

Assignment 10

Samira Zarandioon

April 1, 2017

1. First we download the data:

```
library(downloader)
download("http://stat.columbia.edu/~rachel/datasets/nyt1.csv", destfile="nyt1.csv")
```

2. Add new variable ageGroup for the following groups: < 18, 18-24, 25-34, 35-44, 45-54, 55-64 and 65+

```
nyt_dataframe <- read.csv("nyt1.csv", stringsAsFactors=FALSE, header=TRUE)
nyt_dataframe$ageGroup[nyt_dataframe$Age<18] <- "<18"
nyt_dataframe$ageGroup[nyt_dataframe$Age>=18 & nyt_dataframe$Age<=24 ] <- "18-24"
nyt_dataframe$ageGroup[nyt_dataframe$Age>=25 & nyt_dataframe$Age<=34 ] <- "25-34"
nyt_dataframe$ageGroup[nyt_dataframe$Age>=35 & nyt_dataframe$Age<=44 ] <- "35-44"
nyt_dataframe$ageGroup[nyt_dataframe$Age>=45 & nyt_dataframe$Age<=54 ] <- "45-54"
nyt_dataframe$ageGroup[nyt_dataframe$Age>=55 & nyt_dataframe$Age<=64 ] <- "55-64"
nyt_dataframe$ageGroup[nyt_dataframe$Age>=65] <- "65+"
```

3. Create a sub data set called ImpSub that contains rows with Impressions>0:

```
r ImpSub<-nyt_dataframe[nyt_dataframe$Impressions>0,]
```

```
ImpSub$CTR = ImpSub$Clicks / ImpSub$Impressions
```

```
#install.packages("ggplot2")
library(ggplot2)
ggplot(ImpSub, aes(x=Impressions, fill=ageGroup)) +
  geom_histogram(binwidth=1, position="dodge")
```

```
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '18-24' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '18-24' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '18-24' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '35-44' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '35-44' in 'mbcsToSbcs': dot substituted for <80>
```

```

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '35-44' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '45-54' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '45-54' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '45-54' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '55-64' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '55-64' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '55-64' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-24' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-24' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-24' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for <e2>

```

[illegible]

```
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for
## <e2>

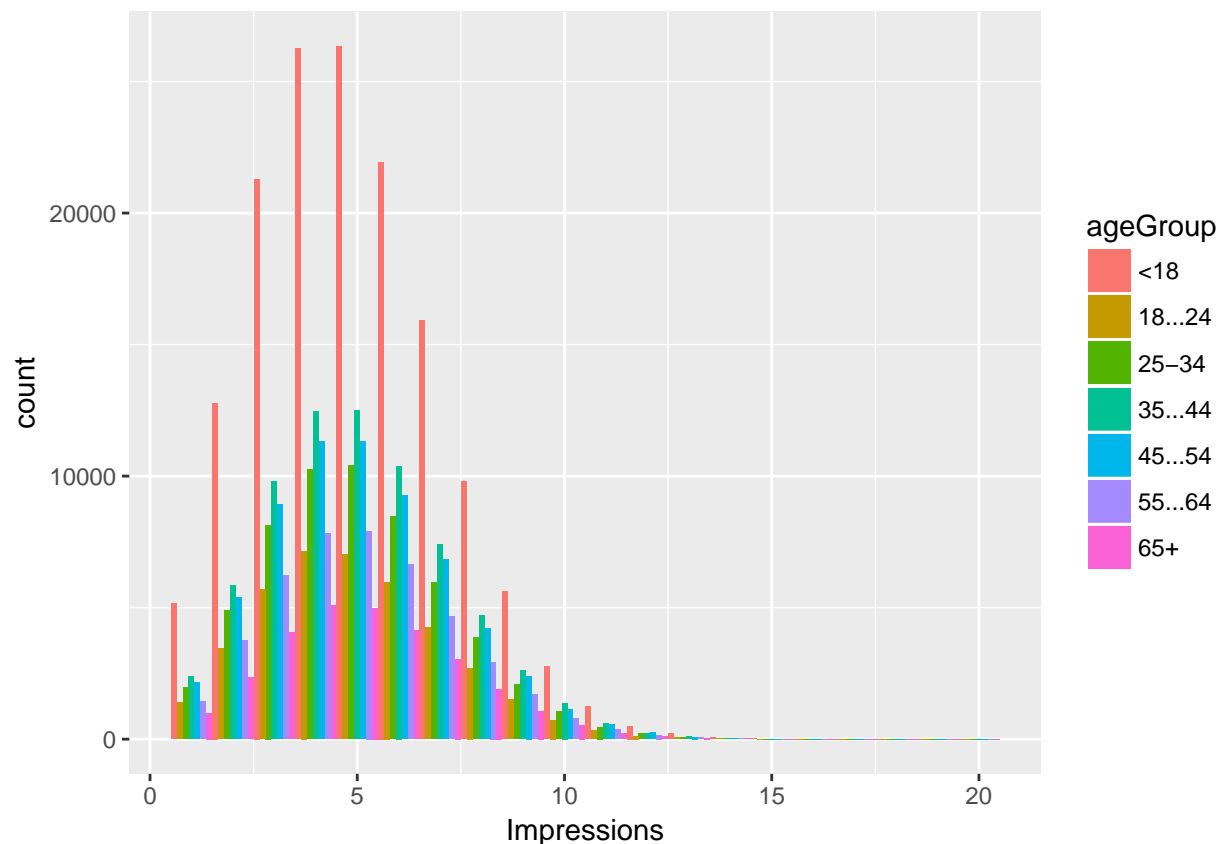
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for
## <93>
```



```
#install.packages("ggplot2")
library(ggplot2)
```

```
ggplot(ImpSub, aes(x=Impressions, fill=ageGroup)) +
  geom_histogram(binwidth=1, alpha=.5, position="identity")
```

```
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '18-24' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '18-24' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '18-24' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '35-44' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '35-44' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '35-44' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '45-54' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '45-54' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '45-54' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '55-64' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '55-64' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '55-64' in 'mbcsToSbcs': dot substituted for <93>
```

[illegible]

```

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '18-24' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '18-24' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '18-24' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for
## <80>

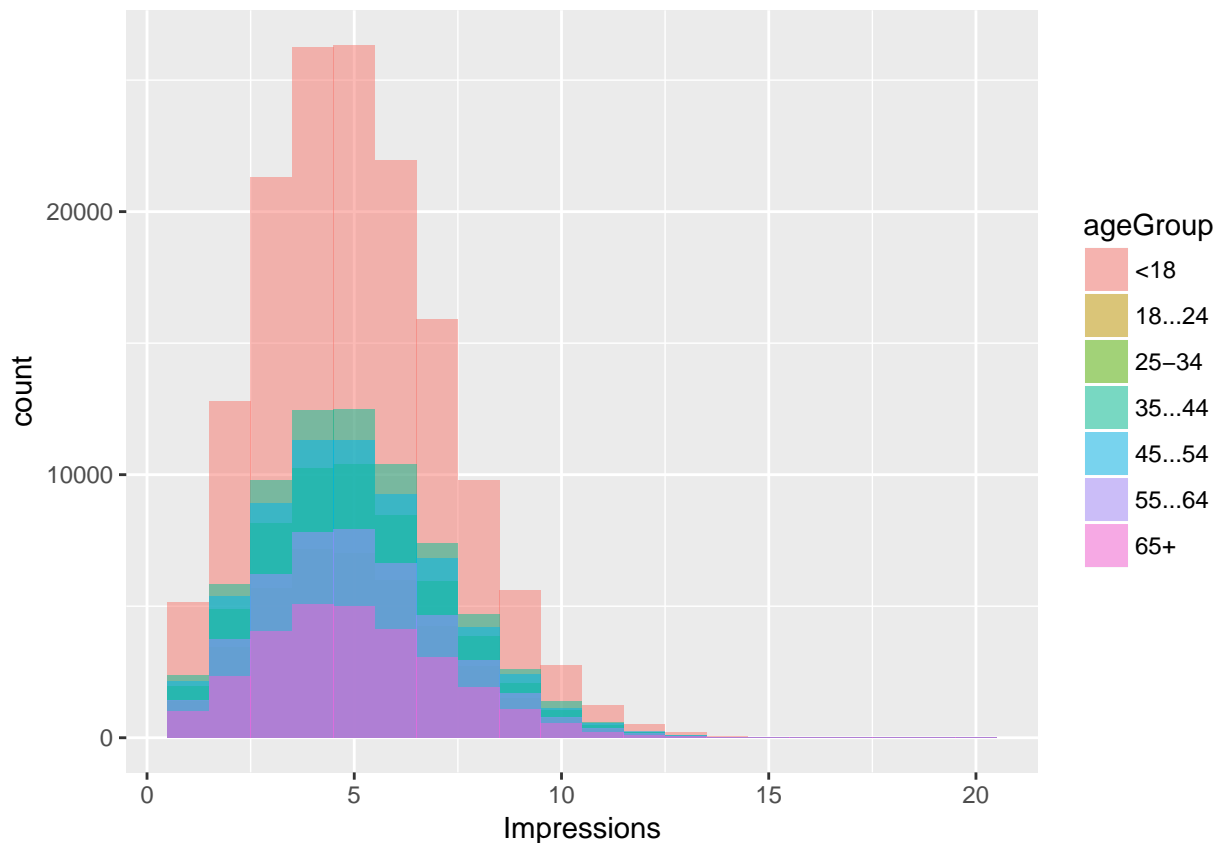
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for
## <93>

```



```
#install.packages("ggplot2")
library(ggplot2)
ggplot(ImpSub, aes(x=CTR, fill=ageGroup)) +
  geom_histogram(binwidth=.1, position="dodge")
```

```
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '18-24' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '18-24' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '18-24' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '35-44' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '35-44' in 'mbcsToSbcs': dot substituted for <80>
```



```

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '35-44' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '45-54' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '45-54' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '45-54' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '55-64' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '55-64' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '55-64' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-24' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-24' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '18-24' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for <e2>

```

[illegible]

```
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for
## <e2>

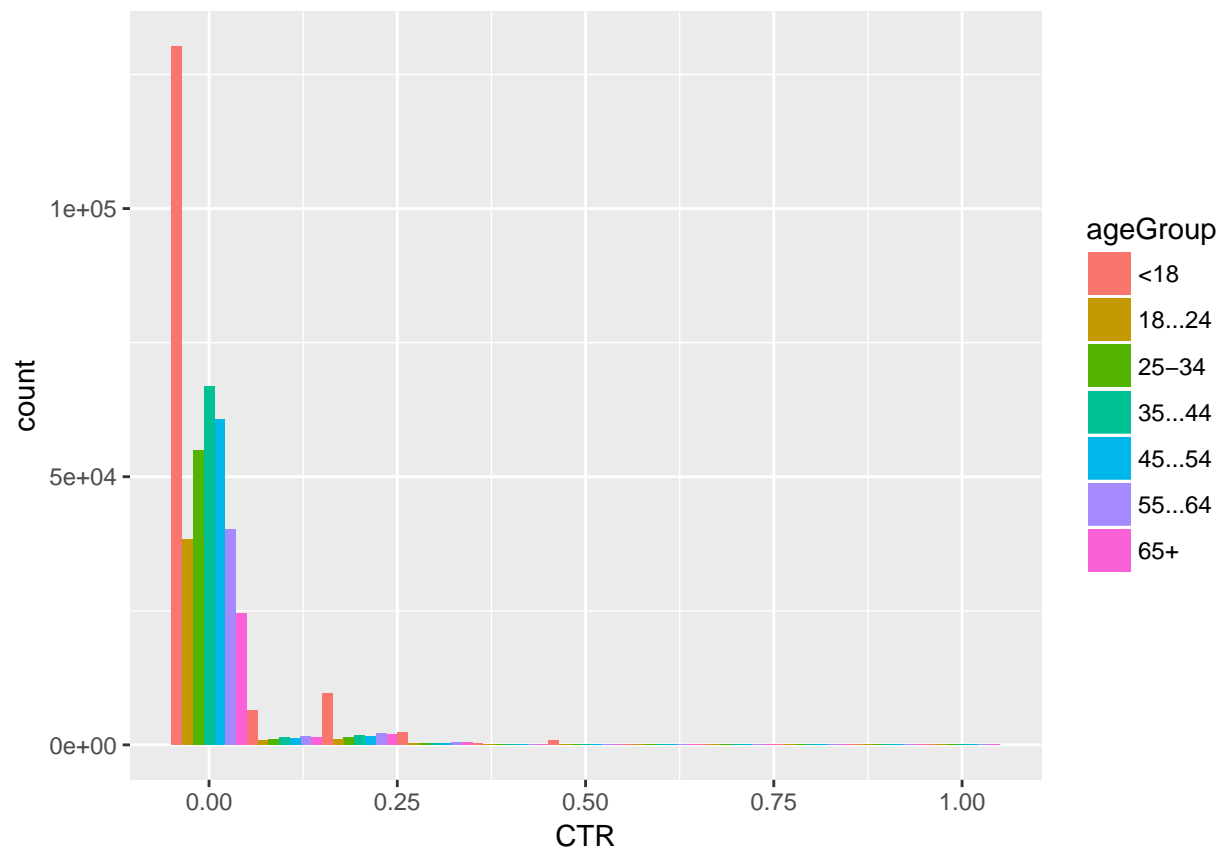
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for
## <93>
```



```
#install.packages("ggplot2")
library(ggplot2)
```

```
ggplot(ImpSub, aes(x=CTR, fill=ageGroup)) +
  geom_histogram(binwidth=.1, alpha=.5, position="identity")
```

```
## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '18-24' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '18-24' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '18-24' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '35-44' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '35-44' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '35-44' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '45-54' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '45-54' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '45-54' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x1e

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): font
## metrics unknown for character 0x80

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '55-64' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '55-64' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_stringMetric, as.graphicsAnnot(x$label)): conversion
## failure on '55-64' in 'mbcsToSbcs': dot substituted for <93>
```

[illegible]

```

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for <e2>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for <80>

## Warning in grid.Call(L_textBounds, as.graphicsAnnot(x$label), x$x, x$y, :
## conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for <93>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '18-24' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '18-24' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '18-24' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '35-44' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for
## <80>

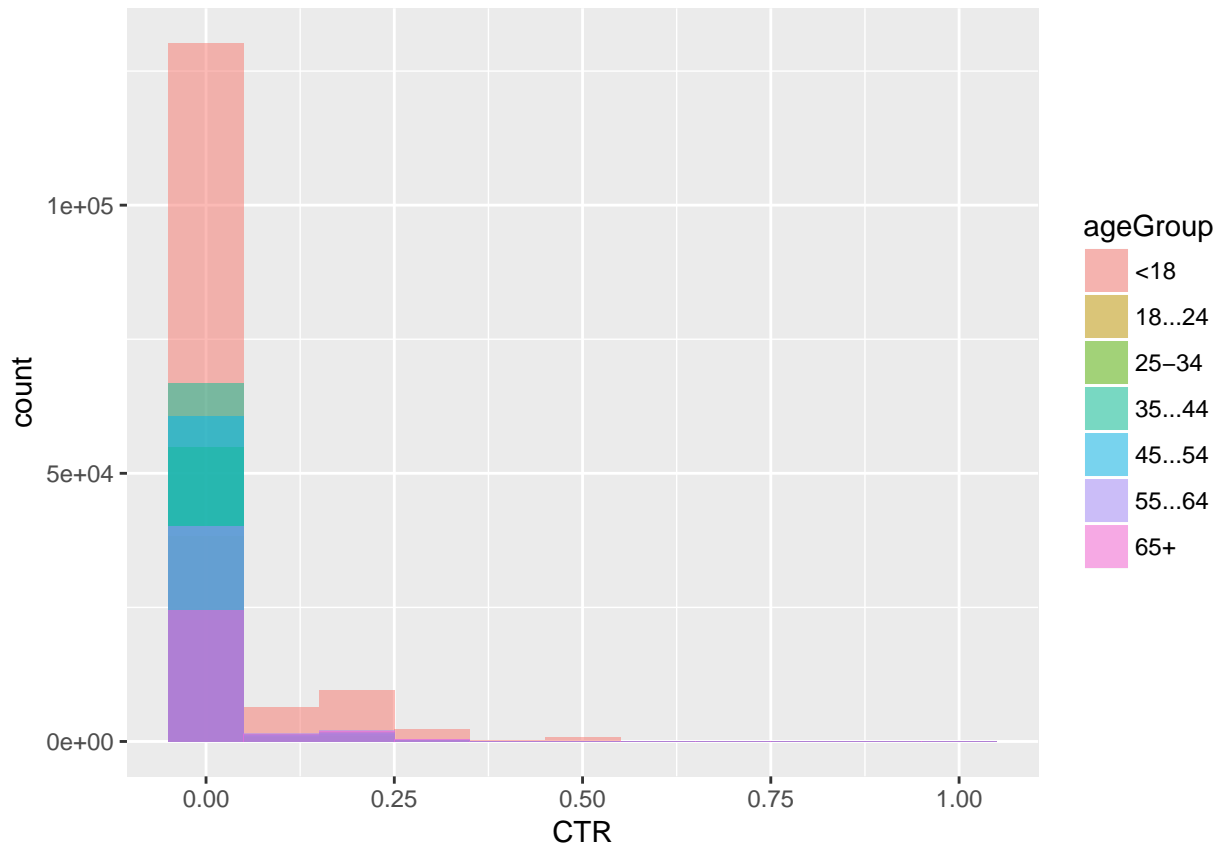
## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '45-54' in 'mbcsToSbcs': dot substituted for
## <93>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for
## <e2>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for
## <80>

## Warning in grid.Call.graphics(L_text, as.graphicsAnnot(x$label), x$x, x
## $y, : conversion failure on '55-64' in 'mbcsToSbcs': dot substituted for
## <93>

```



CTR < 0.2, 0.2 <= CTR < 0.4, 0.4 <= CTR < 0.6, 0.6 <= CTR < 0.8, CTR > 0.8

```
ImpSub$CTRGroup[ImpSub$CTR<.2] <- "CTR<0.2"
ImpSub$CTRGroup[ImpSub$CTR>=.2 & ImpSub$CTR<.4] <- ".2<=CTR<0.4"
ImpSub$CTRGroup[ImpSub$CTR>=.4 & ImpSub$CTR<.6] <- ".4<=CTR<0.6"
ImpSub$CTRGroup[ImpSub$CTR>=.6 & ImpSub$CTR<.8] <- ".6<=CTR<0.8"
ImpSub$CTRGroup[ImpSub$CTR>=.8] <- "CTR>=0.8"
```

```
cat("Number of Male: ",sum(ImpSub$Gender),"\n")
```

```
## Number of Male: 167146
```

```
cat("Number of Impressions: ",sum(ImpSub$Impressions),"\n")
```

```
## Number of Impressions: 2295559
```

```
cat("Number of Clicks: ",sum(ImpSub$Clicks),"\n")
```

```
## Number of Clicks: 42449
```

```
cat("Number of Signed In: ",sum(ImpSub$Signed_In))
```

```
## Number of Signed In: 319198
```

Age, Impressions, Clicks, CTR

```
cat("Average of Age: ",mean(ImpSub$Age),"\n")
```

```
## Average of Age: 29.48401
```

```
cat("Average of Impressions: ",mean(ImpSub$Impressions),"\n")
```

```
## Average of Impressions: 5.04103
```

```
cat("Average of Clicks: ",mean(ImpSub$Clicks),"\n")
```

```
## Average of Clicks: 0.09321768
```

```
cat("Average of CTR: ",mean(ImpSub$CTR),"\n")
```

```
## Average of CTR: 0.01847053
```

```
cat("Percentage of Males: ", 100 * sum(ImpSub$Gender)/nrow(ImpSub),"\n")
```

```
## Percentage of Males: 36.70513
```

```
cat("Percentage of Signed In: ", 100 * sum(ImpSub$Signed_In)/nrow(ImpSub),"\n")
```

```
## Percentage of Signed In: 70.09564
```

```
aggregate(ImpSub[,c("Age", "Impressions", "Clicks", "CTR")], list(ImpSub$ageGroup), mean)
```

##	Group.1	Age	Impressions	Clicks	CTR
## 1	<18	1.397948	5.034534	0.14372052	0.028530995
## 2	18-24	20.832187	5.038235	0.05362799	0.010733764
## 3	25-34	29.503573	5.026055	0.05081227	0.010146329
## 4	35-44	39.493195	5.054749	0.05202148	0.010286330
## 5	45-54	49.492928	5.045172	0.05062260	0.009957612
## 6	55-64	59.497459	5.053484	0.10246952	0.020306816
## 7	65+	72.988409	5.046925	0.15233226	0.029802702

```
aggregate(list(male_perc=ImpSub$Gender, signed_in_perc=ImpSub$Signed_In), by=list(age_group=ImpSub$ageGr
```

##	age_group	male_perc	signed_in_perc
## 1	<18	6.270471	9.160224
## 2	18-24	53.395367	100.000000
## 3	25-34	53.199772	100.000000
## 4	35-44	53.170725	100.000000
## 5	45-54	52.921920	100.000000
## 6	55-64	53.596329	100.000000
## 7	65+	36.325116	100.000000

```
aggregate(list(count=ImpSub$ageGroup), by=list(age_group=ImpSub$ageGroup, ctr_group=ImpSub$CTRGroup), F
```

##	age_group	ctr_group	count
## 1	<18	.2<=CTR<0.4	8724
## 2	18-24	.2<=CTR<0.4	858
## 3	25-34	.2<=CTR<0.4	1203
## 4	35-44	.2<=CTR<0.4	1489
## 5	45-54	.2<=CTR<0.4	1331
## 6	55-64	.2<=CTR<0.4	1793
## 7	65+	.2<=CTR<0.4	1742
## 8	<18	.4<=CTR<0.6	1143
## 9	18-24	.4<=CTR<0.6	99
## 10	25-34	.4<=CTR<0.6	110
## 11	35-44	.4<=CTR<0.6	124
## 12	45-54	.4<=CTR<0.6	125
## 13	55-64	.4<=CTR<0.6	201
## 14	65+	.4<=CTR<0.6	195
## 15	<18	.6<=CTR<0.8	83
## 16	18-24	.6<=CTR<0.8	3
## 17	25-34	.6<=CTR<0.8	7


```
## 18      35-44 .6<=CTR<0.8      4
## 19      55-64 .6<=CTR<0.8      8
## 20        65+ .6<=CTR<0.8     11
## 21        <18   CTR<0.2 139801
## 22      18-24   CTR<0.2  39431
## 23      25-34   CTR<0.2  56462
## 24      35-44   CTR<0.2  68749
## 25      45-54   CTR<0.2  62369
## 26      55-64   CTR<0.2  42424
## 27        65+   CTR<0.2  26587
## 28        <18   CTR>=0.8    158
## 29      18-24   CTR>=0.8     17
## 30      25-34   CTR>=0.8     19
## 31      35-44   CTR>=0.8     28
## 32      45-54   CTR>=0.8     20
## 33      55-64   CTR>=0.8     36
## 34        65+   CTR>=0.8     21
```

```
#install.packages("ggplot2")
library(ggplot2)
ggplot(ImpSub, aes(x=Impressions, fill=CTRGroup)) +
  geom_histogram(binwidth=1, position="dodge")
```

