

bar plots

Subramani.M

18 May 2018

Bar Charts

Vertical and Horizontal.

Day 1

```
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.4.3

library(dplyr)

## Warning: package 'dplyr' was built under R version 3.4.3

##
## Attaching package: 'dplyr'

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

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

data = read.csv("C:/Users/Administrator/Desktop/EDA/EDA TEST/odi-
batting.csv")
## Grouping by ground and selecting top 10 grounds in terms of number of
matches.

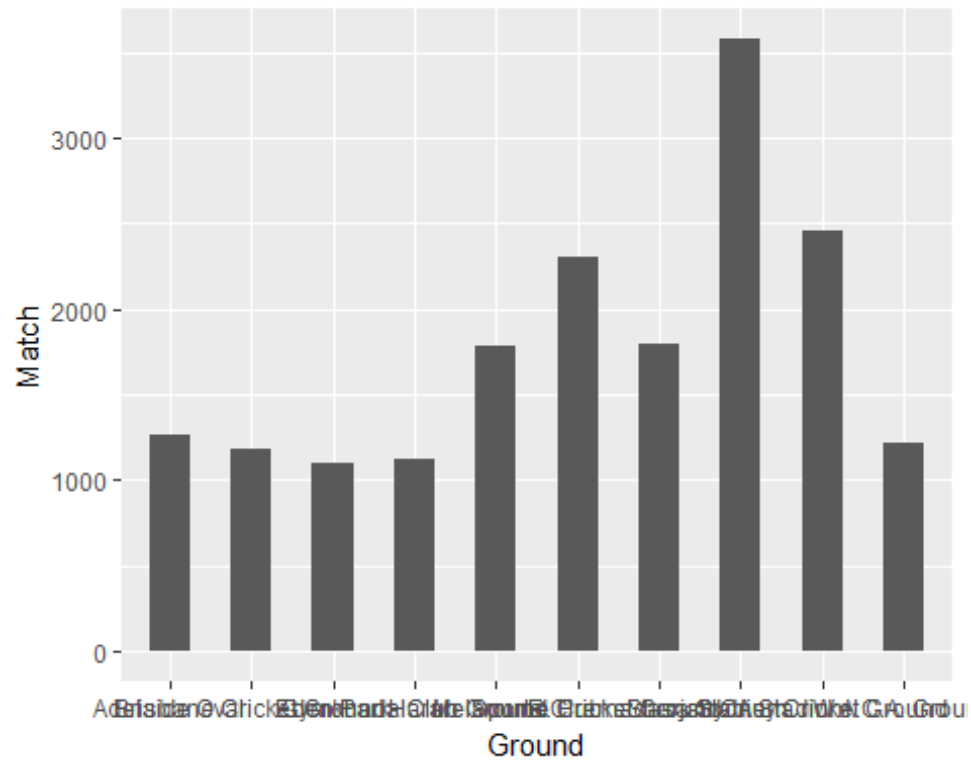
ground = data %>% group_by(Ground) %>% summarise(Match = n()) %>%
arrange(desc(Match))

## Warning: package 'bindrcpp' was built under R version 3.4.3

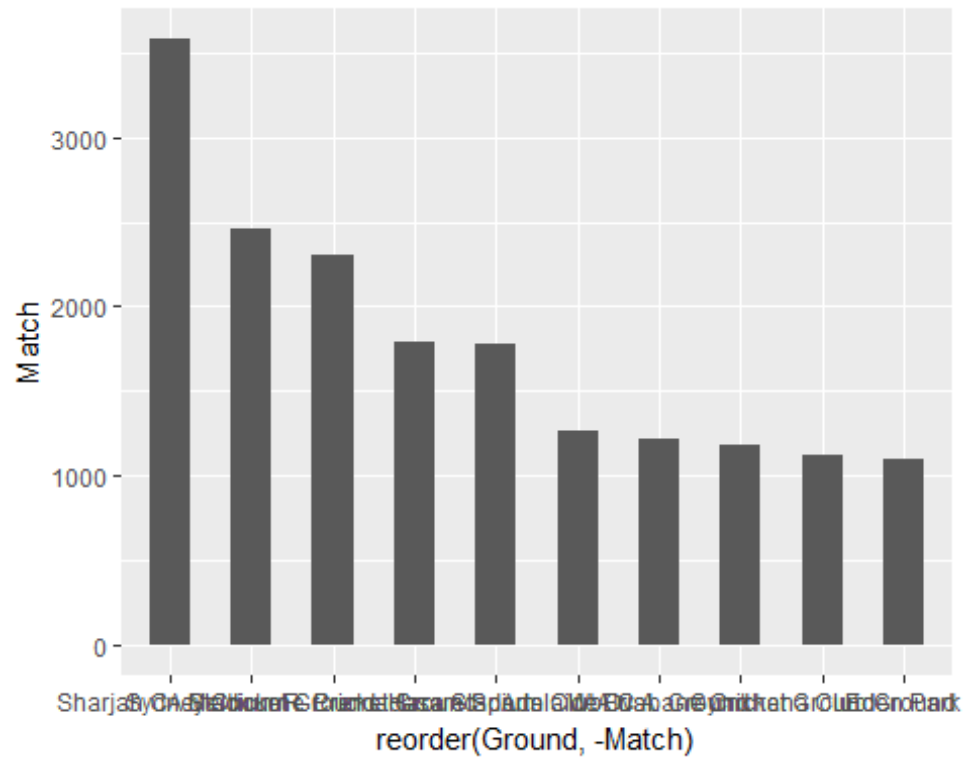
top_10 = head(ground, 10)
```

Plotting a vertical bar chart

```
P1 = ggplot(top_10, aes(x = Ground, y = Match)) + geom_bar(stat = "Identity",
width = .5)
P1
```



```
## Arranging the bars in Order and Bar width and giving a common color
P2 = ggplot(top_10, aes(x = reorder(Ground, -Match), y = Match)) +
  geom_bar(stat = "Identity",
width = .5)
P2
```

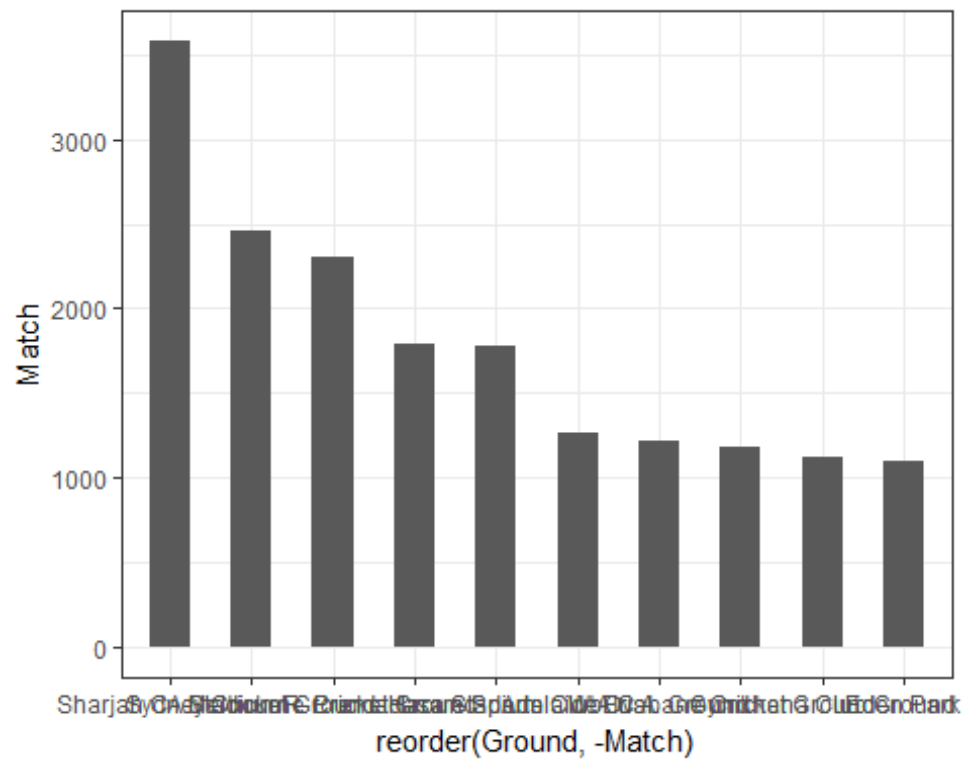


```
## Lets start the make - up of chart
```

```
# Removing the background grid
```

```
P3 = P2 + theme_bw()
```

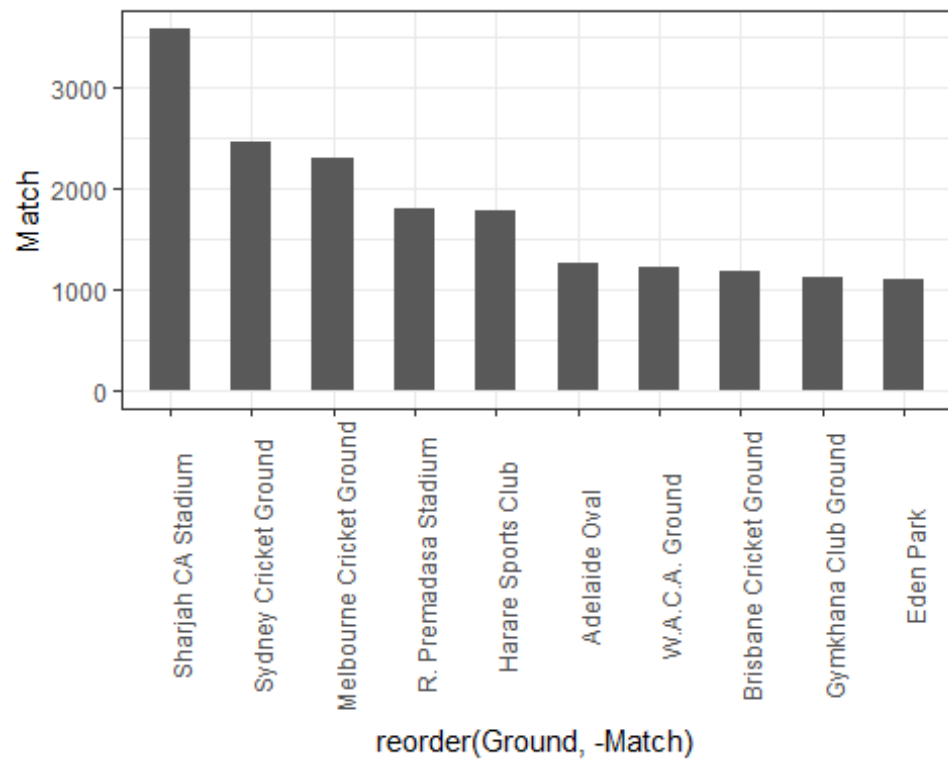
```
P3
```



```
# arranging the x - axis test
```

```
P4 = P3 + theme(axis.text.x = element_text(angle = 90))
```

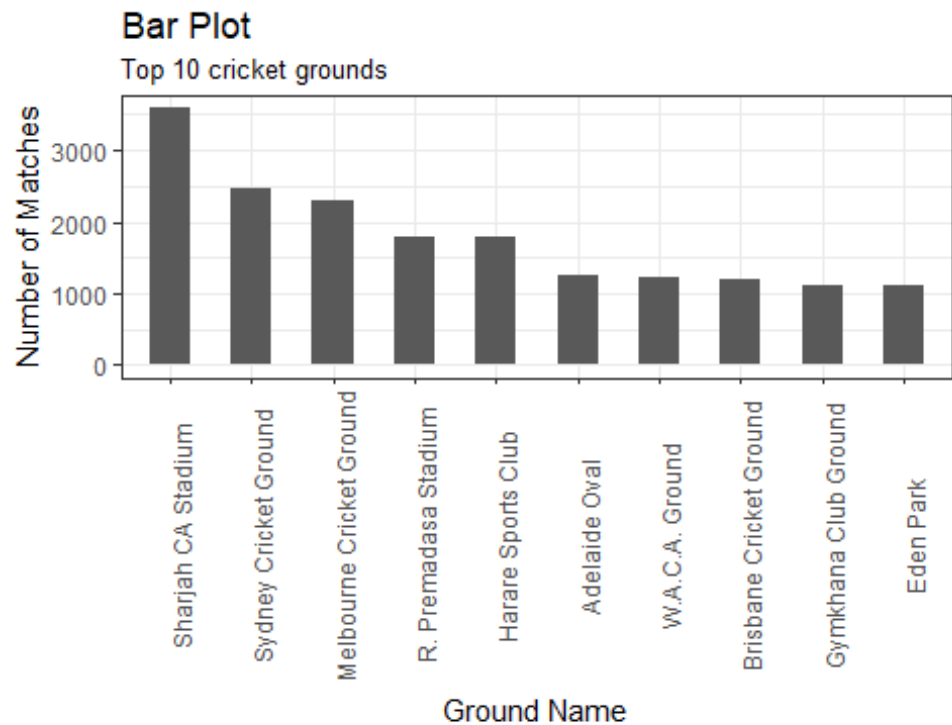
```
P4
```



```
# Renaming x-axis and Y-axis and plot name
```

```
P5 = P4 + labs(subtitle="Top 10 cricket grounds",
               y="Number of Matches", x="Ground Name", title="Bar Plot",
               caption = "Source: ESPN")
```

```
P5
```



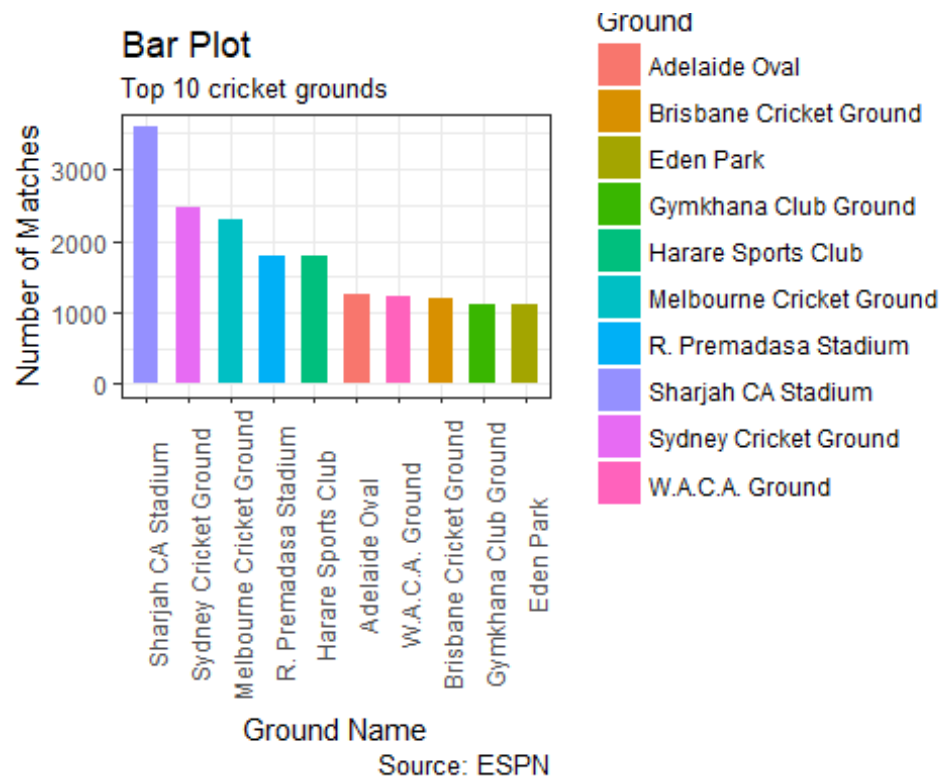
Source: ESPN

```
# Adding col and removing x axis
```

```
# we use stat = identity when y axis is specified
```

```
P6 = P5 + geom_bar(stat = "Identity", aes(fill = Ground), width = .6)
```

```
P6
```



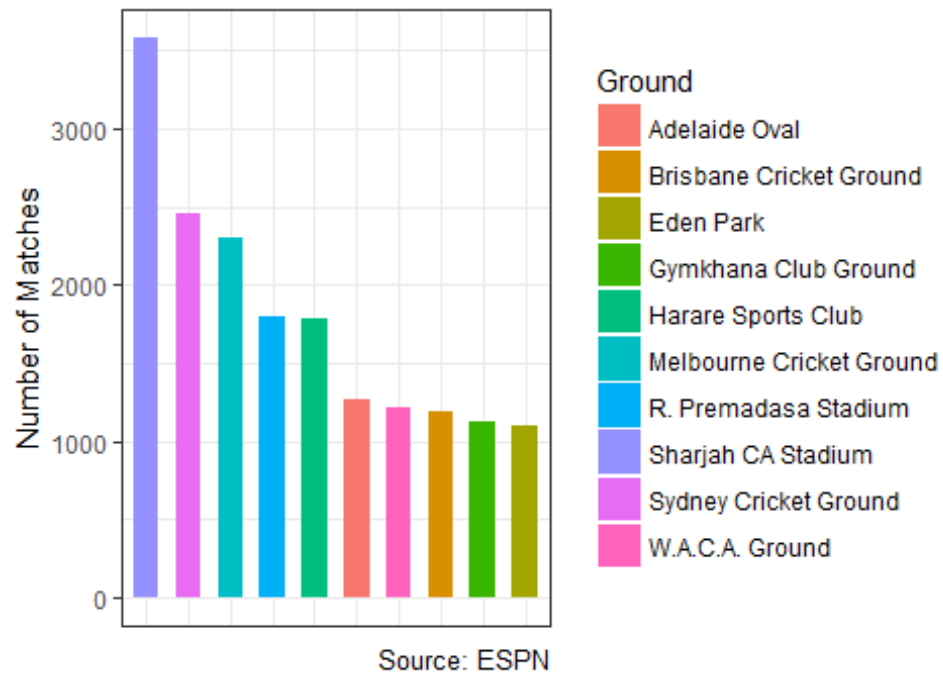
Now I don't need the X - Axis text

```
P7 = P6 + theme(axis.text.x=element_blank(), axis.title.x=element_blank(),
axis.ticks.x = element_blank())
```

P7

Bar Plot

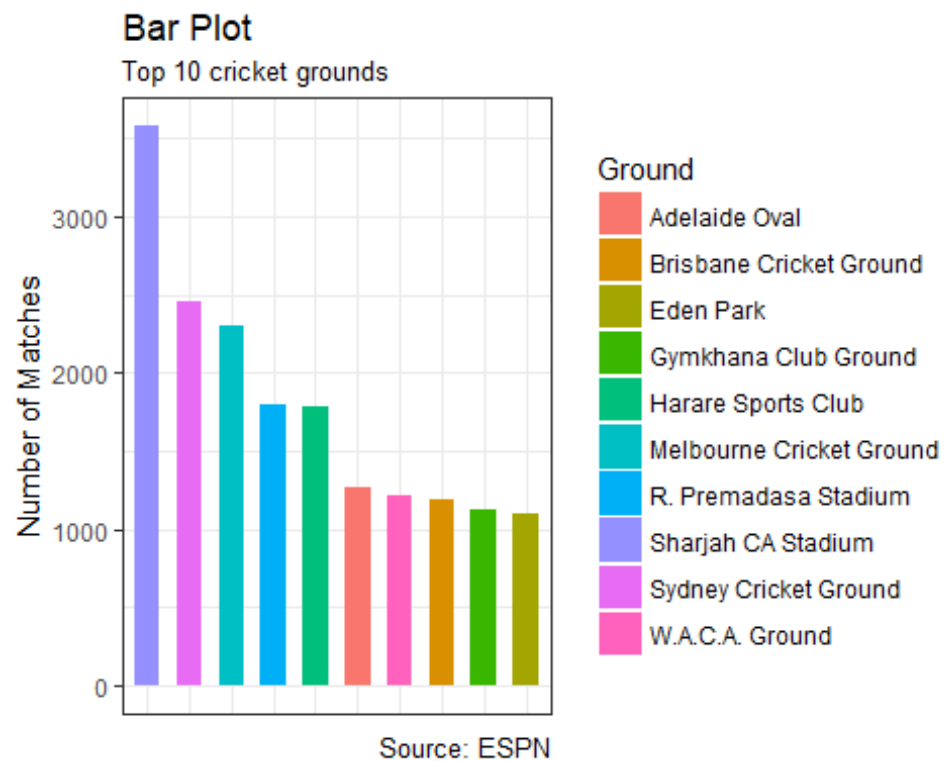
Top 10 cricket grounds



```
# Removing ticks from X - axis
```

```
P8 = P7 + theme(axis.ticks.x=element_blank())
```

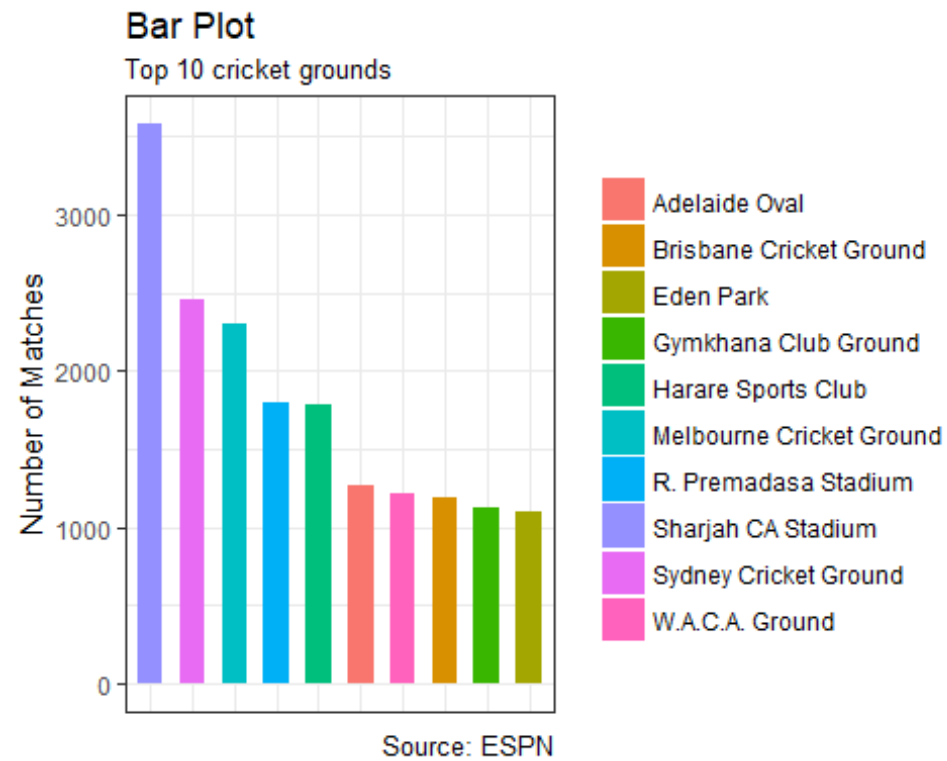
```
P8
```

```
## Adjusting the legends
# Hiding the legend title

P9 = P8 + theme(legend.title=element_blank())

P9
```



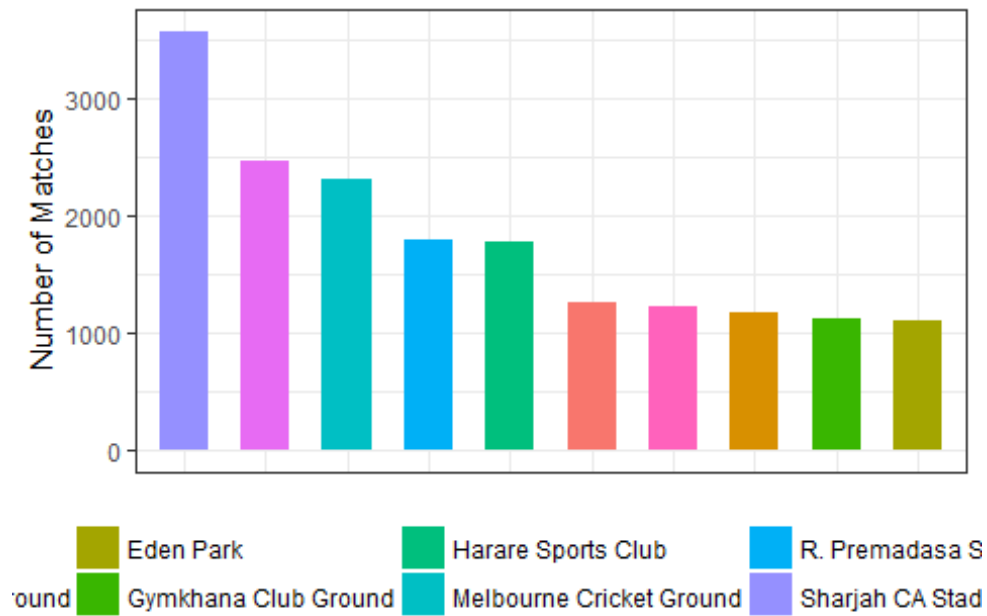
Changing the position of Legends

```
P10 = P9 + theme(legend.position = "bottom")
```

P10

Bar Plot

Top 10 cricket grounds



Source: ESPN

Changing the Legends size (text Plus blocks)

```
P11 = P10 + theme(legend.key.width = unit(.2, "cm"))
```

P11

Bar Plot

Top 10 cricket grounds



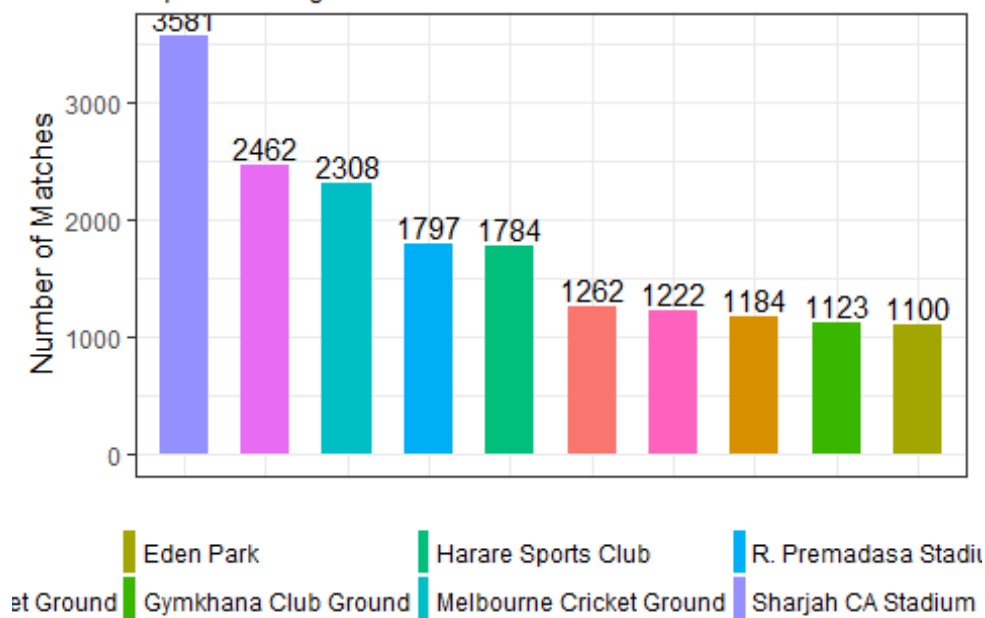
Source: ESPN

creating the data labels

```
P12 = P11 + geom_text(aes(label = Match), position = position_dodge(width =  
.9), vjust = -0.25)  
P12
```

Bar Plot

Top 10 cricket grounds



Source: ESPN

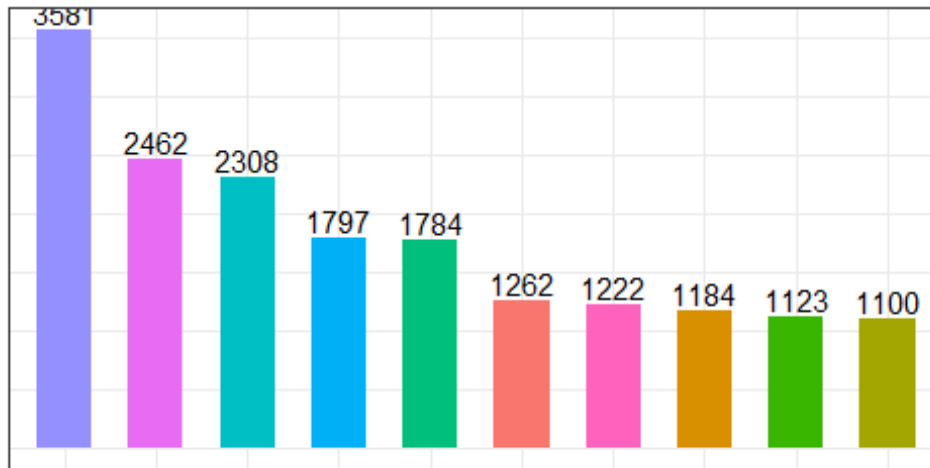
Removing Y - Axis

```
P13 = P12 + theme(axis.title.y = element_blank(), axis.text.y =  
element_blank(), axis.ticks.y =element_blank())
```

P13

Bar Plot

Top 10 cricket grounds



Eden Park Harare Sports Club R. Premadasa Stadium
Gymkhana Club Ground Melbourne Cricket Ground Sharjah CA Stadium

Source: ESPN

Reducing the size of Legends

```
P14 = P13 + theme(legend.text = element_text(colour = 'black', size = 8,  
hjust = 3, vjust = 3, face = 'bold'))
```

Using gridExtra

```
#install.packages("gridExtra")
```

```
library(gridExtra)
```

```
## Warning: package 'gridExtra' was built under R version 3.4.4
```

```
##
```

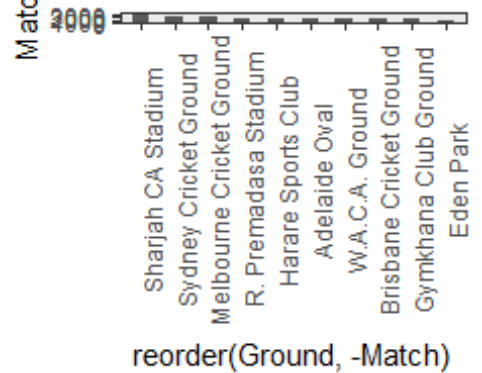
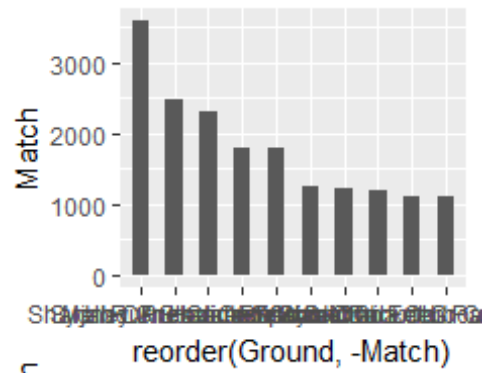
```
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
```

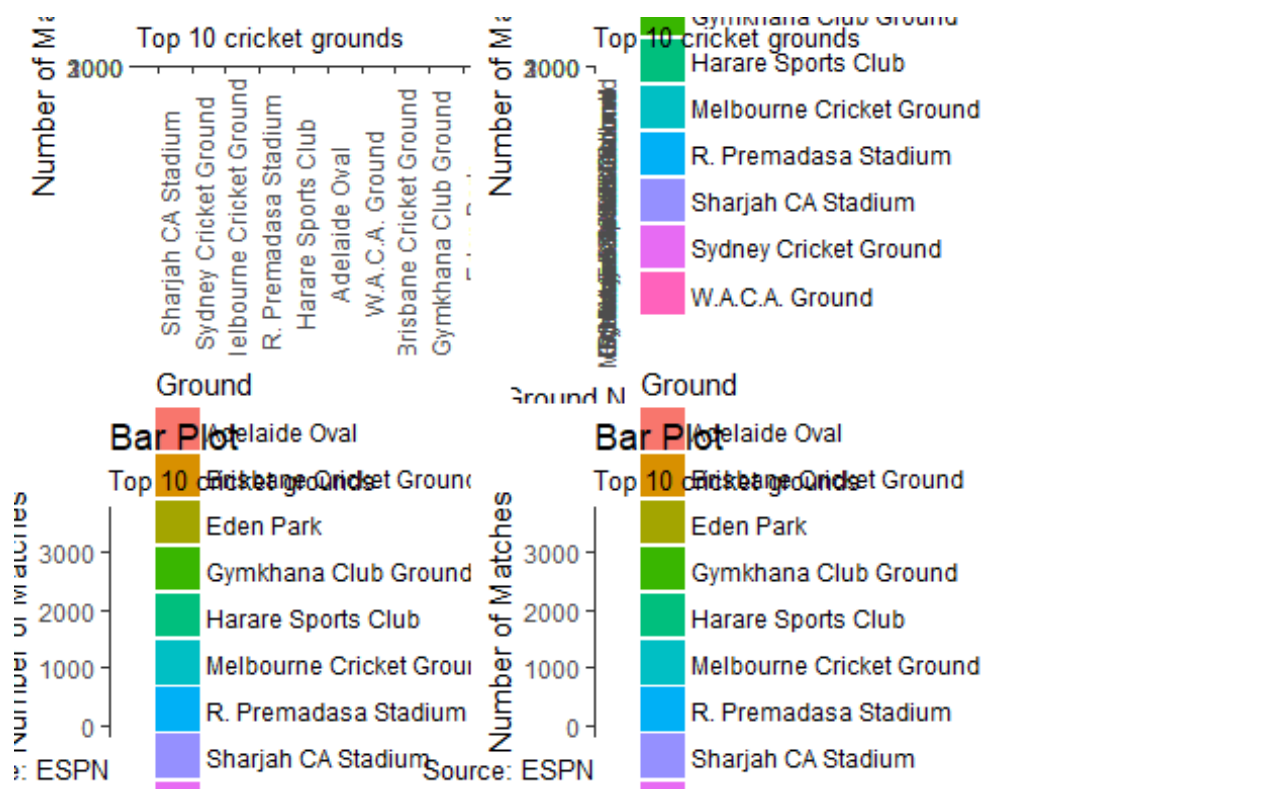
```
##
```

```
## combine
```

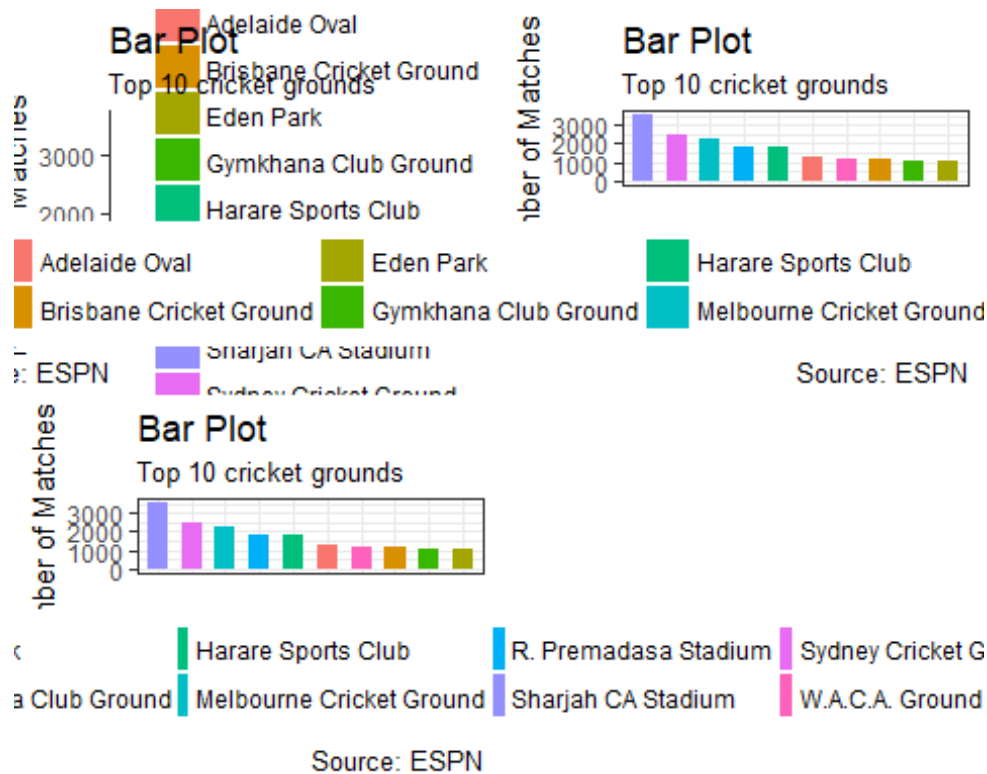
```
PlotA = grid.arrange(P1, P2,P3,P4,nrow = 2, ncol = 2)
```



```
PlotB = grid.arrange(P5,P6,P7,P8, nrow = 2, ncol = 2)
```



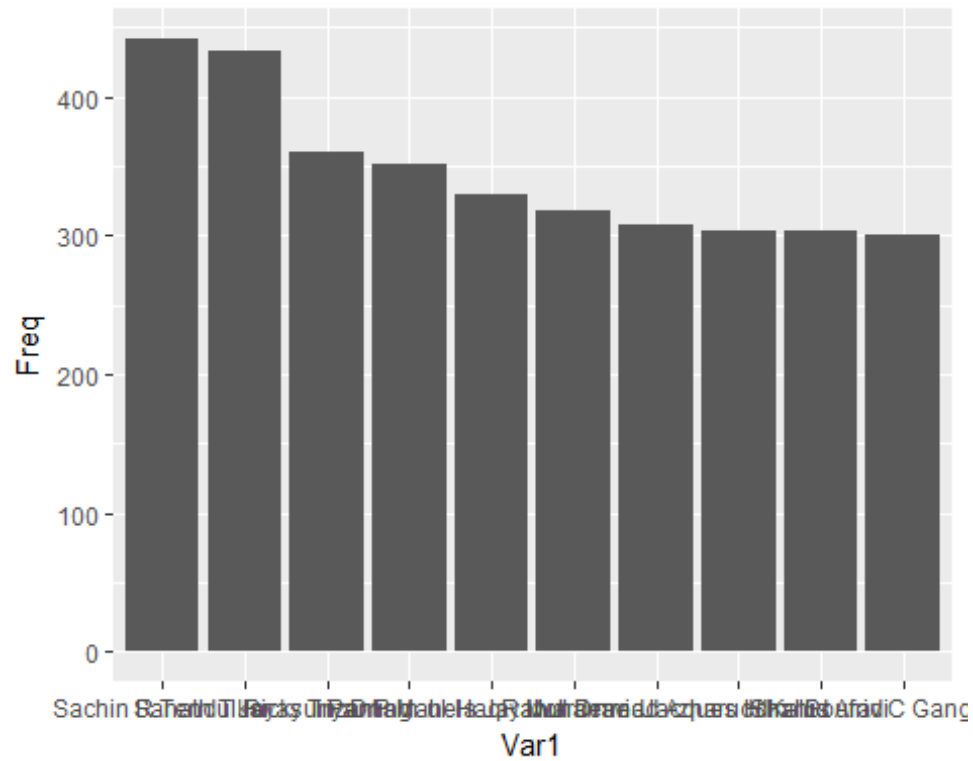
```
PlotC = grid.arrange(P9,P10,P11, nrow = 2, ncol = 2)
```



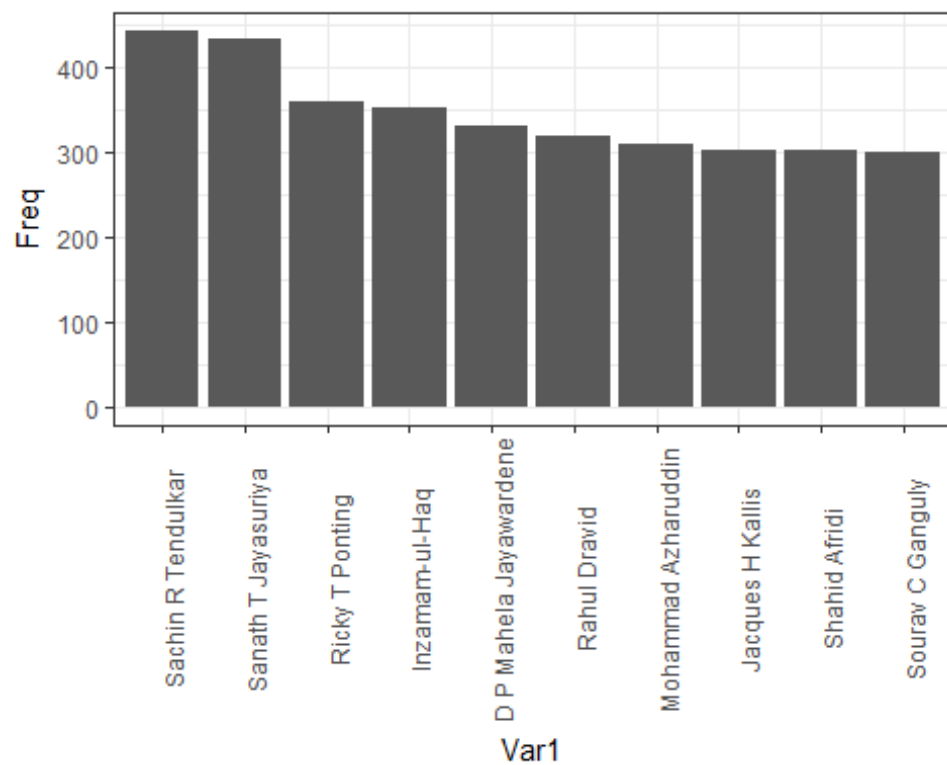
```
#####
#####
library(dplyr)
library(ggplot2)

top_players <- data %>% group_by(Player) %>% summarise(count=n()) %>%
  arrange(desc(count))

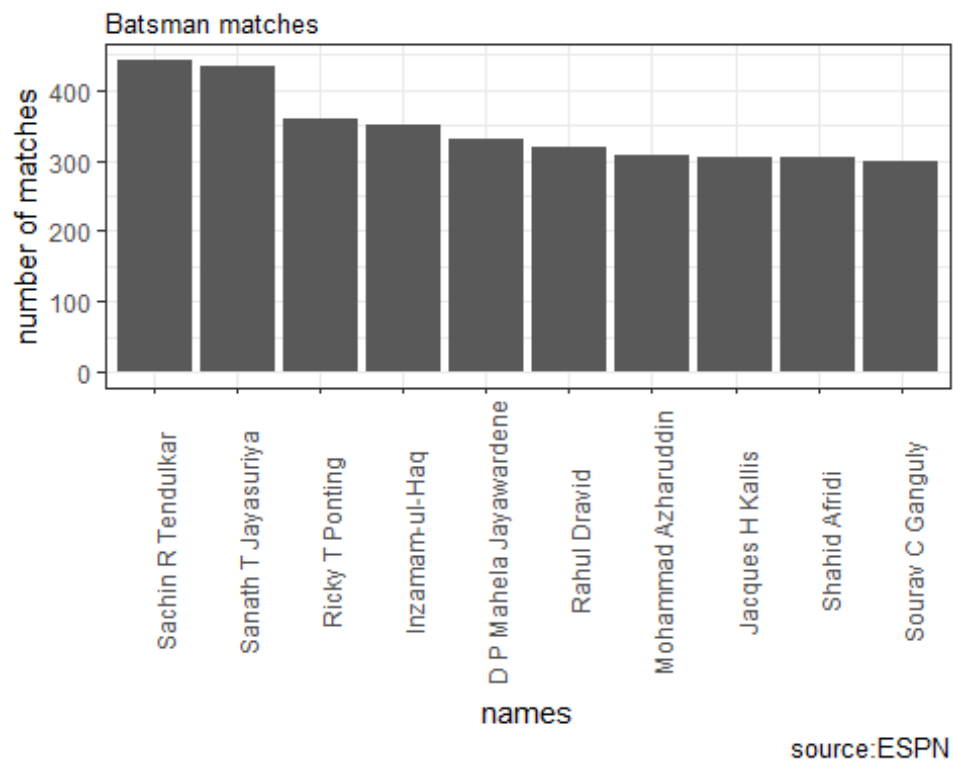
top <- as.data.frame(sort(table(data$Player),decreasing = TRUE))
r <- head(top,10)
p1 <- ggplot(r,aes(x=Var1,y=Freq)) + geom_bar(stat = "Identity")
p1
```

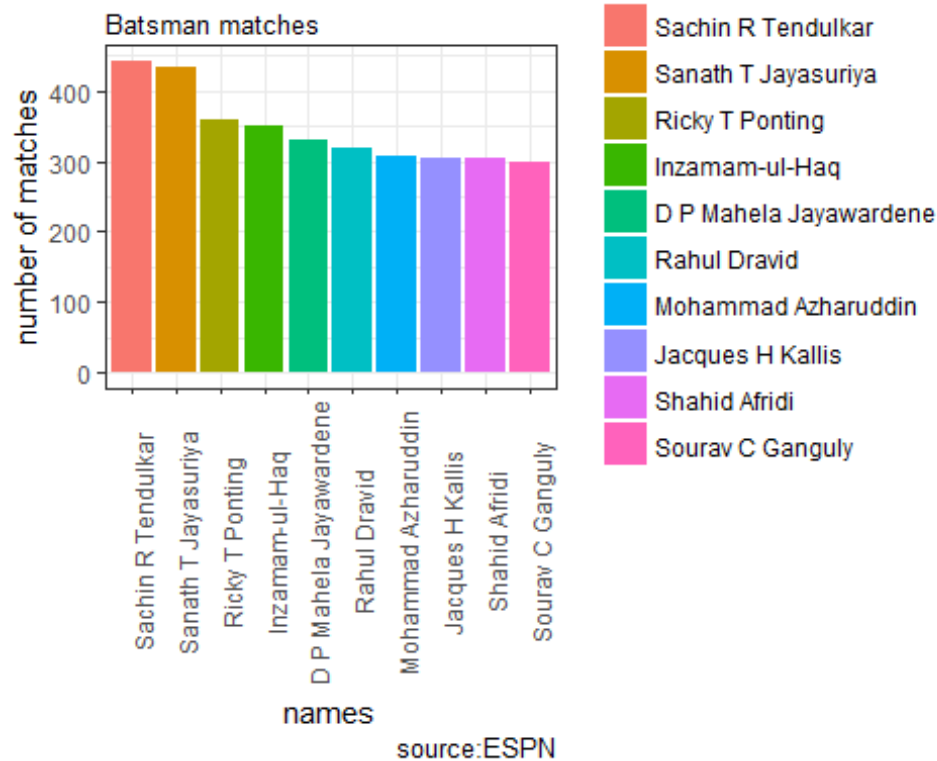
```
p <- p1 + theme_bw()
p2 <- p + theme(axis.text.x = element_text(angle = 90))
p2
```



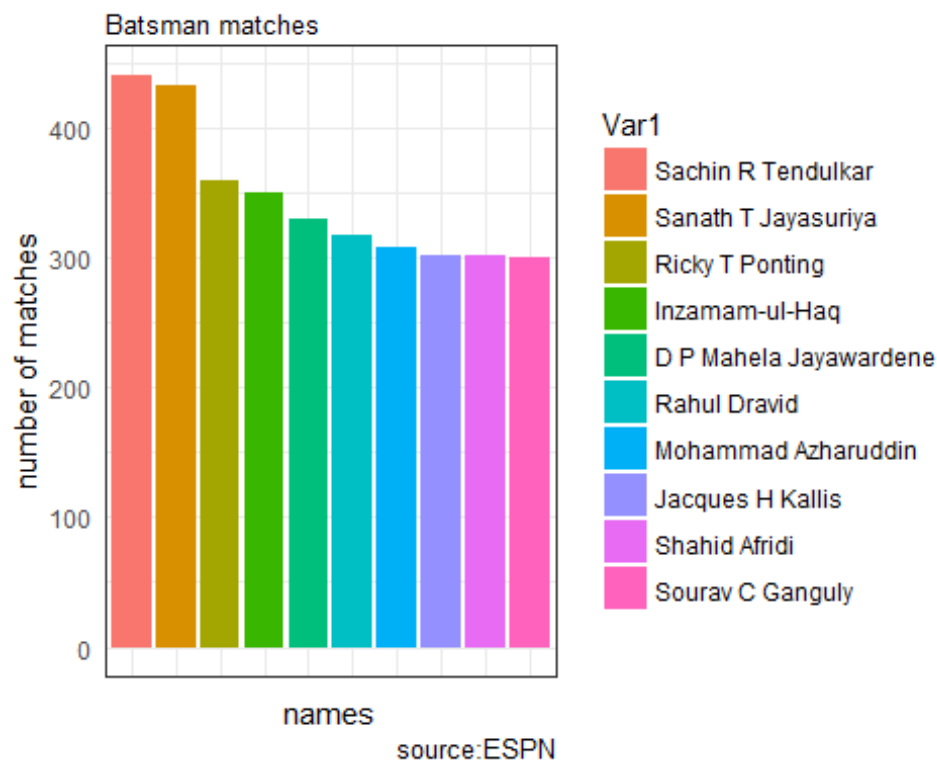
```
p3 <- p2 + labs(subtitle = "Batsman matches", x="names",y="number of matches",caption="source:ESPN")
p3
```



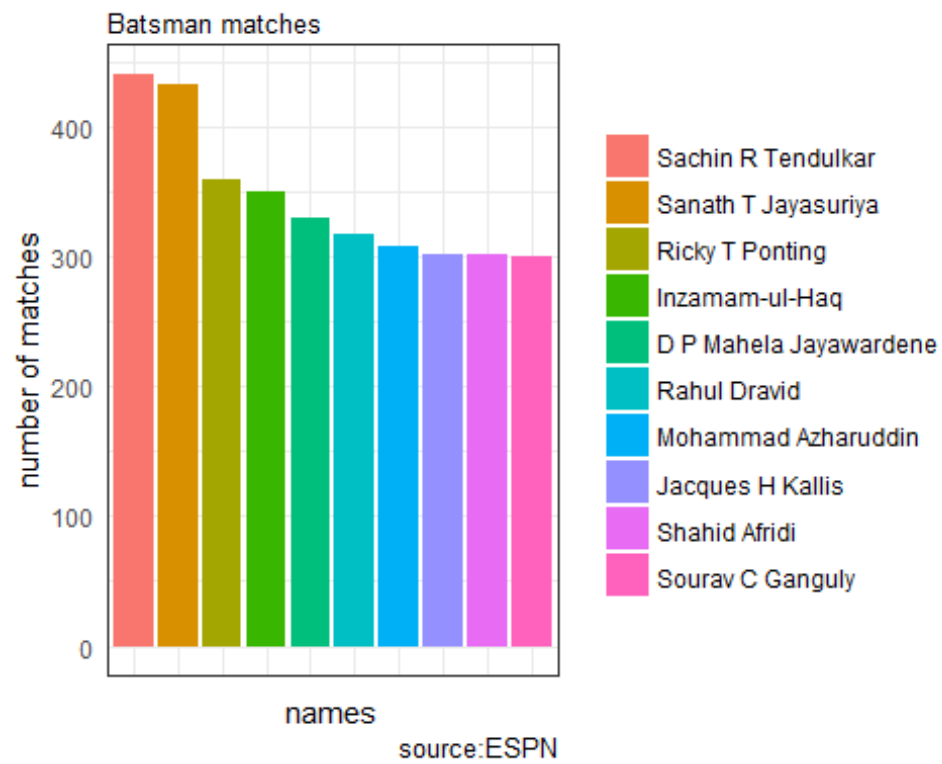
```
p4 <- p3 + geom_bar(stat = "Identity",aes(fill= Var1),Width = .6)
## Warning: Ignoring unknown parameters: Width
p4
```



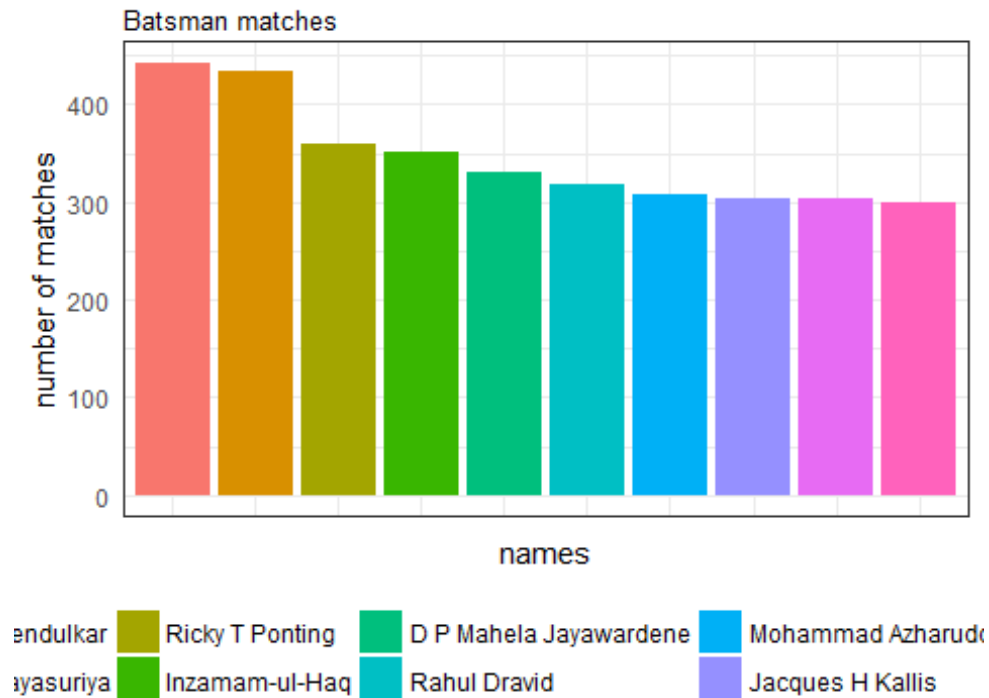
```
p5 <- p4 + theme(axis.text.x = element_blank(),axis.ticks = element_blank())
p5
```



```
p6 <- p5 + theme(legend.title = element_blank())
p6
```

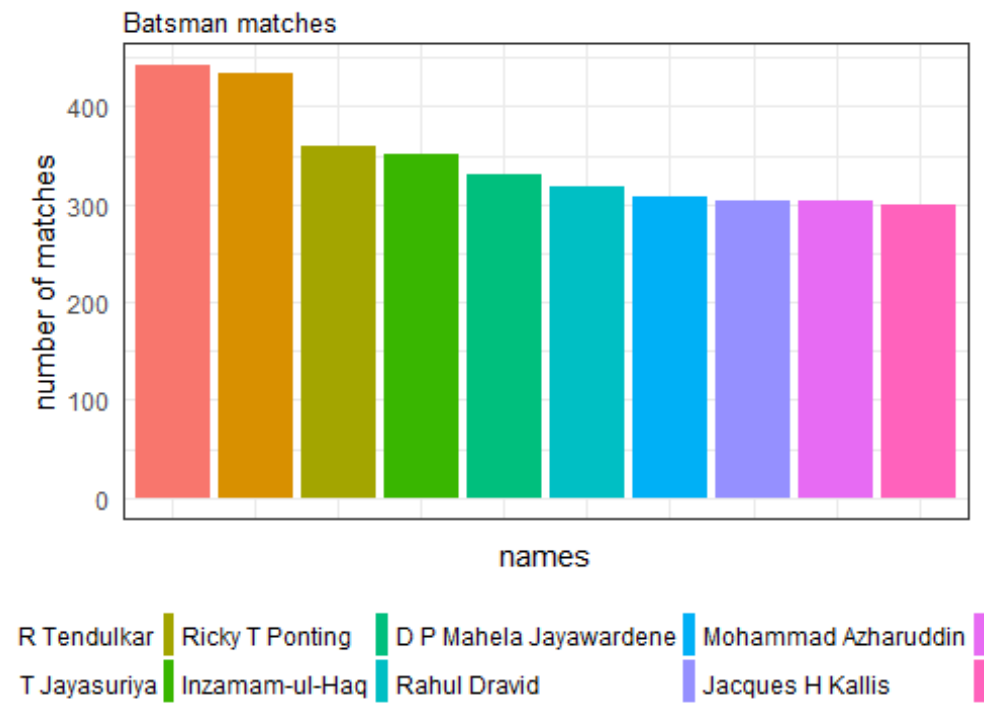


```
p7 <- p6 + theme(legend.position = "bottom")
p7
```



source:ESPN

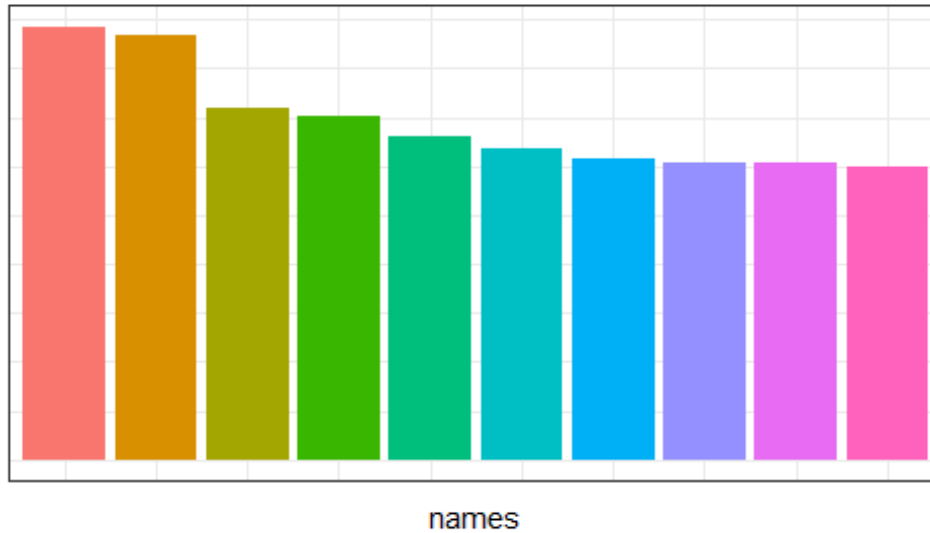
```
p8 <- p7 + theme(legend.key.width = unit(0.2,"cm"))
p8
```



source:ESPN

```
p9 <- p8 + theme(axis.text.y = element_blank(),axis.title.y =
element_blank())
p9
```

Batsman matches

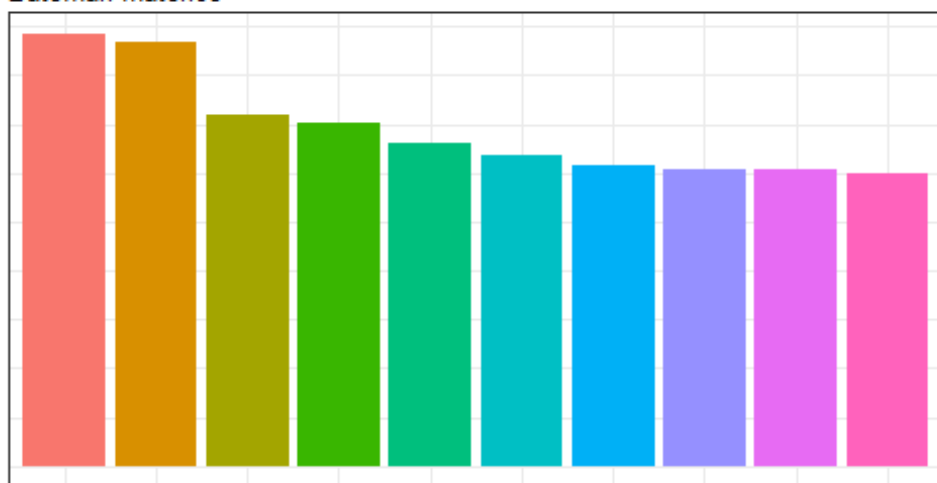


endulkar Ricky T Ponting D P Mahela Jayawardene Mohammad Azharuddin Shi
ayasuriya Inzamam-ul-Haq Rahul Dravid Jacques H Kallis Soi

source:ESPN

```
p10 <- p9 + theme(legend.text = element_text(colour = 'black',size = 8,hjust
= 3,vjust = 3,face = 'bold'))
p10
```

Batsman matches



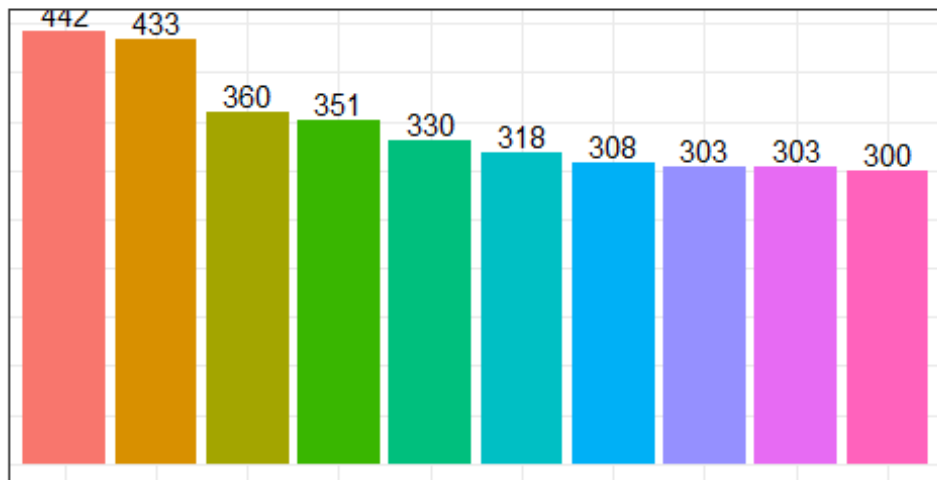
names

endulkar Ricky T Ponting D P Mahela Jayawardene Mohammad Azharuddin Sha
 ayasuriya Inzamam-ul-Haq Rahul Dravid Jacques H Kallis Sou

source:ESPN

```
p11 <- p10 + geom_text(aes(label = Freq), position = position_dodge(width = 0.9), vjust = -0.25)
p11
```

Batsman matches



names

endulkar Ricky T Ponting D P Mahela Jayawardene Mohammad Azharuddin Sha
 ayasuriya Inzamam-ul-Haq Rahul Dravid Jacques H Kallis Sou

source:ESPN