ggplot(newRSNet1718AprNeighborhoodsRetention,aes(x = Dates, y = Participation,
                                        group = variable, colour = variable)) +
  geom_line() +
  theme(legend.position="none")

# rsNet1718 by Neighborhoods during Block
block1 <- ggplot(newRSNet1718AprNeighborhoodsBlock,aes(x = Dates, y = Participation,
                                        group = variable, colour = variable)) +
  geom_line() +
  theme(legend.position="none")

# rsNet1718 by Neighborhoods during General Selection
general1 <- ggplot(newRSNet1718AprNeighborhoodsGeneral,aes(x = Dates, y = Participation,
                                        group = variable, colour = variable)) +
  geom_line() +
  theme(legend.position="none")

# rsNet1718 by Neighborhoods during Open Assignment
open1 <- ggplot(newRSNet1718AprNeighborhoodsOpen,aes(x = Dates, y = Participation,
                                        group = variable, colour = variable)) +
  geom_line()

grid.arrange(retain1, block1, general1, open1, ncol = 2)
```



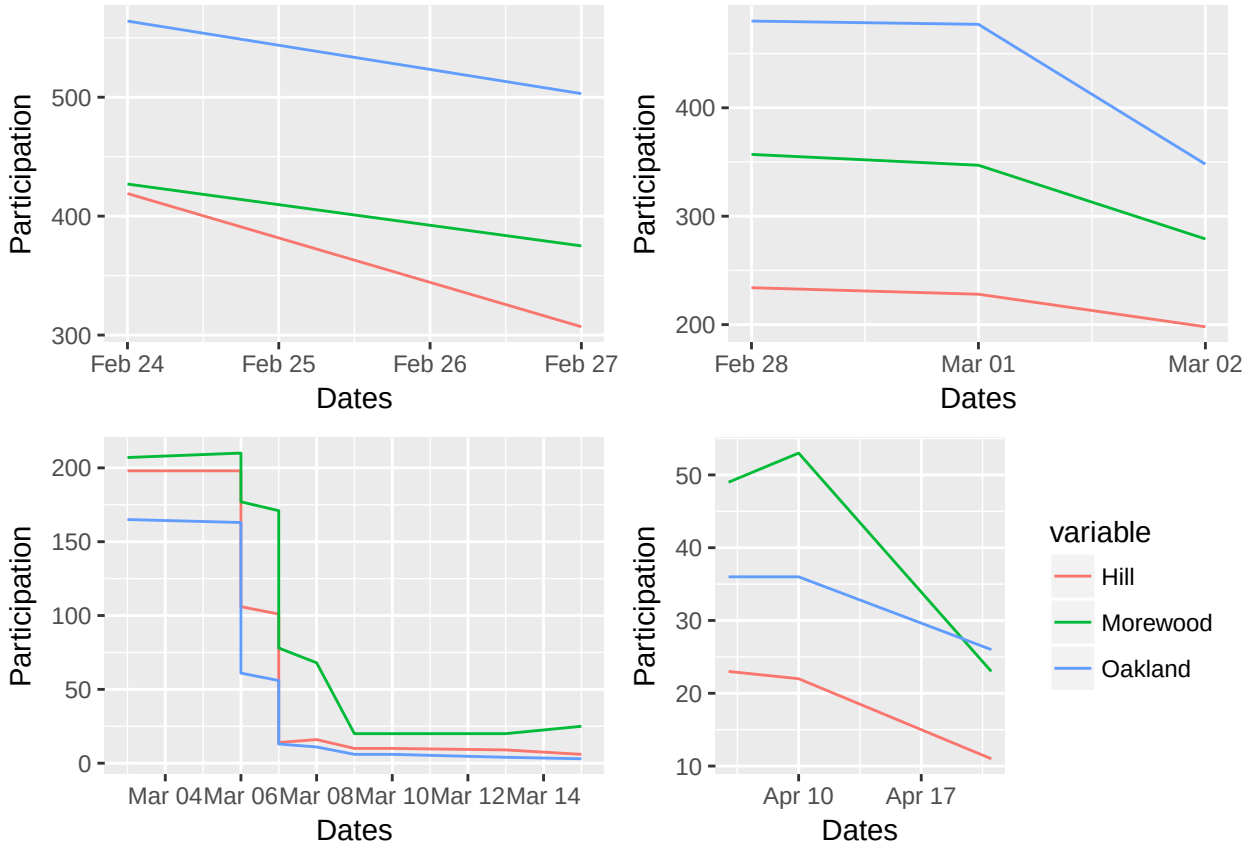
2017-18 Vacancies

```
# vacancies1718 by neighborhood
ggplot(newVacancies1718AprNeighborhoods,aes(x = Dates, y = Participation,
                                            group = variable, colour = variable)) +
  geom_line()
```




```
# vacancies1718 by neighborhood during Retention
retain2 <- ggplot(newVacancies1718AprNeighborhoodsRetention,aes(x = Dates,
                                                                y = Participation,
                                                                group = variable,
                                                                colour = variable)) +
                  geom_line() +
                  theme(legend.position="none")

# vacancies1718 by neighborhood during Block
block2 <- ggplot(newVacancies1718AprNeighborhoodsBlock,aes(x = Dates, y = Participation,
                                           group = variable, colour = variable)) +
                  geom_line() +
                  theme(legend.position="none")

# vacancies1718 by neighborhood during General Selection
general2 <- ggplot(newVacancies1718AprNeighborhoodsGeneral,aes(x = Dates,
                                                               y = Participation,
                                                               group = variable,
                                                               colour = variable)) +
                    geom_line() +
                    theme(legend.position="none")

# vacancies1718 by neighborhood during Open Assignment
open2 <- ggplot(newVacancies1718AprNeighborhoodsOpen,aes(x = Dates, y = Participation,
                                           group = variable, colour = variable)) +
                geom_line()
```

```
grid.arrange(retain2, block2, general2, open2)
```



2018-19 Room Selection

```
# Descriptive Stats
colSums(occupancy1819Neighborhoods[,-1])
```

```
##     Hill Morewood  Oakland
##      420      422      592
```

```
colSums(occupancy1819NeighborhoodsRetention[,-1])
```

```
##     Hill Morewood  Oakland
##      163       27      111
```

```
colSums(occupancy1819NeighborhoodsBlock[,-1])
```

```
##     Hill Morewood  Oakland
##       38      185      429
```

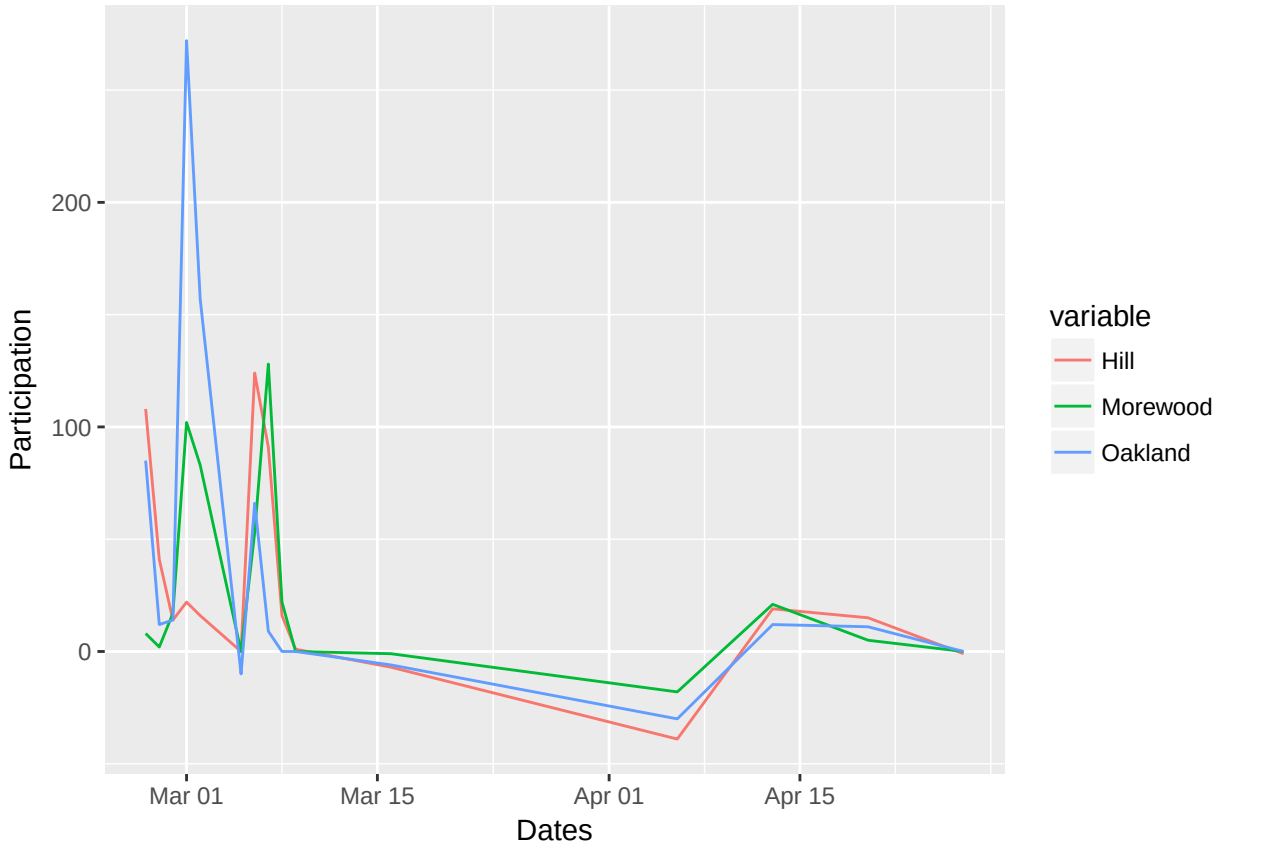
```
colSums(occupancy1819NeighborhoodsGeneral[,-1])
```

```
##     Hill Morewood  Oakland
##      186      184       29
```

```
colSums(occupancy1819NeighborhoodsOpen[,-1])
```

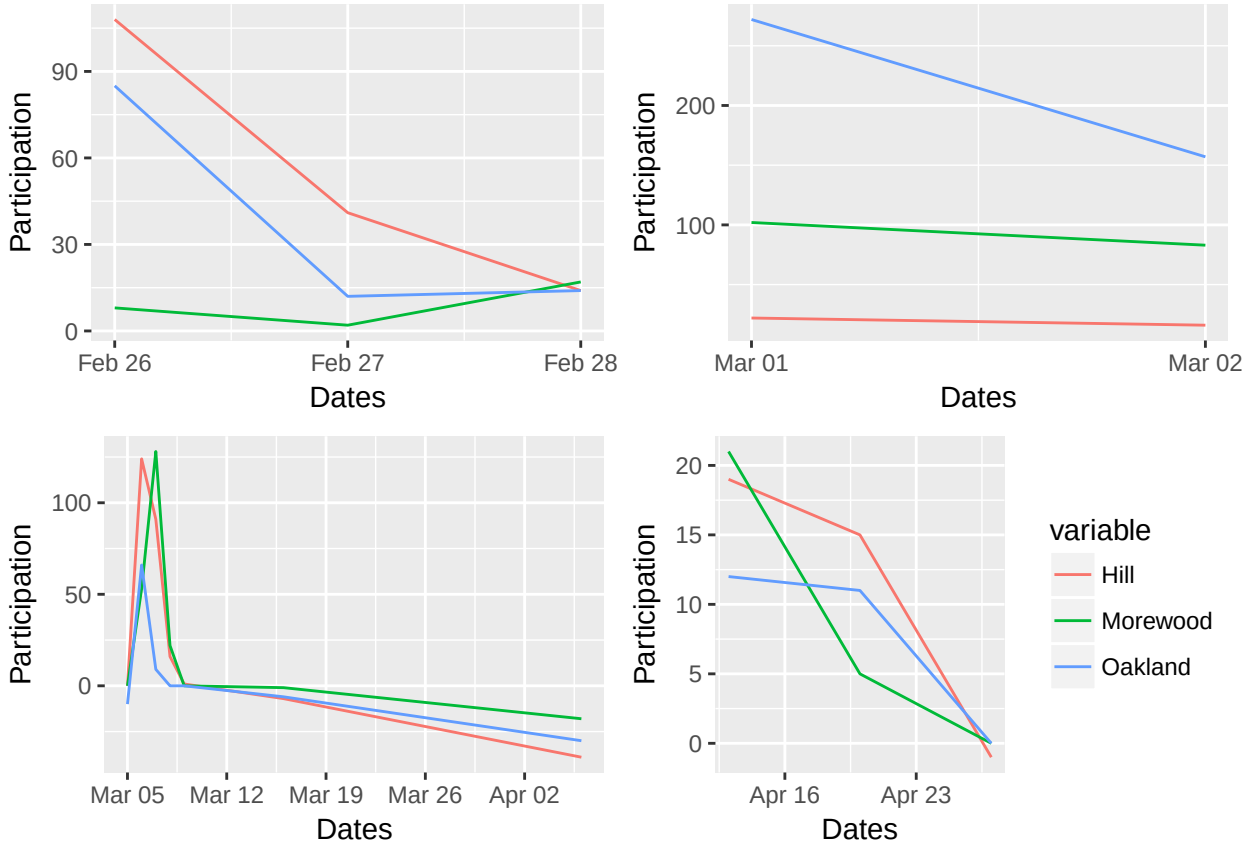
```
##     Hill Morewood  Oakland
##       33       26       23
```

```
# occupancy1819 by neighborhood
ggplot(newOccupancy1819Neighborhoods,aes(x = Dates, y = Participation,
                                         group = variable, colour = variable)) +
  geom_line()
```




```
# occupancy1819 by neighborhood during Retention
retain3 <- ggplot(newOccupancy1819NeighborhoodsRetention,aes(x = Dates, y = Participation,
                                         group = variable, colour = variable)) +
  geom_line() +
  theme(legend.position="none")

# occupancy1819 by neighborhood during Block
block3 <- ggplot(newOccupancy1819NeighborhoodsBlock,aes(x = Dates, y = Participation,
                                         group = variable, colour = variable)) +
  geom_line() +
  theme(legend.position="none")

# occupancy1819 by neighborhood during General Selection
general3 <- ggplot(newOccupancy1819NeighborhoodsGeneral,aes(x = Dates, y = Participation,
                                         group = variable, colour = variable)) +
  geom_line() +
  theme(legend.position="none")

# occupancy1819 by neighborhood during Open Assignment
open3 <- ggplot(newOccupancy1819NeighborhoodsOpen,aes(x = Dates, y = Participation,
                                         group = variable, colour = variable)) +
  geom_line()
```

```
grid.arrange(retain3, block3, general3, open3)
```



By Building/Room Type

- Apartment: CLY, DOH, FAX, HIGH, MMA, NEV, ROSE, SHA, SP, WEB, WOOD
- Residence_Hall: HEN, MCG, MG, RES, WEL, WW

2017-18 Room Selection

```
# Descriptive Stats
colSums(rsNet1718AprBuildings[,-1])
```

```
##      Apartment Residence_Hall
##            829            530
```

```
colSums(rsNet1718AprBuildingsRetention[,-1])
```

```
##      Apartment Residence_Hall
##            181            158
```

```
colSums(rsNet1718AprBuildingsBlock[,-1])
```

```
##      Apartment Residence_Hall
##            330            171
```

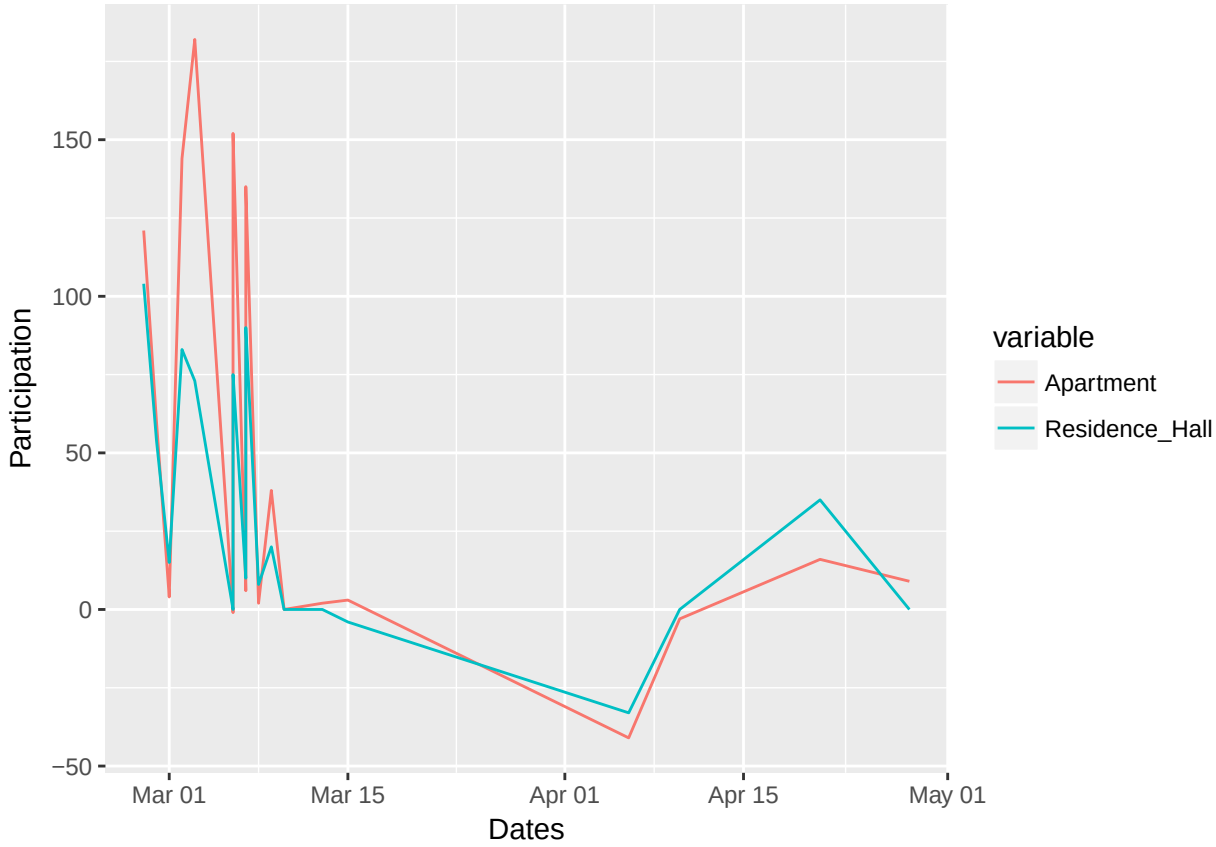
```
colSums(rsNet1718AprBuildingsGeneral[,-1])
```

```
##      Apartment Residence_Hall
##            296            166
```

```
colSums(rsNet1718AprBuildingsOpen[,-1])
```

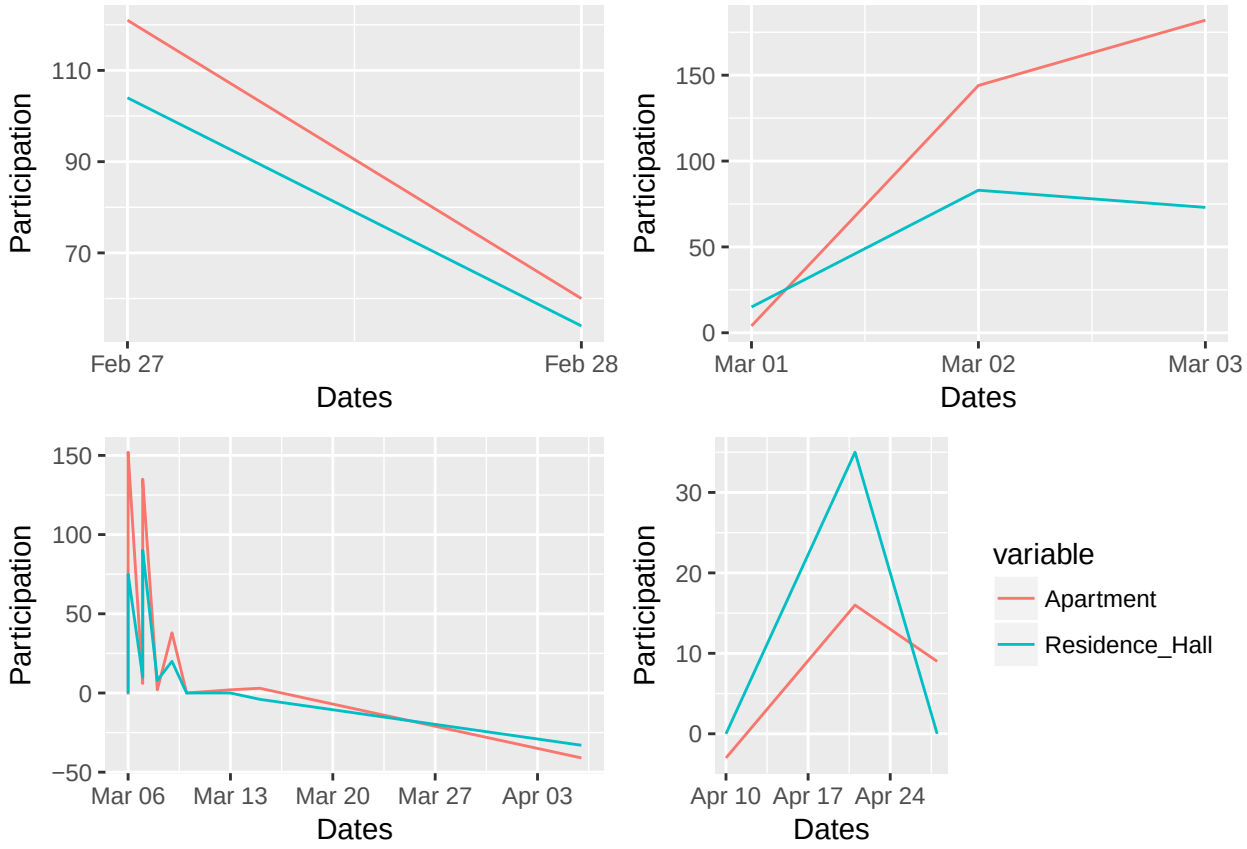
```
##      Apartment Residence_Hall
##             22             35
```

```
# rsNet1718 by buildings
ggplot(newRSNet1718AprBuildings,aes(x = Dates, y = Participation, group = variable,
                                    colour = variable)) + geom_line()
```



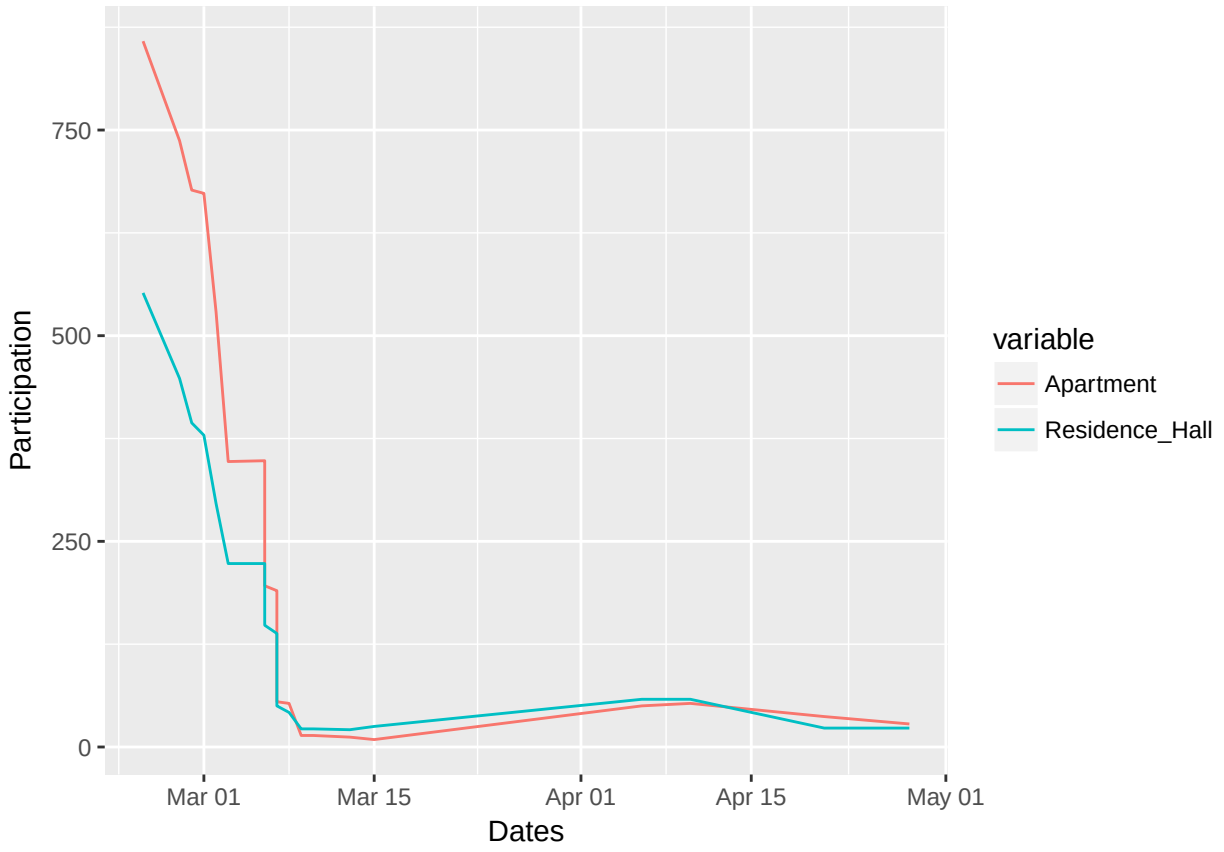

```
# rsNet1718 by buildings during Retention
retain4 <- ggplot(newRSNet1718AprBuildingsRetention,aes(x = Dates,
                                                        y = Participation,
                                                        group = variable,
                                                        colour = variable)) +
                  geom_line() +
                  theme(legend.position="none")

# rsNet1718 by buildings during Block
block4 <- ggplot(newRSNet1718AprBuildingsBlock,aes(x = Dates, y = Participation,
                                                   group = variable,
                                                   colour = variable)) +
                  geom_line() +
                  theme(legend.position="none")

# rsNet1718 by buildings during General Selection
general4 <- ggplot(newRSNet1718AprBuildingsGeneral,aes(x = Dates, y = Participation,
```

```
                                                    group = variable,
                                                    colour = variable)) +
                    geom_line() +
                    theme(legend.position="none")

# rsNet1718 by buildings during Open Assignment
open4 <- ggplot(newRSNet1718AprBuildingsOpen,aes(x = Dates, y = Participation,
                                                 group = variable,
                                                 colour = variable)) +
                 geom_line()

grid.arrange(retain4, block4, general4, open4, ncol = 2)
```

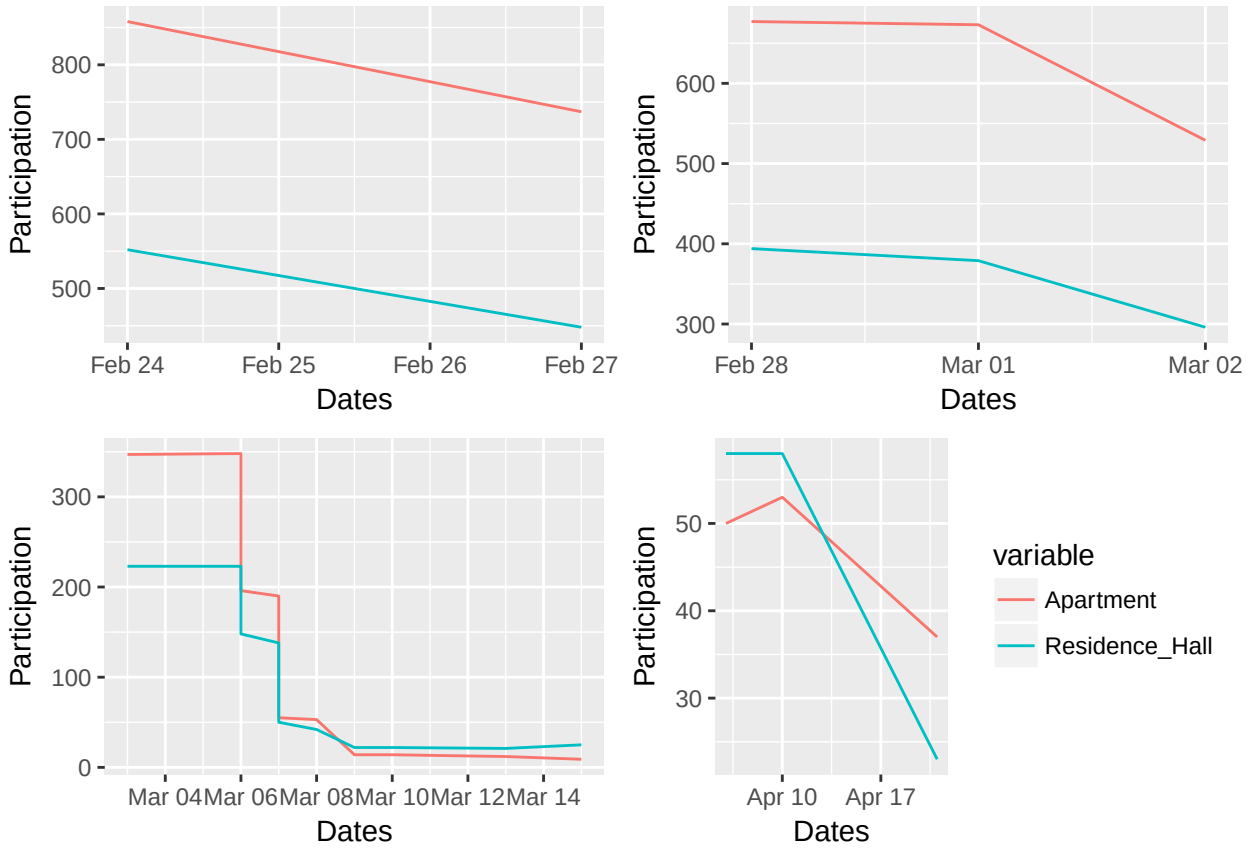


2017-18 Vacancies

```
# vacancies1718 by buildings
ggplot(newVacancies1718AprBuildings,aes(x = Dates, y = Participation,
                                        group = variable, colour = variable)) +
  geom_line()
```

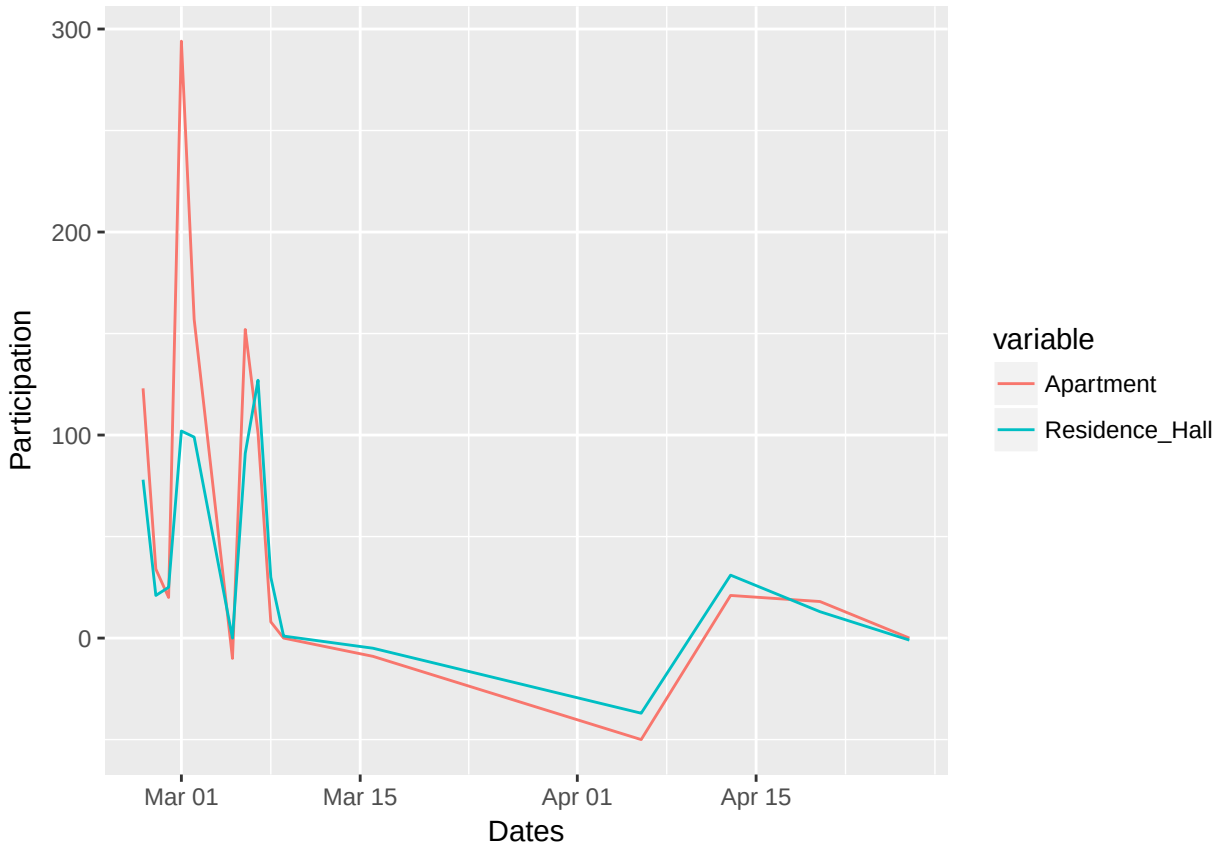



```
# vacancies1718 by buildings during Retention
retain5 <- ggplot(newVacancies1718AprBuildingsRetention,aes(x = Dates, y = Participation,
                                    group = variable, colour = variable)) +
  geom_line() +
  theme(legend.position="none")

# vacancies1718 by buildings during Block
block5 <- ggplot(newVacancies1718AprBuildingsBlock,aes(x = Dates, y = Participation,
                                    group = variable, colour = variable)) +
  geom_line() +
  theme(legend.position="none")

# vacancies1718 by buildings during General Selection
general5 <- ggplot(newVacancies1718AprBuildingsGeneral,aes(x = Dates, y = Participation,
                                    group = variable, colour = variable)) +
  geom_line() +
  theme(legend.position="none")

# vacancies1718 by buildings during Open Assignment
open5 <- ggplot(newVacancies1718AprBuildingsOpen,aes(x = Dates, y = Participation,
                                    group = variable, colour = variable)) +
  geom_line()

grid.arrange(retain5, block5, general5, open5, ncol = 2)
```



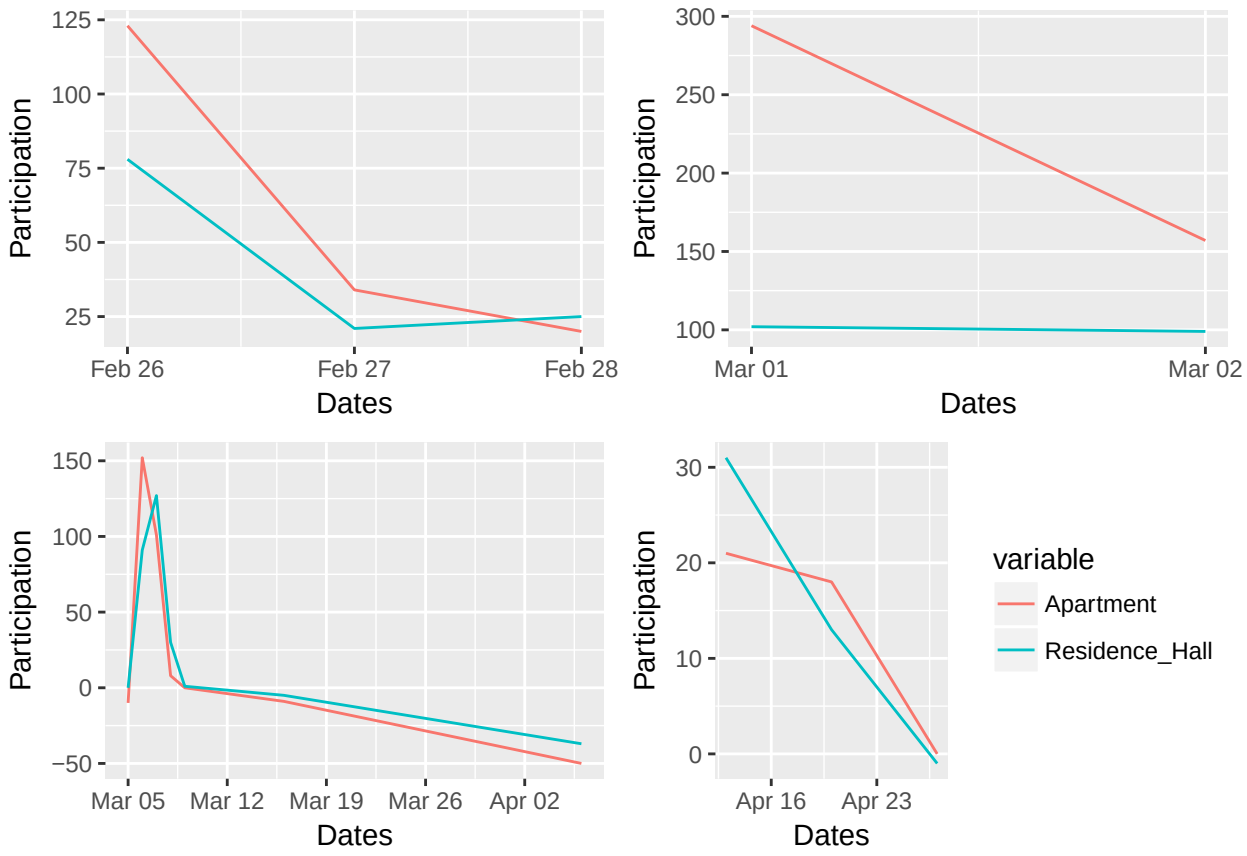
2018-19 Room Selection

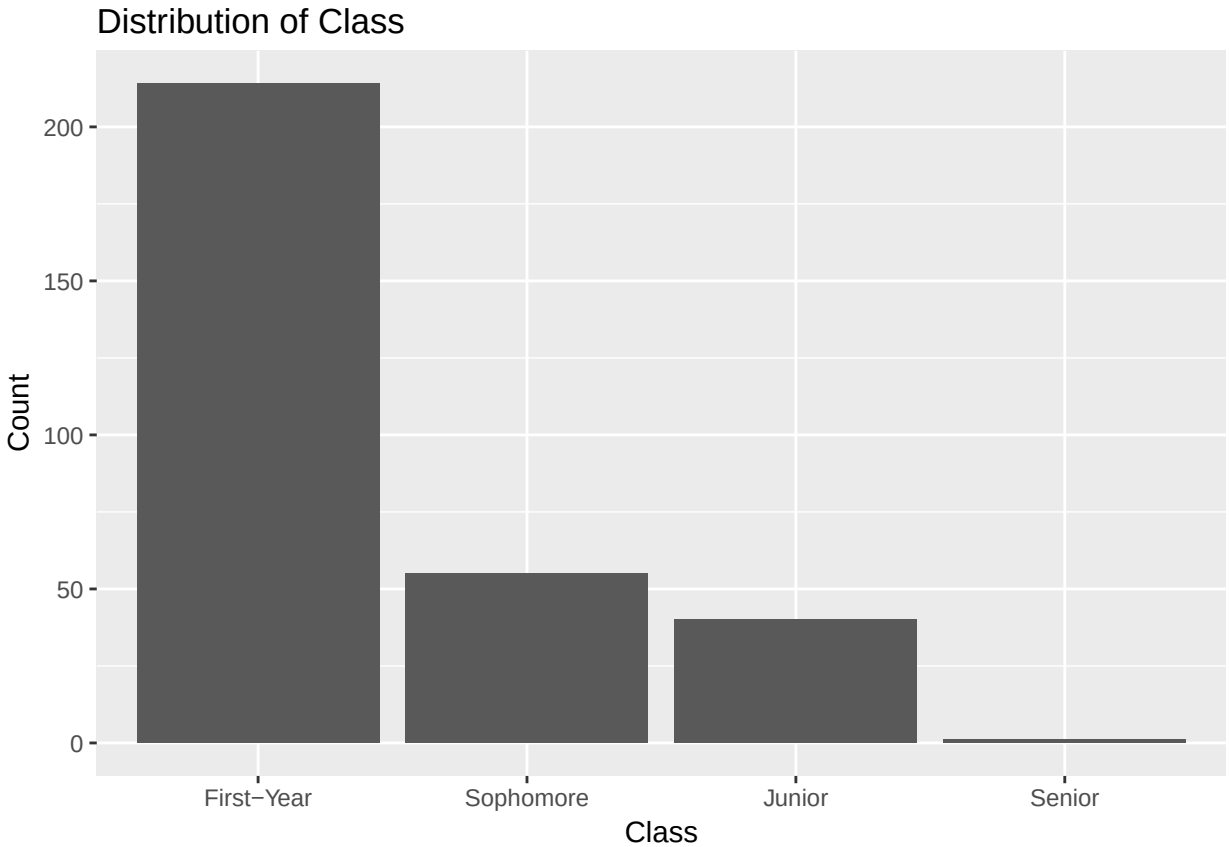
```
# occupancy1819 by buildings
ggplot(newOccupancy1819Buildings,aes(x = Dates, y = Participation, group = variable,
                                     colour = variable)) + geom_line()
```

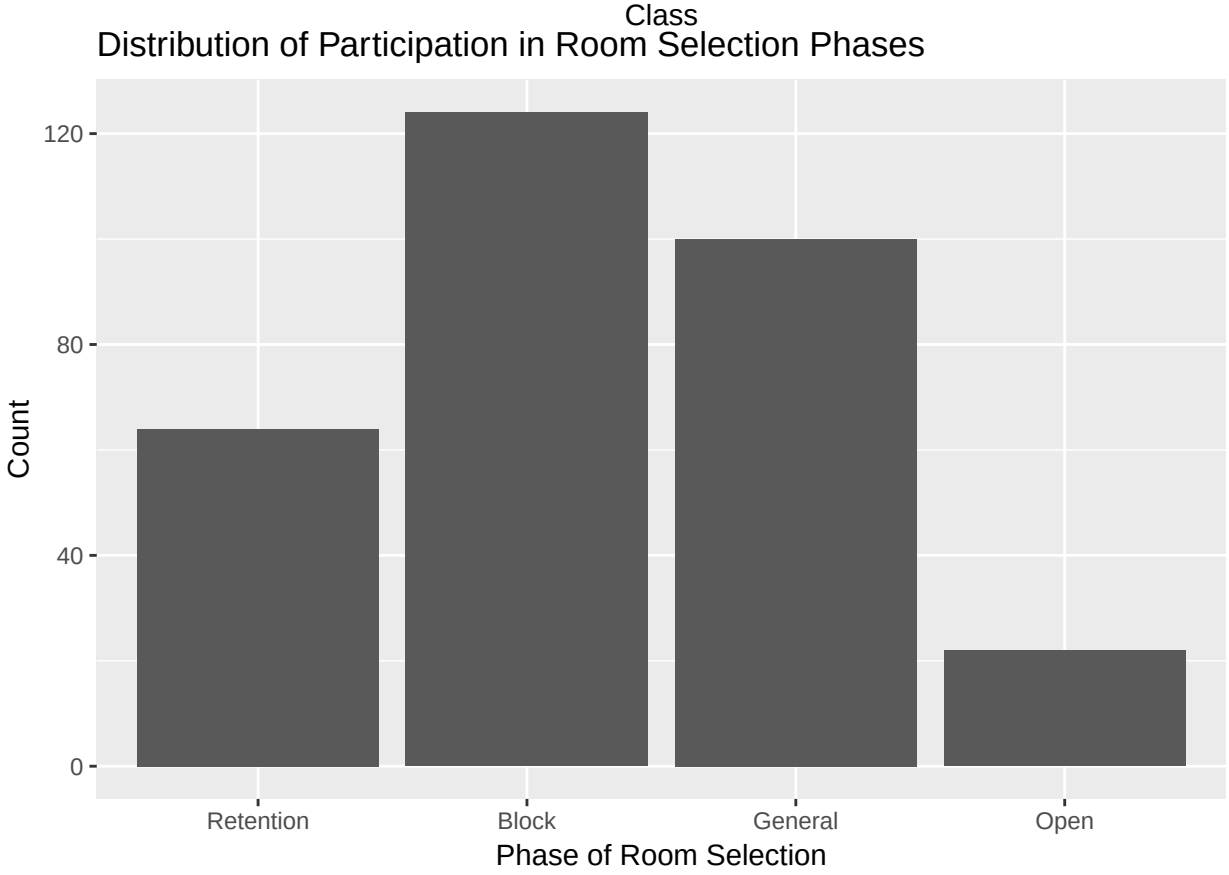


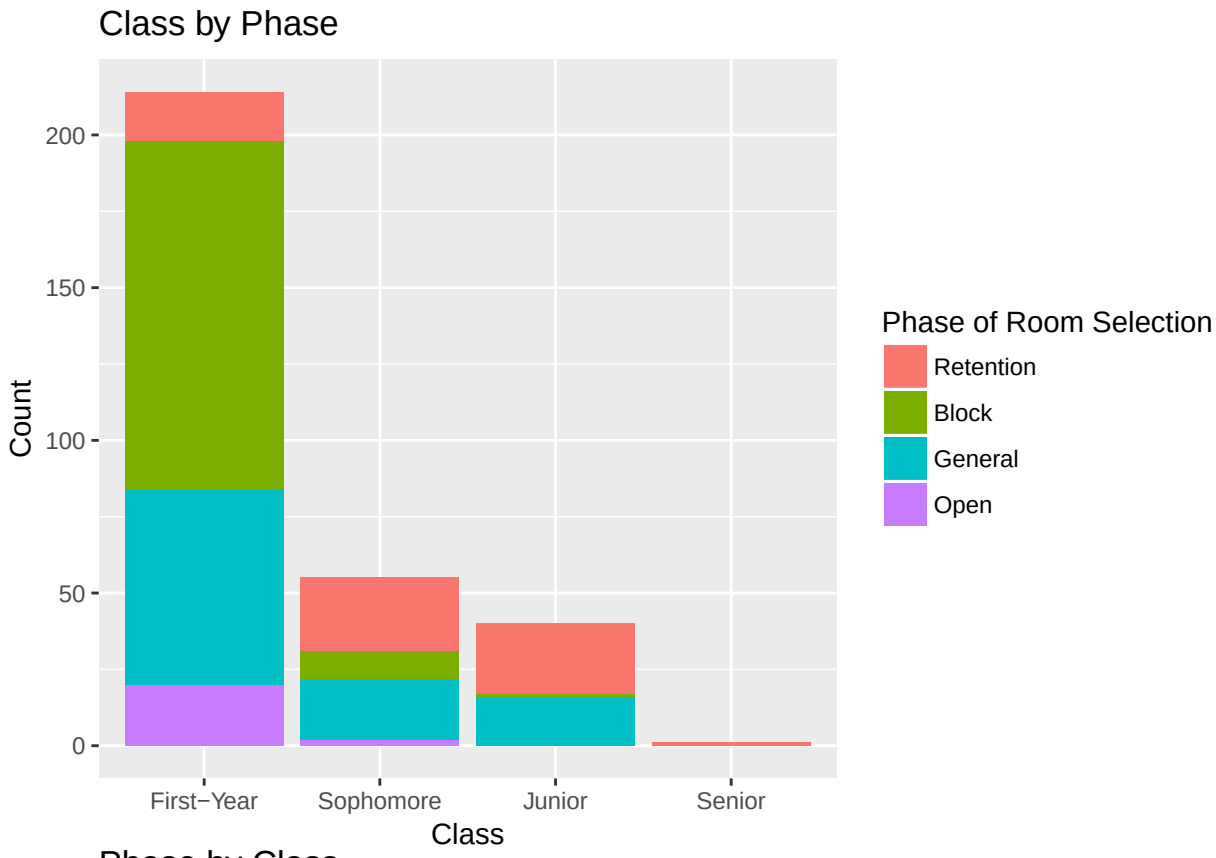

```
# occupancy1819 by buildings during Retention
retain6 <- ggplot(newOccupancy1819BuildingsRetention,aes(x = Dates, y = Participation,
                                                         group = variable,
                                                         colour = variable)) +
                  geom_line() +
                  theme(legend.position="none")

# occupancy1819 by buildings during Block
block6 <- ggplot(newOccupancy1819BuildingsBlock,aes(x = Dates, y = Participation,
                                                    group = variable,
                                                    colour = variable)) +
                  geom_line() +
                  theme(legend.position="none")

# occupancy1819 by buildings during General Selection
general6 <- ggplot(newOccupancy1819BuildingsGeneral,aes(x = Dates, y = Participation,
                                                        group = variable,
                                                        colour = variable)) +
                  geom_line() +
                  theme(legend.position="none")

# occupancy1819 by buildings during Open Assignment
open6 <- ggplot(newOccupancy1819BuildingsOpen,aes(x = Dates, y = Participation,
                                                  group = variable,
                                                  colour = variable)) +
  geom_line()
```

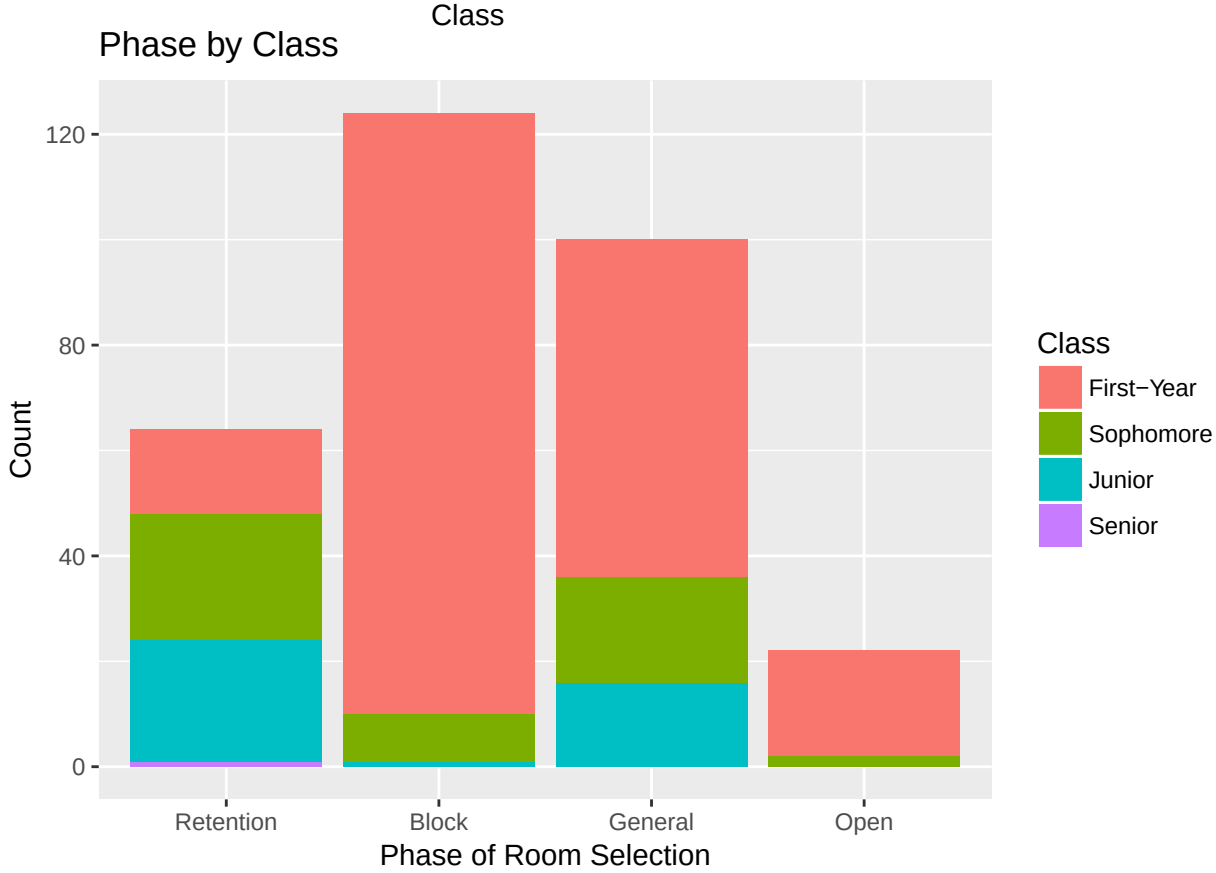
```
grid.arrange(retain6, block6, general6, open6, ncol = 2)
```

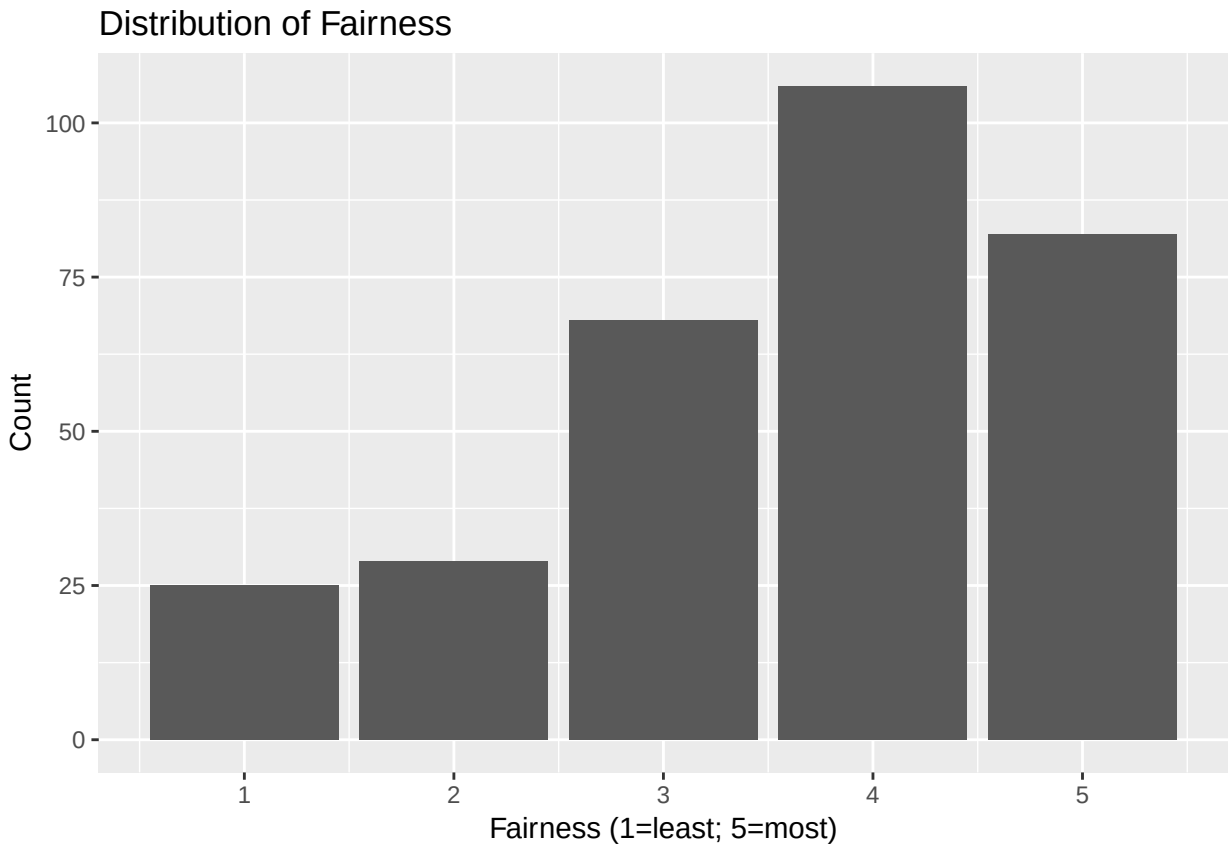


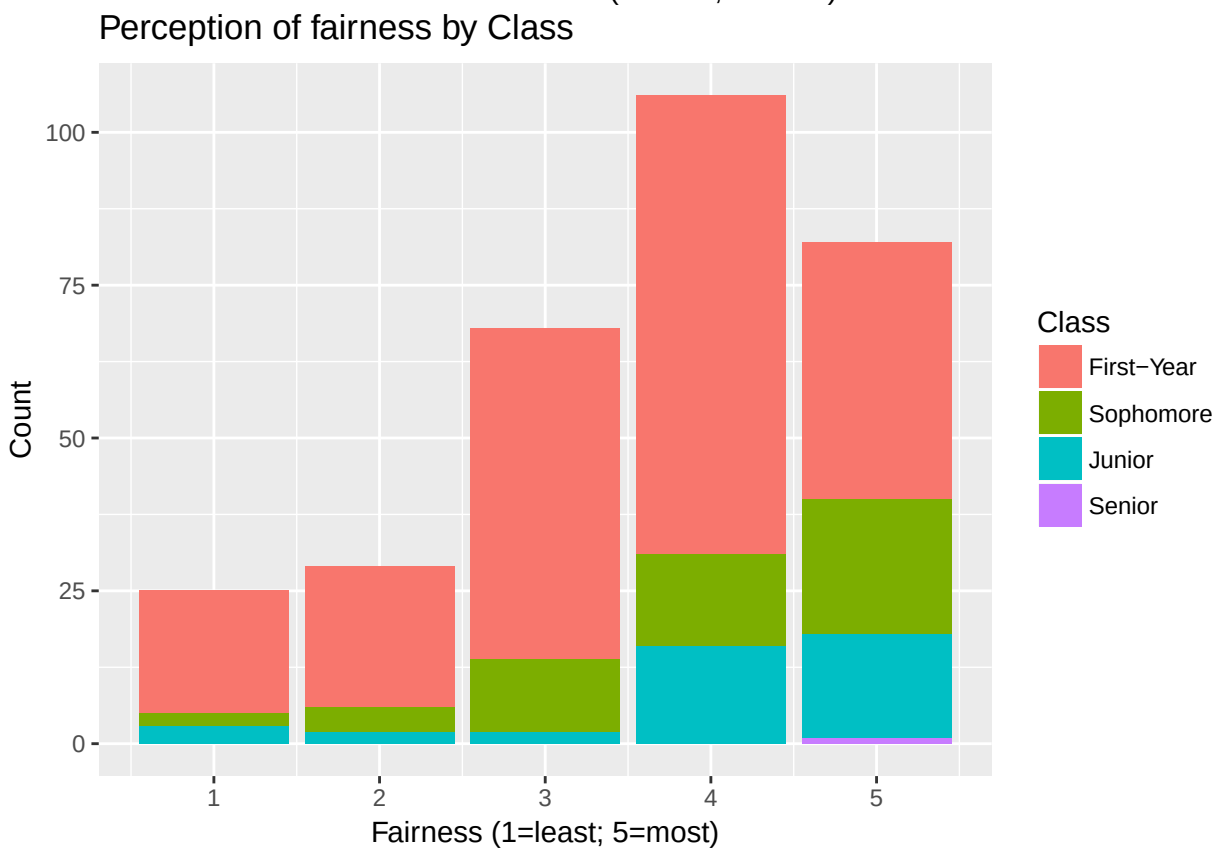

Distribution of Class
Count
200
150
100
50
0
First−Year
Sophomore
Junior
Senior
Class

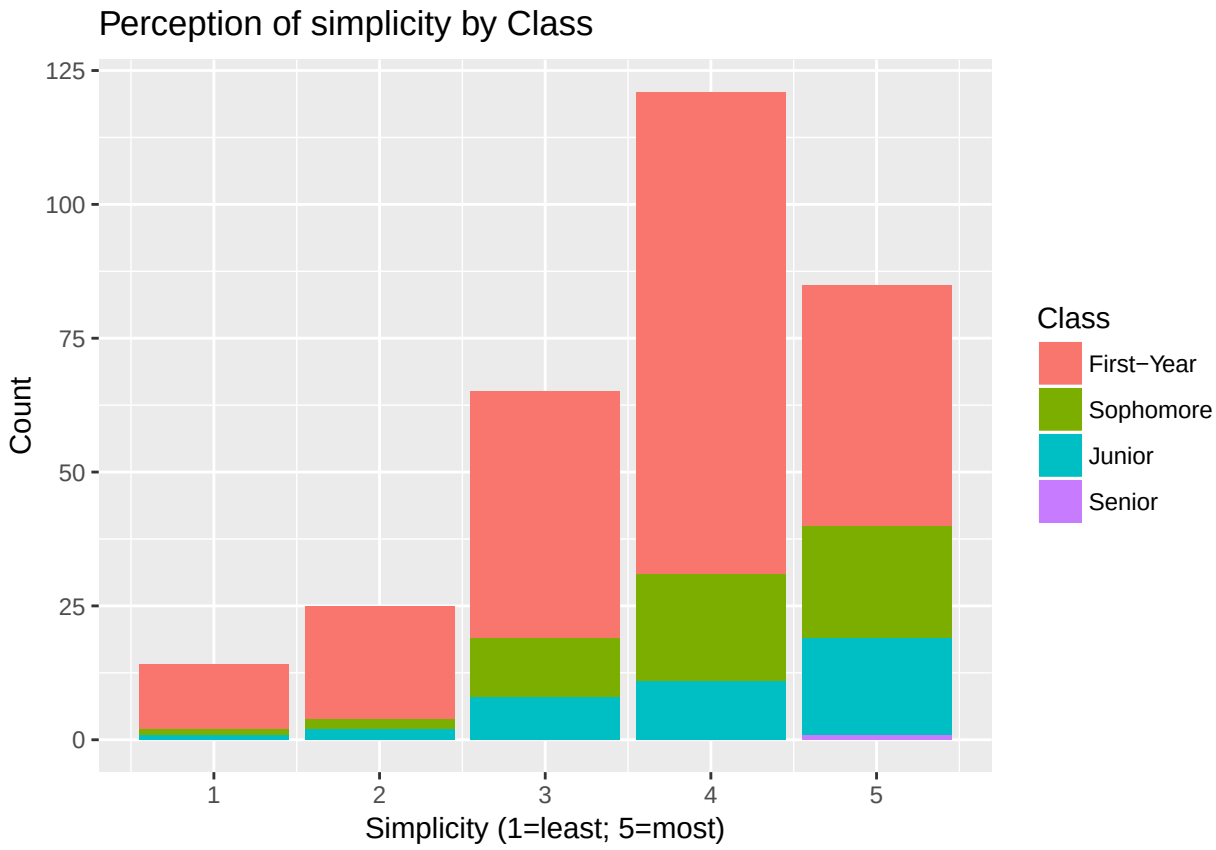

Distribution of Participation in Room Selection Phases
Count
120
80
40
0
Retention
Block
General
Open
Phase of Room Selection

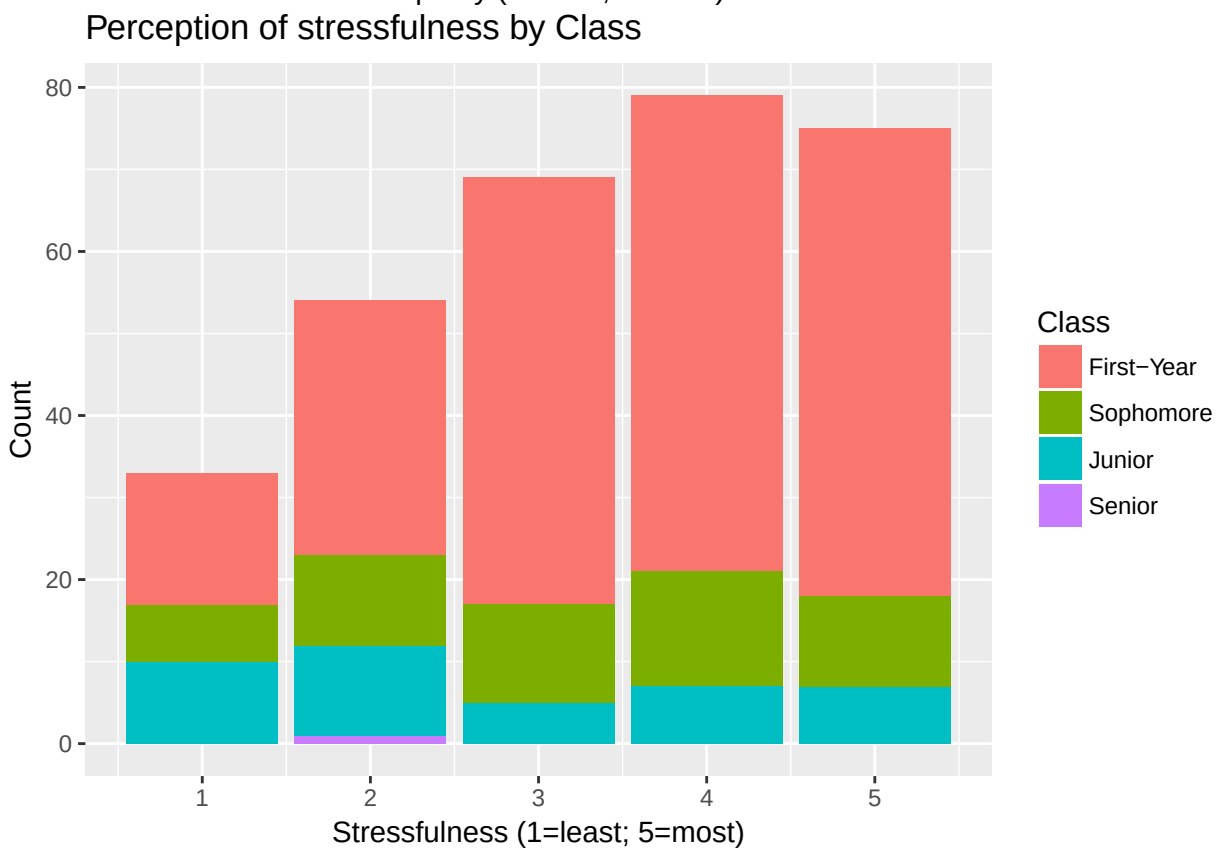

Class by Phase
200
150
100
50
0
Count
First−Year
Sophomore
Junior
Senior
Class
Phase of Room Selection
Retention
Block
General
Open

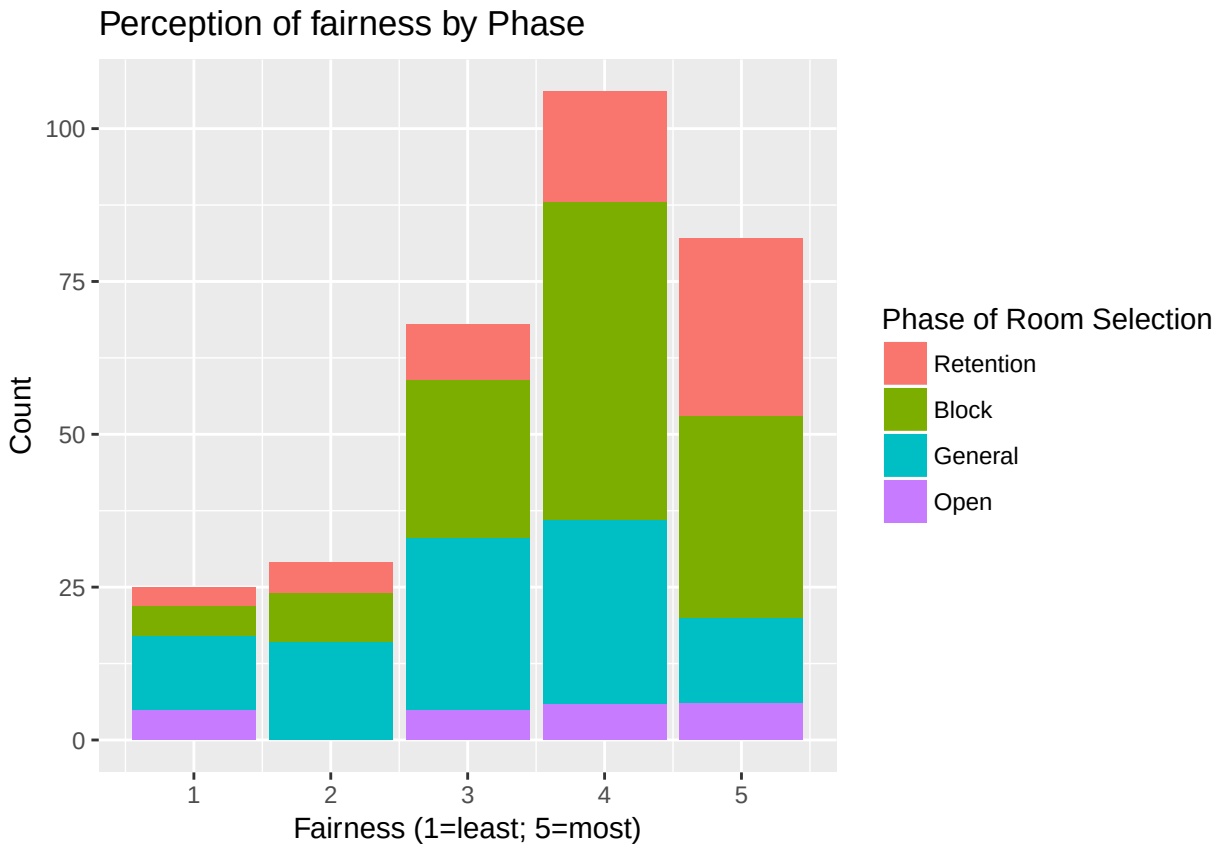

Phase by Class
120
80
40
0
Count
Retention
Block
General
Open
Phase of Room Selection
Class
First−Year
Sophomore
Junior
Senior

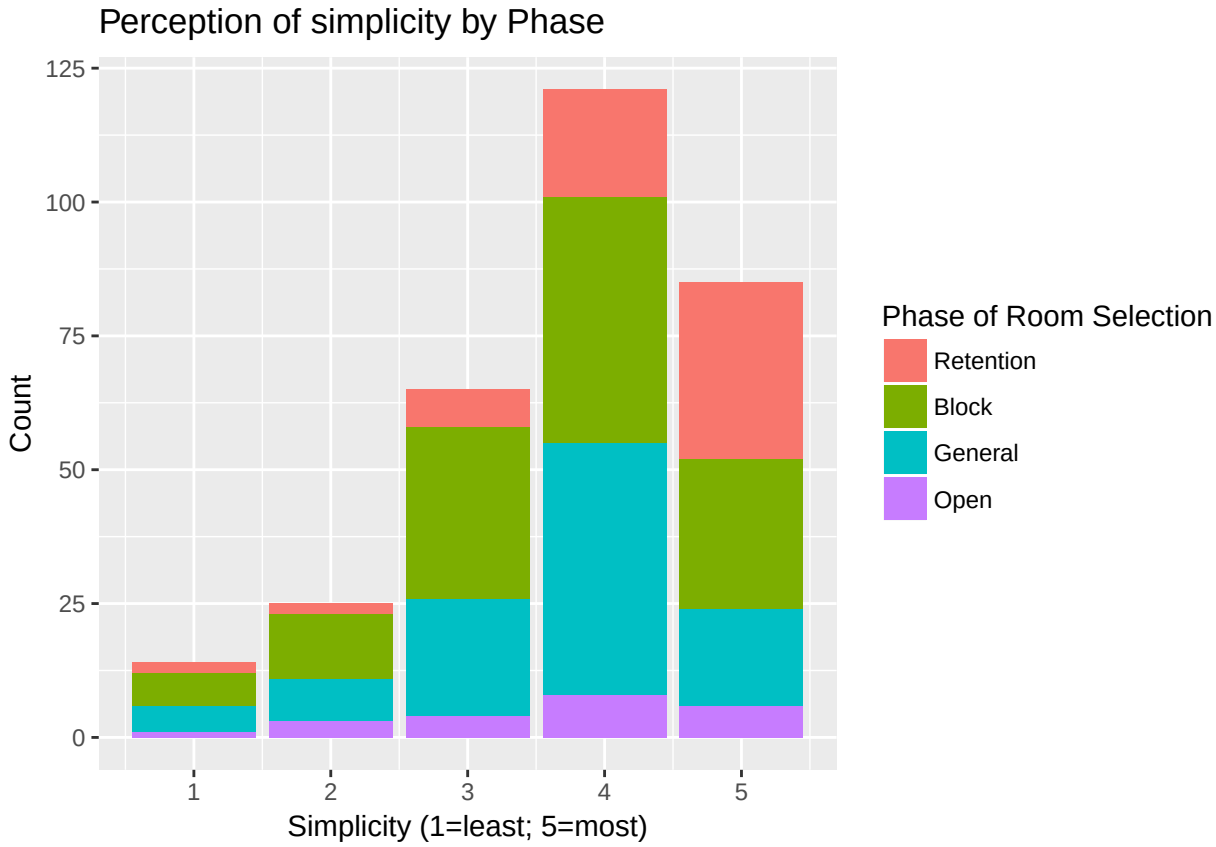

Distribution of Fairness
Count
100
75
50
25
0
1
2
3
4
5
Fairness (1=least; 5=most)

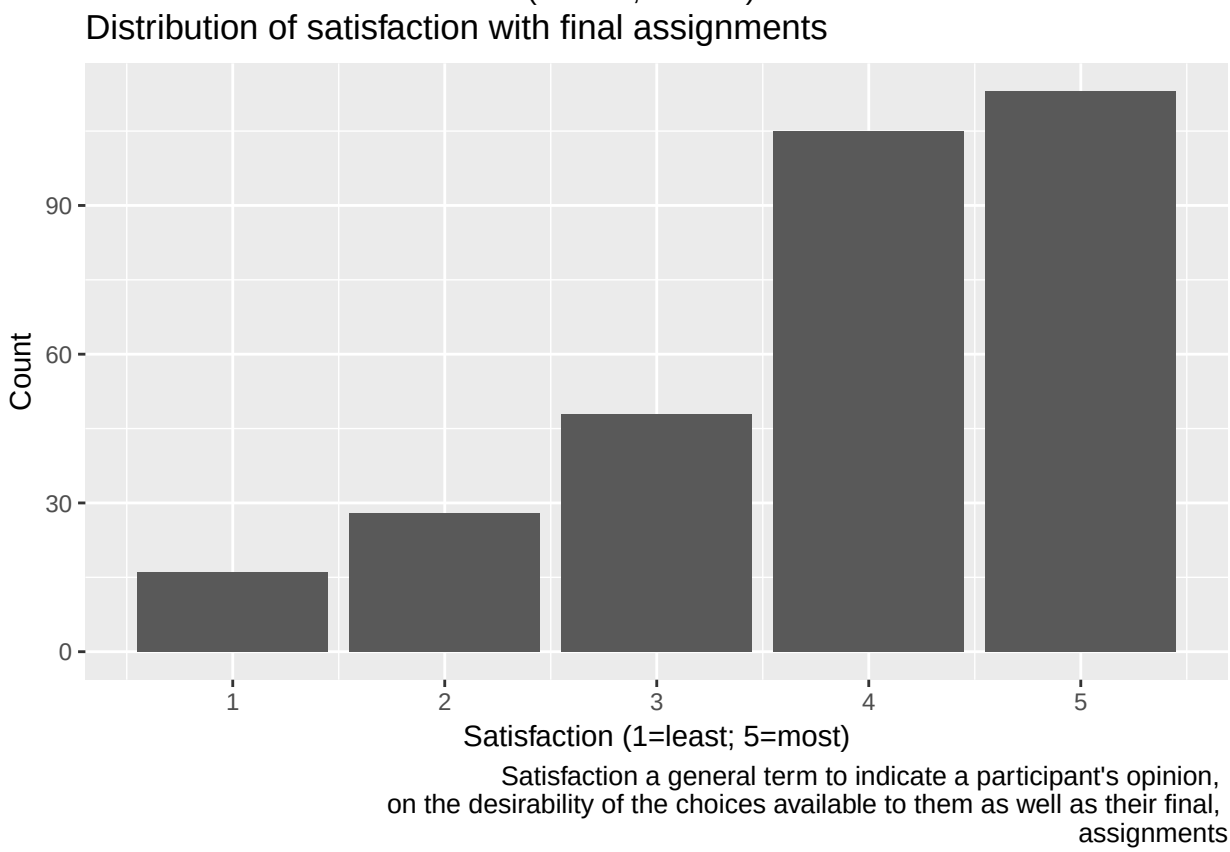

Perception of fairness by Class
Count
100
75
50
25
0
1
2
3
4
5
Fairness (1=least; 5=most)
Class
First−Year
Sophomore
Junior
Senior


Perception of simplicity by Class
125
100
75
50
25
0
Count
1
2
3
4
5
Simplicity (1=least; 5=most)
Class
First−Year
Sophomore
Junior
Senior

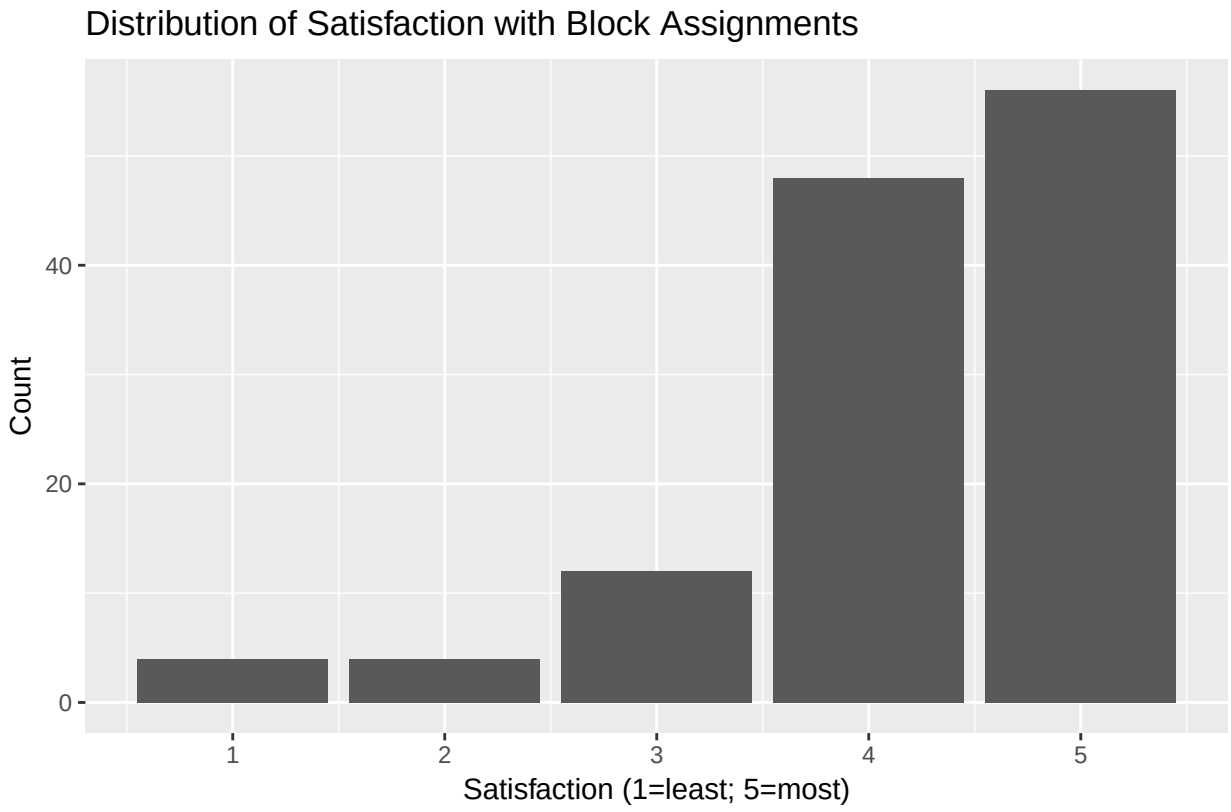

Perception of stressfulness by Class
80
60
40
20
0
Count
1
2
3
4
5
Stressfulness (1=least; 5=most)
Class
First−Year
Sophomore
Junior
Senior


Perception of fairness by Phase
Count
100
75
50
25
0
1
2
3
4
5
Fairness (1=least; 5=most)
Phase of Room Selection
Retention
Block
General
Open

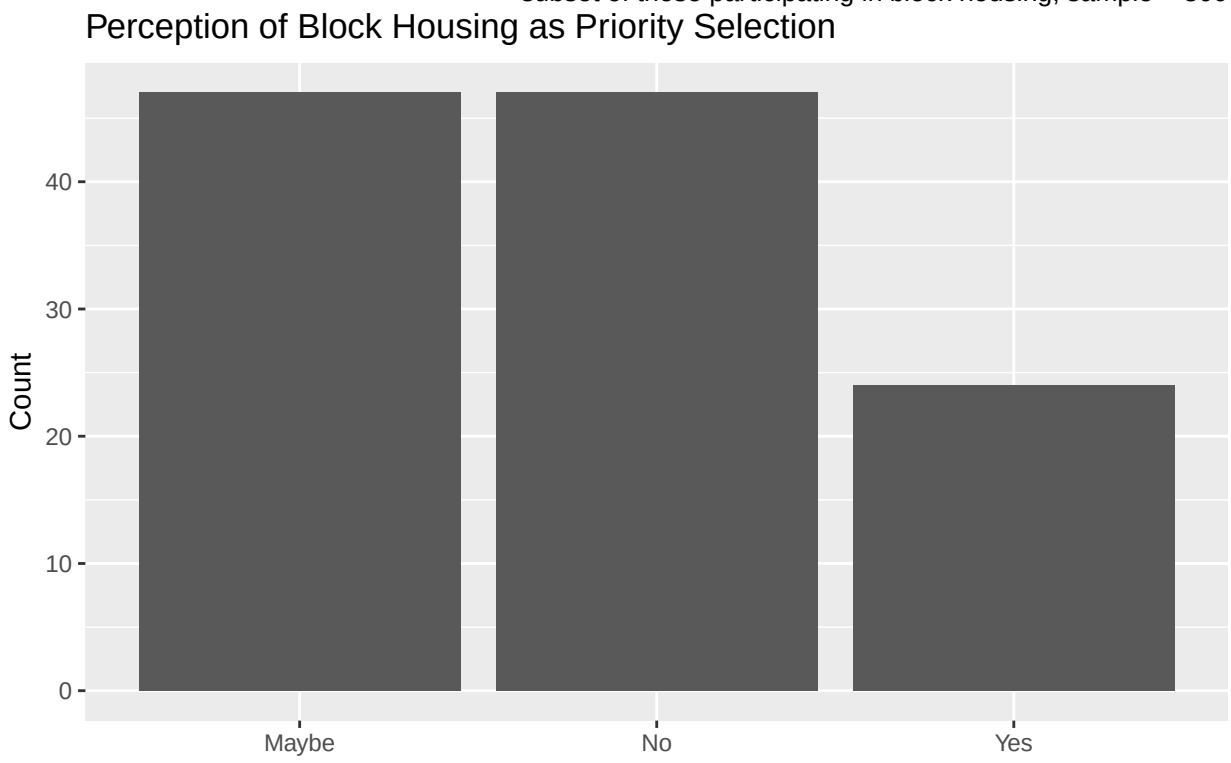

Perception of simplicity by Phase
Count
125
100
75
50
25
0
1
2
3
4
5
Simplicity (1=least; 5=most)
Phase of Room Selection
Retention
Block
General
Open



Satisfaction a general term to indicate a participant's opinion, on the desirability of the choices available to them as well as their final, assignments



subset of those participating in block housing, sample < 300

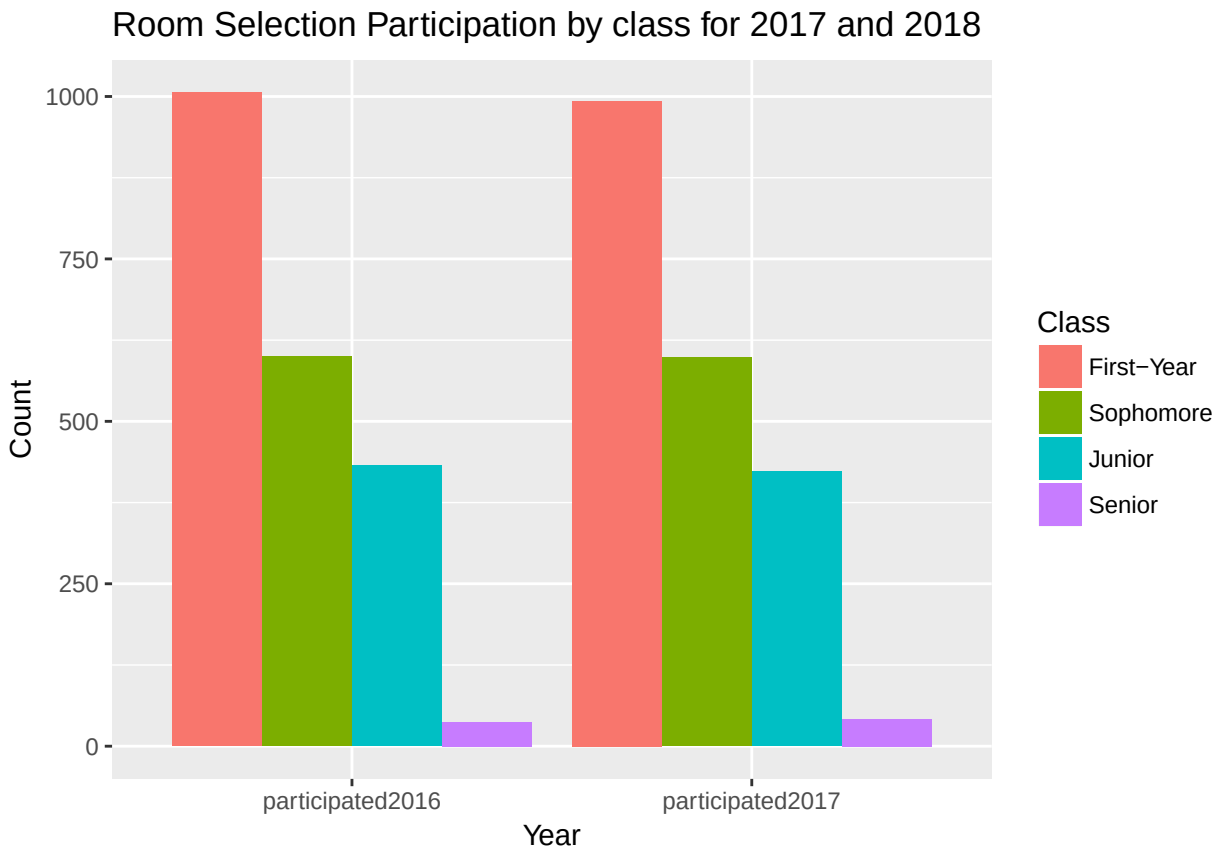


Actual wording of question: 'Would one still do block',
'if block were after General Selection?'

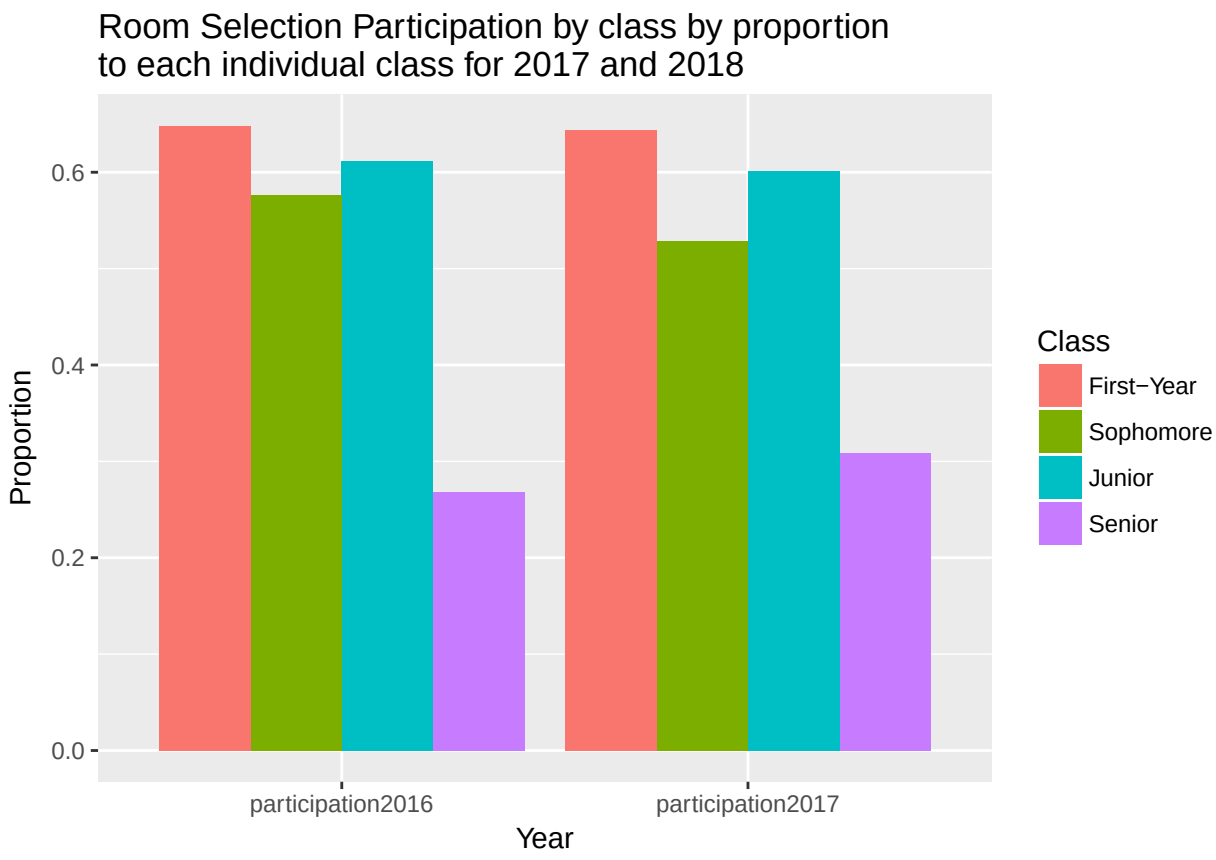
```
byClass <- read.csv("~/Desktop/thesisDocuments/rsParticipationByClass.csv")
byClass <- byClass[1:5,]

byClass$participation2016 <- byClass$participated2016 / byClass$eligible2016
byClass$participation2017 <- byClass$participated2017 / byClass$eligible2017

colnames(byClass) <- c("Year", "eligible2016", "participated2016",
                       "eligible2017", "participated2017", "participation2016",
                       "participation2017")

newByClass <- reshape2::melt(byClass, value.name = "classParticipationRS",
                             varnames = c("Class", "participation"))
```

```
## Using Year as id variables
```

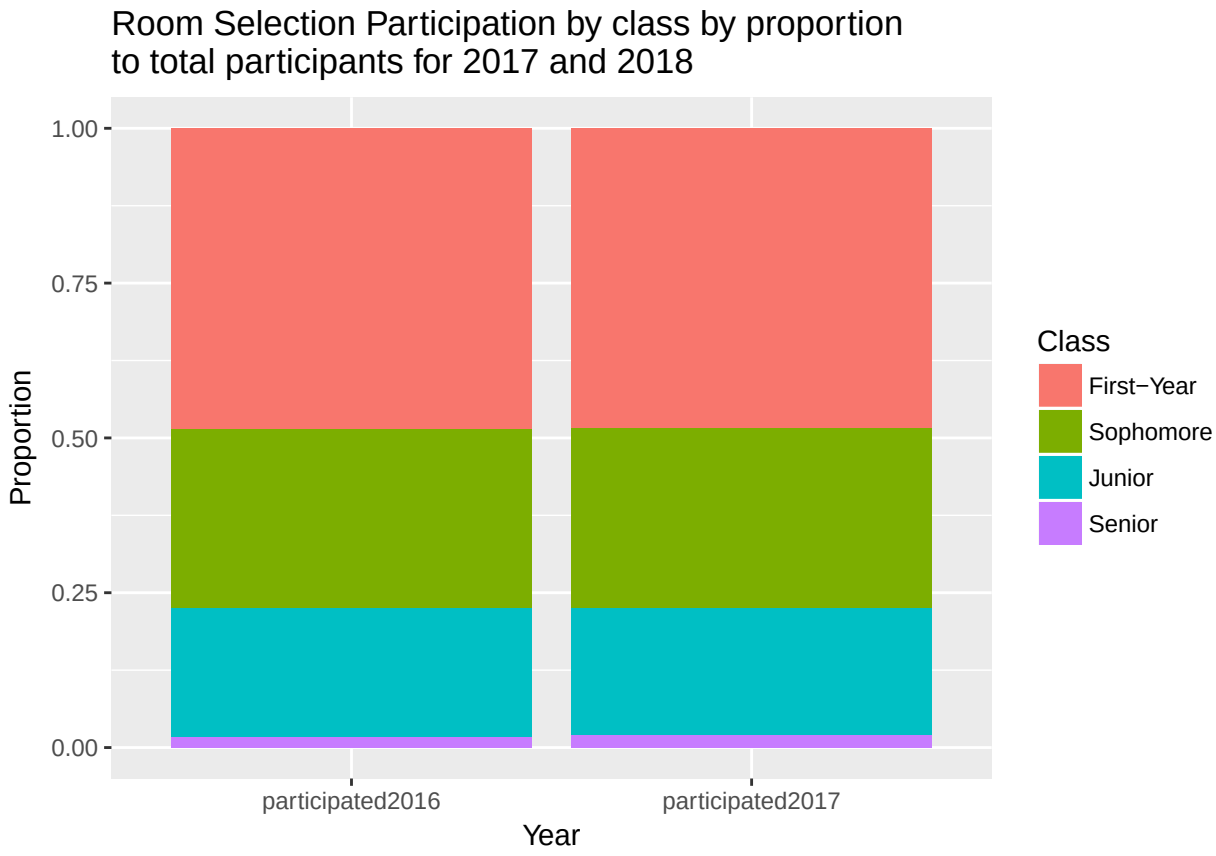
```
newClassParticipationCount <- newByClass[c(6:9,16:19),]
newClassParticipationProportion <- newByClass[c(21:24,26:29),]

Year_ordered <- ordered(newClassParticipationCount$Year, levels = c("First-Year", "Sophomore", "Junior"

ggplot(newClassParticipationCount,aes(x = variable, y = classParticipationRS)) +
                  geom_bar(aes(fill = Year_ordered), stat = "identity",
                           position = "dodge") +
                  labs(title = "Room Selection Participation by class for 2017 and 2018",
                       x = "Year", y = "Count",
                       fill = "Class")
```



```
ggplot(newClassParticipationProportion,aes(x = variable, y = classParticipationRS)) +
                  geom_bar(aes(fill = Year_ordered), stat = "identity",
                           position = "dodge") +
                  labs(title = "Room Selection Participation by class by proportion \nto each individual
                       x = "Year", y = "Proportion",
                       fill = "Class")
```




```
ggplot(newClassParticipationCount,aes(x = variable, y = classParticipationRS)) +
                  geom_bar(aes(fill = Year_ordered), stat = "identity",
                           position = "fill") +
                  labs(title = "Room Selection Participation by class by proportion \nto total participa
                       x = "Year", y = "Proportion",
                       fill = "Class")
```


Room Selection Participation by class by proportion to total participants for 2017 and 2018
1.00
0.75
0.50
0.25
0.00
Proportion
participated2016
participated2017
Year
Class
First−Year
Sophomore
Junior
Senior