Player Anaylsis

\*导入数据，

setwd("/Users/sun/Desktop/NBA/")  
player <- read.csv("/Users/sun/Desktop/NBA/player\_data.csv",header = T,na.strings = "")#用na.strings=""来将数据中的空缺值变成NA,方便之后的处理  
player$height <- gsub('-','.',player$height)  
player$height <- as.numeric(player$height)  
head(player)

## name year\_start year\_end position height weight  
## 1 Alaa Abdelnaby 1991 1995 F-C 6.1 240  
## 2 Zaid Abdul-Aziz 1969 1978 C-F 6.9 235  
## 3 Kareem Abdul-Jabbar 1970 1989 C 7.2 225  
## 4 Mahmoud Abdul-Rauf 1991 2001 G 6.1 162  
## 5 Tariq Abdul-Wahad 1998 2003 F 6.6 223  
## 6 Shareef Abdur-Rahim 1997 2008 F 6.9 225  
## birth\_date college  
## 1 June 24, 1968 Duke University  
## 2 April 7, 1946 Iowa State University  
## 3 April 16, 1947 University of California, Los Angeles  
## 4 March 9, 1969 Louisiana State University  
## 5 November 3, 1974 San Jose State University  
## 6 December 11, 1976 University of California

summary(player)

## name year\_start year\_end position   
## Charles Jones : 3 Min. :1947 Min. :1947 G :1574   
## Charles Smith : 3 1st Qu.:1969 1st Qu.:1973 F :1290   
## George Johnson: 3 Median :1986 Median :1992 C : 502   
## Bill Bradley : 2 Mean :1985 Mean :1989 F-C : 388   
## Bob Duffy : 2 3rd Qu.:2003 3rd Qu.:2009 G-F : 360   
## Bobby Jones : 2 Max. :2018 Max. :2018 (Other): 435   
## (Other) :4535 NA's : 1   
## height weight birth\_date   
## Min. :5.100 Min. :114.0 August 26, 1947 : 3   
## 1st Qu.:6.200 1st Qu.:190.0 August 31, 1952 : 3   
## Median :6.500 Median :210.0 August 8, 1979 : 3   
## Mean :6.459 Mean :208.9 December 3, 1985 : 3   
## 3rd Qu.:6.700 3rd Qu.:225.0 December 31, 1923: 3   
## Max. :7.700 Max. :360.0 (Other) :4504   
## NA's :1 NA's :6 NA's : 31   
## college   
## University of Kentucky : 99   
## University of California, Los Angeles: 91   
## University of North Carolina : 87   
## University of Kansas : 72   
## Duke University : 71   
## (Other) :3828   
## NA's : 302

player数据集中共有9个变量，分别为运动员名字（字符型变量），职业生涯起点（数值型变量），职业生涯结尾（数值型变量）（由于数据涵盖的是1947-2019年中所有运动员的信息，所以运动员的职业生涯起点最晚记录到2018，截止日期也是2018），球队位置（字符型变量），球员身高（英尺-英寸），体重（单位磅），出生日期（字符型），就读学校（字符型变量）。

#球员在球队的位置分布如图， 1号位控球后卫，也叫组织后卫:簡稱“控衛”或“控後”，俗稱“一號位置”，有時也稱組織後衛，是籃球比赛陣容中的一個固定位置。控球後衛往往是全隊進攻的组織者，透過對球的控制來决定在適當的时機传球给适合的球员。控球后卫典型的进攻模式就是在對手得分之後，由控球后卫從底线运球，开始一轮新的进攻。这个位置要求球员具有良好的传球技术和敏锐的比赛观察能力。人们往往通过助攻次数而不是得分的高低来衡量一名控球后卫的成功与否。同时，一流的控球后卫往往也能够有效地跳投，并能够藉由外线投篮威胁对手。一般控球后卫都是各队阵容中身材最小巧的球员，篮球比赛在发球后，通常由控球后卫通过运球将球运过半场并将球传给队友来发动进攻。所以一般担任控球后卫的队员都是队内控球能力最好，助攻能力最强的球员。著名的控球后卫包括以賽亞·湯瑪斯、加里·佩顿、魔術強森、约翰·斯托克顿、史提夫·奈許、傑森·基德。目前，聯盟著名的控球後衛有本·西蒙斯、克里斯·保羅、拉塞尔·威斯布鲁克、德里克·羅斯、肯巴·沃克、凯里·欧文、約翰·沃爾、達米恩·李拉德以及史蒂芬·柯瑞

2号位得分后卫：是籃球比赛陣容中，五個固定位置之一。傳統的得分後衛的主要責任，就是盡可能的投進球。一般擔任該位置的球員對體形、體能的要求較嚴格，在速度和爆發力上具有優勢。他們經常是由隊伍裡投射能力最強的球員擔任，處理多數的中遠距離出手，在外圍防守上也責任重大。許多得分後衛都可以兼任小前锋，稱為鋒衛搖擺人。迈克尔·乔丹被認為是NBA歷史上最出色的得分後衛，使這個位置的功能更加明確。

3号位小前锋：

3小前锋；4大前锋（力量前锋）抢篮板，防守；5中锋；Point Guard(控球后卫)，简称PG。 Shooting Guard(得分后卫)，简称SG。 Small Forward(小前锋)，简称SF。 Power Forward(强力前锋)，简称PF。 Center(中锋)，简称C，在这里控球后卫和得分后卫简称G(后卫)，小前锋和大前锋简称(F),若某球员既可以做前锋，也可以做中峰，将其位置表示为F-C.

**数据可视化**首先将球星所在的大学进行统计，以table的形式展示，通过词云函数得到

library(wordcloud2) library(dplyr) college\_cloud <- as.data.frame(table(player$college)) wc = wordcloud2(college\_cloud,size=0.1,shuffle=2) #由于我们的字符串较长，所以需要调整比例才能让全部的字符显示在图片中 wc+WCtheme(2)

``` 每年的选秀之夜，都会有一批新的NBA新秀听到自己的名字，希望借此机会开始自己的职业篮球生涯。许多像杜克大学这样的强队项目帮助大学生为他们的职业生涯做准备。还有很多排名在杜克大学之前的顶尖大学，如肯塔基大学、加州大学洛杉矶分校和北卡罗来纳大学等也提供培养篮球运动员的项目。上图统计得到培养了最多的篮球运动员的大学有：

肯塔基大学：拥有99名篮球运动员，野猫队不仅培养了最多的NBA球员，而且在历史上也是美国顶尖的大学篮球项目之一。肯塔基大学已经赢得了8次全国冠军，这是历史上排名第二的冠军，在学校历史上有17次进入决赛。自2009年教练约翰·卡利帕里(John Calipari)上任以来，肯塔基州的战绩是305胜71负。像安东尼·戴维斯、德马库斯·考辛斯、埃里克·布莱索和德文·布克这样的顶级球员代表着今天的NBA。

加州大学洛杉矶分校：拥有91名球员 由于传奇教练约翰·伍德和他强有力的招募，加州大学洛杉矶分校以11个总冠军保持着全美大学生体育协会冠军的记录。一些NBA的传奇人物，如卡里姆·阿卜杜尔·贾巴尔和比尔·沃顿，在成为职业球员之前都在伍登手下打球。

杜克大学：拥有71球员 杜克已经参加了41次大型赛事，并进入了16次决赛。在传奇教练迈克·沙舍夫斯基的带领下，蓝魔已经赢得了他们所有的五次全国冠军。著名的杜克球员包括凯里·欧文、格兰特·希尔、埃尔顿·布兰德、马文·巴格利和杰伊·威廉姆斯。

堪萨斯大学：拥有72的球员 堪萨斯州也有着惊人的历史，46次参加NCAA比赛，14次进入四强。松鹰队已经赢得了三次全国总冠军，几位著名的NBA球星的职业生涯也从此起步。包括威尔特·张伯伦，保罗·皮尔斯，安德鲁·威金斯，丹尼·曼宁和乔尔·恩比德。

印第安纳大学拥有60名球星，圣母大学有65名球星，路易斯维尔大学有55名球员。。。