# Etsy Data After Review

*Ramya Balasubramaniam*

*4/3/2017*

## Loading the data used after treatment

```
etsy_data1 <- read.csv(file = "~/downloads/etsy_data/etsy_aft_treat1.csv")
etsy_data2 <- read.csv(file = "~/downloads/etsy_data/etsy_aft_treat2.csv")
names(etsy_data1)
```

```
##  [1] "Sale.Date"          "Item.Name"          "Buyer"
##  [4] "Quantity"           "Price"              "Coupon.Code"
##  [7] "Coupon.Details"     "Coupon.Discount"    "Order.Shipping"
## [10] "Order.Sales.Tax"    "Item.Total"         "Currency"
## [13] "Transaction.ID"     "Listing.ID"         "Date.Paid"
## [16] "Date.Shipped"       "Ship.Name"          "Ship.Address1"
## [19] "Ship.Address2"      "Ship.City"          "Ship.State"
## [22] "Ship.Zipcode"       "Ship.Country"       "Order.ID"
## [25] "Variations"         "Order.Type"         "Listings.Type"
## [28] "Payment.Type"       "InPerson.Discount"  "InPerson.Location"
## [31] "VAT.Paid.by.Buyer"
```

```
names(etsy_data2)
```

```
##  [1] "Sale.Date"                   "Order.ID"
##  [3] "Buyer.User.ID"               "Full.Name"
##  [5] "First.Name"                  "Last.Name"
##  [7] "Number.of.Items"             "Payment.Method"
##  [9] "Date.Shipped"                "Street.1"
## [11] "Street.2"                    "Ship.City"
## [13] "Ship.State"                  "Ship.Zipcode"
## [15] "Ship.Country"                "Currency"
## [17] "Order.Value"                 "Coupon.Code"
## [19] "Coupon.Details"              "Shipping"
## [21] "Sales.Tax"                   "Order.Total"
## [23] "Status"                      "Card.Processing.Fees"
## [25] "Order.Net"                   "Adjusted.Order.Total"
## [27] "Adjusted.Card.Processing.Fees" "Adjusted.Net.Order.Amount"
## [29] "Buyer"                       "Order.Type"
## [31] "Payment.Type"                "InPerson.Discount"
## [33] "InPerson.Location"
```

```
etsysales1 <- etsy_data1[, c("Sale.Date", "Quantity", "Price",
    "Buyer", "Item.Total")]
unique(etsysales1$Sale.Date)  ##Ensuring data is from treatment week
```

```
## [1] 3/26/17  3/25/17  3/24/17  3/23/17  3/22/17  3/21/17  3/20/17  03/20/17
## 8 Levels: 03/20/17 3/20/17 3/21/17 3/22/17 3/23/17 3/24/17 ... 3/26/17
```

```
table(etsysales1$Price)
```

```
##
## 5.59 6.99 7.99 9.59 9.99   25   29
```

```
##   47    4    4    4    1    1    1
```

```
aggregate(x = etsysales1[, c(2:3, 5)], by = list(etsysales1$Buyer),
    FUN = sum)  ##number of items & order value per buyer
```

```
##                              Group.1 Quantity  Price Item.Total
## 1                     aimeenicole525        1   5.59       5.59
## 2        Becca Oronoz (beccaoronoz)        1   9.59       9.59
## 3                         cherylnoel        1   5.59       5.59
## 4            Cristina (cristina81ii)        1   9.59       9.59
## 5   Daniel Ornelas (danielornelas2)        3  36.18      36.18
## 6  Daniella J. Ferraro (dferraro86)        1   5.59       5.59
## 7                         falkwagner        1  29.00      29.00
## 8   Frederic Lehman (XenoMorphious)        1   5.59       5.59
## 9                            jashm94        1   5.59       5.59
## 10       Jessie Anne (TigerShredds)        2  15.98      15.98
## 11                           lasoude        1   5.59       5.59
## 12     Laura Wagner (lauriwagner23)        2  13.98      13.98
## 13      Nathalie Palmans (npalmans)        3  19.17      19.17
## 14                    pascalealessa        1   9.99       9.99
## 15                      sabinevivoda        1   5.59       5.59
## 16               sayig sayig (sayig)       30 152.95     180.90
## 17                        yumenashi8       15  83.85      75.45
## 18                         yvonnezpt        1   5.59       5.59
```

```
etsy1_totsale_dist <- sum(etsysales1[, 2] * etsysales1[, 3])  ## total sales in terms of discounted pri
etsy1_totsale_dist
```

```
## [1] 452.95
```

```
etsy1_totsale_real <- etsy1_totsale_dist * (100/80)  ## total sales in terms of actual price
etsy1_totsale_real
```

```
## [1] 566.1875
```

### Loading the data used before treatment

```
etsy_data3 <- read.csv(file = "~/downloads/etsy_data/etsy_bef_treat1.csv")
etsy_data4 <- read.csv(file = "~/downloads/etsy_data/etsy_bef_treat2.csv")
names(etsy_data3)
```

```
##  [1] "Sale.Date"         "Item.Name"         "Buyer"
##  [4] "Quantity"          "Price"             "Coupon.Code"
##  [7] "Coupon.Details"    "Coupon.Discount"   "Order.Shipping"
## [10] "Order.Sales.Tax"   "Item.Total"        "Currency"
## [13] "Transaction.ID"    "Listing.ID"        "Date.Paid"
## [16] "Date.Shipped"      "Ship.Name"         "Ship.Address1"
## [19] "Ship.Address2"     "Ship.City"         "Ship.State"
## [22] "Ship.Zipcode"      "Ship.Country"      "Order.ID"
## [25] "Variations"        "Order.Type"        "Listings.Type"
## [28] "Payment.Type"      "InPerson.Discount" "InPerson.Location"
## [31] "VAT.Paid.by.Buyer"
```

```
names(etsy_data4)
```

```
##  [1] "Sale.Date"                     "Order.ID"
```

```
##  [3] "Buyer.User.ID"                  "Full.Name"
##  [5] "First.Name"                     "Last.Name"
##  [7] "Number.of.Items"                "Payment.Method"
##  [9] "Date.Shipped"                   "Street.1"
## [11] "Street.2"                       "Ship.City"
## [13] "Ship.State"                     "Ship.Zipcode"
## [15] "Ship.Country"                   "Currency"
## [17] "Order.Value"                    "Coupon.Code"
## [19] "Coupon.Details"                 "Shipping"
## [21] "Sales.Tax"                      "Order.Total"
## [23] "Status"                         "Card.Processing.Fees"
## [25] "Order.Net"                      "Adjusted.Order.Total"
## [27] "Adjusted.Card.Processing.Fees" "Adjusted.Net.Order.Amount"
## [29] "Buyer"                          "Order.Type"
## [31] "Payment.Type"                   "InPerson.Discount"
## [33] "InPerson.Location"
```

```
date.int <- c("03/19/17", "03/18/17", "03/17/17", "03/16/17",
    "03/15/17", "03/14/17", "03/13/17")
etsysales2 <- etsy_data3[which(etsy_data3$Sale.Date %in% date.int),
    c("Sale.Date", "Quantity", "Price", "Buyer", "Item.Total")]
unique(etsysales2$Sale.Date)
```

```
## [1] 03/19/17 03/18/17 03/17/17 03/16/17 03/15/17 03/14/17 03/13/17
## 26 Levels: 03/02/17 03/03/17 03/05/17 03/06/17 03/07/17 ... 03/28/17
```

```
table(etsysales2$Price)
```

```
##
##  6.99  9.99 11.99 19.99    29
##    23     1     4     1     3
```

```
aggregate(x = etsysales2[, c(2:3, 5)], by = list(etsysales2$Buyer),
    FUN = sum)  ##number of items & order value per buyer
```

```
##                               Group.1 Quantity Price Item.Total
## 1    Aileen Rosario (aileenrosario)        1  6.99       6.99
## 2           Anne Huang (annehuang1)        1 11.99      11.99
## 3       Caitlin Hine (caitlinahine)        1  6.99       6.99
## 4                   carolineyusuf98        1  6.99       6.99
## 5        Emiliano Colosimo (exc220)        1  6.99       6.99
## 6      Emily Dahlquist (edahlquist)        1  6.99       6.99
## 7                          grieder1        1 29.00      29.00
## 8      Jesse Fischer (jestersixty7)        1 29.00      29.00
## 9          Juan Banda S (JuanBandaS)        1 11.99      11.99
## 10     Lauren Ackert (laurenackert)        1  6.99       6.99
## 11                     lauren141822        1  6.99       6.99
## 12                           luschn        2 18.98      18.98
## 13                     meirekimberly        1  6.99       6.99
## 14 Michael LaFleur (lafleurmichael)        1 29.00      29.00
## 15                         micheleqj        1 11.99      11.99
## 16                   nathaneman1992        1  6.99       6.99
## 17                          nephelys        2 13.98      13.98
## 18   Sharis Arikian (sharisarikian)        2 13.98      11.18
## 19           Telos Snow (telossnow)        9 78.91      78.91
## 20                  veronicaqg2011        1  6.99       6.99
```

```
## 21                        vsavage1980        1  6.99       6.99
```

```r
etsy2_totsale_real <- sum(etsysales2[, 2] * etsysales2[, 3])  ## total sales in terms of actual price
etsy2_totsale_real
```

```
## [1] 325.71
```

**Observations:** 1. Sales of items of lower price, in terms of the number has increased noticibly.

2. The total sales during the control week was 325.71 and the total sales during the treatment week was 566.19. This is calculated in terms of actual prices.

3. The number of sales of the items that were not discounted remains the same (These are the set of already discounted items).

### Analysis of another control week

```r
date.int <- c("03/12/17", "03/11/17", "03/10/17", "03/09/17",
    "03/08/17", "03/07/17", "03/06/17")
etsysales3 <- etsy_data3[which(etsy_data3$Sale.Date %in% date.int),
    c("Sale.Date", "Quantity", "Price", "Buyer", "Item.Total")]
table(etsysales3$Price)
```

```
##
##  6.99  9.99 11.99    25    29
##    27     3     2     1     1
```

```r
unique(etsysales3$Sale.Date)
```

```
## [1] 03/12/17 03/11/17 03/10/17 03/09/17 03/08/17 03/07/17 03/06/17
## 26 Levels: 03/02/17 03/03/17 03/05/17 03/06/17 03/07/17 ... 03/28/17
```

```r
aggregate(x = etsysales3[, c(2:3, 5)], by = list(etsysales3$Buyer),
    FUN = sum)  ##number of items & order value per buyer
```

```
##                             Group.1 Quantity Price Item.Total
## 1                            arhaaz        1  6.99       6.99
## 2                          bkirk735        1 11.99      11.99
## 3                       BustyMelons        2 13.98      13.98
## 4                     eleonoravilli        1  6.99       6.99
## 5               Esra Qambar (esra37)        1  9.99       9.99
## 6                            fara7h        6 30.96      47.94
## 7                         mebwoyner        4 27.96      27.96
## 8                          nageek63        1 25.00      25.00
## 9                 pascalbaumgartner1        4 32.96      32.96
## 10 Paula Smathers (paulasmathers1)        1  6.99       6.99
## 11    valarie king (valarieking88)        1  6.99       6.99
## 12                          vtapken        2 29.00      58.00
## 13                       yumenashi8       13 86.88      96.87
```

```r
etsy3_totsale_real <- sum(etsysales3[, 2] * etsysales3[, 3])  ## total sales in terms of actual price
etsy3_totsale_real
```

```
## [1] 352.65
```

**Observations:** The trend in the control weeks remains the same.

### Analysis of age data over control and treatment weeks

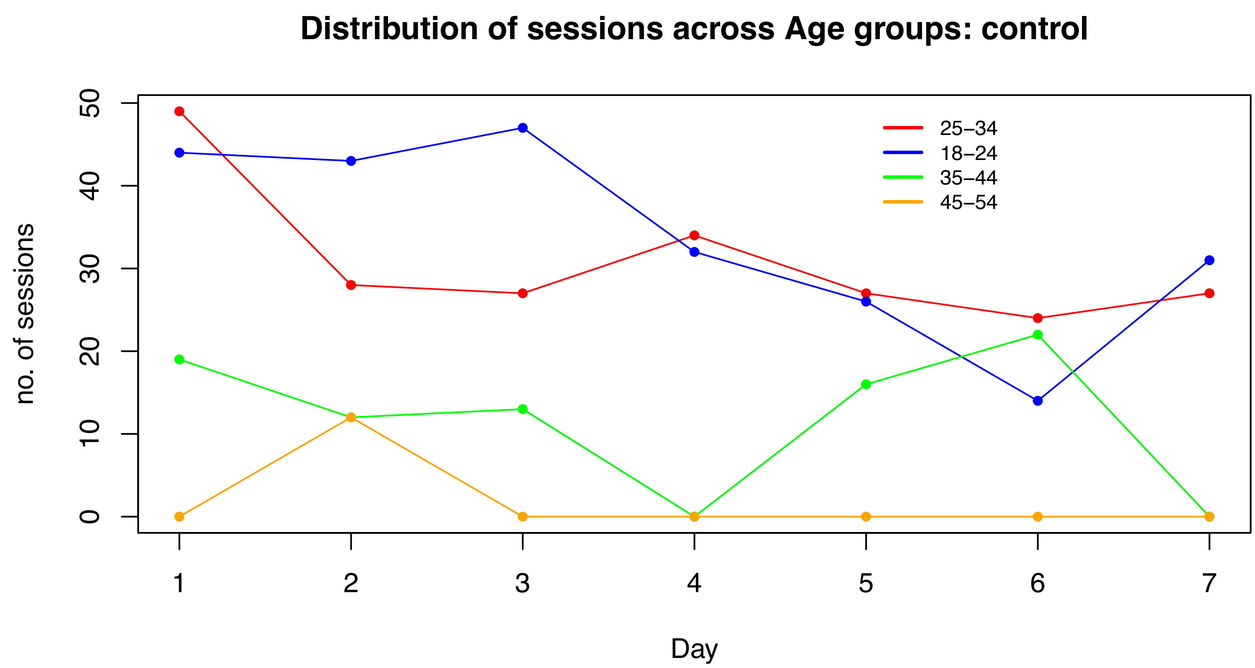
```
data.age <- read.csv(file = "~/downloads/etsy_data/GA_age.csv")
head(data.age)
```

```
##     Age  X0 X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12 X13 X14
## 1 25-34  49 28 27 34 27 24 27 40 36 27  49  36  34  32  35
## 2 18-24  44 43 47 32 26 14 31 39 23 32  34  17  27  38  27
## 3 35-44  19 12 13  0 16 22  0 11 14 13  17  18  14  11  13
## 4 45-54   0 12  0  0  0  0  0  0  0  0  14  15   0  11   0
## 5 Total 112 95 87 66 69 60 58 90 73 72 114  86  75  92  75
```

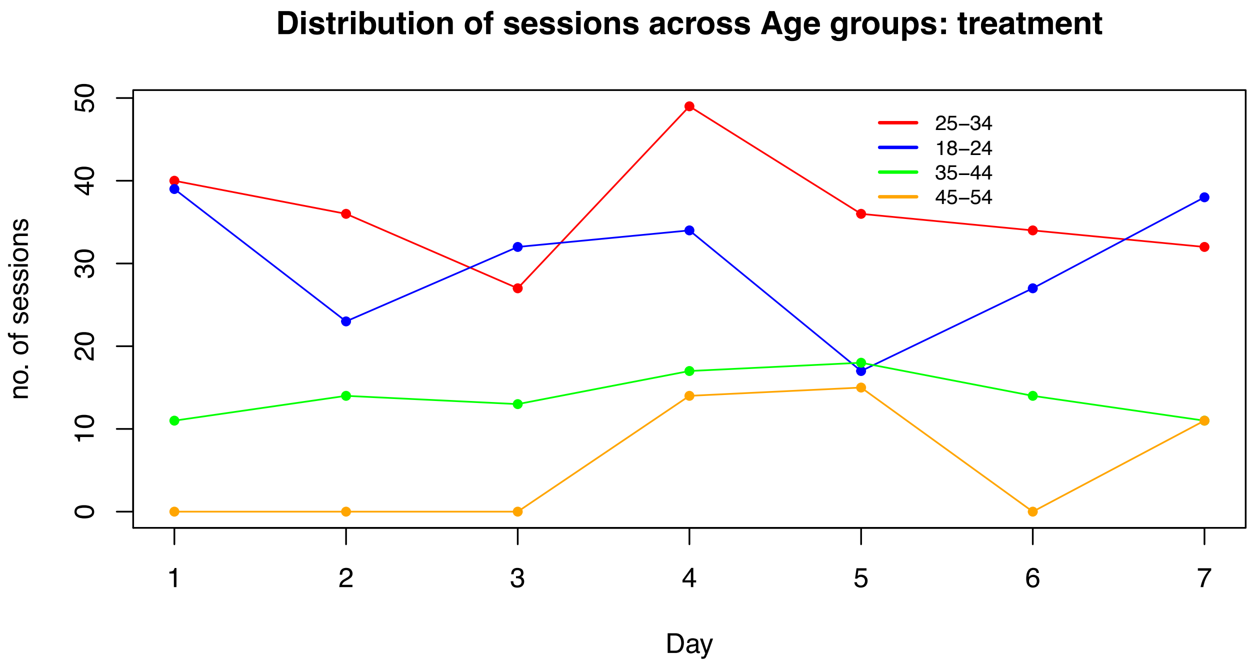
```
rownames(data.age) <- data.age[, 1]
data.age <- data.age[, -c(1)]
head(data.age)
```

```
##        X0 X1 X2 X3 X4 X5 X6 X7 X8 X9 X10 X11 X12 X13 X14
## 25-34  49 28 27 34 27 24 27 40 36 27  49  36  34  32  35
## 18-24  44 43 47 32 26 14 31 39 23 32  34  17  27  38  27
## 35-44  19 12 13  0 16 22  0 11 14 13  17  18  14  11  13
## 45-54   0 12  0  0  0  0  0  0  0  0  14  15   0  11   0
## Total 112 95 87 66 69 60 58 90 73 72 114  86  75  92  75
```

```
data.age.control <- data.age[, c(1:7)]
data.age.treat <- data.age[, c(8:14)]
plot(c(1:7), data.age.control[1, ], type = "l", main = "Distribution of sessions across Age groups: con
    xlab = "Day", ylab = "no. of sessions", ylim = c(min(data.age.control[1:4,
        ]), max(data.age.control[1:4, ])), col = "red")
points(c(1:7), data.age.control[1, ], pch = 20, col = "red")
lines(c(1:7), data.age.control[2, ], type = "l", col = "blue")
points(c(1:7), data.age.control[2, ], pch = 20, col = "blue")
lines(c(1:7), data.age.control[3, ], type = "l", col = "green")
points(c(1:7), data.age.control[3, ], pch = 20, col = "green")
lines(c(1:7), data.age.control[4, ], type = "l", col = "orange")
points(c(1:7), data.age.control[4, ], pch = 20, col = "orange")
legend(x = 5, y = 50, legend = c("25-34", "18-24", "35-44", "45-54"),
    col = c("red", "blue", "green", "orange"), lwd = 2, bty = "n",
    cex = 0.75)
```




```
plot(c(1:7), data.age.treat[1, ], type = "l", main = "Distribution of sessions across Age groups: treatm
    xlab = "Day", ylab = "no. of sessions", ylim = c(min(data.age.treat[1:4,
        ]), max(data.age.treat[1:4, ])), col = "red")
points(c(1:7), data.age.treat[1, ], pch = 20, col = "red")
lines(c(1:7), data.age.treat[2, ], type = "l", col = "blue")
points(c(1:7), data.age.treat[2, ], pch = 20, col = "blue")
lines(c(1:7), data.age.treat[3, ], type = "l", col = "green")
points(c(1:7), data.age.treat[3, ], pch = 20, col = "green")
lines(c(1:7), data.age.treat[4, ], type = "l", col = "orange")
points(c(1:7), data.age.treat[4, ], pch = 20, col = "orange")
legend(x = 5, y = 50, legend = c("25-34", "18-24", "35-44", "45-54"),
    col = c("red", "blue", "green", "orange"), lwd = 2, bty = "n",
    cex = 0.75)
```




We check the agewise distributions in the treatment and control groups by synthesizing a dataset that reflects the above informattion imported from Google Analytics.

### Creating Two vectors representing age distribuions

We create two vectors having the age distributions observed during the control and treatment weeks and conducting Wilcox.test for the different age groups. We use Wilcoxon test as it is independent of degree of freedom and underlying distribution.

```
data.age <- rbind(data.age[2, ], data.age[1, ], data.age[3:5,
    ])
data.age.control <- data.age[, c(1:7)]
data.age.treat <- data.age[, c(8:14)]
wilcox.test(as.numeric(data.age.control[1, ]), as.numeric(data.age.treat[1,
    ]))  ## Distribution among visitors in age group 18-24 yrs
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  as.numeric(data.age.control[1, ]) and as.numeric(data.age.treat[1, ])
## W = 29.5, p-value = 0.5649
## alternative hypothesis: true location shift is not equal to 0
```

```
wilcox.test(as.numeric(data.age.control[2, ]), as.numeric(data.age.treat[2,
    ]))  ## Distribution among visitors in age group 25-34 yrs
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  as.numeric(data.age.control[2, ]) and as.numeric(data.age.treat[2, ])
## W = 11.5, p-value = 0.1051
## alternative hypothesis: true location shift is not equal to 0
```

```r
wilcox.test(as.numeric(data.age.control[3, ]), as.numeric(data.age.treat[3,
    ]))  ## Distribution among visitors in age group 35-44 yrs
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  as.numeric(data.age.control[3, ]) and as.numeric(data.age.treat[3, ])
## W = 23.5, p-value = 0.9488
## alternative hypothesis: true location shift is not equal to 0
```

```r
wilcox.test(as.numeric(data.age.control[4, ]), as.numeric(data.age.treat[4,
    ]))  ## Distribution among visitors in age group 45-54 yrs
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  as.numeric(data.age.control[4, ]) and as.numeric(data.age.treat[4, ])
## W = 17, p-value = 0.2626
## alternative hypothesis: true location shift is not equal to 0
```

```r
wilcox.test(as.numeric(data.age.control[5, ]), as.numeric(data.age.treat[5,
    ]))  ## Distribution among visitors in all age groups
```

```
##
##  Wilcoxon rank sum test
##
## data:  as.numeric(data.age.control[5, ]) and as.numeric(data.age.treat[5, ])
## W = 16, p-value = 0.3176
## alternative hypothesis: true location shift is not equal to 0
```

Thus,the above tests conducted on each age group reveals that the distribution of the visitors in the age-groups remains identical during the control and treatment weeks.

### Analysis of gender data over control and treatment weeks

```r
data.sex <- read.csv(file = "~/downloads/etsy_data/GA_gender.csv")
head(data.sex)
```

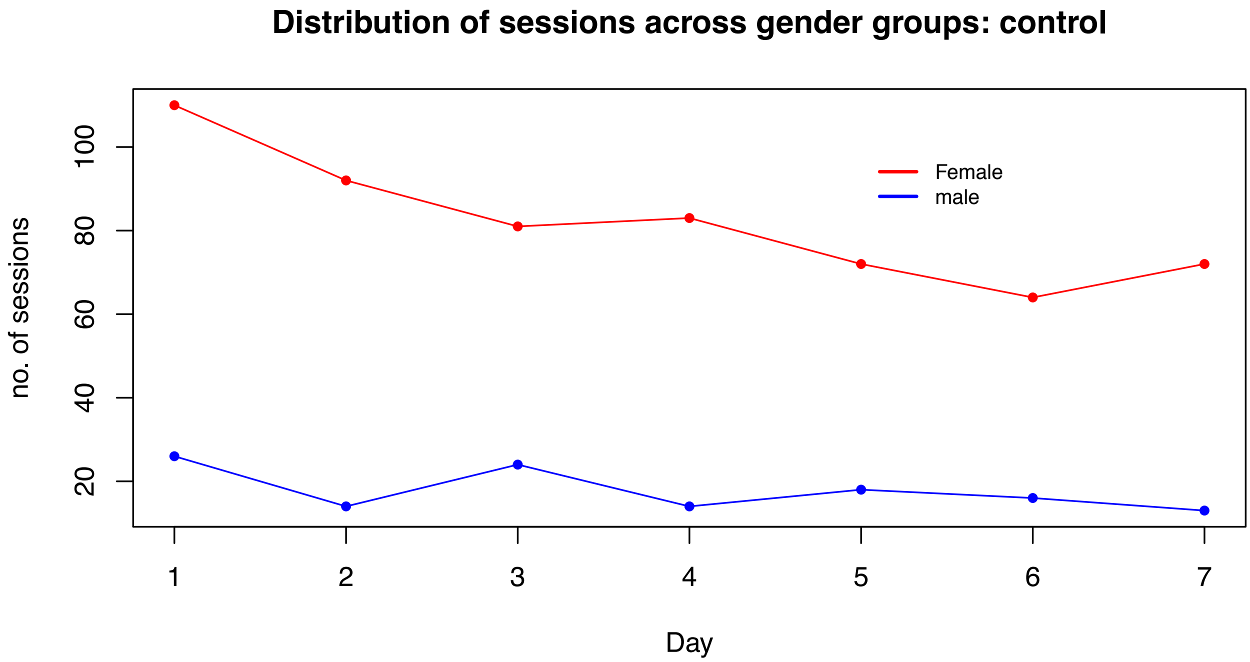
```
##   gender  X0  X1  X2 X3 X4 X5 X6  X7 X8 X9 X10 X11 X12 X13 X14
## 1 female 110  92  81 83 72 64 72  96 65 73 109  86  87  92  70
## 2   male  26  14  24 14 18 16 13  11 20 13  20  13  17   0  22
## 3  total 136 106 105 97 90 80 85 107 85 86 129  99 104  92  92
```

```r
rownames(data.sex) <- data.sex[, 1]
data.sex <- data.sex[, -c(1)]
head(data.sex)
```

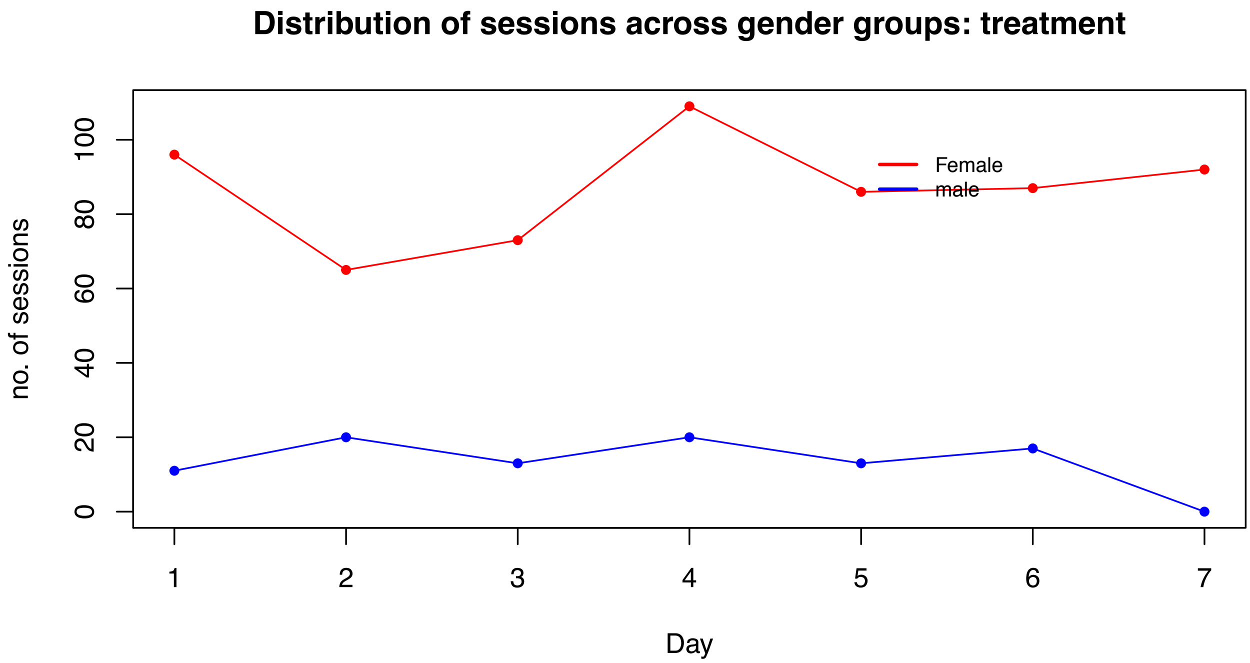
```
##         X0  X1  X2 X3 X4 X5 X6  X7 X8 X9 X10 X11 X12 X13 X14
## female 110  92  81 83 72 64 72  96 65 73 109  86  87  92  70
## male    26  14  24 14 18 16 13  11 20 13  20  13  17   0  22
## total  136 106 105 97 90 80 85 107 85 86 129  99 104  92  92
```

```r
data.sex.control <- data.sex[, c(1:7)]
data.sex.treat <- data.sex[, c(8:14)]
plot(c(1:7), data.sex.control[1, ], type = "l", main = "Distribution of sessions across gender groups: 
    xlab = "Day", ylab = "no. of sessions", ylim = c(min(data.sex.control[1:2,
```

```
        ]), max(data.sex.control[1:2, ])), col = "red")
points(c(1:7), data.sex.control[1, ], pch = 20, col = "red")
lines(c(1:7), data.sex.control[2, ], type = "l", col = "blue")
points(c(1:7), data.sex.control[2, ], pch = 20, col = "blue")
legend(x = 5, y = 100, legend = c("Female", "male"), col = c("red",
    "blue"), lwd = 2, bty = "n", cex = 0.75)
```




```
plot(c(1:7), data.sex.treat[1, ], type = "l", main = "Distribution of sessions across gender groups: tre
    xlab = "Day", ylab = "no. of sessions", ylim = c(min(data.sex.treat[1:2,
        ]), max(data.sex.treat[1:2, ])), col = "red")
points(c(1:7), data.sex.treat[1, ], pch = 20, col = "red")
lines(c(1:7), data.sex.treat[2, ], type = "l", col = "blue")
points(c(1:7), data.sex.treat[2, ], pch = 20, col = "blue")
legend(x = 5, y = 100, legend = c("Female", "male"), col = c("red",
    "blue"), lwd = 2, bty = "n", cex = 0.75)
```




**Creating Two vectors representing gender distribuions**

We create two vectors having the gender distributions observed during the control and treatment weeks and conducting Wilcox.test for the different gender groups. We use Wilcoxon test as it is independent of degree of freedom and underlying distribution.

```
data.sex.control <- data.sex[, c(1:7)]
data.sex.treat <- data.sex[, c(8:14)]
wilcox.test(as.numeric(data.sex.control[1, ]), as.numeric(data.sex.treat[1,
    ]))  ## Distribution among females
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  as.numeric(data.sex.control[1, ]) and as.numeric(data.sex.treat[1, ])
## W = 17.5, p-value = 0.4052
## alternative hypothesis: true location shift is not equal to 0
```

```
wilcox.test(as.numeric(data.sex.control[2, ]), as.numeric(data.sex.treat[2,
    ]))  ## Distribution among males
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  as.numeric(data.sex.control[2, ]) and as.numeric(data.sex.treat[2, ])
## W = 34, p-value = 0.247
## alternative hypothesis: true location shift is not equal to 0
```

```
wilcox.test(as.numeric(data.sex.control[3, ]), as.numeric(data.sex.treat[3,
    ]))  ## Distribution among all visitors
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
```

```
## data:  as.numeric(data.sex.control[3, ]) and as.numeric(data.sex.treat[3, ])
## W = 22.5, p-value = 0.8478
## alternative hypothesis: true location shift is not equal to 0
```

Thus,the above tests conducted on each gender group reveals that the distribution of females and males remains identical during the control and treatment weeks.

### Analysis of data country wise for treatment & control weeks

Since we are having 67 different countries/ categories we use Matched pair Wilconxon test to see if the distribution is identical.

```
data.country.control <- read.csv(file = "~/downloads/etsy_data/GA_country_control.csv")
data.country.treat <- read.csv(file = "~/downloads/etsy_data/GA_country_treat.csv")
country <- intersect(data.country.control$Country, data.country.treat$Country)
data1 <- subset(data.country.control, data.country.control$Country %in%
    country)
data2 <- subset(data.country.treat, data.country.treat$Country %in%
    country)
data1 <- data1[order(data1$Country), 1:2]
data2 <- data2[order(data2$Country), 1:2]
data.country <- cbind(data1, data2$Sessions)
names(data.country) <- c("Country", "Sess.control", "Sess.treat")
head(data.country)
```

```
##      Country Sess.control Sess.treat
## 27 Argentina            9          5
## 13 Australia           17         17
## 12   Austria           20         11
## 30   Belgium            8          9
## 51   Bolivia            2          3
## 5     Brazil           43         59
```

```
wilcox.test(data.country$Sess.control, data.country$Sess.treat,
    paired = T)
```

```
##
##  Wilcoxon signed rank test with continuity correction
##
## data:  data.country$Sess.control and data.country$Sess.treat
## V = 869, p-value = 0.9066
## alternative hypothesis: true location shift is not equal to 0
```

We observe that the distributions across countries are identical for the two weeks.

### Analysis of data langauge wise for treatment & control weeks

Since we are having 58 different countries/ categories we use Matched pair Wilcoxon test to see if the distribution is identical.

```
data.lang.control <- read.csv(file = "~/downloads/etsy_data/GA_lang_control.csv")
data.lang.treat <- read.csv(file = "~/downloads/etsy_data/GA_lang_treat.csv")
lang <- intersect(data.lang.control$Language, data.lang.treat$Language)
data1 <- subset(data.lang.control, data.lang.control$Language %in%
    lang)
```

```r
data2 <- subset(data.lang.treat, data.lang.treat$Language %in%
    lang)
data1 <- data1[order(data1$Language), 1:2]
data2 <- data2[order(data2$Language), 1:2]
data.lang <- cbind(data1, data2$Sessions)
names(data.lang) <- c("Langauge", "Sess.control", "Sess.treat")
head(data.lang)
```

```
##      Langauge Sess.control Sess.treat
## 51 (not set)            1          1
## 24        ar            5          5
## 41        bg            2          4
## 52        cs            1          4
## 35     cs-cz            3          2
## 53     da-dk            1          4
```

```r
wilcox.test(data.lang$Sess.control, data.lang$Sess.treat, paired = T)
```

```
##
##  Wilcoxon signed rank test with continuity correction
##
## data:  data.lang$Sess.control and data.lang$Sess.treat
## V = 602.5, p-value = 0.9242
## alternative hypothesis: true location shift is not equal to 0
```

We observe that the distributions across langauges are identical for the two weeks.