MOOC 1 Week 3 Assignment Part II

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.3     v dplyr   1.0.0
## v tidyr   1.1.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

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

```
library(readxl)
```

```
IPLbyb = read_xlsx("~/Google Drive/Sports Analytics Moocs/MOOC 1 - Foundations of sports analytics/Assig
colnames(IPLbyb)
```

```
##  [1] "home team"      "batting_team"   "bowling_team"   "gameno"
##  [5] "innings_number" "over_no"        "ball_no"        "runs_total_end"
##  [9] "delivery_no"    "delno^2"        "runs*wkts"      "wicket"
## [13] "wkts_total_end"
```

```
# change "home team" to "home_team"
IPLbyb <- IPLbyb %>% dplyr::rename(home_team = "home team")
```

```
# First, the function for plotting the runs and wickets for each team in a game.
plot_runs_wickets <- function(IPLbyb){
    gameno = unique(IPLbyb$gameno)[1]
    inning_list = unique(IPLbyb$innings_number)
    p1 <- ggplot()
    for (inning in inning_list){
        # create separate dataframe for wickets
        data = IPLbyb %>% filter(innings_number==inning)
        wicket = data %>% filter(wicket >0)
        # plots line and  markers
        if(inning == 1){
          color1 = "blue"
        }
        else {
          color1 = "orange"
        }

        p1 <- p1 +
              geom_line(data = data,aes(x = delivery_no, y = runs_total_end),
                        colour = color1)+
              geom_point(data = wicket, aes(x = delivery_no, y = runs_total_end),
                         colour = color1, size = 3)+
              xlab('balls')+
```

```
              ylab('runs')+
              ggtitle(paste("Game", gameno,sep = " "))

    }
  return(p1)
}
```

```
# Second, a function that allows us to plot two or more games at the same time.
plot_runs_wickets_multi_game <- function(list_games){
    n = length(list_games)
    for (i in list_games){
        game = IPLbyb %>% filter(gameno == i)
        p <- plot_runs_wickets(game)
        print(p)
    }
}
```

```
# Identify if the home team batted first
IPLbyb[,'hometeambatsfirst']=
  ifelse((IPLbyb$home_team==IPLbyb$batting_team) & (IPLbyb$innings_number==1),'yes','no')
head(IPLbyb)
```

```
## # A tibble: 6 x 14
##   home_team batting_team bowling_team gameno innings_number over_no ball_no
##   <chr>     <chr>        <chr>         <dbl>          <dbl>   <dbl>   <dbl>
## 1 Mumbai I~ Mumbai Indi~ Chennai Sup~      1              1       0       1
## 2 Mumbai I~ Mumbai Indi~ Chennai Sup~      1              1       0       2
## 3 Mumbai I~ Mumbai Indi~ Chennai Sup~      1              1       0       3
## 4 Mumbai I~ Mumbai Indi~ Chennai Sup~      1              1       0       4
## 5 Mumbai I~ Mumbai Indi~ Chennai Sup~      1              1       0       5
## 6 Mumbai I~ Mumbai Indi~ Chennai Sup~      1              1       0       6
## # ... with 7 more variables: runs_total_end <dbl>, delivery_no <dbl>,
## #   `delno^2` <dbl>, `runs*wkts` <dbl>, wicket <dbl>, wkts_total_end <dbl>,
## #   hometeambatsfirst <chr>
```

```
tail(IPLbyb)
```

```
## # A tibble: 6 x 14
##   home_team batting_team bowling_team gameno innings_number over_no ball_no
##   <chr>     <chr>        <chr>         <dbl>          <dbl>   <dbl>   <dbl>
## 1 Chennai ~ Chennai Sup~ Sunrisers        60              2      17       4
## 2 Chennai ~ Chennai Sup~ Sunrisers        60              2      17       5
## 3 Chennai ~ Chennai Sup~ Sunrisers        60              2      17       6
## 4 Chennai ~ Chennai Sup~ Sunrisers        60              2      18       1
## 5 Chennai ~ Chennai Sup~ Sunrisers        60              2      18       2
## 6 Chennai ~ Chennai Sup~ Sunrisers        60              2      18       3
## # ... with 7 more variables: runs_total_end <dbl>, delivery_no <dbl>,
## #   `delno^2` <dbl>, `runs*wkts` <dbl>, wicket <dbl>, wkts_total_end <dbl>,
## #   hometeambatsfirst <chr>
```

```
# drop duplicates so we just have a list of games

games = IPLbyb %>% group_by(gameno) %>% slice(1)
games
```

```
## # A tibble: 60 x 14
```

```
## # Groups:   gameno [60]
##    home_team batting_team bowling_team gameno innings_number over_no ball_no
##    <chr>     <chr>        <chr>         <dbl>          <dbl>   <dbl>   <dbl>
##  1 Mumbai I~ Mumbai Indi~ Chennai Sup~      1              1       0       1
##  2 Kings XI~ Delhi Dared~ Kings XI Pu~      2              1       0       1
##  3 Kolkata ~ Royal Chall~ Kolkata Kni~      3              1       0       1
##  4 Sunrisers Rajasthan R~ Sunrisers         4              1       0       1
##  5 Chennai ~ Kolkata Kni~ Chennai Sup~      5              1       0       1
##  6 Rajastha~ Rajasthan R~ Delhi Dared~      6              1       0       1
##  7 Sunrisers Mumbai Indi~ Sunrisers         7              1       0       1
##  8 Royal Ch~ Kings XI Pu~ Royal Chall~      8              1       0       1
##  9 Mumbai I~ Mumbai Indi~ Delhi Dared~      9              1       0       1
## 10 Kolkata ~ Kolkata Kni~ Sunrisers        10              1       0       1
## # ... with 50 more rows, and 7 more variables: runs_total_end <dbl>,
## #   delivery_no <dbl>, `delno^2` <dbl>, `runs*wkts` <dbl>, wicket <dbl>,
## #   wkts_total_end <dbl>, hometeambatsfirst <chr>
```

```
# generate list of games

games = games[,c('gameno','home_team','batting_team','bowling_team','hometeambatsfirst')]
games <- games %>% mutate(road_team = ifelse(home_team == bowling_team,batting_team,bowling_team))
sum(games$home_team == games$road_team)
```

```
## [1] 0
```

```
games = games[,c('gameno','home_team','road_team','hometeambatsfirst')]
head(games)
```

```
## # A tibble: 6 x 4
## # Groups:   gameno [6]
##   gameno home_team             road_team                  hometeambatsfirst
##    <dbl> <chr>                 <chr>                      <chr>
## 1      1 Mumbai Indians        Chennai Super Kings        yes
## 2      2 Kings XI Punjab       Delhi Daredevils           no
## 3      3 Kolkata Knight Riders Royal Challengers Bangalore no
## 4      4 Sunrisers             Rajasthan Royals           no
## 5      5 Chennai Super Kings   Kolkata Knight Riders      no
## 6      6 Rajasthan Royals      Delhi Daredevils           yes
```

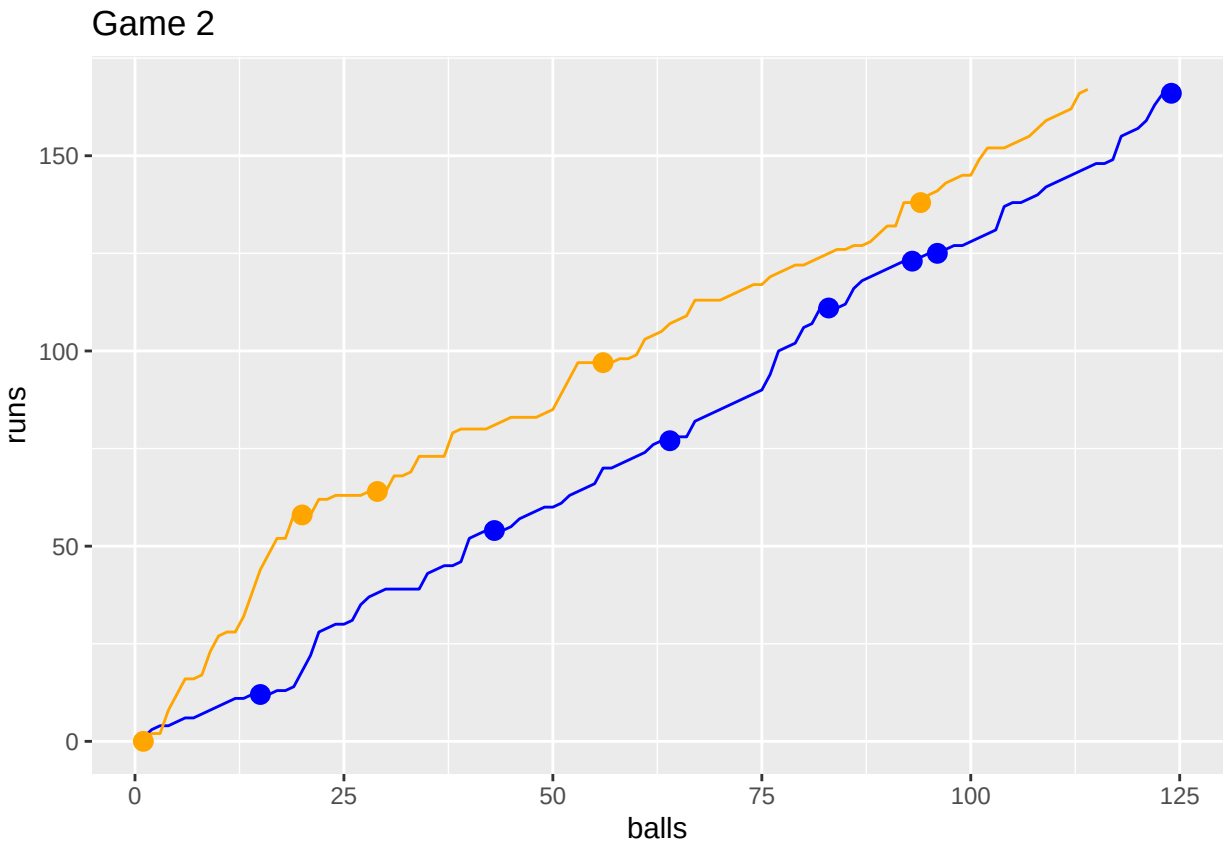
```
tail(games)
```

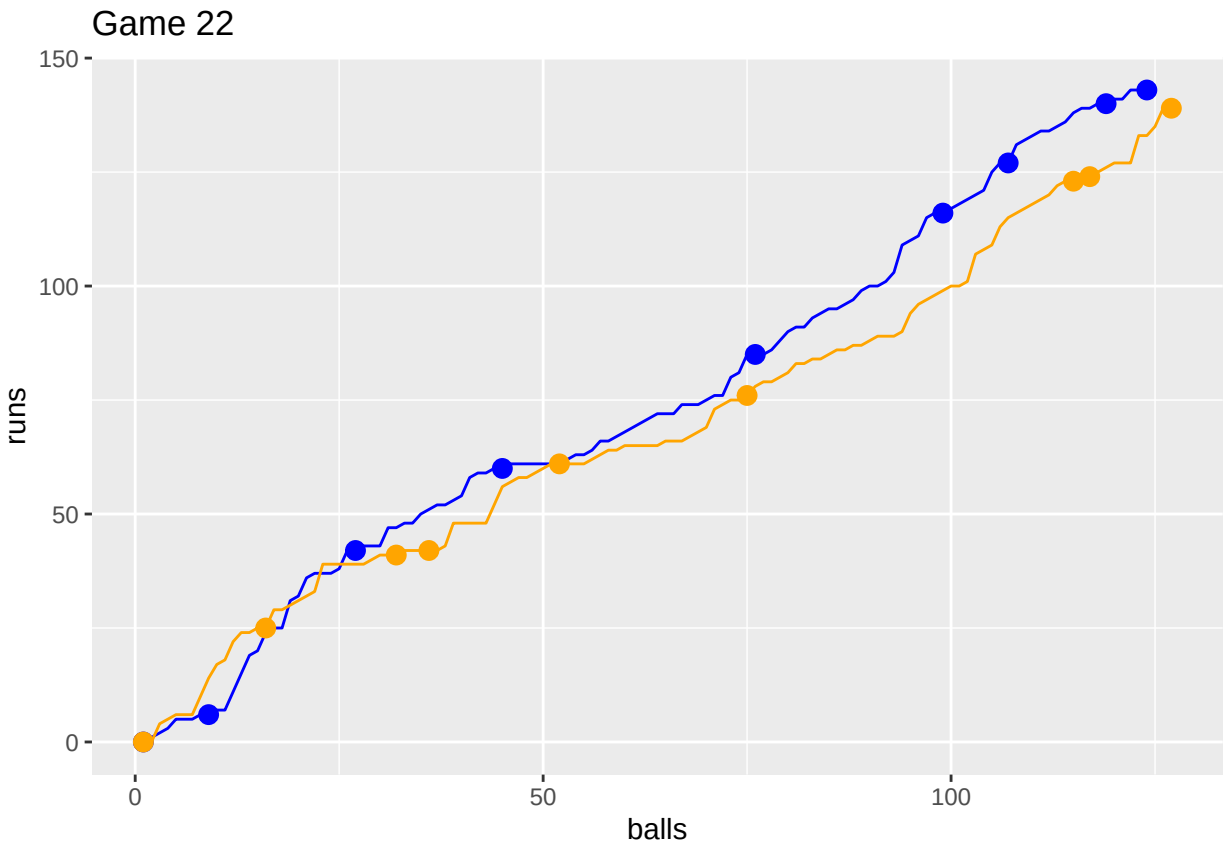
```
## # A tibble: 6 x 4
## # Groups:   gameno [6]
##   gameno home_team             road_team       hometeambatsfirst
##    <dbl> <chr>                 <chr>           <chr>
## 1     55 Delhi Daredevils      Mumbai Indians  yes
## 2     56 Chennai Super Kings   Kings XI Punjab no
## 3     57 Chennai Super Kings   Sunrisers       no
## 4     58 Kolkata Knight Riders Rajasthan Royals yes
## 5     59 Kolkata Knight Riders Sunrisers       no
## 6     60 Chennai Super Kings   Sunrisers       no
```

```
# Inside the square brackets we can type in game numbers, separated by a comma,
# to specaify the games we want to compare

# Kings XI Punjab v Delhi Daredevils
```

```
plot_runs_wickets_multi_game(c(2, 22))
```




Game 22
150
100
50
0
runs
0
50
100
balls