Bundesliga

Gurgen Blbulyan

13/10/19

```
library(ggplot2)
library(dplyr,warn.conflicts = F)
```

```
bund<-read.csv("bundesliga.csv",stringsAsFactors = F)
bund1<-read.csv("bundesliga2.csv",stringsAsFactors = F)

fi<-function(year){
  test<-bund[bund$SEASON==year,]
  tms<-unique(test$HOMETEAM)
  him<-function(r){
    a<-test[test$AWAYTEAM==r,]
    b<-test[test$HOMETEAM==r,]
    TEAM<-r
    M<-nrow(a)+nrow(b)
    W<-nrow(a[a$FTAG>a$FTHG,])+nrow(b[b$FTAG<b$FTHG,])
    L<-nrow(a[a$FTAG<a$FTHG,])+nrow(b[b$FTAG>b$FTHG,])
    D<-M-W-L
    GF<-sum(a$FTAG,b$FTHG)
    GA<-sum(b$FTAG,a$FTHG)
    DIFF<-GF-GA
    POINTS<-3*W+D
    SEASON<-unique(test$SEASON)
    dt <-data.frame(TEAM,M,W,D,L,GF,GA,DIFF,POINTS,SEASON)
    return(dt)
  }
  kku<-c()
  for (i in tms){
    fk<-him(i)
    kku<-rbind(kku,fk)
  }
  kku <-arrange(kku, desc(POINTS), desc(DIFF))
  kku$POSITION<-seq(1,length(tms))
  return(kku)
}
years<-unique(bund$SEASON)
iamdata<-c()
for (i in years){
  ui<-fi(i)
  iamdata<-rbind(iamdata,ui)
}
iamdata<-iamdata[,c(1,2,3,4,5,6,7,8,9,11,10)]
```

2.1 On average how many goals were scored per match in matches between Bayern Munich and Dortmund and is it above or below from average of scored goals in all matches?

```
bvd<-bund[(bund$HOMETEAM=="Bayern Munich"&bund$AWAYTEAM=="Dortmund")|
            (bund$HOMETEAM=="Dortmund"&bund$AWAYTEAM=="Bayern Munich"),]
s<-mean(bvd$FTTG)
```

```r
a<-mean(bund$FTTG)
s-a
```

```
## [1] 0.09426848
```

This is mean that on average 0.0942 more goals scored per match between Bayern Munich and Dortmund than in all matches.

2.2 Does Bayern Munich and Dortmund affect each other?

```r
byr<-iamdata[iamdata$TEAM=="Bayern Munich",]
bvb<-iamdata[iamdata$TEAM=="Dortmund",]
cor(bvb$W,byr$W)
```

```
## [1] 0.2104449
```

```r
cor(bvb$L,byr$L)
```

```
## [1] 0.003336143
```

```r
cor(bvb$POINTS,byr$POINTS)
```

```
## [1] 0.1729224
```

```r
cor(bvb$POSITION,byr$POSITION)
```

```
## [1] -0.1652011
```

```r
cor(bvb$W,byr$L)
```

```
## [1] -0.06846899
```

```r
cor(bvb$L,byr$W)
```

```
## [1] -0.1483153
```

```r
cor(bvb$POSITION,byr$POINTS)
```

```
## [1] -0.06125081
```

```r
cor(bvb$POINTS,byr$POSITION)
```

```
## [1] 0.1706579
```

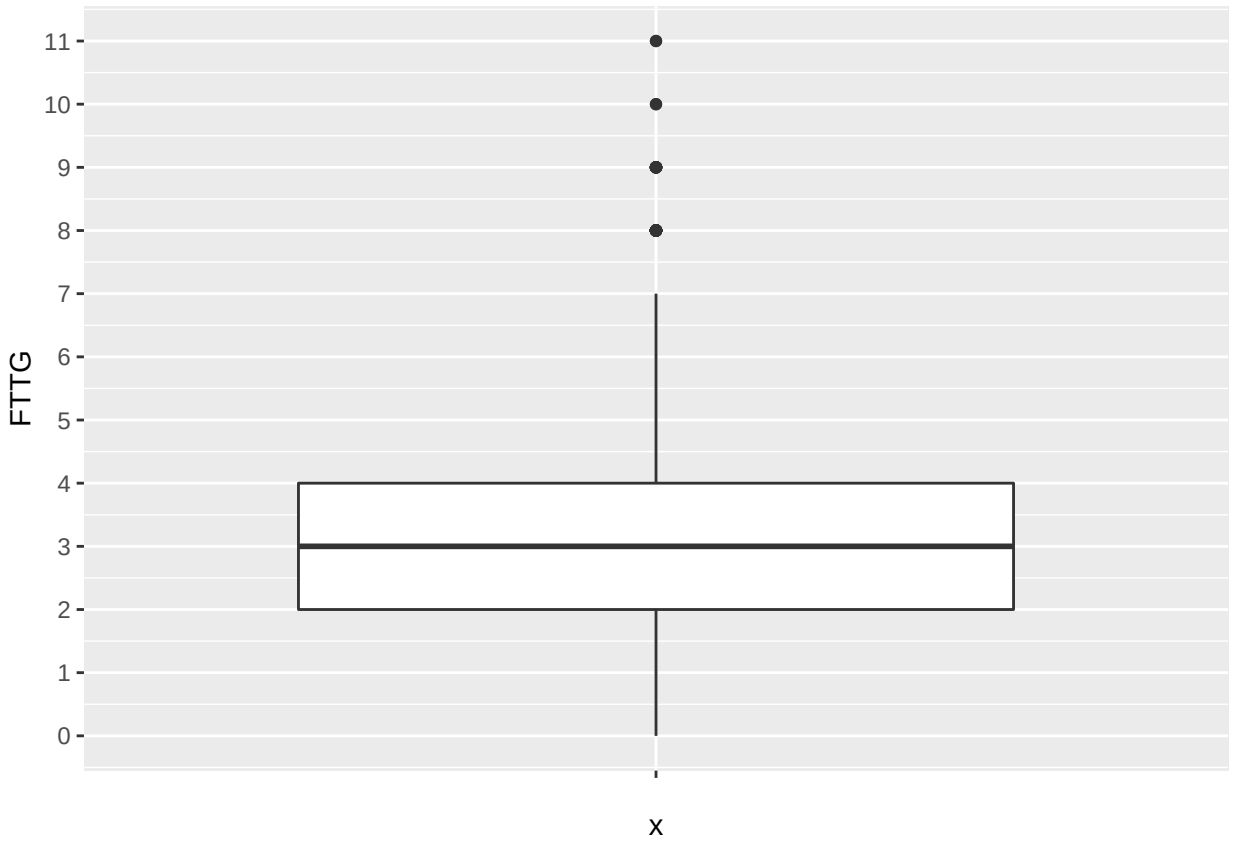
```r
cor(bvb$DIFF,byr$DIFF)
```

```
## [1] 0.3356458
```

As we see this correlations are too small to conclude, that there is relation between Bayern Munich and Dortmund.

2.3 Do some statistical analysis on Full Time Total Goals series.

```r
ggplot(bund,aes(x="",y=FTTG))+geom_boxplot()+scale_y_continuous(breaks = seq(0,15))
```




```
mean(bund$FTTG)
```

```
## [1] 2.886501
```

```
sd(bund$FTTG)
```

```
## [1] 1.708871
```

```
table(bund$FTTG)
```

```
##
##    0    1    2    3    4    5    6    7    8    9   10   11
##  542 1117 1924 1710 1355  716  384  125   62   17    2    2
```

As we see maximum number of goals in one match is 11, 1st quartile is 2, Median is 3, 3rd quartile is 4 and 8,9,10,11 goals in one match is outlier, mean is 2.887, standard deviation is 1.7 and 2 goals in one match appears more often.