ggplotly

Subramani.M

20 May 2018

```
# Problem Stmt : Plot the relation between Rape and kidnapping of women and
girls for
# the state where the number of murder cases is more than 70 from 2001 to
2012

#install.packages("plotly")
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.4.3
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(plotly)
```

```
## Warning: package 'plotly' was built under R version 3.4.4
```

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

```
## Warning: package 'ggplot2' was built under R version 3.4.3
```

```
##
## Attaching package: 'plotly'
```

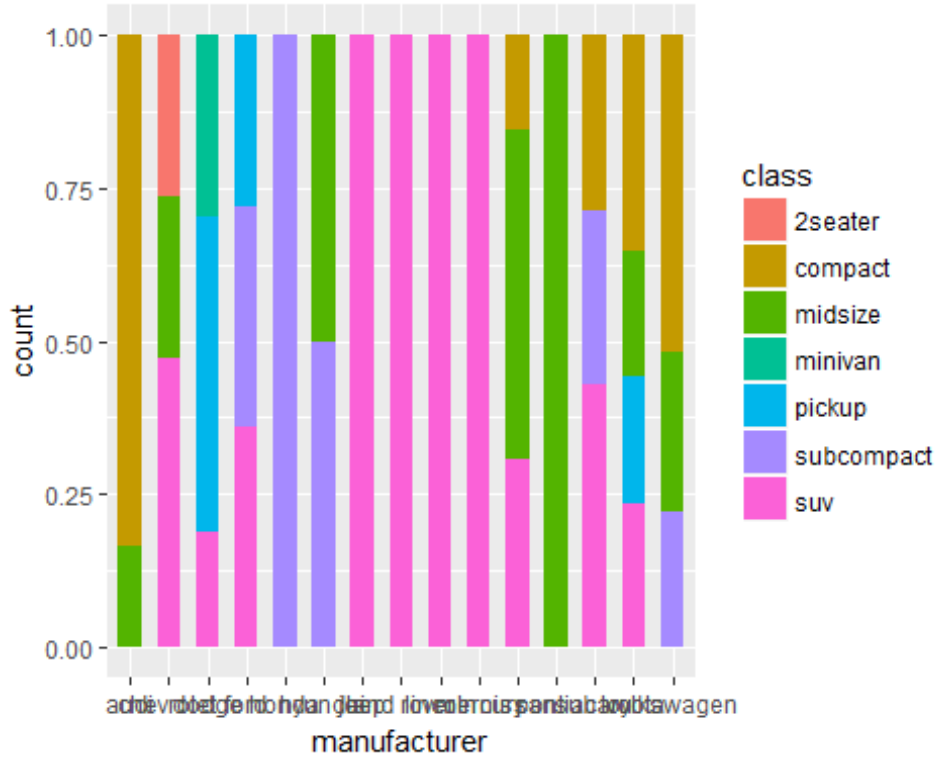
```
## The following object is masked from 'package:ggplot2':
##
##     last_plot
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
## The following object is masked from 'package:graphics':
##
##     layout
```

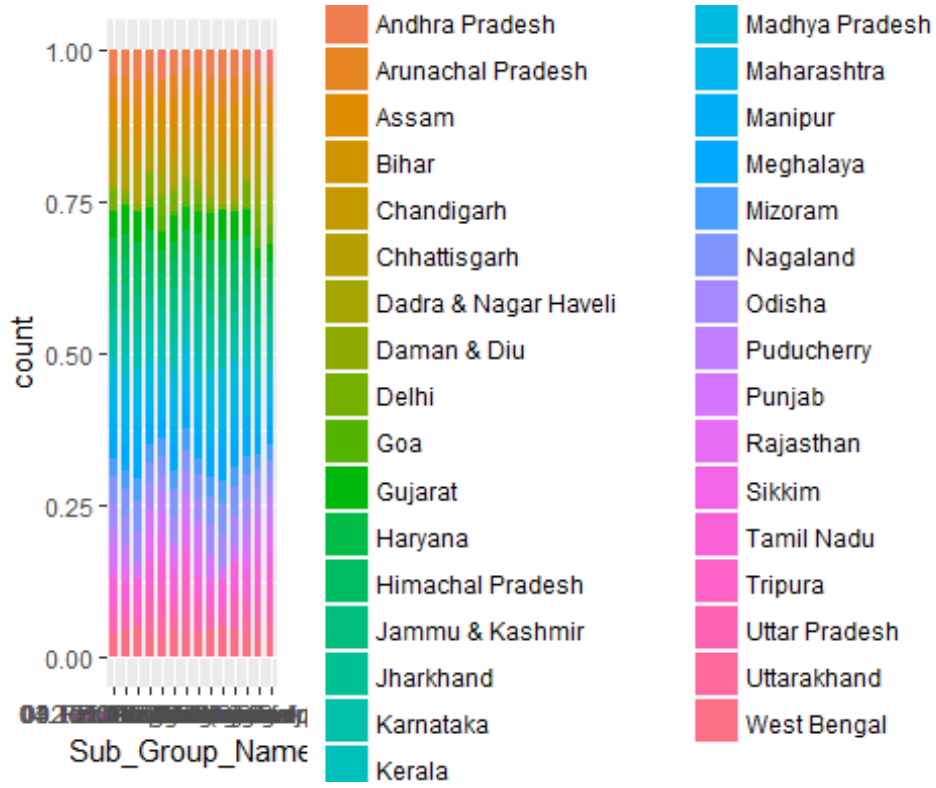
```r
library(ggplot2)

setwd("C:/Users/Administrator/Desktop/Data Visualization/rajanand-crime-in-india")

crime_india <-
read.csv("01_District_wise_crimes_committed_IPC_2001_2012.csv",header = T,sep = ",")
View(crime_india)

crimes <- crime_india %>% group_by(STATE.UT,YEAR,RAPE,
  KIDNAPPING.AND.ABDUCTION.OF.WOMEN.AND.GIRLS,MURDER) %>% filter( DISTRICT == "TOTAL" & MURDER > 70)

## Warning: package 'bindrcpp' was built under R version 3.4.3

head(crimes,20)

## # A tibble: 20 x 33
## # Groups: STATE.UT, YEAR, RAPE,
## #   KIDNAPPING.AND.ABDUCTION.OF.WOMEN.AND.GIRLS, MURDER [20]
##    STAT~ DIST~  YEAR MURD~ ATTE~ CULP~  RAPE CUST~ OTHE~ KIDN~ KIDN~ KIDN~
##    <fct> <fct> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int>
##  1 ANDH~ TOTAL  2001  2602  1555   136   871     0   871  1182   765   417
##  2 ARUN~ TOTAL  2001    83    53     3    33     0    33    83    55    28
##  3 ASSAM TOTAL  2001  1356   481    40   817     0   817  1480  1070   410
##  4 BIHAR TOTAL  2001  3643  3419   250   888     0   888  2159   518  1641
##  5 CHHA~ TOTAL  2001   880   529    45   959     0   959   207   171    36
##  6 GUJA~ TOTAL  2001  1226   537    94   286     0   286   998   857   141
##  7 HARY~ TOTAL  2001   781   467    78   398     0   398   449   297   152
##  8 HIMA~ TOTAL  2001   119    75    11   124     0   124   126   105    21
##  9 JAMM~ TOTAL  2001  1075  1474    33   169     0   169   606   504   102
## 10 JHAR~ TOTAL  2001  1507   866   143   567     0   567   441   279   162
## 11 KARN~ TOTAL  2001  1626  1475    74   293     0   293   559   271   288
## 12 KERA~ TOTAL  2001   472   615   111   562     0   562   183    97    86
## 13 MADH~ TOTAL  2001  2425  2870   125  2851     0  2851   956   668   288
## 14 MAHA~ TOTAL  2001  2839  1454   101  1302     0  1302   985   611   374
## 15 MANI~ TOTAL  2001   209   168     5    20     0    20    94    62    32
## 16 MEGH~ TOTAL  2001   164    47     2    26     0    26    55    11    44
## 17 NAGA~ TOTAL  2001   101    39     9    17     0    17    23     6    17
## 18 ODIS~ TOTAL  2001   987  1151    53   790     0   790   522   434    88
## 19 PUNJ~ TOTAL  2001   738   789    80   298     0   298   513   324   189
## 20 RAJA~ TOTAL  2001  1259  1923    63  1049     0  1049  2718  2165   553
## # ... with 21 more variables: DACOITY <int>,
## #   PREPARATION.AND.ASSEMBLY.FOR.DACOITY <int>, ROBBERY <int>, BURGLARY
## #   <int>, THEFT <int>, AUTO.THEFT <int>, OTHER.THEFT <int>, RIOTS <int>,
## #   CRIMINAL.BREACH.OF.TRUST <int>, CHEATING <int>, COUNTERFIETING <int>,
## #   ARSON <int>, HURT.GREVIOUS.HURT <int>, DOWRY.DEATHS <int>,
## #   ASSAULT.ON.WOMEN.WITH.INTENT.TO.OUTRAGE.HER.MODESTY <int>,
## #   INSULT.TO.MODESTY.OF.WOMEN <int>, CRUELTY.BY.HUSBAND.OR.HIS.RELATIVES
```

```
## #   <int>, IMPORTATION.OF.GIRLS.FROM.FOREIGN.COUNTRIES <int>,
## #   CAUSING.DEATH.BY.NEGLIGENCE <int>, OTHER.IPC.CRIMES <int>,
## #   TOTAL.IPC.CRIMES <int>

chart <-
ggplot(crimes,aes(x=RAPE,y=KIDNAPPING.AND.ABDUCTION.OF.WOMEN.AND.GIRLS)) +
  geom_point(aes(col = STATE.UT,size = MURDER,frame=YEAR))

## Warning: Ignoring unknown aesthetics: frame

chart1 <- ggplotly(chart)

## We recommend that you use the dev version of ggplot2 with `ggplotly()`
## Install it with: `devtools::install_github('hadley/ggplot2')`

################################################################################
############
# problem Stmt : Plot manufacturer wise,class wise count of cars

cars <- mpg %>% group_by(manufacturer,class) %>% summarise(count=n())

ch <- ggplot(cars,aes(x = manufacturer,y = count,fill = class)) +
  geom_bar(stat = "Identity")

p <- ch + geom_text(data = cars,aes(x=manufacturer,y=count,label=count),
                    position = position_stack(vjust=0.5))
r <- p + theme(axis.text.x = element_text(angle = 90))

ggplot(cars,aes(x=manufacturer,y = count,fill=class)) +
  geom_bar(stat="Identity",width = 0.6,position = "fill")
```



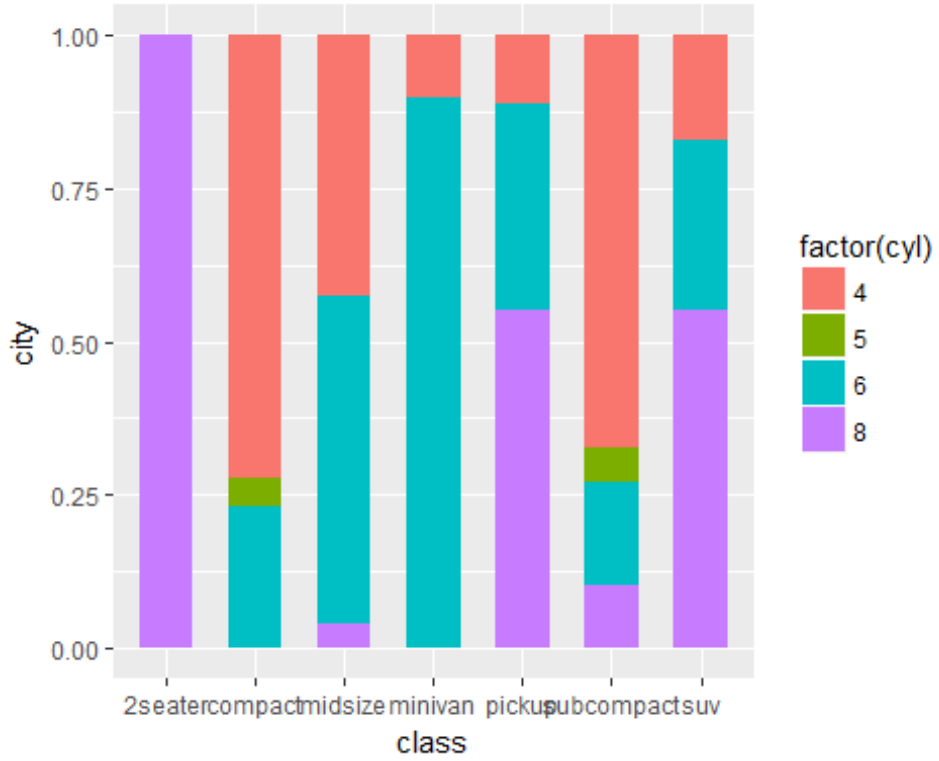

```
################################################################################
#############
?boxplot

## starting httpd help server ...

##  done

#problem Stmt: Major reason people being kidnapped in each state

kidnap <-
read.csv("39_Specific_purpose_of_kidnapping_and_abduction.csv",header =
TRUE,sep = ",")
View(kidnap)
colnames(kidnap)

##  [1] "ï..Area_Name"              "Year"
##  [3] "Group_Name"                "Sub_Group_Name"
##  [5] "K_A_Cases_Reported"        "K_A_Female_10_15_Years"
##  [7] "K_A_Female_15_18_Years"    "K_A_Female_18_30_Years"
##  [9] "K_A_Female_30_50_Years"    "K_A_Female_Above_50_Years"
## [11] "K_A_Female_Total"          "K_A_Female_Upto_10_Years"
## [13] "K_A_Grand_Total"           "K_A_Male_10_15_Years"
## [15] "K_A_Male_15_18_Years"      "K_A_Male_18_30_Years"
## [17] "K_A_Male_30_50_Years"      "K_A_Male_Above_50_Years"
## [19] "K_A_Male_Total"            "K_A_Male_Upto_10_Years"
```

```r
reason <- kidnap %>% group_by(ï..Area_Name,Sub_Group_Name) %>%
summarise(count=n())

ggplot(reason,aes(x=Sub_Group_Name,y=count,fill=ï..Area_Name )) +
  geom_bar(stat = "Identity",width = 0.6,position = "fill")
```

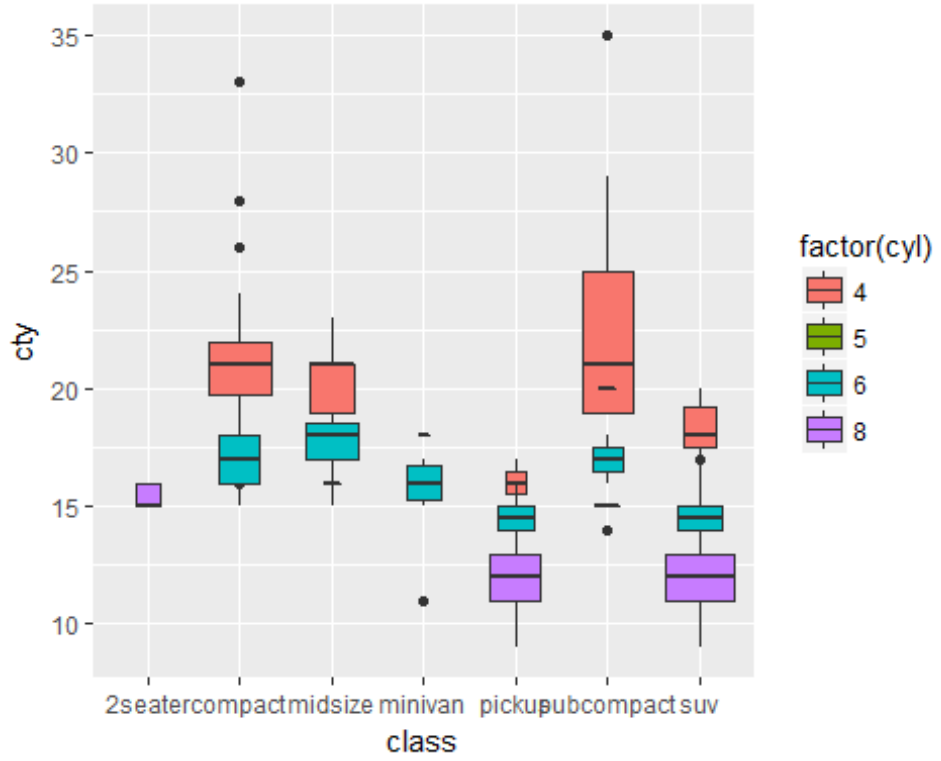



```r
################################################################################
#############
# Heat maps
#heatmap(as.matrix(mtcars),scale = "column",col = heat.colors(256),
 #       main = "characteristics of car models")

plot1 <- ggplot(data = reason,aes(x = Sub_Group_Name, y = ï..Area_Name)) +
  geom_tile(aes(fill = count),col="white") + theme(axis.text.x =
element_text(angle = 90))

################################################################################
#############

# problem stmt : analyse city mileage of various classes of cars for each
cylinder

m <- mpg %>% group_by(class,cyl) %>% summarise(city=sum(cty))
```

```
ggplot(m,aes(x=class,y=city,fill = factor(cyl))) + geom_bar(stat =
"Identity",width = 0.6,position = "fill")
```

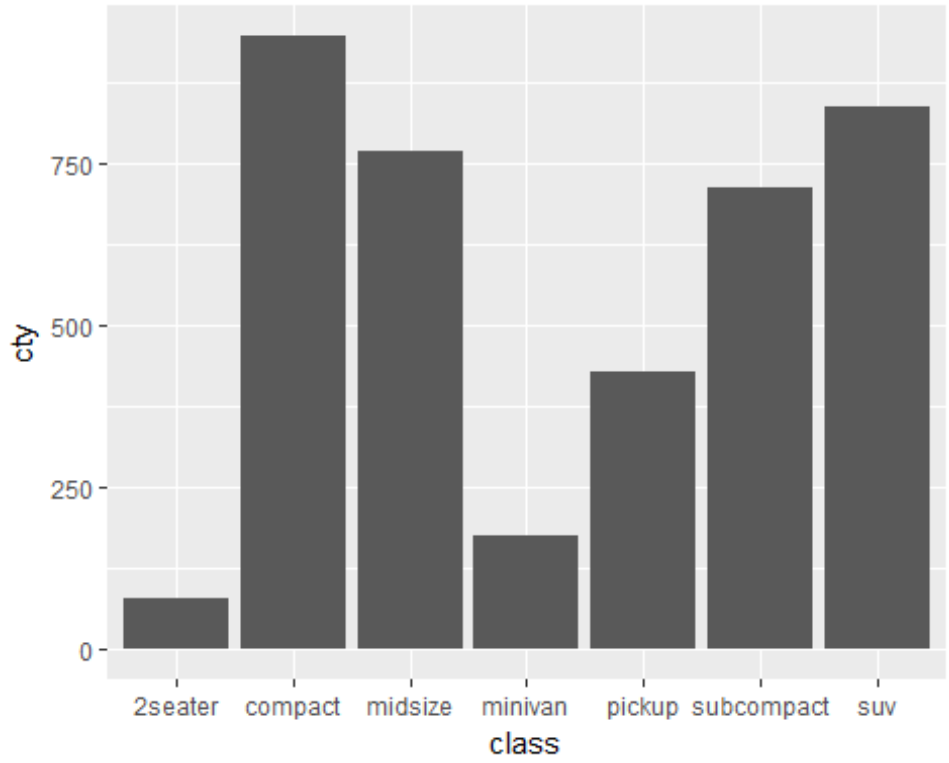



```
# Box plot is used when you want to plot distribution of data

ggplot(mpg,aes(x=class,y=cty)) + geom_boxplot(aes(fill=factor(cyl)),varwidth
= TRUE)

## Warning: position_dodge requires non-overlapping x intervals
```

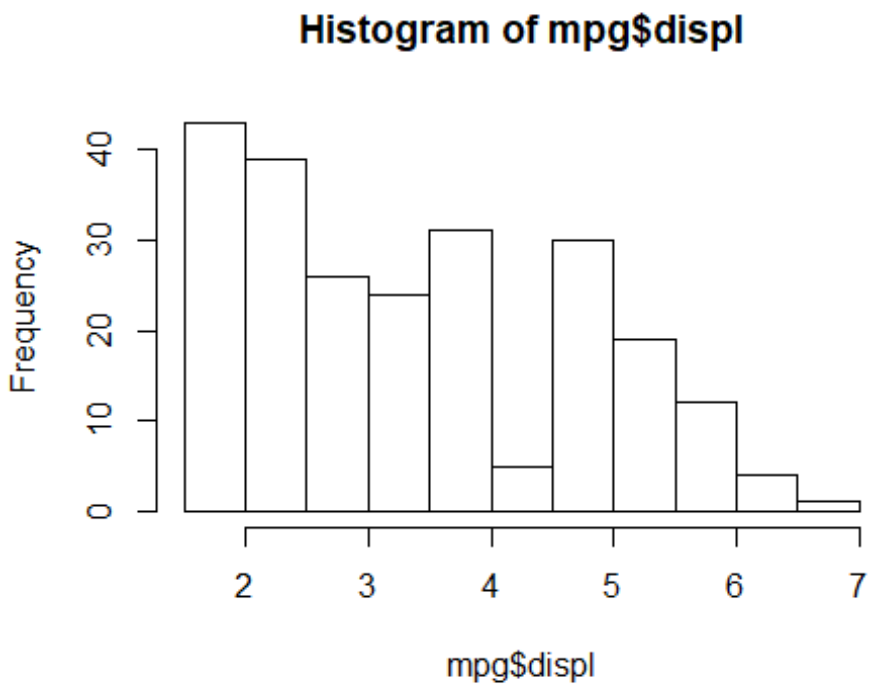



```
ggplot(mpg,aes(x=class,y=cty))+geom_histogram(stat = "Identity")
```

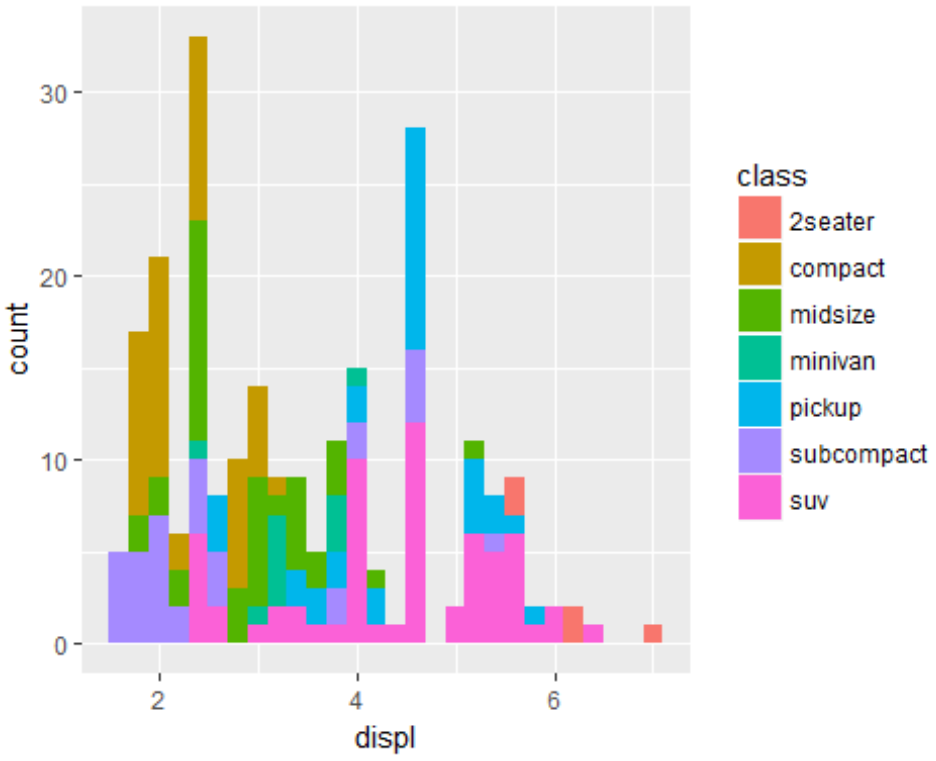
```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



```r
# problem stmt : to create a histogram of displacement of various car's
class.

?hist
hist(mpg$displ,breaks = "Sturges")
```




```r
ggplot(mpg,aes(x=displ,fill=class)) + geom_histogram(binwidth = 0.2)
```


30
20
10
0
count
2
4
6
displ
class
2seater
compact
midsize
minivan
pickup
subcompact
suv