Lab 6

This data set consists of observations from 276 speed dates. It was obtained from the blackboard site for BTRY 6010.

The source of the `SpeedDating` data is "Gelman, A. and Hill, J., *Data analysis using regression and multilevel/hierarchical models*, Cambridge University Press: New York, 2007."

```
SpeedDating <- read.csv("SpeedDating.csv")
```

We will start by looking at our data before we proceed with any other analysis.

```
dim(SpeedDating)
```

```
## [1] 276  22
```

```
library(Hmisc)
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'Hmisc'
```

```
## The following objects are masked from 'package:base':
##
##     format.pval, round.POSIXt, trunc.POSIXt, units
```

```
describe(SpeedDating)
```

```
## SpeedDating
##
##  22  Variables      276  Observations
## ---------------------------------------------------------------------------
## DecisionM
##       n missing  unique    Info     Sum    Mean
##     276       0       2    0.75     146   0.529
## ---------------------------------------------------------------------------
## DecisionF
##       n missing  unique    Info     Sum    Mean
##     276       0       2    0.75     127  0.4601
## ---------------------------------------------------------------------------
## LikeM
##       n missing  unique    Info    Mean     .05     .10     .25     .50
```

```
##     274       2      12    0.97   6.682    4.00    4.15    6.00    7.00
##     .75     .90     .95
##    8.00    9.00    9.35
##
##           1 2 3  4 4.5  5  6 6.5  7  8  9 10
## Frequency 2 4 7 15   1 33 54   3 68 46 27 14
## %         1 1 3  5   0 12 20   1 25 17 10  5
## ------------------------------------------------------------------------
## LikeF
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     272       4      11    0.96   6.366       3       4       5       7
##     .75     .90     .95
##       8       8       9
##
##           1 2  3  4  5  6  7  8 8.5  9 10
## Frequency 1 7 14 13 41 54 72 48   1 13  8
## %         0 3  5  5 15 20 26 18   0  5  3
## ------------------------------------------------------------------------
## PartnerYesM
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     272       4      11    0.98   5.757    2.00    3.00    5.00    6.00
##     .75     .90     .95
##    7.00    8.00    9.45
##
##           0  1  2  3  4  5  6  7  8  9 10
## Frequency 1 12 14 19 19 49 54 46 34 10 14
## %         0  4  5  7  7 18 20 17 12  4  5
## ------------------------------------------------------------------------
## PartnerYesF
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     272       4      10    0.97   5.835    1.00    2.10    5.00    6.00
##     .75     .90     .95
##    7.00    9.00    9.45
##
##           1  2  3  4  5  6  7  8  9 10
## Frequency 15 13 15 10 66 44 50 28 17 14
## %          6  5  6  4 24 16 18 10  6  5
## ------------------------------------------------------------------------
## AgeM
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     273       3      21    0.99    26.6    22.0    22.0    24.0    27.0
##     .75     .90     .95
##    29.0    30.8    32.4
##
## lowest : 18 19 20 21 22, highest: 34 36 37 39 42
## ------------------------------------------------------------------------
## AgeF
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     271       5      20    0.99   26.19      21      22      23      26
##     .75     .90     .95
##      28      31      34
##
## lowest : 19 20 21 22 23, highest: 34 35 36 38 55
## ------------------------------------------------------------------------
```

```
## RaceM
##       n missing  unique
##     276       0       6
##
##             Asian Black Caucasian Latino Other
## Frequency 2    65    10       161     17    21
## %         1    24     4        58      6     8
## ------------------------------------------------------------------------
## RaceF
##       n missing  unique
##     276       0       6
##
##             Asian Black Caucasian Latino Other
## Frequency 4    70    15       148     23    16
## %         1    25     5        54      8     6
## ------------------------------------------------------------------------
## AttractiveM
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     273       3      11    0.97   6.687       4       5       5       7
##     .75     .90     .95
##       8       9      10
##
##           1 2 3  4  5  6 6.5  7  8  9 10
## Frequency 1 3 7 12 54 50   1 48 55 23 19
## %         0 1 3  4 20 18   0 18 20  8  7
## ------------------------------------------------------------------------
## AttractiveF
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     274       2      12    0.97   6.274   3.000   4.000   5.000   6.000
##     .75     .90     .95
##   8.000   8.000   9.175
##
##           1 2  3  4  5  6  7 7.5  8  9 9.5 10
## Frequency 3 7 16 15 51 50 56   1 49 12   1 13
## %         1 3  6  5 19 18 20   0 18  4   0  5
## ------------------------------------------------------------------------
## SincereM
##       n missing  unique    Info    Mean
##     271       5       9    0.95   7.856
##
##           1 3 4  5  6  7  8  9 10
## Frequency 1 1 5 11 25 52 87 50 39
## %         0 0 2  4  9 19 32 18 14
## ------------------------------------------------------------------------
## SincereF
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     273       3      10    0.96   7.784       5       5       7       8
##     .75     .90     .95
##       9      10      10
##
##           1 2 3 4  5  6  7  8  9 10
## Frequency 1 1 3 5 19 26 44 78 50 46
## %         0 0 1 2  7 10 16 29 18 17
## ------------------------------------------------------------------------
```

```
## IntelligentM
##       n missing  unique    Info    Mean
##     268       8       8    0.95   7.621
##
##           4  5  6  7 7.5  8  9 10
## Frequency 3 17 33 65   1 82 41 26
## %         1  6 12 24   0 31 15 10
## ------------------------------------------------------------------------
## IntelligentF
##       n missing  unique    Info    Mean
##     273       3       9    0.94   7.923
##
##           2 3 4  5  6  7  8  9 10
## Frequency 1 1 3 14 13 59 92 51 39
## %         0 0 1  5  5 22 34 19 14
## ------------------------------------------------------------------------
## FunM
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     270       6      11    0.97   6.863       4       5       6       7
##     .75     .90     .95
##       8       9      10
##
##           0 1 2 3  4  5  6  7  8  9 10
## Frequency 1 1 3 4 12 37 51 57 61 22 21
## %         0 0 1 1  4 14 19 21 23  8  8
## ------------------------------------------------------------------------
## FunF
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     270       6      10    0.98   6.563       3       4       5       7
##     .75     .90     .95
##       8       9      10
##
##           1 2  3  4  5  6  7  8  9 10
## Frequency 1 7 11 23 34 48 55 49 23 19
## %         0 3  4  9 13 18 20 18  9  7
## ------------------------------------------------------------------------
## AmbitiousM
##       n missing  unique    Info    Mean
##     259      17       9    0.97   6.768
##
##           2 3  4  5  6  7  8  9 10
## Frequency 2 4 13 51 44 55 46 27 17
## %         1 2  5 20 17 21 18 10  7
## ------------------------------------------------------------------------
## AmbitiousF
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     266      10      10    0.97   7.429       5       5       6       8
##     .75     .90     .95
##       9      10      10
##
##           1 2 3 4  5  6  7  8  9 10
## Frequency 1 2 3 4 35 30 48 65 45 33
## %         0 1 1 2 13 11 18 24 17 12
## ------------------------------------------------------------------------
```

```
## SharedInterestsM
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     249      27      12    0.97   5.588       2       3       4       5
##     .75     .90     .95
##       7       8       9
##
##           0 1  2  3  4  5  6 6.5  7  8 9 10
## Frequency 4 5 11 12 31 63 42   1 38 24 7 11
## %         2 2  4  5 12 25 17   0 15 10 3  4
## ------------------------------------------------------------------------------
## SharedInterestsF
##       n missing  unique    Info    Mean     .05     .10     .25     .50
##     246      30      12    0.98    5.47       2       2       4       6
##     .75     .90     .95
##       7       8       9
##
##           0  1  2  3  4  5 5.5  6  7  8  9 10
## Frequency 2 10 20 20 19 50   1 40 38 25 12  9
## %         1  4  8  8  8 20   0 16 15 10  5  4
## ------------------------------------------------------------------------------
```

Note: the variable `Decision` *is categorical. R has read it in as a quantitative value.*

```
table(SpeedDating$DecisionF, SpeedDating$DecisionM)
```

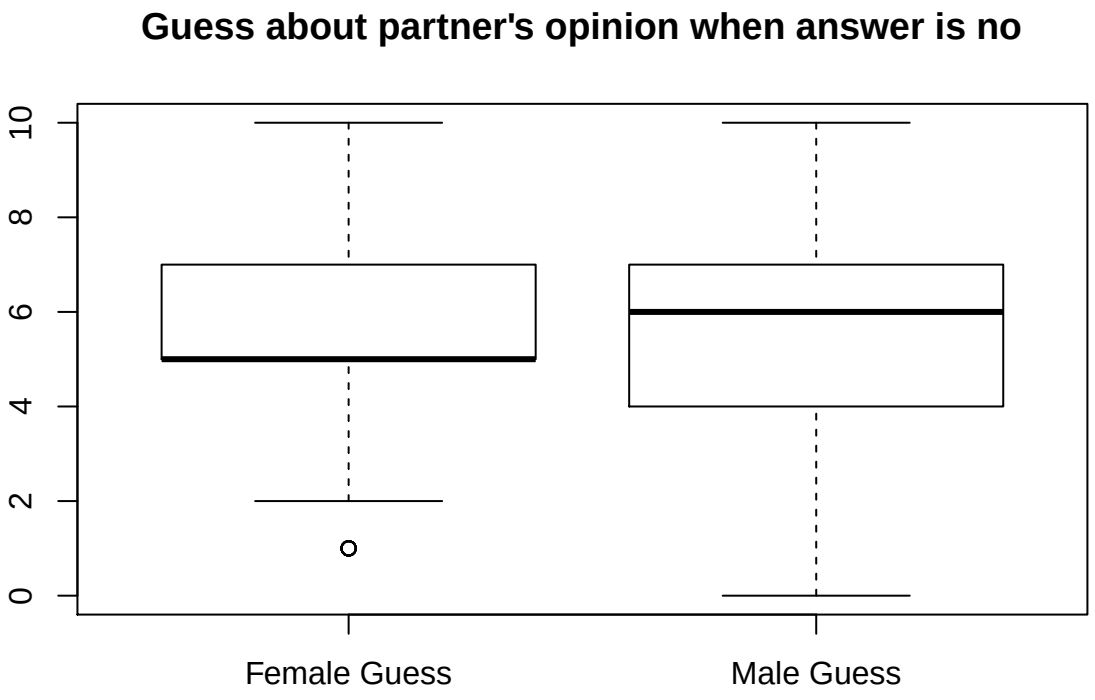
```
##
##      0  1
##   0 66 83
##   1 64 63
```

When `DecisionF` and `DecisionM` are both 0, neither participant wanted another date. When `DecisionF` is 0 and `DecisionM` is 1, the female did not want another date, but the male did want another date. When `DecisionF` is 1 and `DecisionM` is 0, the male did not want another date, but the female did want another date. When `DecisionF` and `DecisionM` are both 1, both participants wanted another date.

```
onlyOne <- subset(SpeedDating, DecisionF + DecisionM == 1)
SympMale = table(onlyOne$DecisionM)/nrow(onlyOne)
barplot(SympMale, names.arg = c("F", "M"), main = "Who Desired the Second Date When One of Them Did Not
```

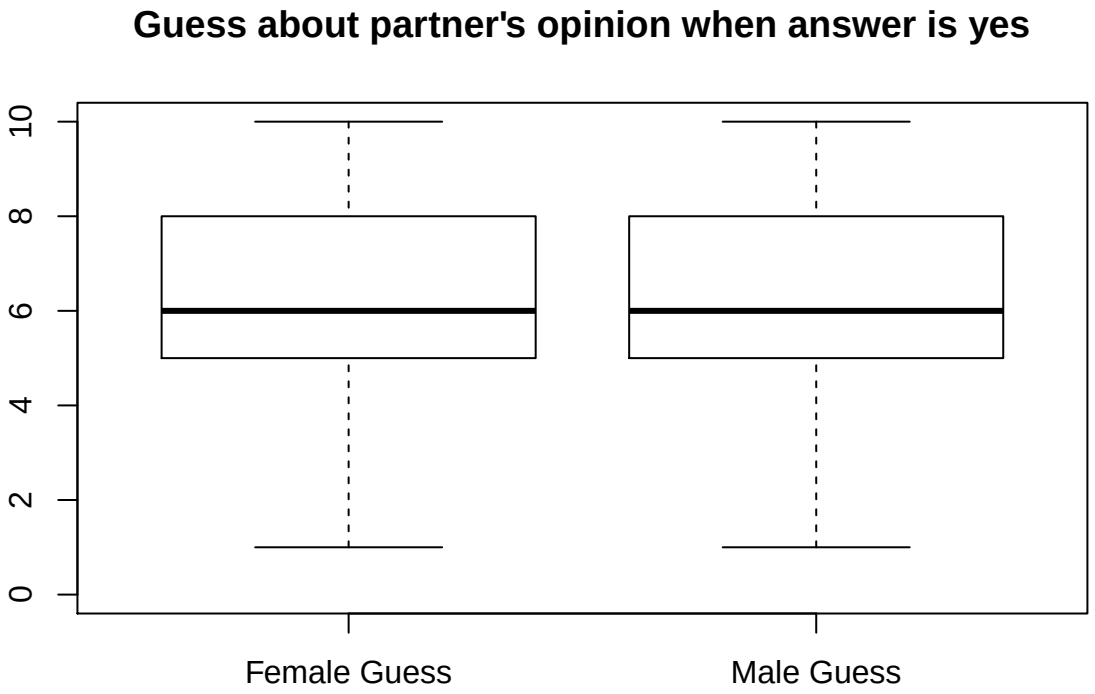
When only one partner was interested in another date, it was more often the male than the female.

```
boxplot(SpeedDating$PartnerYesF[SpeedDating$DecisionM == 0], SpeedDating$PartnerYesM[SpeedDating$Decisi
    0], main = "Guess about partner's opinion when answer is no", ylim = c(0,
    10), names = c("Female Guess", "Male Guess"))
```



This plot contains ratings on a scale of 1 to 10 by females (on the left) and males (on the right) of how likely their "date" would want to see them again for observations where the date did not want to see them again.

```
boxplot(SpeedDating$PartnerYesF[SpeedDating$DecisionM == 1], SpeedDating$PartnerYesM[SpeedDating$Decisi
    1], main = "Guess about partner's opinion when answer is yes", ylim = c(0,
    10), names = c("Female Guess", "Male Guess"))
```

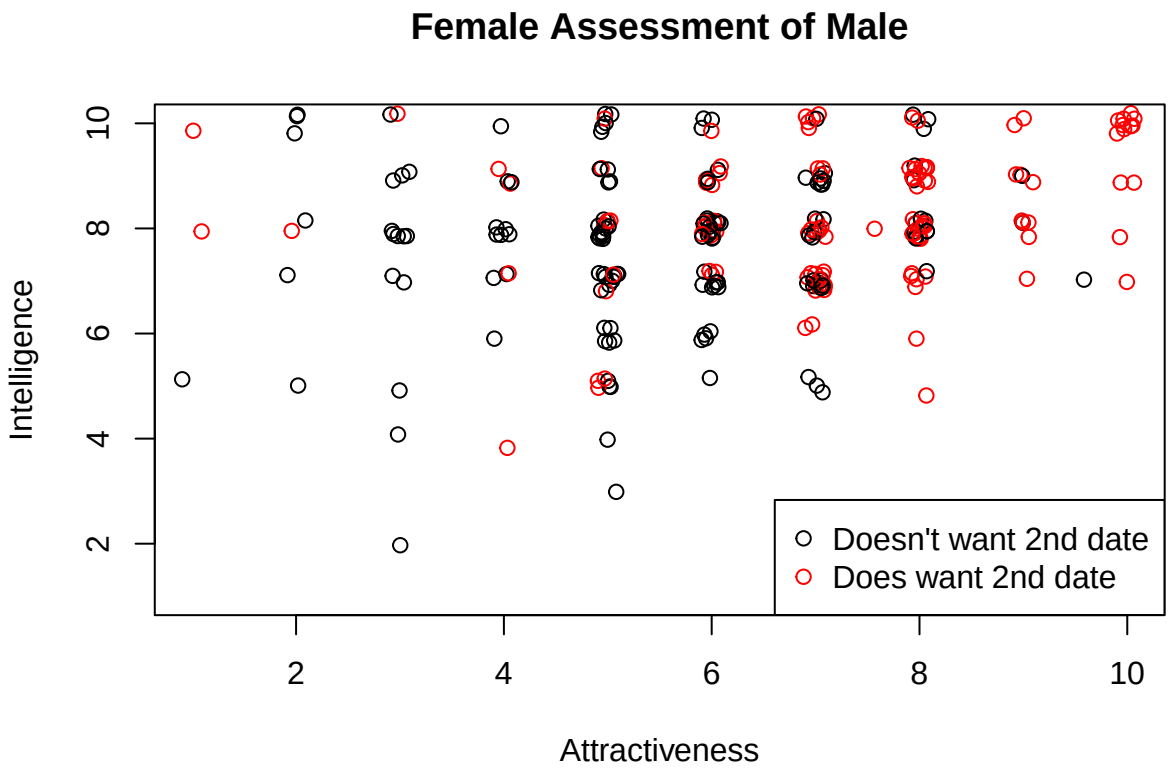


This plot contains ratings on a scale of 1 to 10 by females (on the left) and males (on the right) of how likely their "date" would want to see them again for observations where the date did want to see them again.

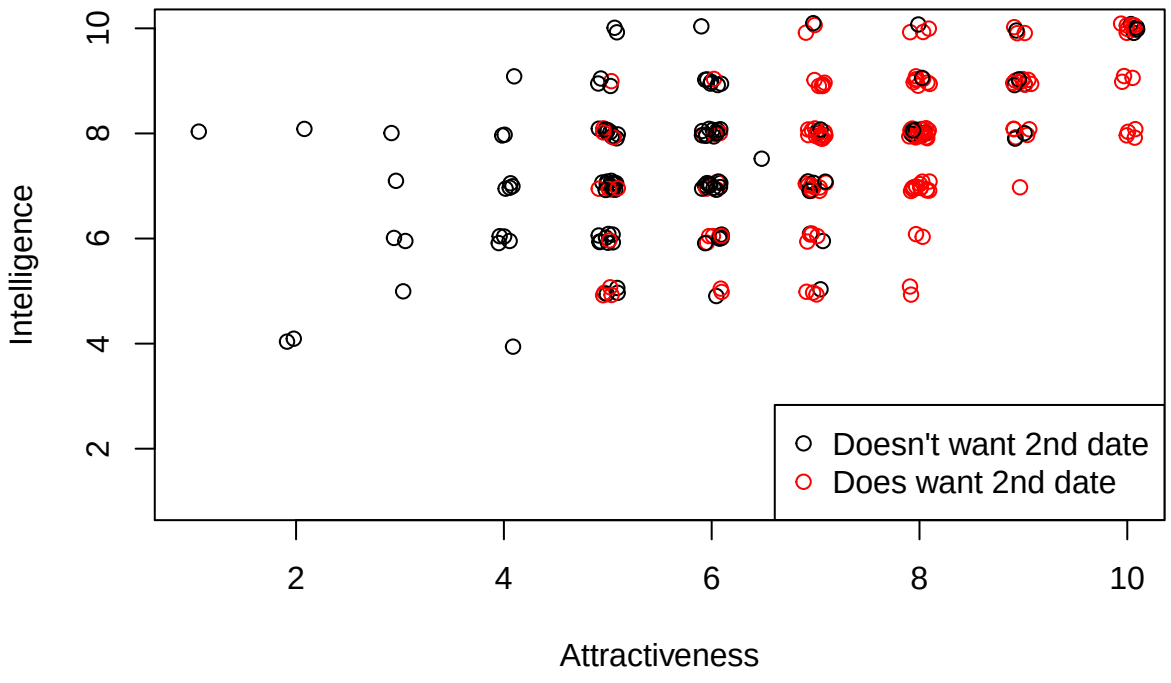
Males and females were equally likely to detect when their partners wanted another date. The distributions look identical.

When the partner did not want another date, the median male's assessment of the situation is unchanged while the median female does notice that he is not interested in another date. That said, the 3rd quartile of the male distribution does drop as much as that of the female distribution.

```
plot(jitter(SpeedDating$AttractiveF), jitter(SpeedDating$IntelligentF), main = "Female Assessment of Ma
    ylim = c(1, 10), xlim = c(1, 10), col = 1 + SpeedDating$DecisionF, xlab = "Attractiveness",
    ylab = "Intelligence")
legend("bottomright", legend = c("Doesn't want 2nd date", "Does want 2nd date"),
    col = 1:2, pch = 1)
```




```
plot(jitter(SpeedDating$AttractiveM), jitter(SpeedDating$IntelligentM), main = "Male Assessment of Fema
    ylim = c(1, 10), xlim = c(1, 10), col = 1 + SpeedDating$DecisionM, xlab = "Attractiveness",
    ylab = "Intelligence")
legend("bottomright", legend = c("Doesn't want 2nd date", "Does want 2nd date"),
    col = 1:2, pch = 1)
```



Males tended to agree to a second date with partners they found highly attractive and intelligent. Females are less predictable. Compared to the males, they were more likely to be interested in a second date with someone they found intelligent, but not very attractive. However, it is hard to generalize because we are only looking at two of the six attributes measured for each date.