STAT796_HW1

Colleen Brents

January 23, 2019

1. Import Data

2. rename and recode

3. summary stats for entire cohort

3a. average, max, min, sc of age (years)

```
## # A tibble: 1 x 5
##       n mean_age max_age min_age   sd_age
##   <int>    <dbl>   <dbl>   <dbl>    <dbl>
## 1   200   57.545      92      16 20.05465
```

3b. count of subjects with each service type

```
##
##  medical surgical
##       93      107
```

3c. count of subjects with procedure type

```
##
##  elective emergency
##        53       147
```

3d. count of subjects of each race and each sex

```
##
##          male female
##   white   108     67
##   black    10      5
##   other     6      4
```

3e. count of subject by consciousness

```
##
## no_coma_or_stupor            stupor              coma
##               185                 5                10
```

4. summary stats for cohort, grouped by by vital status

4a. average, max, min, sd age

```
## # A tibble: 2 x 6
##    Died     n mean_age max_age min_age   sd_age
##   <int> <int>    <dbl>   <dbl>   <dbl>    <dbl>
```

```
## 1     0   160   55.650       91       16 20.42818
## 2     1    40   65.125       92       19 16.64900
```

4b. count by service type

```
##
##  elective emergency
##        53       147
```

4c. count by procedure type

```
##
##  elective emergency
##        53       147
```

4d. count by race and sex

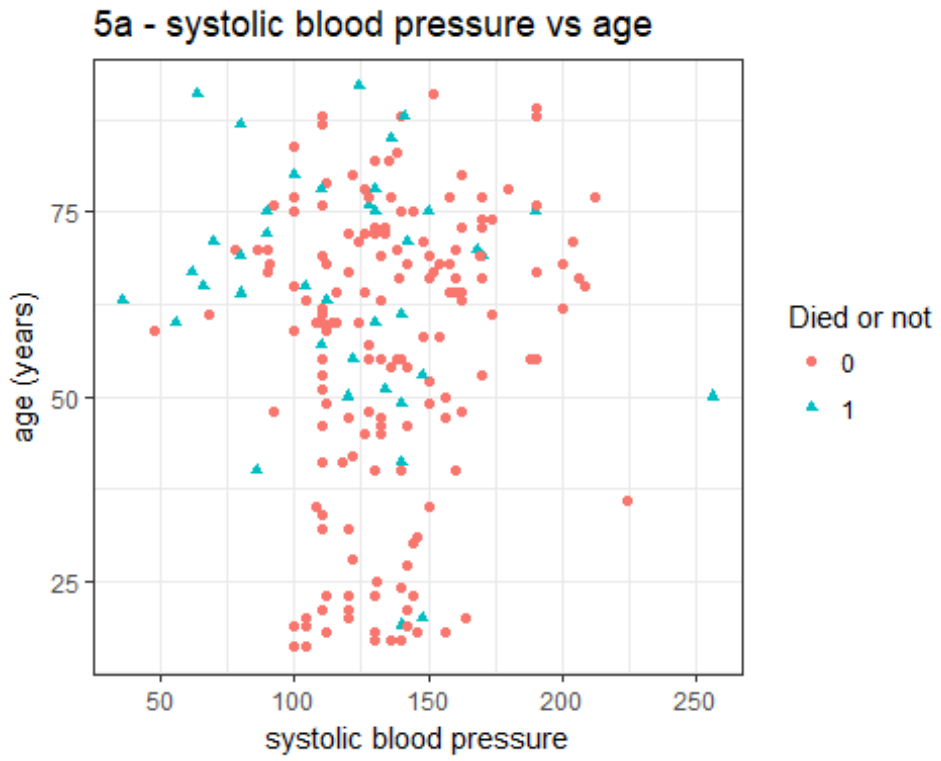
```
##
##          male female
##   white   108     67
##   black    10      5
##   other     6      4
```

4e. count by consciousness

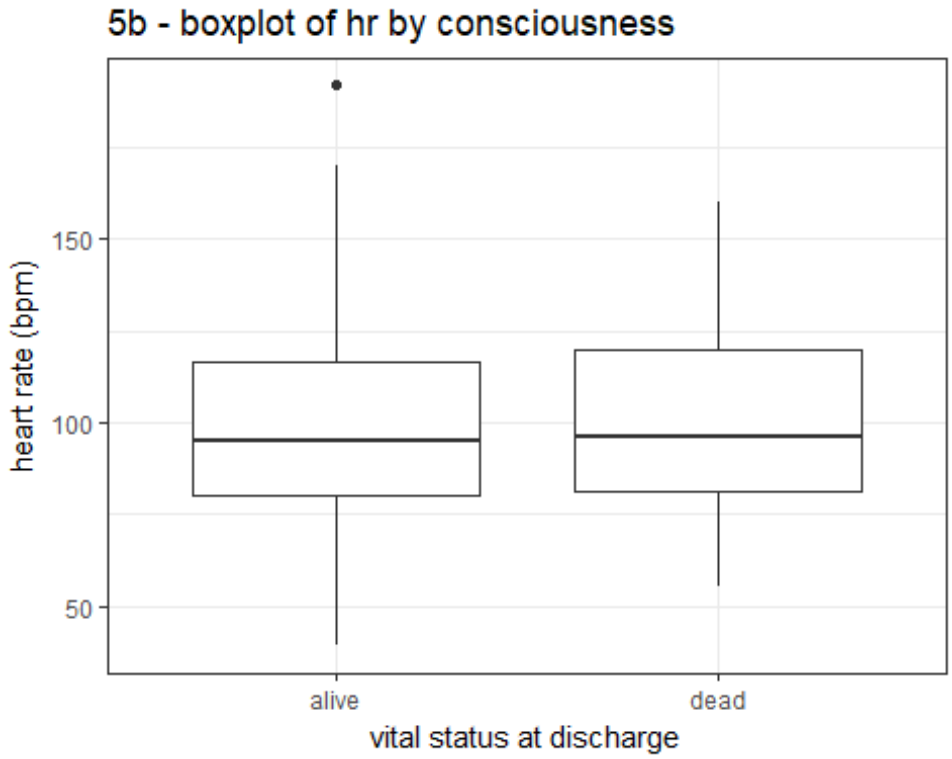
```
##
## no_coma_or_stupor            stupor              coma
##               185                 5                10
```

5. graphical summaries

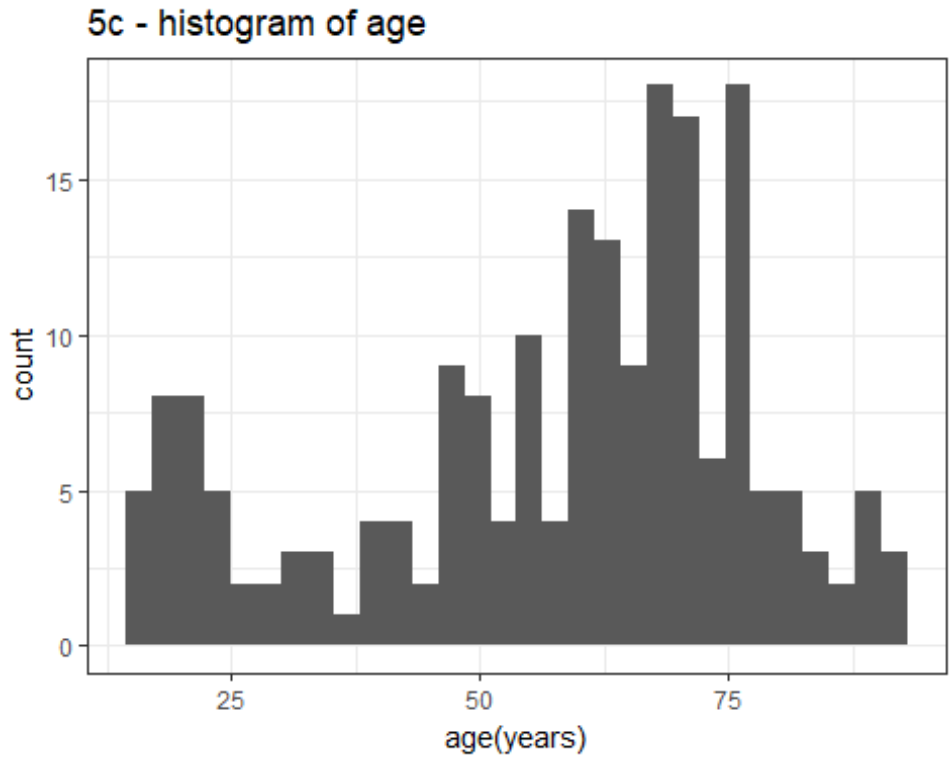
5a. scatterplot of sbp against age



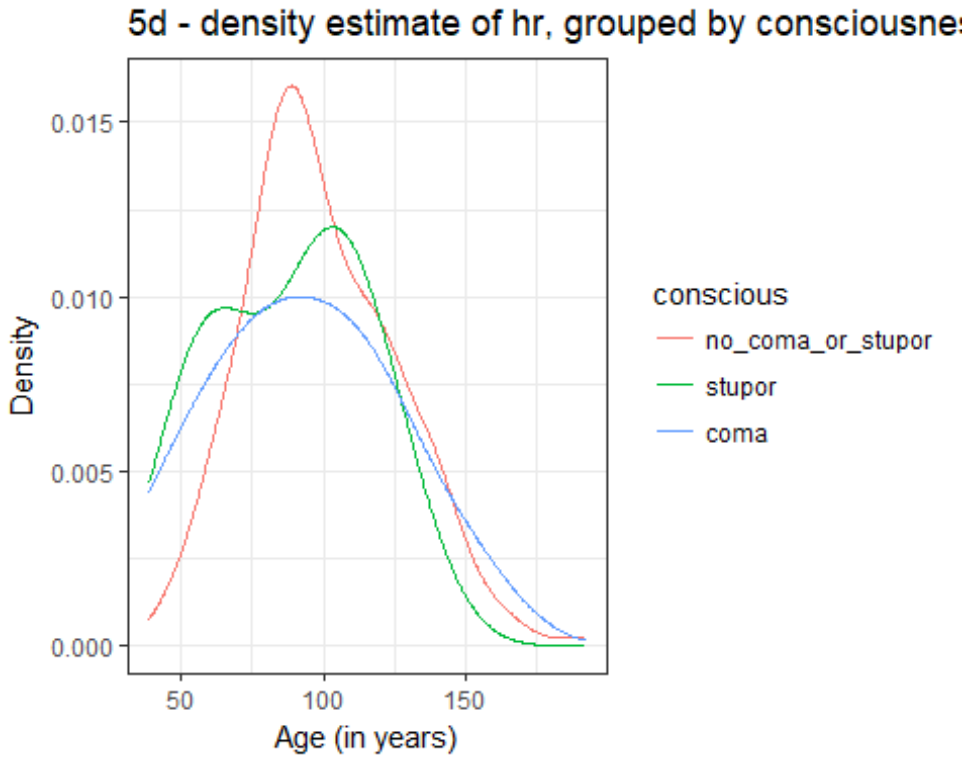
5b. boxplots of hr, grouped by consciousness



5c. histograms of age



5d. densit estimates of hr, grouped by consciousness



STAT796_HW1_code

Colleen Brents

January 23, 2019

```
library(ggplot2)
library(dplyr)
library(tidyr)
library(readr)
library(knitr)
```

1. Import Data

```
url <-
paste0("https://raw.githubusercontent.com/cbrents/stat796logreg/master/hw1/ICU_raw.txt")
icu_raw<-read_table2(url)
```

2. rename and recode

```
icu<- icu_raw %>%
    transmute(id=ID,
              died=STA,
              age=AGE,
              female=GENDER,
              race=factor(RACE,
                          levels=1:3,
                          labels=c("white", "black", "other")),
              service=factor(SER,
                             levels=0:1,
                             labels=c("medical", "surgical")),
              cancer=CAN,
              renal=CRN,
              infect=INF,
              cpr=CPR,
              sbp=SYS,
              hr=HRA,
              prv_admit=PRE,
              type=factor(TYP,
                          levels=0:1,
                          labels=c("elective", "emergency")),
              fracture=FRA,
              po2_g60=PO2,
              '1-PO2'= po2_g60,
              pco2_g45=PCO,
              bicarb_g18=BIC,
              '1 - BIC' = bicarb_g18,
              creat_g2=CRE,
```

```
              conscious=factor(LOC,
                levels=0:2,
                labels=c("no_coma_or_stupor", "stupor", "coma")))
```

3. summary stats for entire cohort

3a. average, max, min, sc of age (years)

```
x<- icu %>%
  summarize(n=n(),
            mean_age = mean(age),
            max_age = max(age),
            min_age = min(age),
            sd_age =sd(age))
print(x)
```

```
## # A tibble: 1 x 5
##       n mean_age max_age min_age   sd_age
##   <int>    <dbl>   <dbl>   <dbl>    <dbl>
## 1   200   57.545      92      16 20.05465
```

3b. count of subjects with each service type

```
a<- table (icu$service)
print(a)
```

```
##
##  medical surgical
##       93      107
```

3c. count of subjects with procedure type

```
b<- table(icu$type)
print(b)
```

```
##
##  elective emergency
##        53       147
```

3d. count of subjects of each race and each sex

```
#count(icu, race, female)
z<- table(icu$race, icu$female)
colnames(z)<- c("male", "female")
print(z)
```

```
##
##         male female
##   white  108     67
##   black   10      5
##   other    6      4
```

3e. count of subject by consciousness

```
#count(icu, conscious)
x<- table(icu$conscious)
print(x)

##
## no_coma_or_stupor            stupor              coma
##               185                 5                10
```

4. summary stats for cohort, grouped by by vital status

4a. average, max, min, sd age

```
age_table <- icu %>%
  group_by(Died=died) %>%
  summarize(n=n(),
            mean_age = mean(age),
             max_age = max(age),
             min_age = min(age),
             sd_age =sd(age))  # died == 0  alive == 1
print(age_table)

## # A tibble: 2 x 6
##    Died     n mean_age max_age min_age   sd_age
##   <int> <int>    <dbl>   <dbl>   <dbl>    <dbl>
## 1     0   160   55.650      91      16 20.42818
## 2     1    40   65.125      92      19 16.64900
```

4b. count by service type

```
vitals <- icu %>%
  group_by(died)
y<- table(vitals$type)
print(y)

##
##  elective emergency
##        53       147
```

4c. count by procedure type

```
a<- table(vitals$type)
print(a)

##
##  elective emergency
##        53       147
```

4d. count by race and sex

```
z<- table(vitals$race, vitals$female)
 colnames(z)<- c("male", "female")
print(z)
```

```
##
##          male female
##   white   108     67
##   black    10      5
##   other     6      4
```

4e. count by consciousness

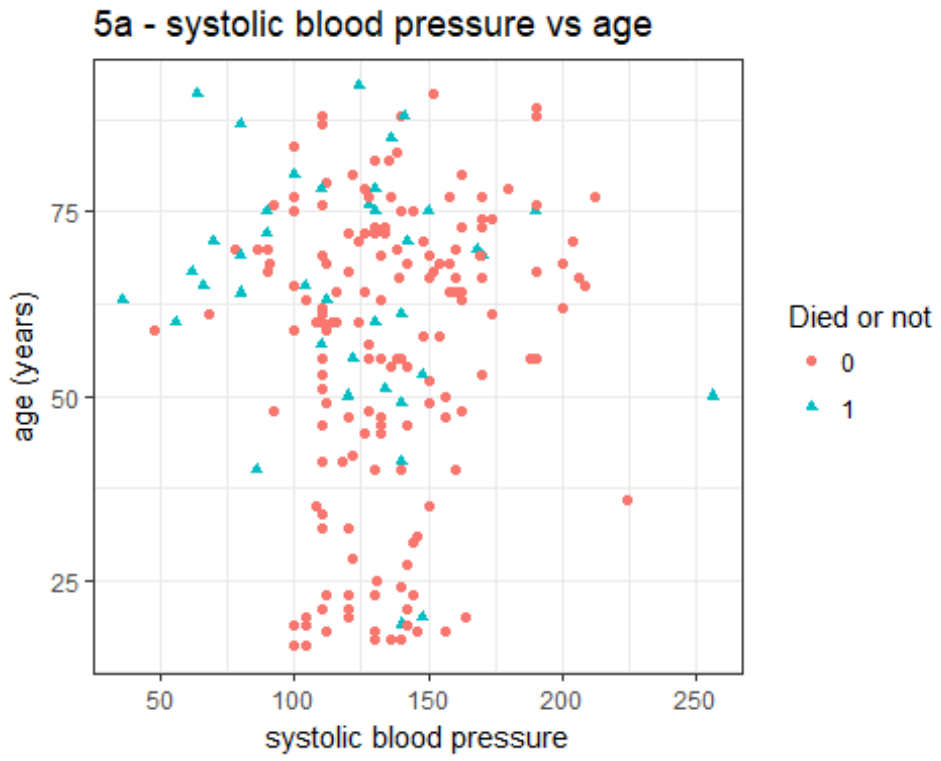
```
x<- table(vitals$conscious)
print(x)

##
## no_coma_or_stupor            stupor              coma
##               185                 5                10
```

5. graphical summaries

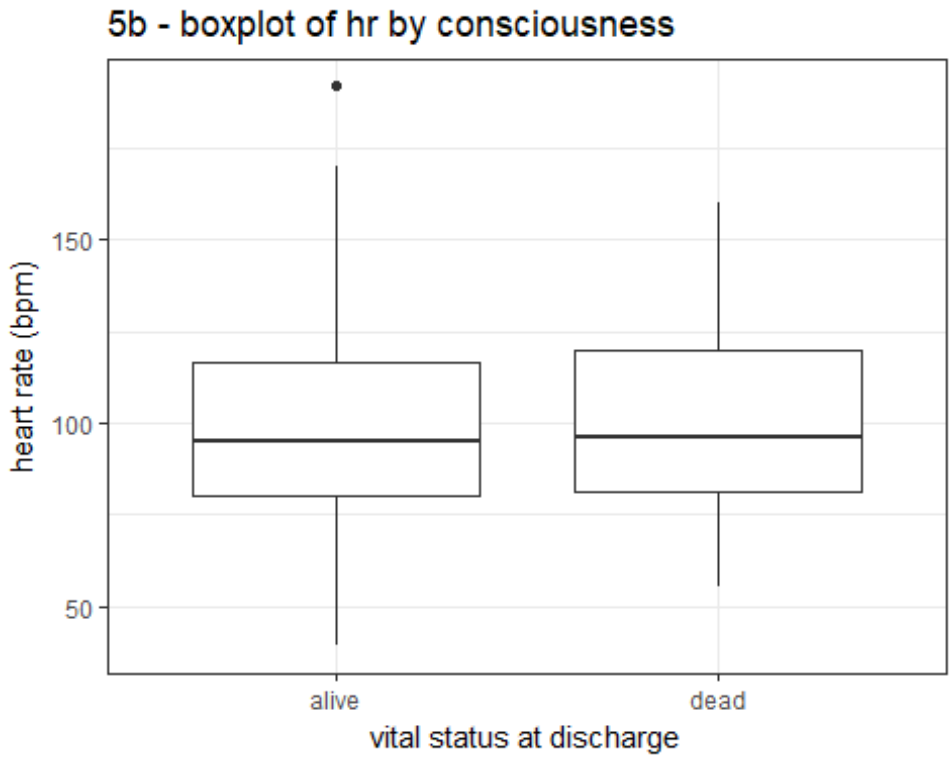
5a. scatterplot of sbp against age

```
ggplot(icu) +
  theme_bw() +
  geom_point(aes(x=sbp,
                 y=age,
                 color= as.factor(died),
                 shape=as.factor(died))) +
  xlab("systolic blood pressure") +
  ylab("age (years)") +
  ggtitle("5a - systolic blood pressure vs age") +
  scale_color_discrete(name="Died or not") +
  scale_shape_discrete(name="Died or not")  #having the same text for legens
combines them
```



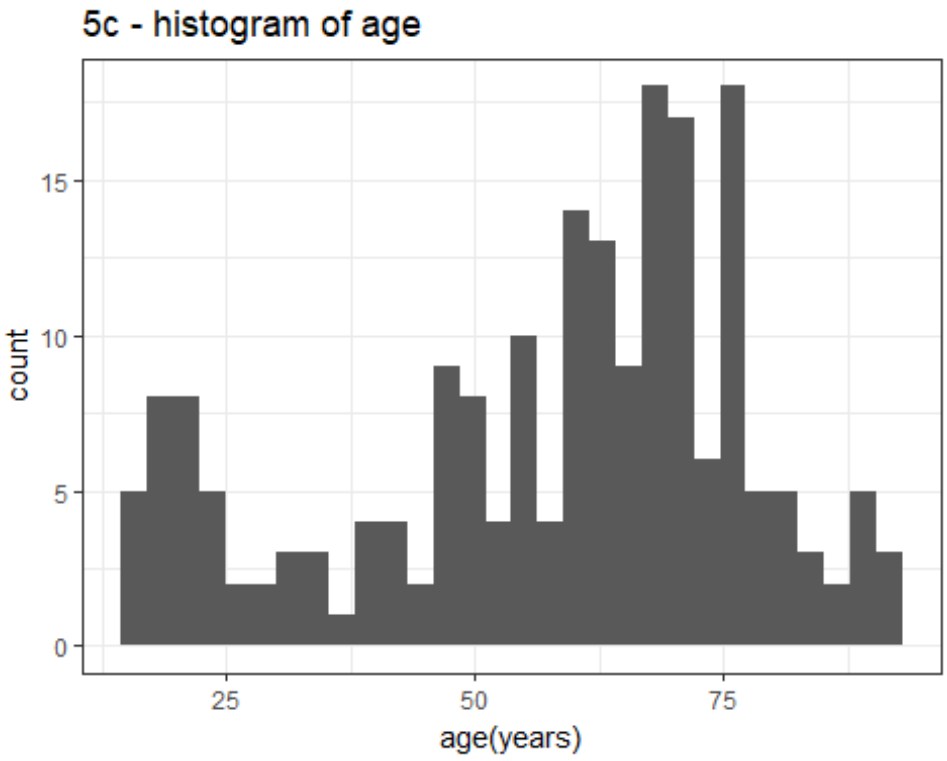
5b. boxplots of hr, grouped by consciousness

```
ggplot(icu) +
  theme_bw() +
  geom_boxplot(aes(group = as.factor(died), x=as.factor(died), y= hr)) +
  xlab("vital status at discharge") +
  ylab("heart rate (bpm)") +
  ggtitle("5b - boxplot of hr by consciousness") +
  scale_x_discrete(labels = c("alive", "dead"))
```



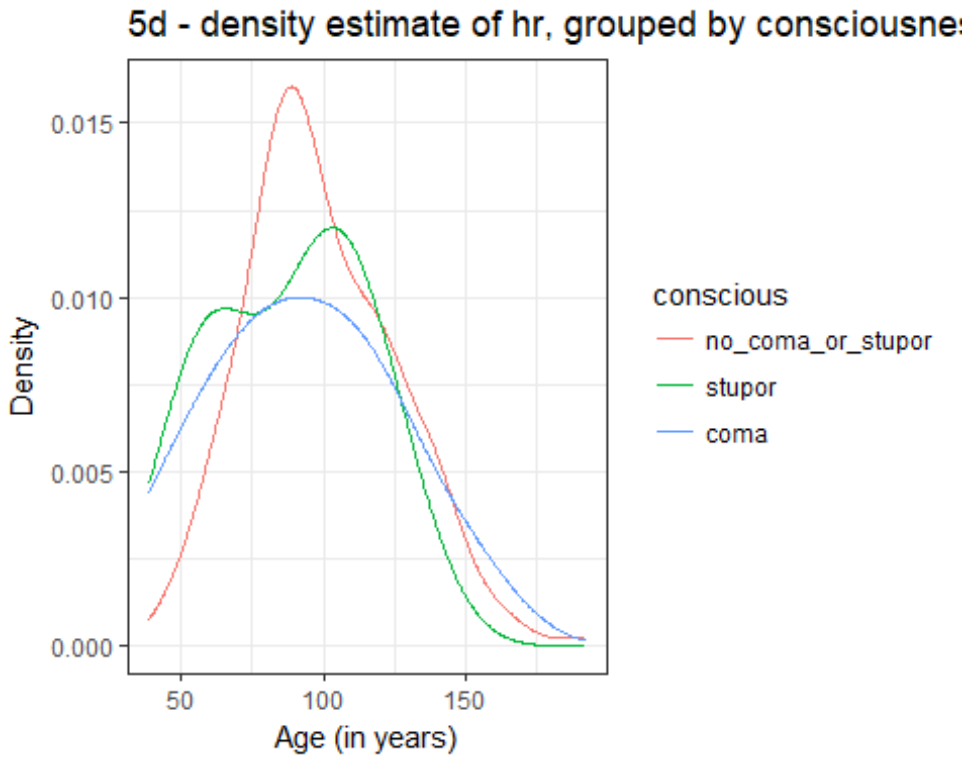
5c. histograms of age

```
ggplot(icu) +
  theme_bw() +
  geom_histogram(aes(x=age)) +
  xlab("age(years)") +
  ylab("count") +
  ggtitle("5c - histogram of age")
```



5d. densit estimates of hr, grouped by consciousness

```
ggplot(icu) +
  theme_bw() +
  geom_line(aes(x=hr, col=as.factor(conscious), group=conscious),
stat="density") +
  xlab("Age (in years)") +
  ylab("Density") +
  scale_color_discrete("conscious") +
  ggtitle("5d - density estimate of hr, grouped by consciousness") #from
notes
```


5d - density estimate of hr, grouped by consciousnes
0.015
0.010
0.005
0.000
Density
50
100
150
Age (in years)
conscious
no_coma_or_stupor
stupor
coma