Final1.R

dhruvang

2022-02-21

```
#install plyr package
r=getOption("repos")
r["CRAN"]="http://cran.us.r-project.org"
options(repos=r)
install.packages("plyr")
```

```
## Installing package into 'C:/Users/dhruvang/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)
```

```
## package 'plyr' successfully unpacked and MD5 sums checked
```

```
## Warning: cannot remove prior installation of package 'plyr'
```

```
## Warning in file.copy(savedcopy, lib, recursive = TRUE): problem copying C:
## \Users\dhruvang\Documents\R\win-library\4.1\00LOCK\plyr\libs\x64\plyr.dll to C:
## \Users\dhruvang\Documents\R\win-library\4.1\plyr\libs\x64\plyr.dll: Permission
## denied
```

```
## Warning: restored 'plyr'
```

```
##
## The downloaded binary packages are in
##  C:\Users\dhruvang\AppData\Local\Temp\RtmpsXA6rB\downloaded_packages
```

```
library(plyr)

#install FSA package
install.packages("FSA")
```

```
## Installing package into 'C:/Users/dhruvang/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)
```

```
##
##   There is a binary version available but the source version is later:
##     binary source needs_compilation
## FSA  0.9.2  0.9.3             FALSE
```

```
## installing the source package 'FSA'
```

```
library(FSA)
```

```
## ## FSA v0.9.3. See citation('FSA') if used in publication.
## ## Run fishR() for related website and fishR('IFAR') for related book.

##
## Attaching package: 'FSA'

## The following object is masked from 'package:plyr':
##
##     mapvalues
```

```
install.packages("plotrix")
```

```
## Installing package into 'C:/Users/dhruvang/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)

## package 'plotrix' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\dhruvang\AppData\Local\Temp\RtmpsXA6rB\downloaded_packages
```

```
library(plotrix)
#install FSAdata package
install.packages("FSAdata")
```

```
## Installing package into 'C:/Users/dhruvang/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)

## package 'FSAdata' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\dhruvang\AppData\Local\Temp\RtmpsXA6rB\downloaded_packages
```

```
library(FSAdata)
```

```
## ## FSAdata v0.3.9. See ?FSAdata to find data for specific fisheries analyses.
```

```
#install magrittr package
install.packages("magrittr")
```

```
## Installing package into 'C:/Users/dhruvang/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)

## package 'magrittr' successfully unpacked and MD5 sums checked

## Warning: cannot remove prior installation of package 'magrittr'
```

```
## Warning in file.copy(savedcopy, lib, recursive = TRUE):
## problem copying C:\Users\dhruvang\Documents\R\win-
## library\4.1\00LOCK\magrittr\libs\x64\magrittr.dll to C:
## \Users\dhruvang\Documents\R\win-library\4.1\magrittr\libs\x64\magrittr.dll:
## Permission denied

## Warning: restored 'magrittr'

##
## The downloaded binary packages are in
##  C:\Users\dhruvang\AppData\Local\Temp\RtmpsXA6rB\downloaded_packages
```

```
library(magrittr)

#install dplyr package
install.packages("dplyr")
```

```
## Installing package into 'C:/Users/dhruvang/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)

## package 'dplyr' successfully unpacked and MD5 sums checked

## Warning: cannot remove prior installation of package 'dplyr'

## Warning in file.copy(savedcopy, lib, recursive = TRUE): problem copying C:
## \Users\dhruvang\Documents\R\win-library\4.1\00LOCK\dplyr\libs\x64\dplyr.dll
## to C:\Users\dhruvang\Documents\R\win-library\4.1\dplyr\libs\x64\dplyr.dll:
## Permission denied

## Warning: restored 'dplyr'

##
## The downloaded binary packages are in
##  C:\Users\dhruvang\AppData\Local\Temp\RtmpsXA6rB\downloaded_packages
```

```
library(dplyr)
```

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

## The following objects are masked from 'package:plyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize

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

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

```r
#install tidyverse package
install.packages("tidyverse")
```

```
## Installing package into 'C:/Users/dhruvang/Documents/R/win-library/4.1'
## (as 'lib' is unspecified)

## package 'tidyverse' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\dhruvang\AppData\Local\Temp\RtmpsXA6rB\downloaded_packages
```

```r
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v stringr 1.4.0
## v tidyr   1.2.0     v forcats 0.5.1
## v readr   2.1.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::arrange()   masks plyr::arrange()
## x purrr::compact()   masks plyr::compact()
## x dplyr::count()     masks plyr::count()
## x tidyr::extract()   masks magrittr::extract()
## x dplyr::failwith()  masks plyr::failwith()
## x dplyr::filter()    masks stats::filter()
## x dplyr::id()        masks plyr::id()
## x dplyr::lag()       masks stats::lag()
## x dplyr::mutate()    masks plyr::mutate()
## x dplyr::rename()    masks plyr::rename()
## x purrr::set_names() masks magrittr::set_names()
## x dplyr::summarise() masks plyr::summarise()
## x dplyr::summarize() masks plyr::summarize()
```

```r
#install tidyr package
install.packages("tidyr")
```

```
## Warning: package 'tidyr' is in use and will not be installed
```

```r
library(tidyr)

install.packages("ggplot2")
```

```
## Warning: package 'ggplot2' is in use and will not be installed
```

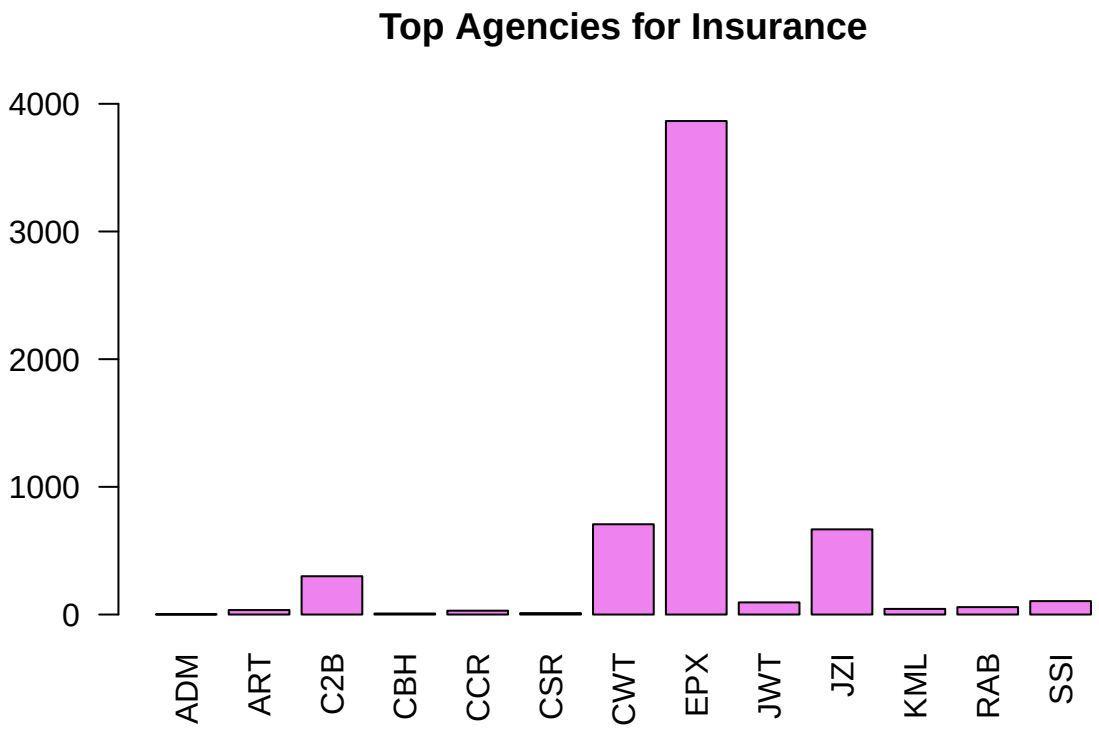
```r
library(ggplot2)

library(readxl)
insurance <- read_excel("C:/Users/dhruvang/Downloads/insurance.xlsx")
insurance
```

```
## # A tibble: 5,927 x 11
##    Agency `Agency Type` `Distribution Channel` `Product Name`     Claim Duration
##    <chr>  <chr>         <chr>                  <chr>              <chr>    <dbl>
##  1 CBH    Travel Agency Offline                Comprehensive Plan No         186
##  2 CBH    Travel Agency Offline                Comprehensive Plan No         186
##  3 CWT    Travel Agency Online                 Rental Vehicle Ex~ No          65
##  4 CWT    Travel Agency Online                 Rental Vehicle Ex~ No          60
##  5 CWT    Travel Agency Online                 Rental Vehicle Ex~ No          79
##  6 JZI    Airlines      Online                 Value Plan         No          66
##  7 CWT    Travel Agency Online                 Rental Vehicle Ex~ No          47
##  8 CWT    Travel Agency Online                 Rental Vehicle Ex~ No          63
##  9 CWT    Travel Agency Online                 Rental Vehicle Ex~ No          57
## 10 CWT    Travel Agency Online                 Rental Vehicle Ex~ No         186
## # ... with 5,917 more rows, and 5 more variables: Destination <chr>,
## #   Net Sales <dbl>, Commision (in value) <dbl>, Gender <chr>, Age <dbl>
```

```
agency<-table(insurance$Agency)
agency
```

```
##
##  ADM  ART  C2B  CBH  CCR  CSR  CWT  EPX  JWT  JZI  KML  RAB  SSI
##    2   35  300    8   30   11  707 3865   95  667   44   58  105
```

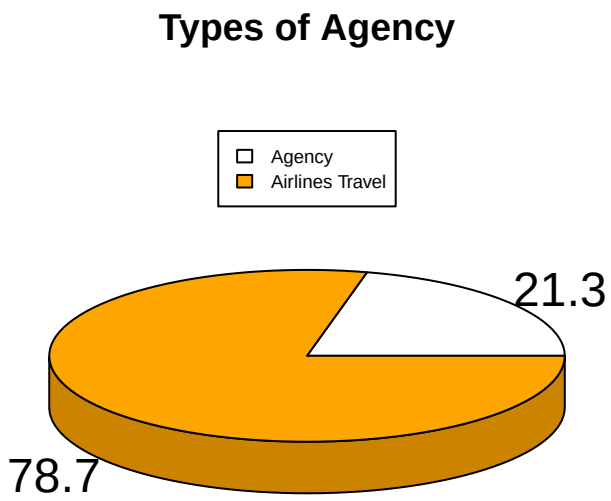
```
barplot(agency, las=2, col = "violet", main = "Top Agencies for Insurance", ylim = c(0,4000))
```



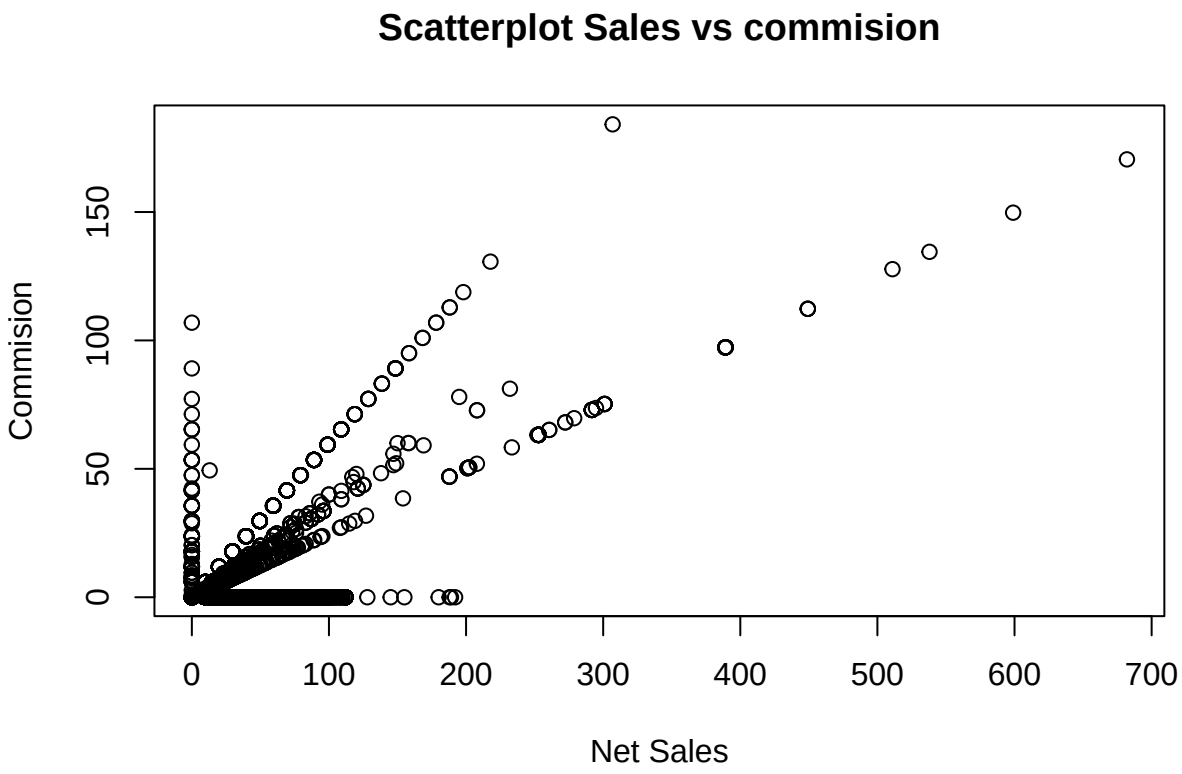
```r
agencytype<-table(insurance$`Agency Type`)
agencytype
```

```
##
##      Airlines Travel Agency
##           1260          4667
```

```r
#piechart
lbls<-round(agencytype/sum(agencytype) * 100, 1)
pie3D(agencytype, labels = lbls, col = c("white","orange"), main="Types of Agency")
legend("top",inset = 0.05, c("Agency","Airlines Travel"), cex = 0.6,
       fill = c("white","orange"))
```



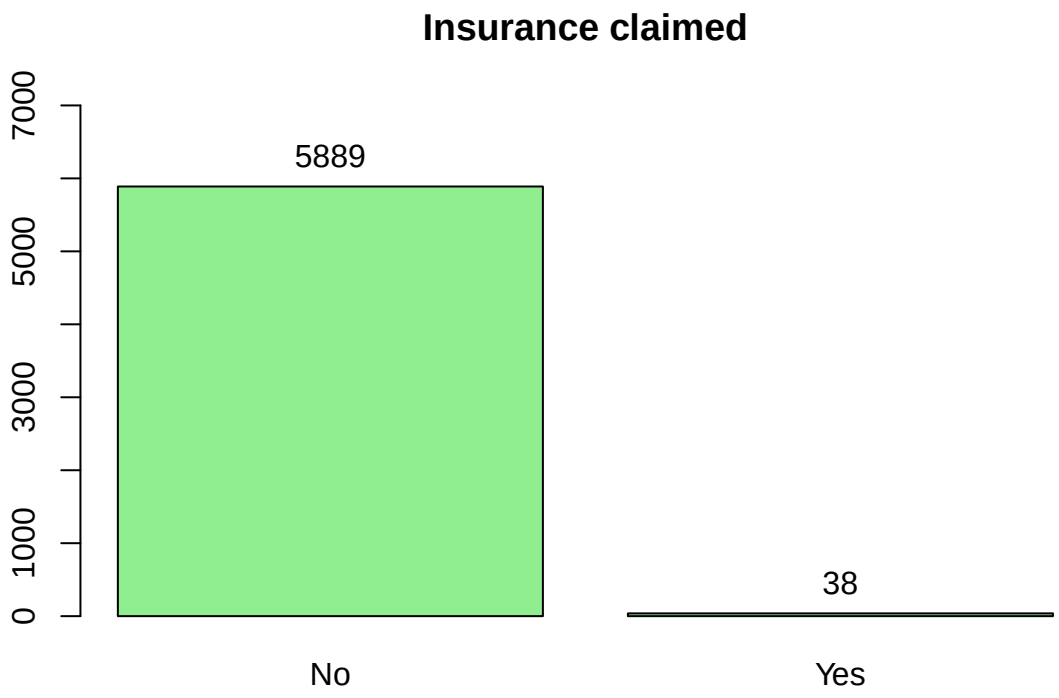

```r
plot(insurance$`Net Sales`,insurance$`Commision (in value)`,
     ylab = "Commision",
     xlab = "Net Sales",
     main="Scatterplot Sales vs commision")
```




```
claim<-table(insurance$Claim)
claim
```

```
##
##   No  Yes
## 5889   38
```

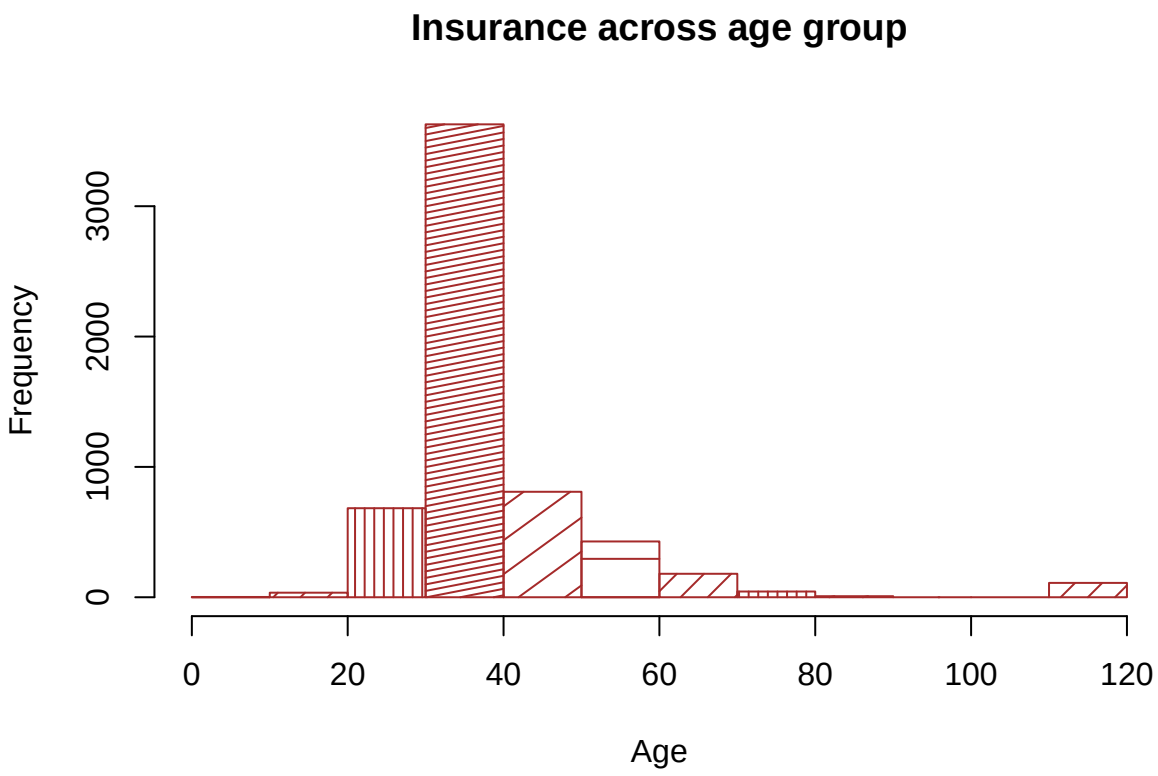
```
pl<-barplot(claim, ylim = c(0,7000), col = "lightgreen", main = "Insurance claimed")
text(pl,claim+3,  labels = claim, pos = 3)
```




```
age<-table(insurance$Age)
age
```

```
##
##    8   10   12   13   14   16   18   19   20   21   22   23   24   25   26   27
##    1    1    1    4    1    3    2    1   23   30   34   37   47   58   83   79
##   28   29   30   31   32   33   34   35   36   37   38   39   40   41   42   43
##  110  109   96  215   89   80   84   92 2731   92   92   74   79   66   74   63
##   44   45   46   47   48   49   50   51   52   53   54   55   56   57   58   59
##   57   75   58   69  238   52   57   52   49   52   41   34   43   46   46   32
##   60   61   62   63   64   65   66   67   68   69   70   71   72   73   74   75
##   33   31   20   19   17   15   11   28    9   15   15    9    8    4    3    7
##   76   77   78   79   80   81   83   84   85   87  118
##    2    2    3    4    2    3    1    1    1    2  110
```

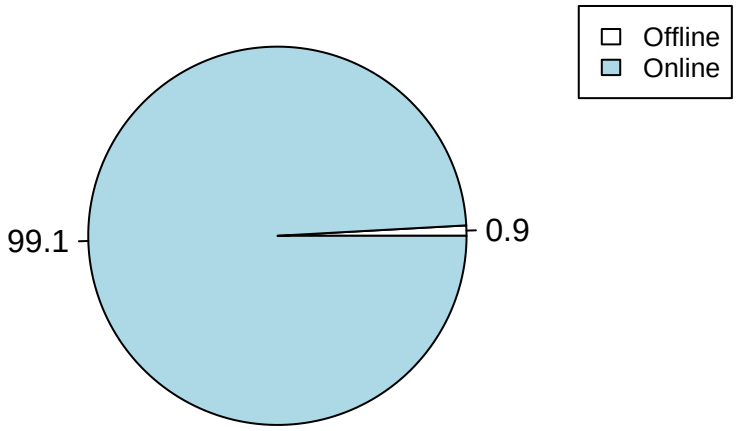
```
hist(insurance$Age,xlab = "Age",main = "Insurance across age group", density=c(5,10,20,30,7) , angle=c(
```




```
channel<-table(insurance$`Distribution Channel`)
channel
```

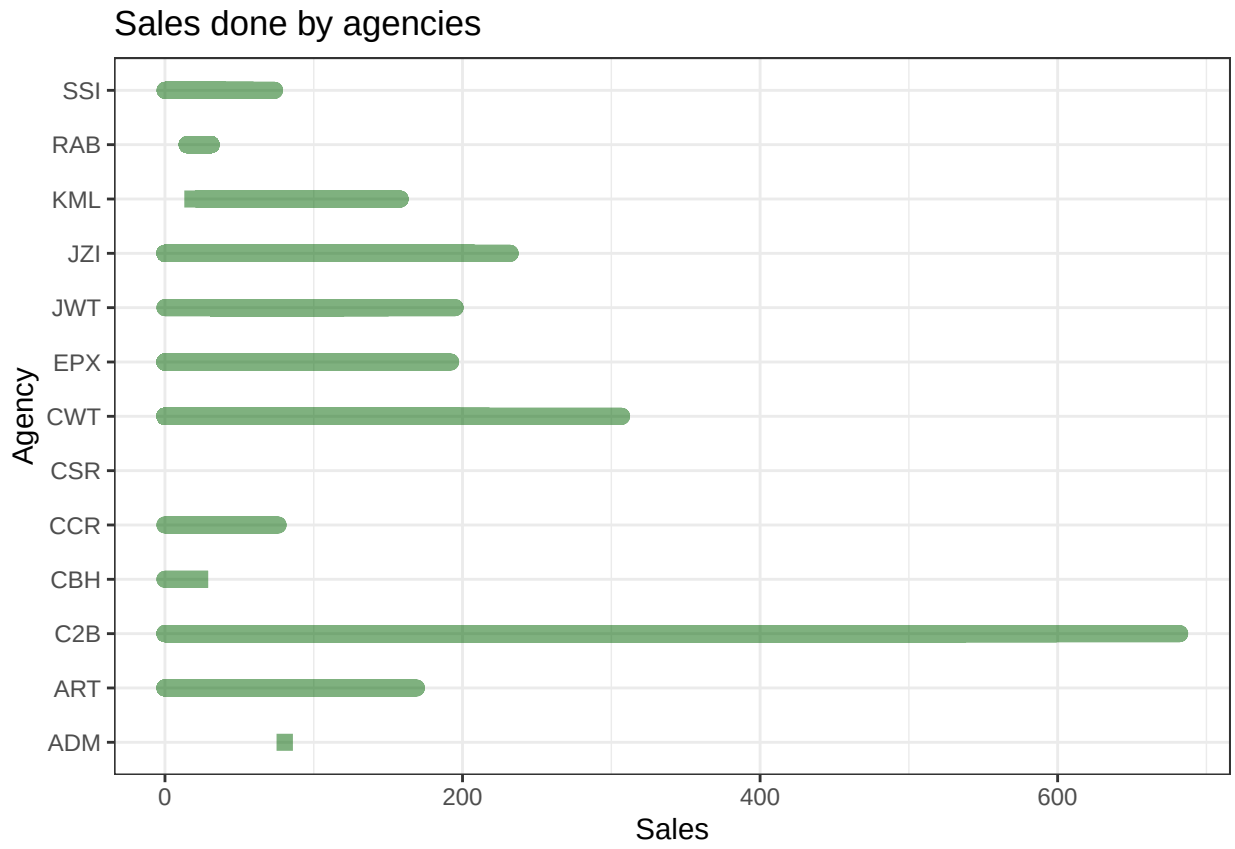
```
##
## Offline  Online
##      52    5875
```

```
ll<-round(channel/sum(channel)*100,1)
pie(channel,labels = ll, main = "Mode of distribution channel")
legend("topright",inset = 0.05, c("Offline","Online"), cex = 0.8,
       fill = c("white","lightblue"))
```



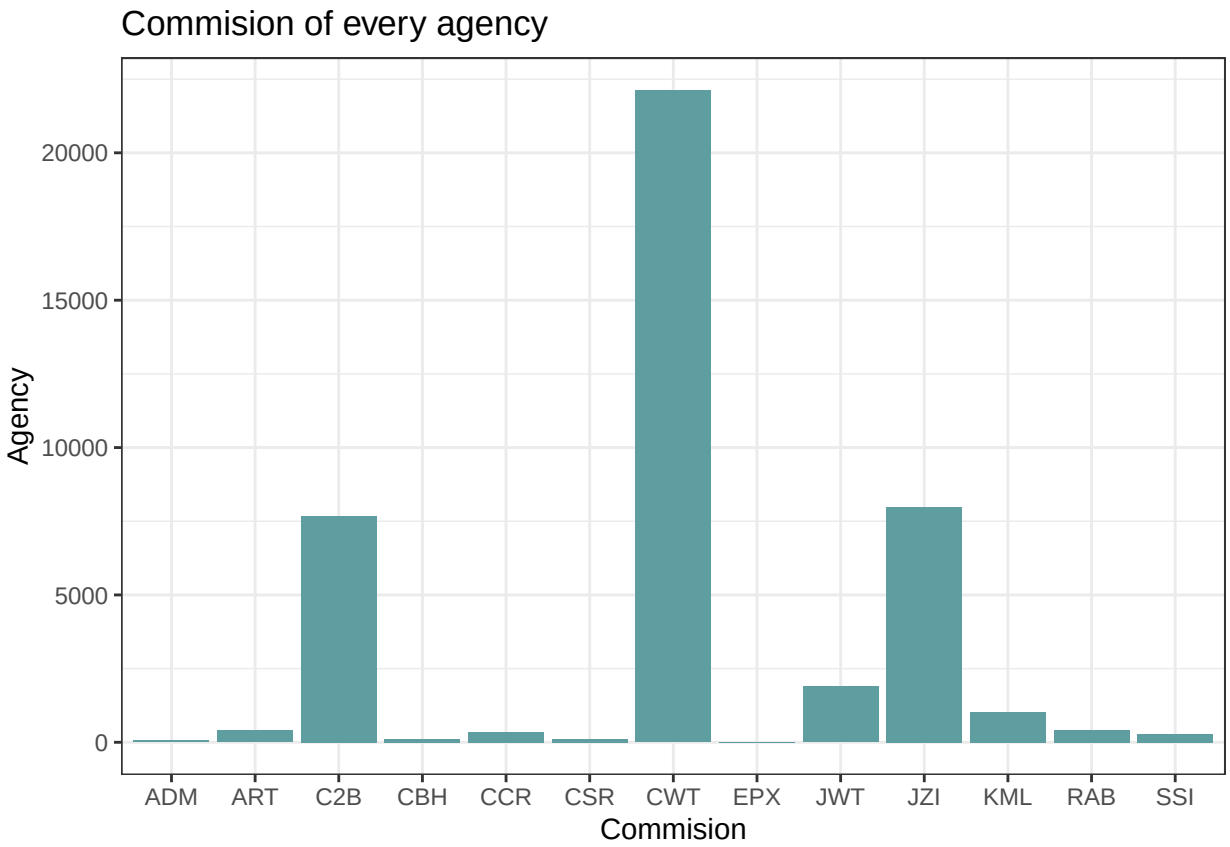

```
#plot1
ggplot(data = insurance,
       aes(x = insurance$Agency,
           y = insurance$`Net Sales`))+
  geom_line(colour = 'darkgreen',
            size = 3,
            stat = 'identity',
            alpha = 0.5)+
  geom_smooth()+
  coord_flip()+
  theme_bw()+
  labs(title = 'Sales done by agencies',
       x = 'Agency',
       y = 'Sales')
```

```
## Warning: Use of ‘insurance$Agency‘ is discouraged. Use ‘Agency‘ instead.

## Warning: Use of ‘insurance$‘Net Sales‘‘ is discouraged. Use ‘Net Sales‘ instead.

## Warning: Use of ‘insurance$Agency‘ is discouraged. Use ‘Agency‘ instead.

## Warning: Use of ‘insurance$‘Net Sales‘‘ is discouraged. Use ‘Net Sales‘ instead.

## ‘geom_smooth()‘ using method = ’gam’ and formula ’y ~ s(x, bs = "cs")’
```




```
#plot2
pfg = ggplot(data = insurance,
             aes(y = insurance$Agency,
                 x = insurance$`Commision (in value)`))+
  geom_bar(fill = 'cadetblue',
           stat = 'identity',
           position = 'stack',)+
  xlab("Agency")+
  coord_flip()+
  theme_bw()+
  ylab("Commision")+
  labs(title = 'Commision of every agency')
pfg
```

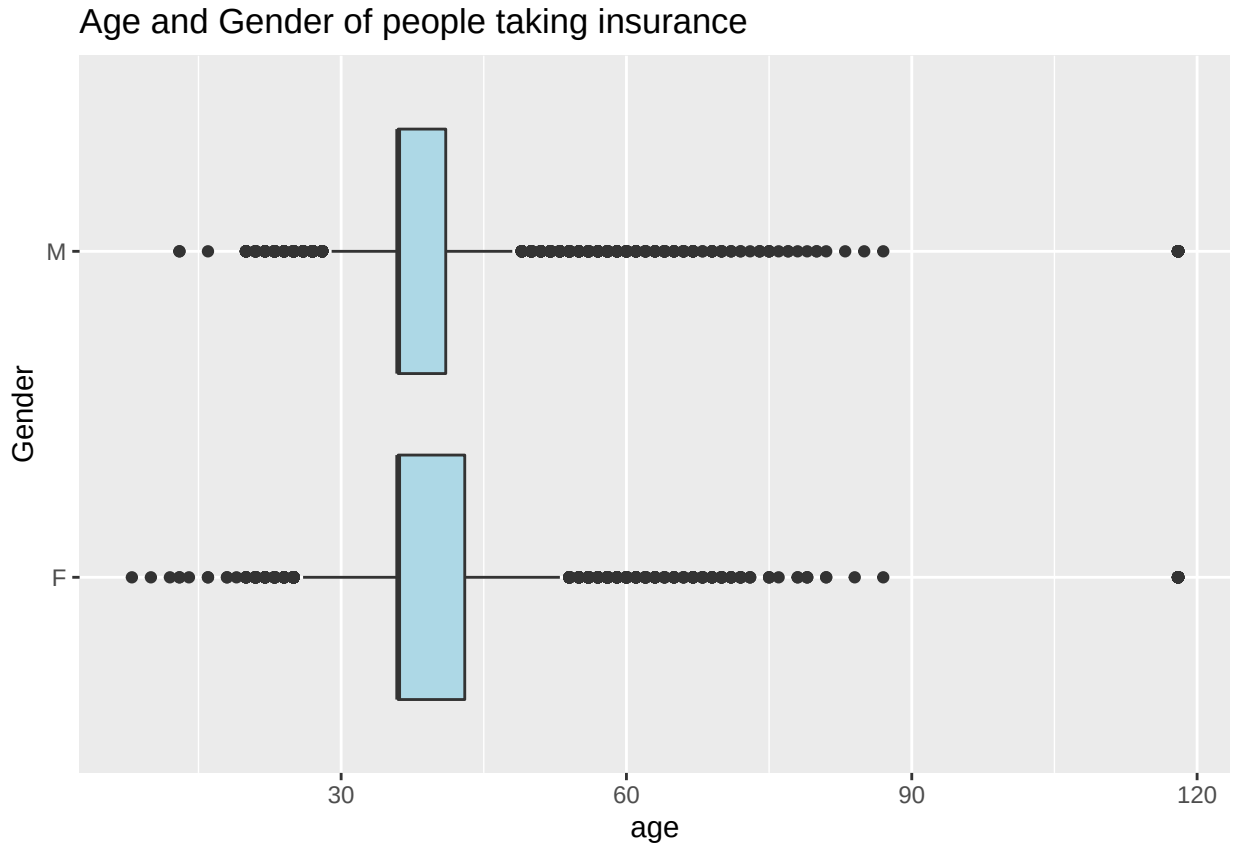
```
## Warning: Use of ‘insurance$‘Commision (in value)‘‘ is discouraged. Use
## ‘Commision (in value)‘ instead.

## Warning: Use of ‘insurance$Agency‘ is discouraged. Use ‘Agency‘ instead.
```




```
#boxplot
ggplot(insurance, aes(x=insurance$Age, y=insurance$Gender)) +
  geom_boxplot(fill="lightblue") +
  xlab("age")+ ylab("Gender")+
  labs(title = 'Age and Gender of people taking insurance')
```

```
## Warning: Use of ‘insurance$Age‘ is discouraged. Use ‘Age‘ instead.
```

```
## Warning: Use of ‘insurance$Gender‘ is discouraged. Use ‘Gender‘ instead.
```




```
summary(insurance)
```

```
##     Agency          Agency Type        Distribution Channel Product Name
##  Length:5927        Length:5927        Length:5927          Length:5927
##  Class :character   Class :character   Class :character     Class :character
##  Mode  :character   Mode  :character   Mode  :character     Mode  :character
##
##
##
##     Claim             Duration    Destination          Net Sales
##  Length:5927        Min.   :   0   Length:5927        Min.   :  0.00
##  Class :character   1st Qu.:  10   Class :character   1st Qu.: 18.00
##  Mode  :character   Median :  24   Mode  :character   Median : 26.00
##                     Mean   :  47                      Mean   : 37.57
##                     3rd Qu.:  55                      3rd Qu.: 45.00
##                     Max.   :4881                      Max.   :682.00
##  Commision (in value)    Gender               Age
##  Min.   :  0.000      Length:5927        Min.   :  8.00
##  1st Qu.:  0.000      Class :character   1st Qu.: 36.00
##  Median :  0.000      Mode  :character   Median : 36.00
##  Mean   :  7.142                         Mean   : 40.11
##  3rd Qu.:  7.700                         3rd Qu.: 42.00
##  Max.   :184.140                         Max.   :118.00
```

```r
mean(insurance$Age)
```

```
## [1] 40.10849
```

```r
mean(insurance$Duration)
```

```
## [1] 46.99899
```

```r
mean(insurance$`Net Sales`)
```

```
## [1] 37.57097
```

```r
mean(insurance$`Commision (in value)`)
```

```
## [1] 7.141893
```