hw6-nathan-mehta

Nathan Mehta

4/26/2022

```
library(tidyverse)
```

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

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.4     v dplyr   1.0.7
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   2.0.1     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

NBA players data

```
dat<-read.csv("Downloads/nba2018-players.csv")
```

```
#dat
```

```
#names(dat)
```

2

a)

```
dat%>%filter(height<70)
```

```
##          player team position height weight age experience
## 1 Isaiah Thomas  BOS       PG     69    185  27          5
## 2    Kay Felder  CLE       PG     69    176  21          0
##                    college  salary games minutes points points3 points2 points1
## 1 University of Washington 6587132    76    2569   2199     245     437     590
## 2       Oakland University  543471    42     386    166       7      55      35
```

b)

```
dat%>%filter(team == "GSW" & position == "C")
```

```
##             player team position height weight age experience
## 1 Anderson Varejao  GSW        C     82    273  34         12
## 2     Damian Jones  GSW        C     84    245  21          0
## 3       David West  GSW        C     81    250  36         13
## 4     JaVale McGee  GSW        C     84    270  29          8
## 5     Kevon Looney  GSW        C     81    220  20          1
## 6    Zaza Pachulia  GSW        C     83    270  32         13
##                                  college  salary games minutes points points3
## 1                                        1551659    14      92     18       0
## 2                  Vanderbilt University 1171560    10      85     19       0
## 3                      Xavier University 1551659    68     854    316       3
## 4             University of Nevada, Reno 1403611    77     739    472       0
## 5 University of California, Los Angeles 1182840    53     447    135       2
## 6                                        2898000    70    1268    426       0
##   points2 points1
## 1       5       8
## 2       8       3
## 3     132      43
## 4     208      56
## 5      54      21
## 6     164      98
```

c)

```
dat%>%filter(experience > 10 & salary >= 15000000) %>% select(c("player","age","team"))
```

```
##            player age team
## 1    LeBron James  32  CLE
## 2   Dwight Howard  31  ATL
## 3     Dwyane Wade  35  CHI
## 4 Carmelo Anthony  32  NYK
## 5       Pau Gasol  36  SAS
## 6      Chris Paul  31  LAC
## 7   Dirk Nowitzki  38  DAL
## 8       Luol Deng  31  LAL
```

d)

```
dat%>%filter(experience == 0 & age == 20) %>% select(c("player","team","height","weight")) %>% arrange(
```

```
##              player team height weight
## 1 Stephen Zimmerman  ORL     84    240
## 2    Henry Ellenson  DET     83    245
## 3  Domantas Sabonis  OKC     83    240
## 4     Deyonta Davis  MEM     83    237
## 5   Skal Labissiere  SAC     83    225
```

e)

```
dat%>%arrange(desc(salary/minutes))%>% select(c("player","points"))%>%slice_head(n=5)
```

```
##             player points
## 1     Danuel House      0
## 2      Roy Hibbert      4
## 3     Andrew Bogut      0
## 4 Chris McCullough      1
## 5    Brice Johnson      4
```

f)

```
gsw_mpg<-dat%>%filter(team == "GSW")%>%mutate(min_per_game = minutes/games)%>%select(c("player","experi
gsw_mpg
```

```
##                  player experience min_per_game
## 1         Klay Thompson          5    33.961538
## 2         Stephen Curry          7    33.392405
## 3          Kevin Durant          9    33.387097
## 4        Draymond Green          4    32.513158
## 5        Andre Iguodala         12    26.289474
## 6           Matt Barnes         13    20.500000
## 7         Zaza Pachulia         13    18.114286
## 8      Shaun Livingston         11    17.697368
## 9         Patrick McCaw          0    15.126761
## 10            Ian Clark          3    14.766234
## 11           David West         13    12.558824
## 12         JaVale McGee          8     9.597403
## 13 James Michael McAdoo          2     8.788462
## 14         Damian Jones          0     8.500000
## 15         Kevon Looney          1     8.433962
## 16     Anderson Varejao         12     6.571429
```

g)

```
dat%>%group_by(position)%>% summarise(pos_avg_pts = mean(points), pos_std_pts= sd(points))
```

```
## # A tibble: 5 x 3
##   position pos_avg_pts pos_std_pts
##   <chr>          <dbl>       <dbl>
## 1 C               477.        447.
## 2 PF              430.        397.
## 3 PG              573.        600.
## 4 SF              541.        516.
## 5 SG              536.        462.
```

h)

```
dat%>%group_by(team)%>%summarise(med_of_team = median(points))
```

```
## # A tibble: 30 x 2
##    team  med_of_team
##    <chr>       <dbl>
##  1 ATL          335
##  2 BOS          515
##  3 BRK          428
##  4 CHI          306.
##  5 CHO          458.
##  6 CLE          265
##  7 DAL          200.
##  8 DEN          587
##  9 DET          365
## 10 GSW          408.
## # ... with 20 more rows
```

i)

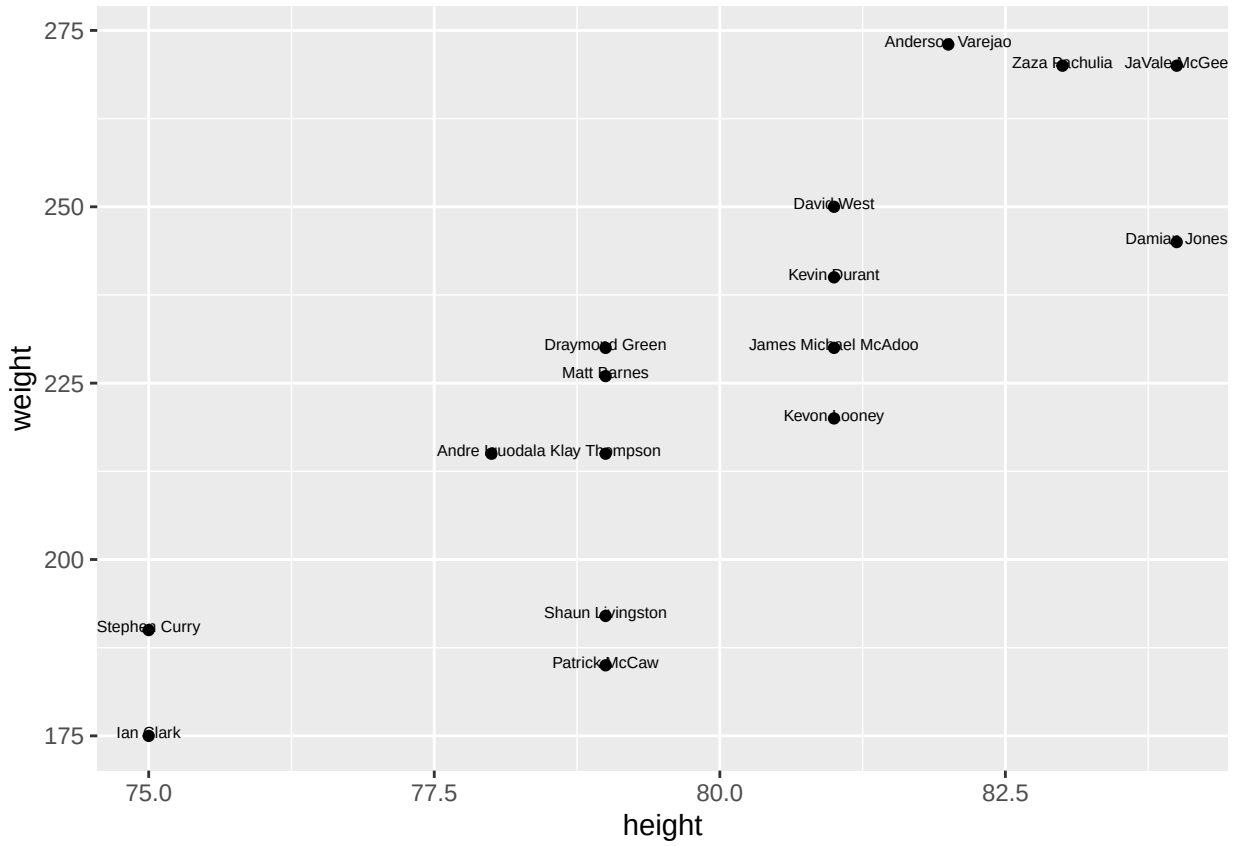
```
dat%>%group_by(team)%>%mutate(sum_of_team = sum(points)) %>%ungroup()%>%mutate(percentage_of_team_pts =
```

```
## # A tibble: 5 x 2
##   player              salary
##   <chr>                <dbl>
## 1 Anthony Davis     22116750
## 2 Russell Westbrook 26540100
## 3 James Harden      26540100
## 4 Isaiah Thomas      6587132
## 5 DeMar DeRozan     26540100
```

3

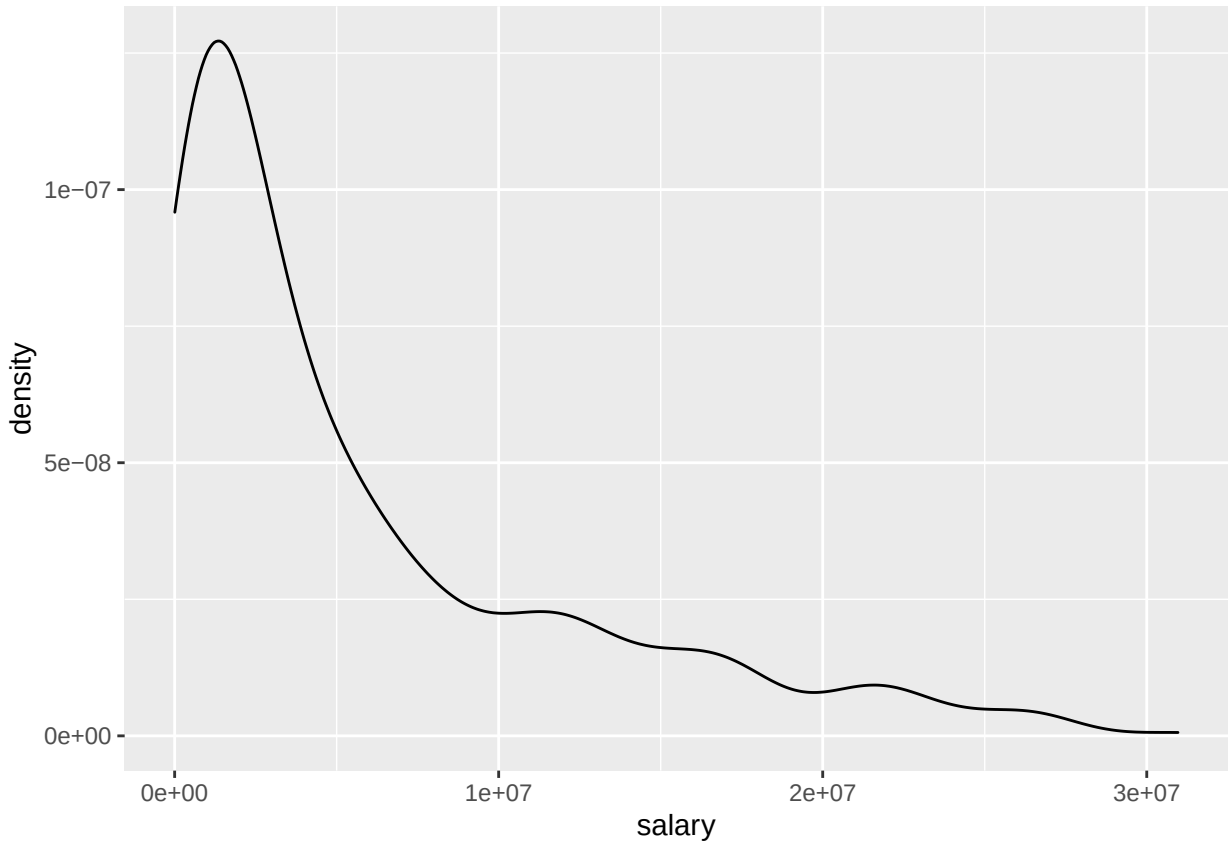
a)

```
ggplot(data = dat%>%filter(team == "GSW"), aes(x = height, y = weight)) +
geom_point() + geom_text(size=2, aes(label=player), nudge_y = .5)
```



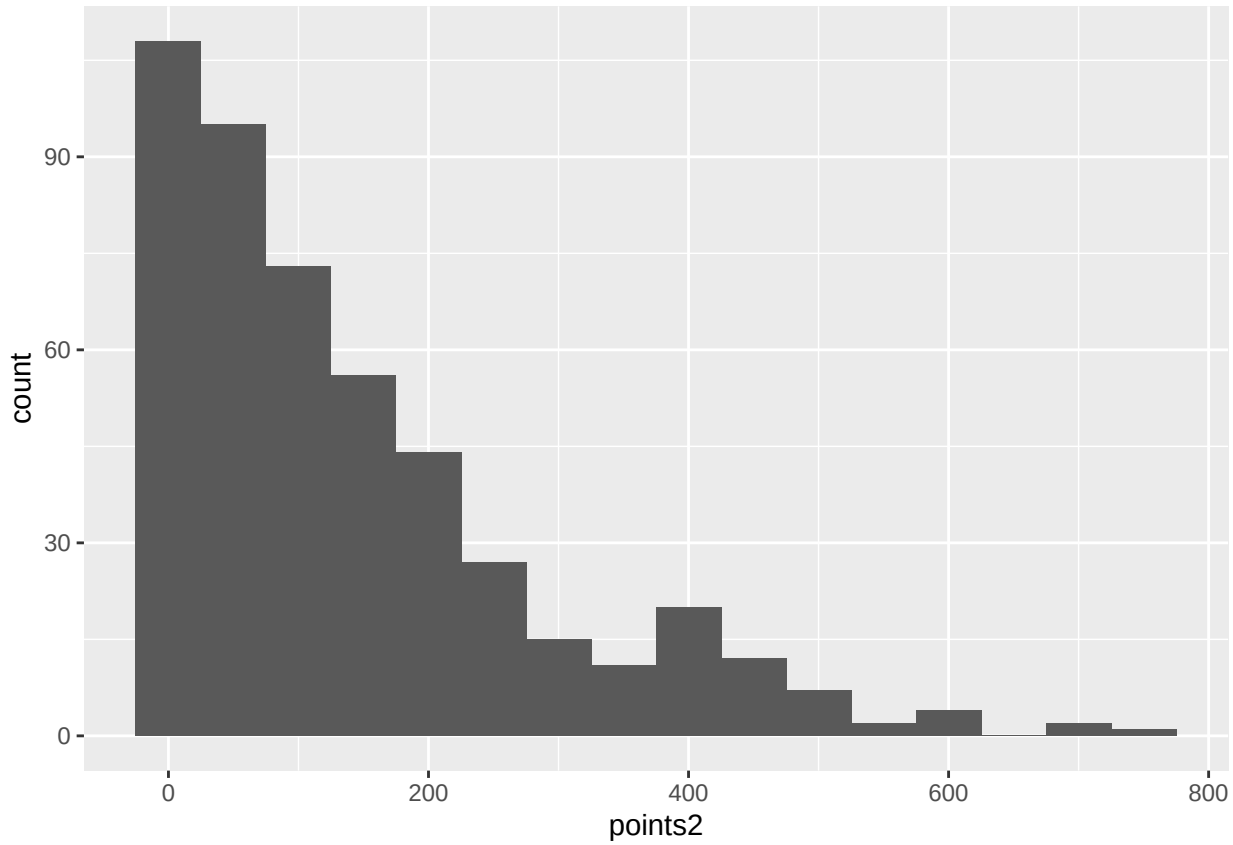
b)

```
ggplot(data = dat, aes(x = salary)) +
geom_density()
```



c)

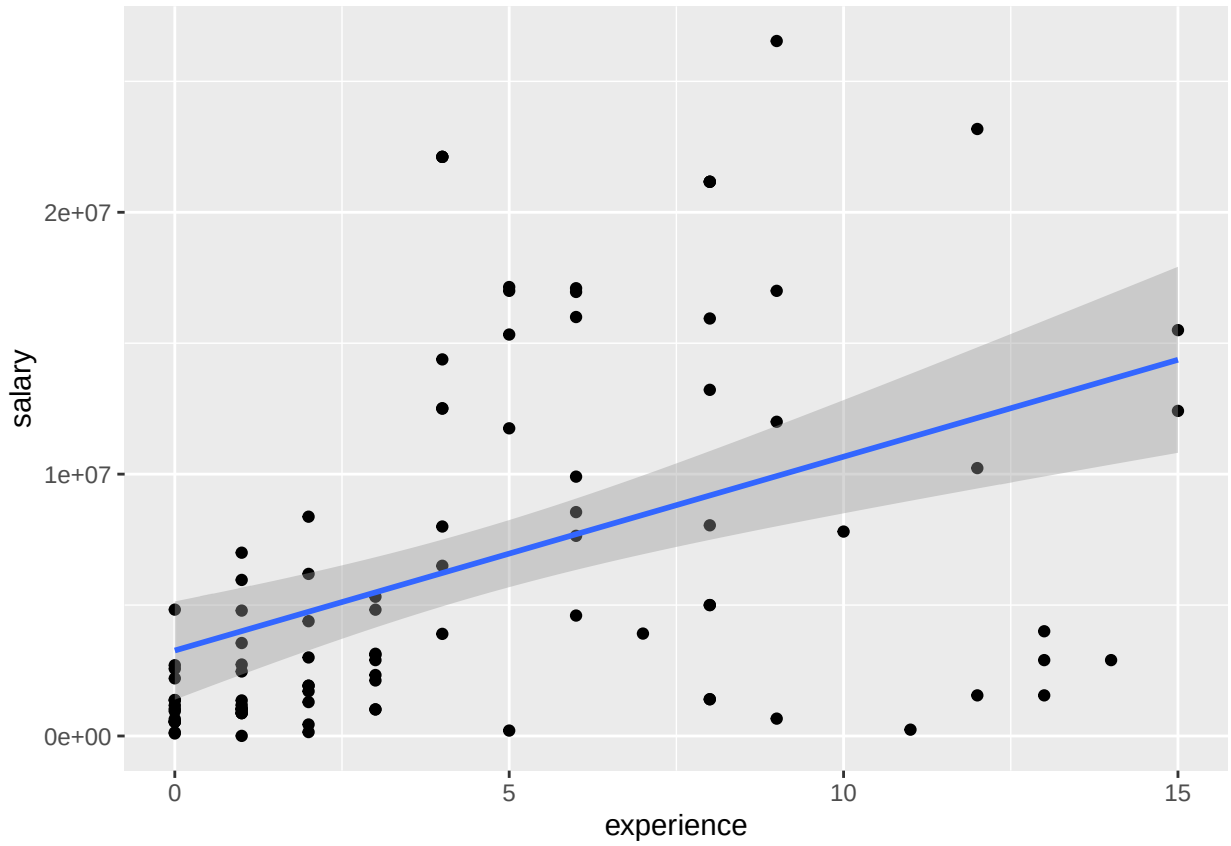
```
ggplot(data = dat, aes(x = points2)) +
geom_histogram(binwidth = 50)
```



d)

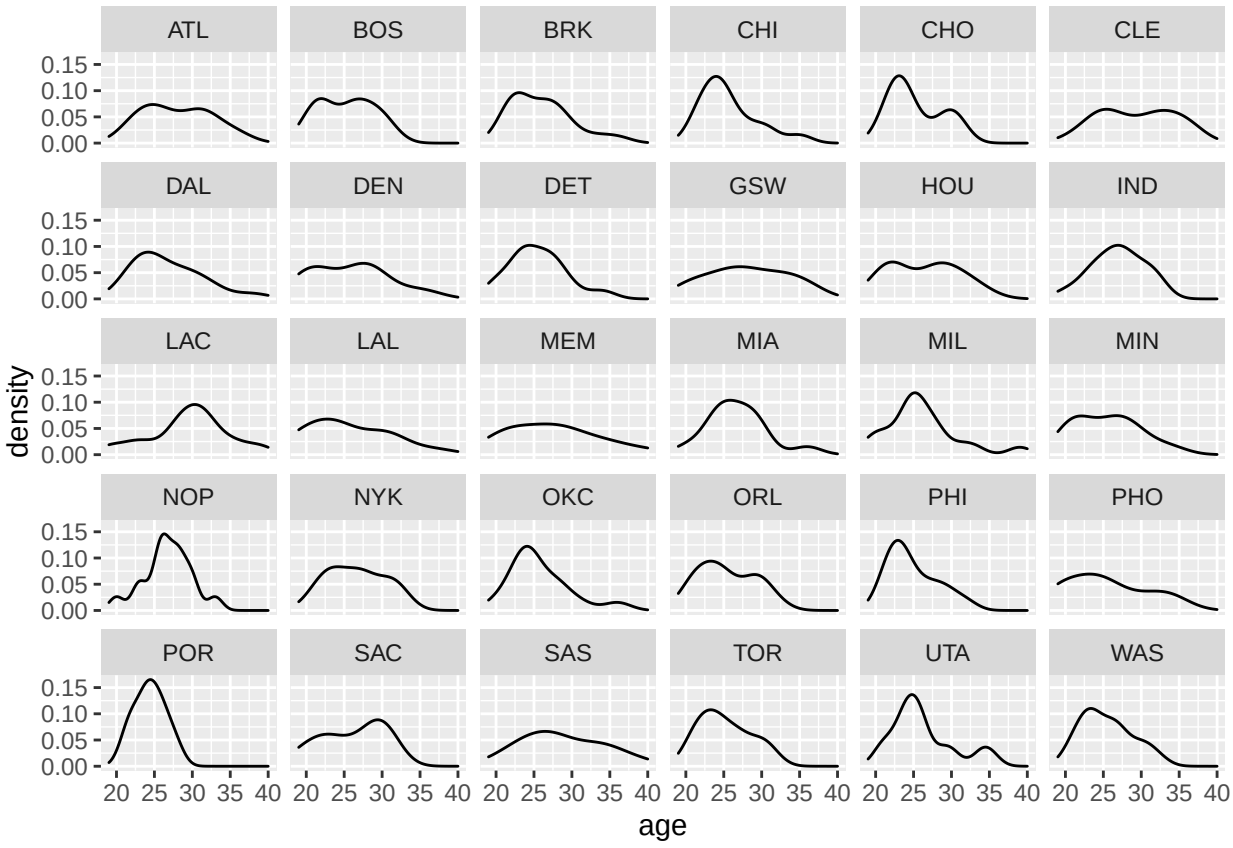
```
ggplot(data = dat%>%filter(position == "C"), aes(x = experience, y = salary)) +
geom_point() +
geom_smooth(method = 'lm')
```

```
## `geom_smooth()` using formula 'y ~ x'
```



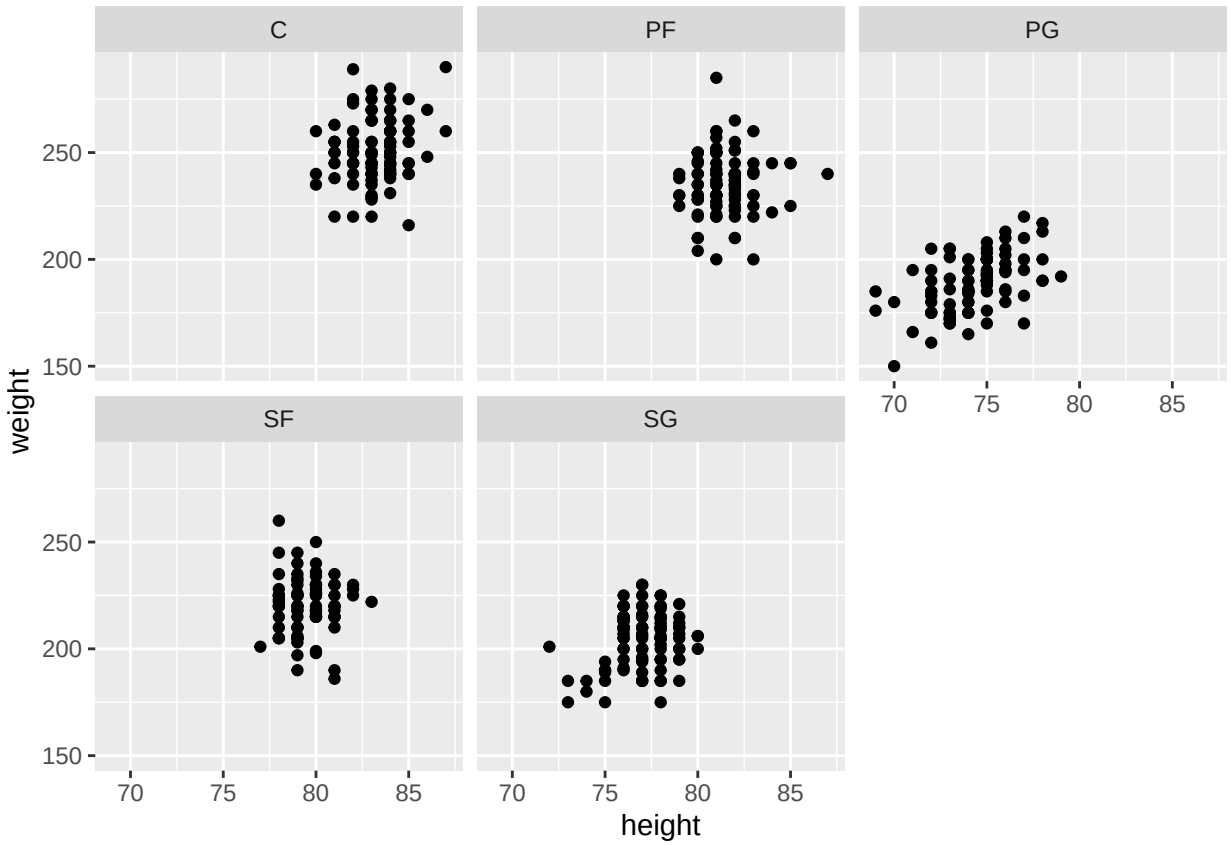
e)

```
ggplot(data = dat, aes(x=age)) + geom_density() +
facet_wrap(~ team)
```



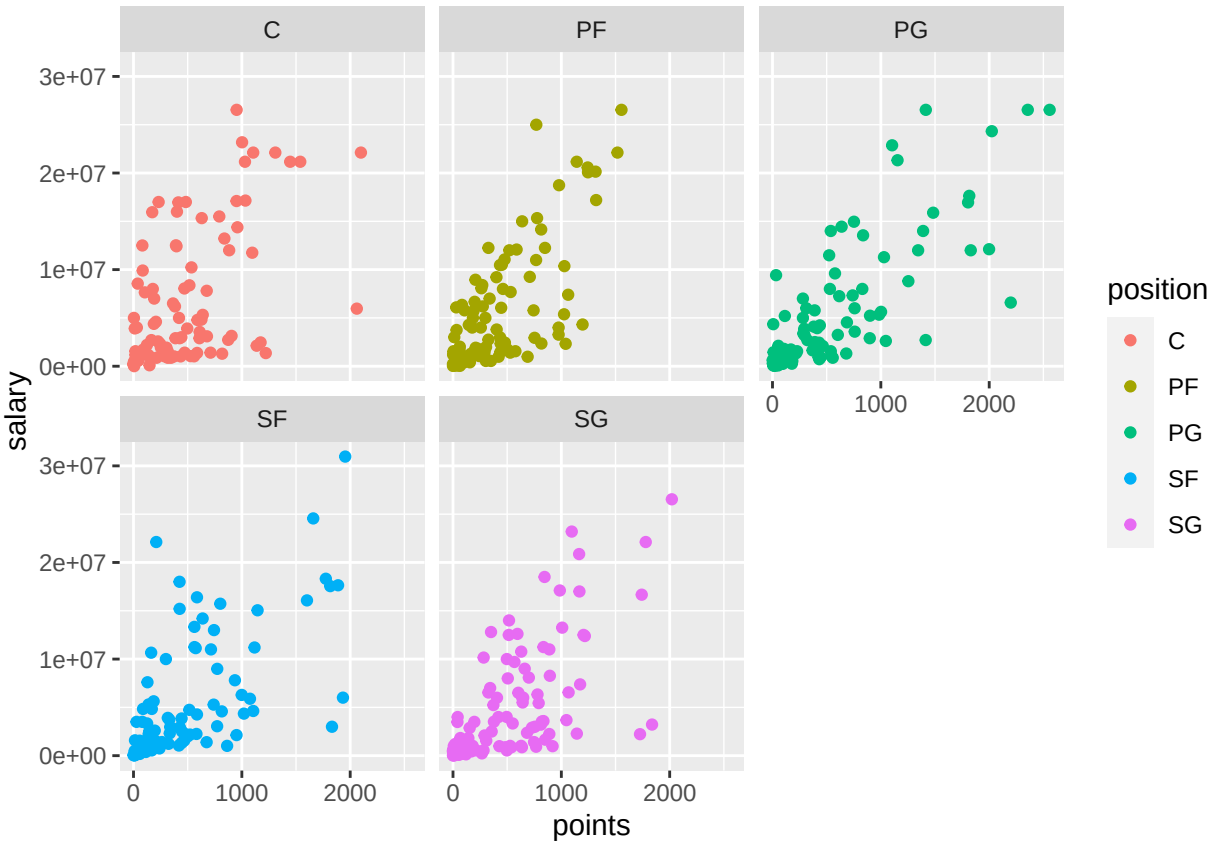
f)

```
ggplot(data = dat, aes(x = height, y = weight)) +
geom_point() +
facet_wrap(~ position)
```



g)

```
ggplot(data = dat, aes(x = points, y = salary, color = position)) +
geom_point() +
facet_wrap(. ~ position)
```


C
PF
PG
SF
SG
salary
points
3e+07
2e+07
1e+07
0e+00
0
1000
2000
position
C
PF
PG
SF
SG