Problem Set 1:Height~Position for Indiana Pacers (2017-2018)

Sumeet Mishra

1/12/2019

Discussed With Team Members:Kimball Wu,Pei-Yi Cheng

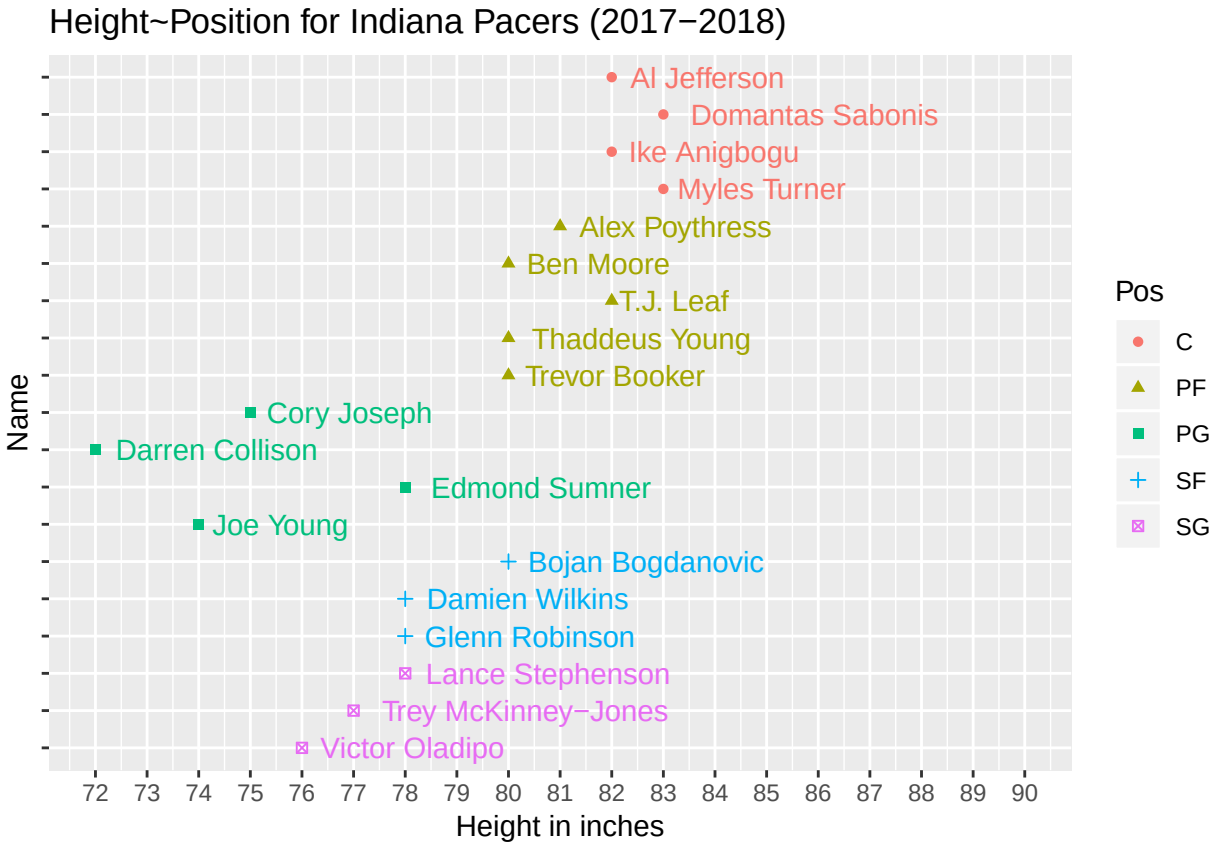
```
library(ggplot2)
pacers=read.csv('pacers2017-18.csv', header = TRUE)
nba=read.csv('nba2017-18.csv', header = TRUE)
#View(pacers)
#View(nba)
```

The ggplot2 package was read and the datasets 'pacers2017-18.csv','nba2017-18.csv' were loaded to variables pacers,nba respectively.

```
library(tidyr)
library(dplyr, warn.conflicts = FALSE)
pacers<-pacers %>%
  separate(Player,c("Name","USERID"),sep="\\\\")
pacers<-pacers %>%
  separate(Ht,c("Feet","Inches"),sep="-")
pacers$Height<-as.numeric(pacers$Feet)*12+as.numeric(pacers$Inches)
```

The tidyr and dplyr package was loaded for data cleaning and processing. The Player column in pacers dataset was separated into 2 columns named as Name and UserID and the Ht column was separated into Feet and Inches. Then the feet column was converted to inches by multiplying 12 to it and then Height column was created on pacers dataset by adding inches column values to the previously converted feet into inches values.

```
o=order(pacers$Pos,pacers$Name,decreasing = TRUE)
pacers$Name=factor(pacers$Name,levels=pacers$Name[o])
ggplot(pacers,aes(x=Height,y=Name,color=Pos))+geom_point(aes(shape=Pos,color=Pos))+
  geom_text(aes(label=Name),position=position_dodge(width=0.2),hjust=-0.1,
            show.legend = FALSE)+
  scale_x_continuous(limits=c(72,90),breaks=seq(72,90))+xlab('Height in inches')+
  ggtitle('Height~Position for Indiana Pacers (2017-2018)')+
  theme(axis.text.y=element_blank())
```



The pacers dataset was ordered in decreasing order of position(Pos) and names. In the plot, different shapes and color were used for different positions. The names of the players were displayed along with the points which each represents a player.