

Playoffs

Rubens Cortelazzi Roncato

2024-05-07

```
source("dados_playoffs.R")
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr    1.5.1
## v lubridate  1.9.3      v tibble     3.2.1
## v purrr      1.0.2      v tidyr      1.3.1
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
## Loading required package: splines
##
## Loading required package: gamlss.data
##
##
## Attaching package: 'gamlss.data'
##
##
## The following object is masked from 'package:datasets':
##
##     sleep
##
## Loading required package: gamlss.dist
##
## Loading required package: nlme
##
##
## Attaching package: 'nlme'
##
##
## The following object is masked from 'package:dplyr':
##
##     collapse
##
## Loading required package: parallel
##
## ***** GAMLSS Version 5.4-22 *****
##
## For more on GAMLSS look at https://www.gamlss.com/
##
```

```
## Type gamlssNews() to see new features/changes/bug fixes.
##
##
## Loading required package: carData
##
##
## Attaching package: 'car'
##
##
## The following object is masked from 'package:dplyr':
##
##     recode
##
##
## The following object is masked from 'package:purrr':
##
##     some
##
##
## Loading required package: zoo
##
##
## Attaching package: 'zoo'
##
##
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
##### Análise descritiva Playoffs #####
```

```
#Análise de mínimo, máximo, média, mediana, 1 e 3 quartil de todas as variáveis.
summary(dados_playoffs)
```

```
##      Posicao      TEAM      GP      W
## Min.   : 1   Boston Celtics   : 14   Min.   : 4.00   Min.   : 0.000
## 1st Qu.: 4   Atlanta Hawks    : 12   1st Qu.: 5.75   1st Qu.: 1.750
## Median : 8   Miami Heat      : 12   Median : 8.00   Median : 3.500
## Mean   : 8   Portland Trail Blazers: 11   Mean   :10.46   Mean   : 5.229
## 3rd Qu.:11   San Antonio Spurs    : 11   3rd Qu.:14.00   3rd Qu.: 7.250
## Max.   :16   Dallas Mavericks    : 10   Max.   :24.00   Max.   :16.000
##           (Other)      :170
##      L      WINP      MIN      PTS
## Min.   : 1.000   Min.   :0.0000   Min.   :48.00   Min.   : 78.0
## 1st Qu.: 4.000   1st Qu.:0.2998   1st Qu.:48.00   1st Qu.: 93.8
## Median : 4.000   Median :0.4290   Median :48.00   Median :101.1
## Mean   : 5.229   Mean   :0.4031   Mean   :48.33   Mean   :100.6
## 3rd Qu.: 6.000   3rd Qu.:0.5450   3rd Qu.:48.50   3rd Qu.:107.7
## Max.   :11.000   Max.   :0.9410   Max.   :53.00   Max.   :119.5
##
##      FGM      FGA      FGP      3PM
## Min.   :27.80   Min.   :69.00   Min.   :38.20   Min.   : 2.300
## 1st Qu.:34.30   1st Qu.:78.47   1st Qu.:42.70   1st Qu.: 6.600
## Median :36.70   Median :82.85   Median :44.50   Median : 8.800
## Mean   :36.63   Mean   :82.55   Mean   :44.38   Mean   : 9.068
```

```
## 3rd Qu.:38.83 3rd Qu.:86.03 3rd Qu.:46.20 3rd Qu.:11.300
## Max. :44.20 Max. :96.80 Max. :50.30 Max. :18.000
##
## 3PA 3PP FTM FTA
## Min. :10.80 Min. :20.00 Min. :11.30 Min. :14.40
## 1st Qu.:19.30 1st Qu.:32.08 1st Qu.:16.38 1st Qu.:21.30
## Median :25.55 Median :34.40 Median :18.30 Median :23.70
## Mean :26.17 Mean :34.39 Mean :18.29 Mean :23.95
## 3rd Qu.:32.73 3rd Qu.:37.30 3rd Qu.:20.02 3rd Qu.:26.23
## Max. :46.80 Max. :46.50 Max. :28.20 Max. :36.80
##
## FTP OREB DREB REB
## Min. :60.80 Min. : 5.20 Min. :24.20 Min. :33.50
## 1st Qu.:73.08 1st Qu.: 8.90 1st Qu.:30.00 1st Qu.:39.77
## Median :76.70 Median :10.00 Median :31.70 Median :42.05
## Mean :76.47 Mean :10.19 Mean :31.83 Mean :42.01
## 3rd Qu.:79.92 3rd Qu.:11.60 3rd Qu.:33.80 3rd Qu.:44.30
## Max. :88.60 Max. :16.70 Max. :41.80 Max. :51.90
##
## AST TOV STL BLK
## Min. :12.60 Min. : 8.40 Min. : 3.400 Min. :2.100
## 1st Qu.:18.30 1st Qu.:12.38 1st Qu.: 6.200 1st Qu.:3.800
## Median :20.45 Median :13.25 Median : 6.900 Median :4.600
## Mean :20.56 Mean :13.45 Mean : 6.982 Mean :4.678
## 3rd Qu.:22.52 3rd Qu.:14.50 3rd Qu.: 7.800 3rd Qu.:5.500
## Max. :28.40 Max. :19.00 Max. :10.800 Max. :8.100
##
## BLKA PF PFD PlusMinus
## Min. : 2.300 Min. :15.50 Min. :14.80 Min. : -24.200
## 1st Qu.: 4.075 1st Qu.:20.38 1st Qu.:20.00 1st Qu.: -6.800
## Median : 4.800 Median :21.55 Median :21.30 Median : -1.650
## Mean : 5.078 Mean :21.71 Mean :21.43 Mean : -2.743
## 3rd Qu.: 6.000 3rd Qu.:22.90 3rd Qu.:23.00 3rd Qu.: 1.800
## Max. :10.300 Max. :30.80 Max. :27.80 Max. : 13.500
##
## Temporada Conferencia Numero_temporada
## Length:240 Length:240 1 : 16
## Class :character Class :character 2 : 16
## Mode :character Mode :character 3 : 16
## 4 : 16
## 5 : 16
## 6 : 16
## (Other):144
```

```
##### Análise descritiva da variável resposta #####
```

```
#Medidas básicas
```

```
summary((dados_playoffs$WINP))
```

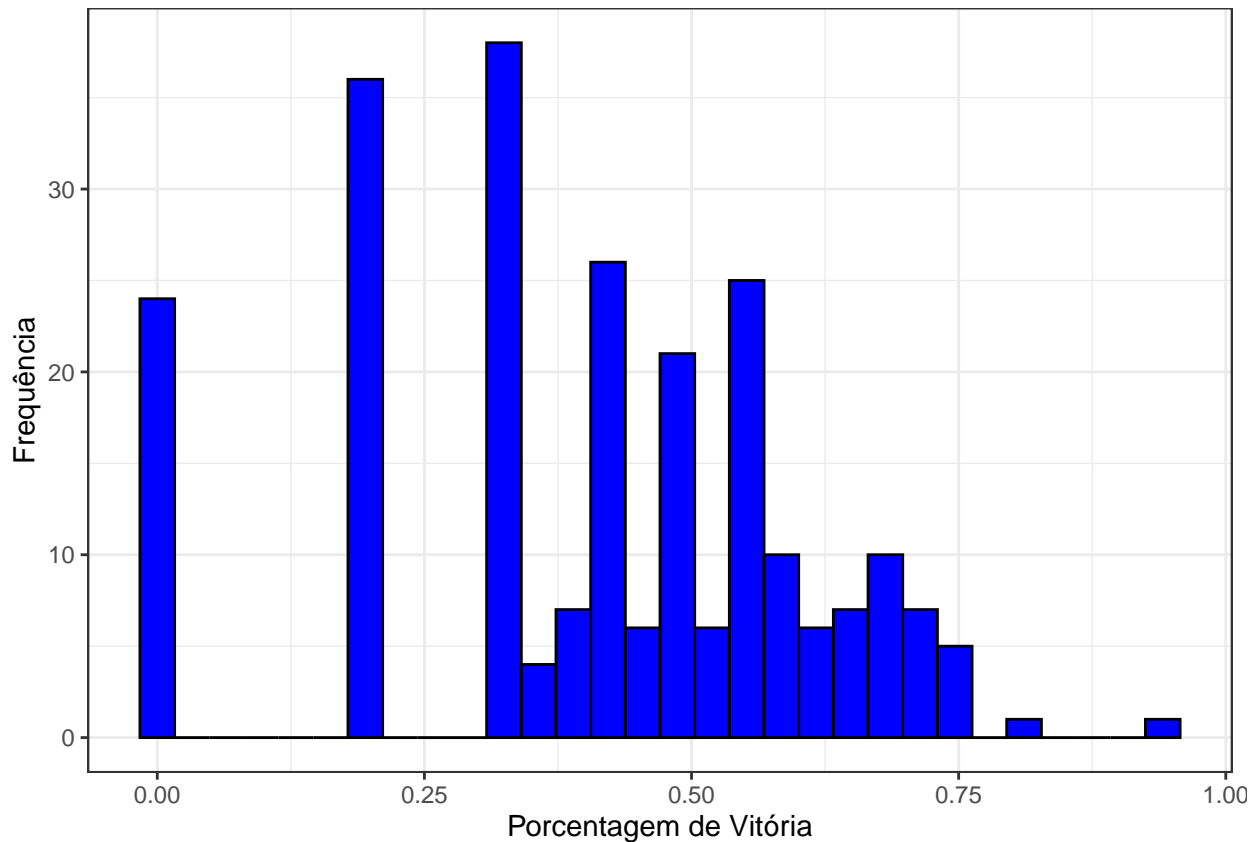
```
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.0000 0.2998 0.4290 0.4031 0.5450 0.9410
```

```
#Histograma da Porcentagem de Vitória
```

```
hist_play <- dados_playoffs %>% ggplot(aes(x = WINP)) +
  geom_histogram(fill = "blue", color = "black") +
```

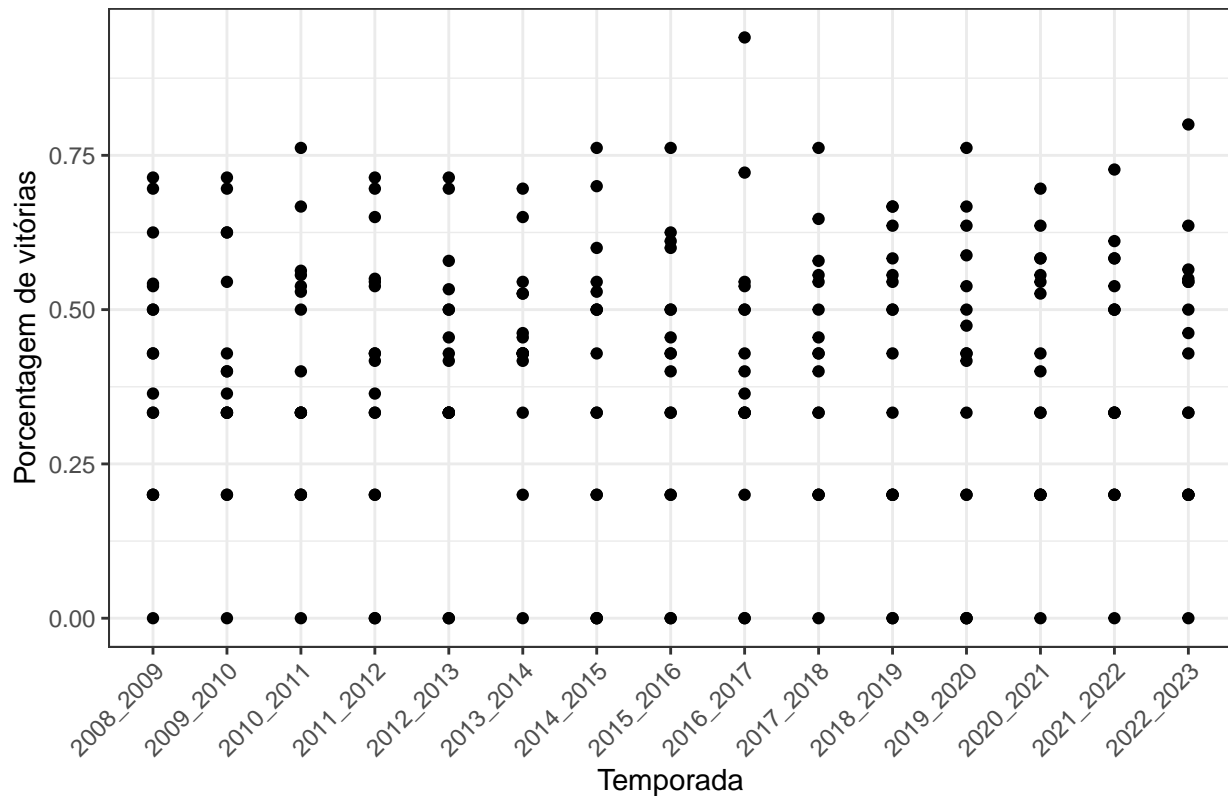
```
labs(x = "Porcentagem de Vitória",
     y = "Frequência"
) +
theme_bw()
hist_play
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



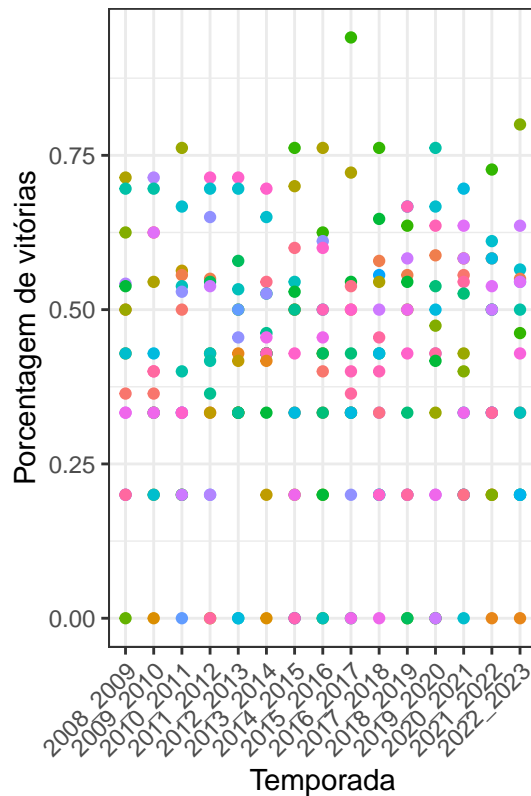
```
#Gráfico de pontos com a Porcentagem de vitórias ao decorrer das temporadas
dados_playoffs %>% ggplot() +
  geom_point(aes(x = Temporada, y = WINP)) +
  theme_bw() +
  labs(title = "Gráfico de pontos dos times",
       y = "Porcentagem de vitórias") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```

Gráfico de pontos dos times



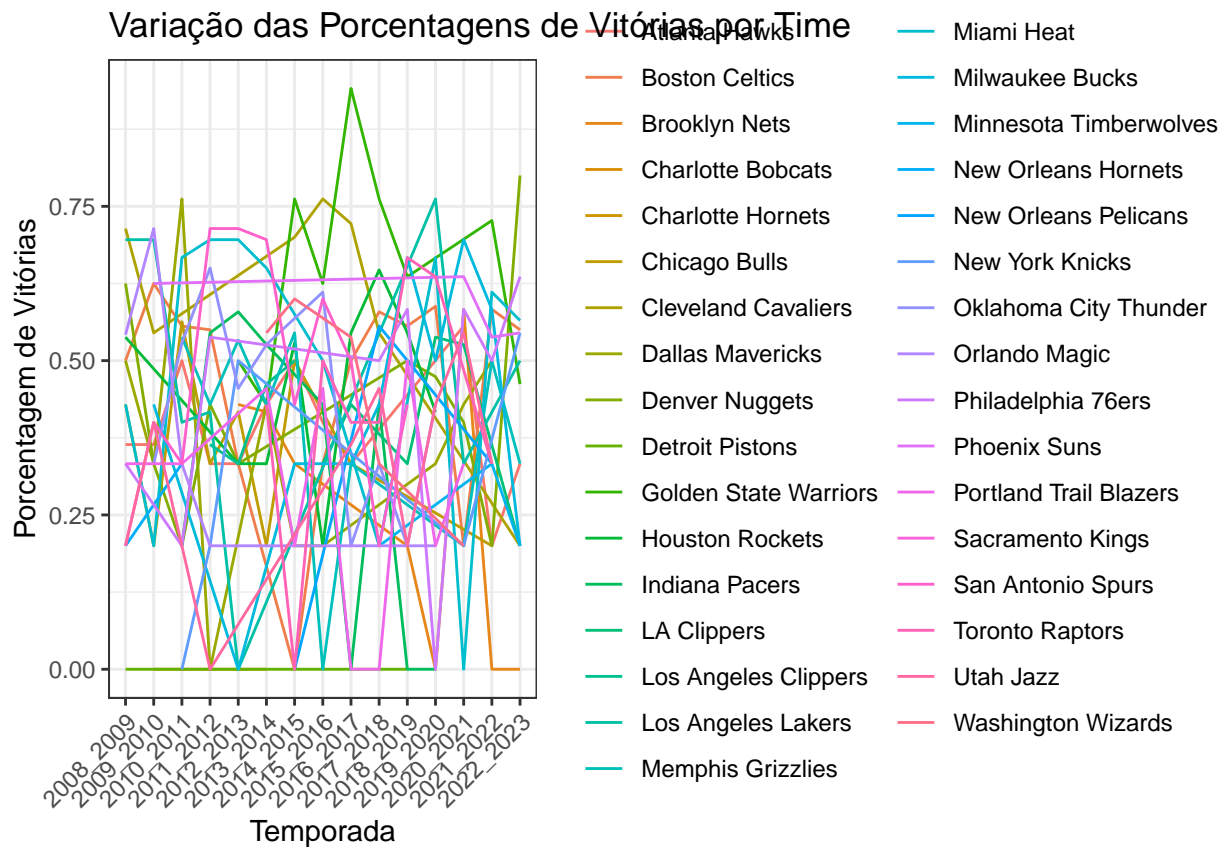
```
#Gráfico de pontos com a Porcentagem de vitórias ao decorrer das temporadas por time
dados_playoffs %>%
  ggplot() +
  geom_point(aes(x = Temporada, y = WINP, color = TEAM)) +
  theme_bw() +
  labs(title = "Gráfico de pontos dos times",
        y = "Porcentagem de vitórias") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```

Gráfico de pontos dos times



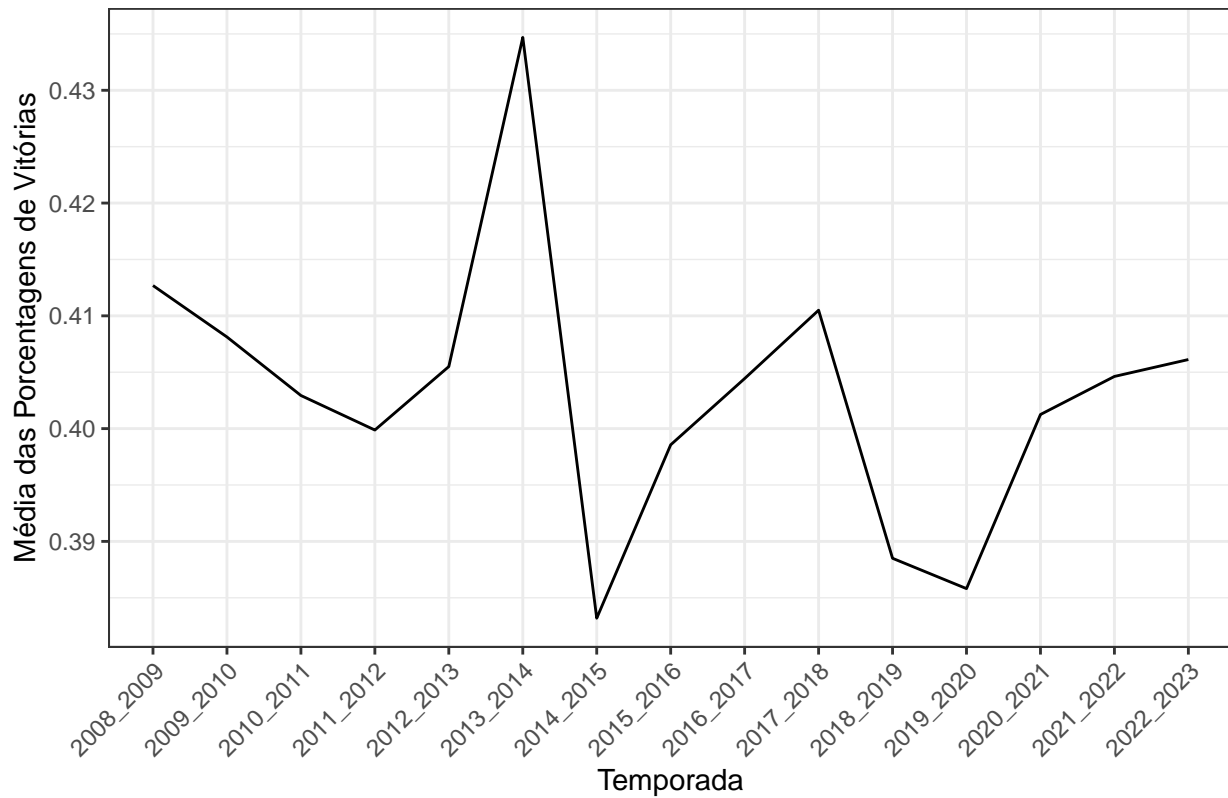
- Atlanta Hawks
- Boston Celtics
- Brooklyn Nets
- Charlotte Bobcats
- Charlotte Hornets
- Chicago Bulls
- Cleveland Cavaliers
- Dallas Mavericks
- Denver Nuggets
- Detroit Pistons
- Golden State Warriors
- Houston Rockets
- Indiana Pacers
- LA Clippers
- Los Angeles Clippers
- Los Angeles Lakers
- Memphis Grizzlies
- Miami Heat
- Milwaukee Bucks
- Minnesota Timberwolves
- New Orleans Hornets
- New Orleans Pelicans
- New York Knicks
- Oklahoma City Thunder
- Orlando Magic
- Philadelphia 76ers
- Phoenix Suns
- Portland Trail Blazers
- Sacramento Kings
- San Antonio Spurs
- Toronto Raptors
- Utah Jazz
- Washington Wizards

```
#Gráfico de linha com a Porcentagem de vitórias ao decorrer das temporadas por time
dados_playoffs %>% ggplot(aes(x = Temporada, y = WINP, color = TEAM, group = TEAM)) +
  geom_line() +
  labs(title = "Variação das Porcentagens de Vitórias por Time",
       x = "Temporada",
       y = "Porcentagem de Vitórias") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```



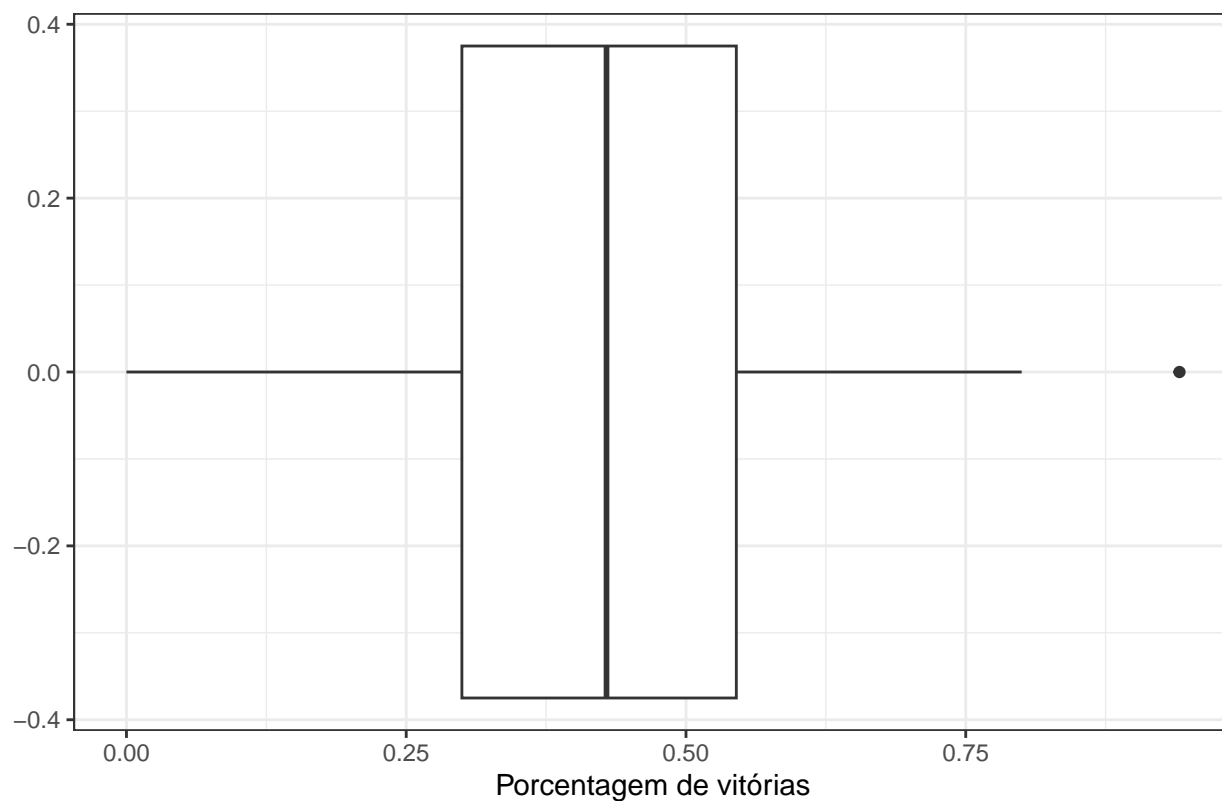
```
#Gráfico de linha com a Porcentagem média de vitórias ao decorrer das temporadas
dados_playoffs %>% dplyr::select(TEAM, Temporada, WINP) %>% group_by(Temporada) %>% summarise(Media = me
ggplot(aes(x = Temporada, y = Media, group = 1)) +
  geom_line() +
  #scale_y_continuous(limits = c(0.493, 0.505)) +
  labs(title = "Média das Porcentagens de Vitórias por Temporada",
        x = "Temporada",
        y = "Média das Porcentagens de Vitórias") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```

Média das Porcentagens de Vitórias por Temporada

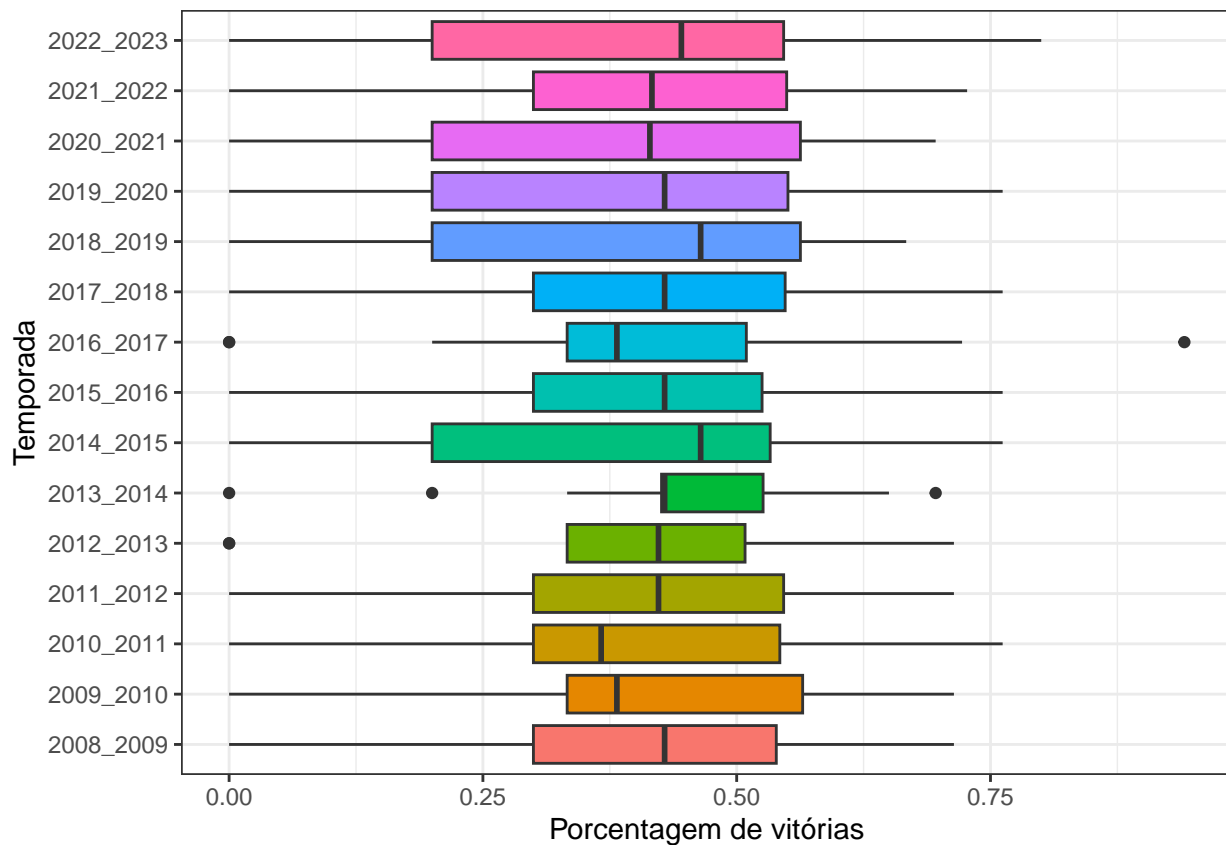


```
#Boxplot das porcentagens de vitória nas temporadas
dados_playoffs %>%
  dplyr::select(WINP) %>%
  ggplot() +
  geom_boxplot(aes(x = WINP)) +
  labs(title = "Boxplot das porcentagens de vitória nas temporadas",
       x = "Porcentagem de vitórias") +
  theme_bw()
```


Boxplot das porcentagens de vitória nas temporadas

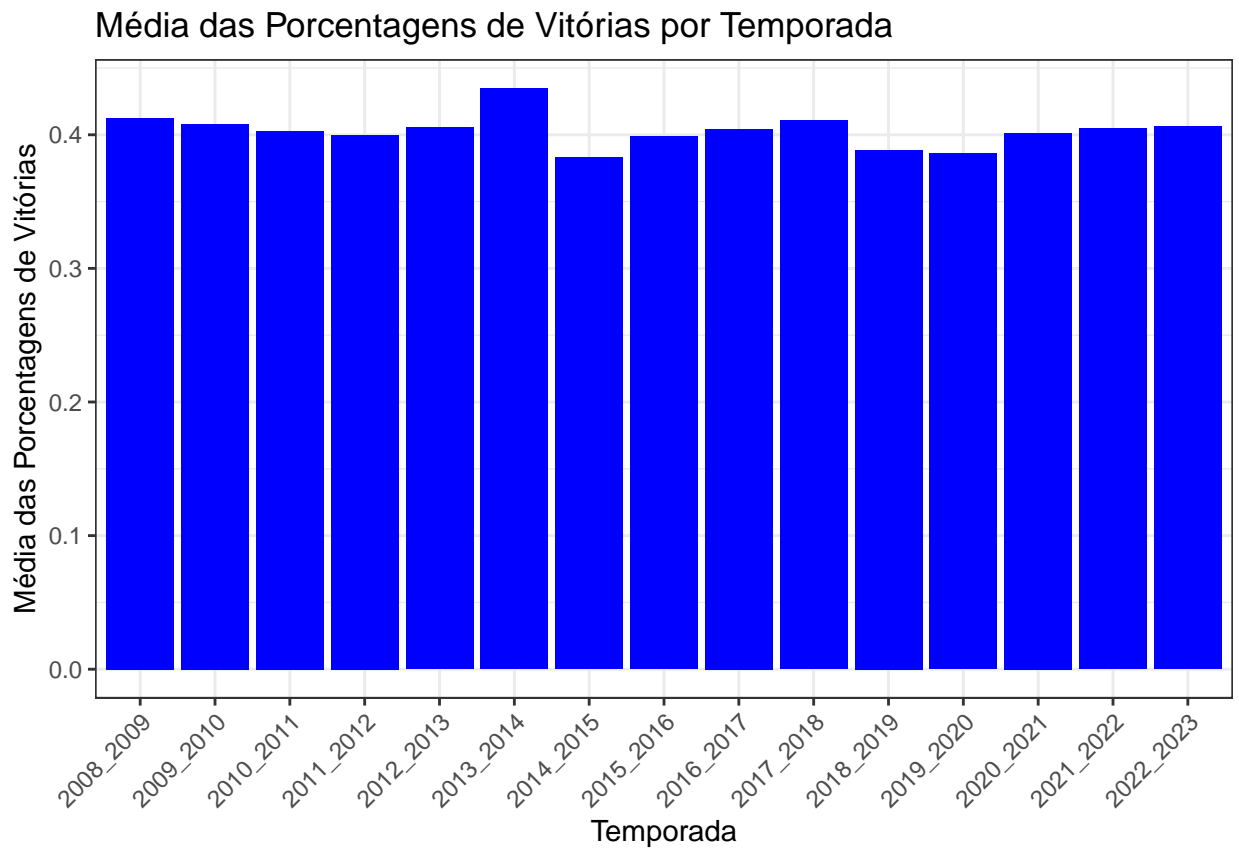


```
#Boxplot da porcentagem de vitórias durante as temporadas
box_play <- dados_playoffs %>%
  ggplot() +
  geom_boxplot(aes(x = WINP, y = Temporada, fill = Temporada), show.legend = FALSE) +
  labs(x = "Porcentagem de vitórias") +
  theme_bw()
box_play
```



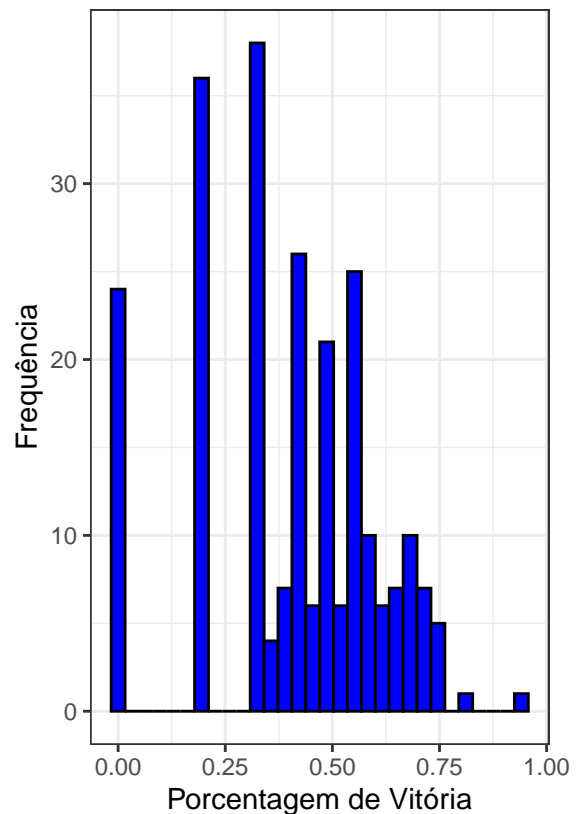
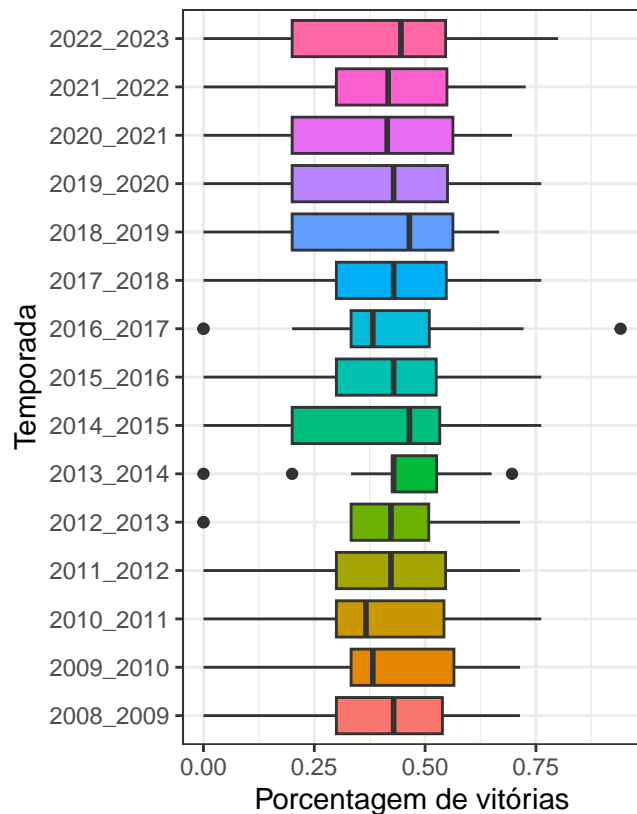
```
#Histograma com a porcentagem média de vitórias por temporada
dados_playoffs %>% dplyr::select(Team, Temporada, WINP) %>% group_by(Temporada) %>% summarise(Media = mean(WINP))

ggplot(aes(x = Temporada, y = Media)) +
  geom_bar(stat = "identity", fill = "blue") + # stat = "identity" para usar os valores diretamente
  labs(title = "Média das Porcentagens de Vitórias por Temporada",
        x = "Temporada",
        y = "Média das Porcentagens de Vitórias") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```



```
(box_play + hist_play)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Análise de 3 pontos

#Gráfico de linha com a variação das bolas de 3 tentadas durante as temporadas

```
tentadas3 <- dados_playoffs %>% group_by(Temporada) %>% summarise(chute = mean(`3PA`)) %>%
  ggplot(aes(x = Temporada, y = chute, group = 1)) +
  geom_line() +
  labs(title = "Tentados de 3 pontos",
       x = "Temporada",
       y = "Média de tentativas") +
  theme_minimal() +
  theme(axis.text.x = element_blank(),
        axis.ticks.x = element_blank())
```

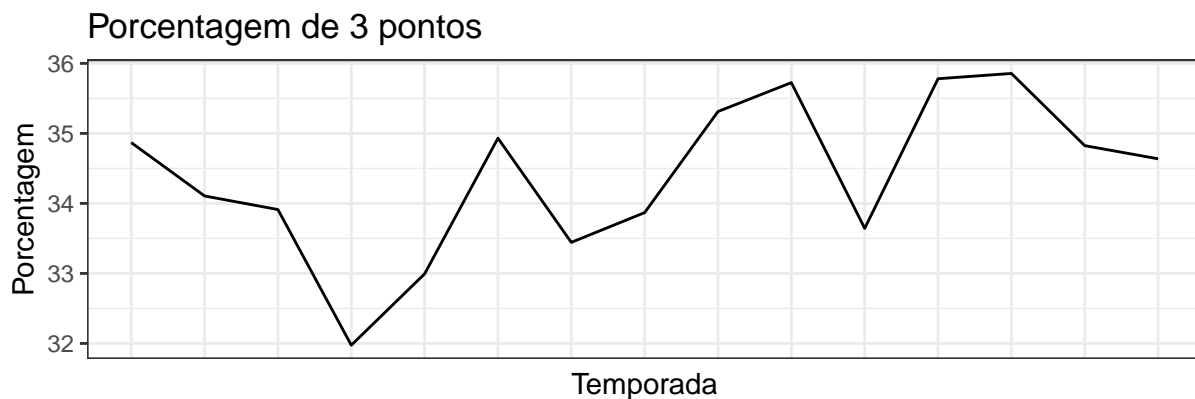
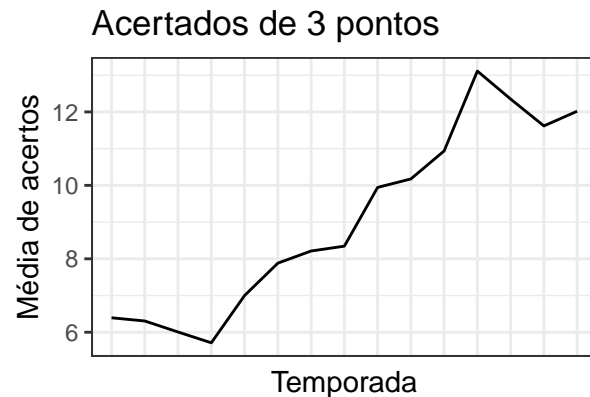
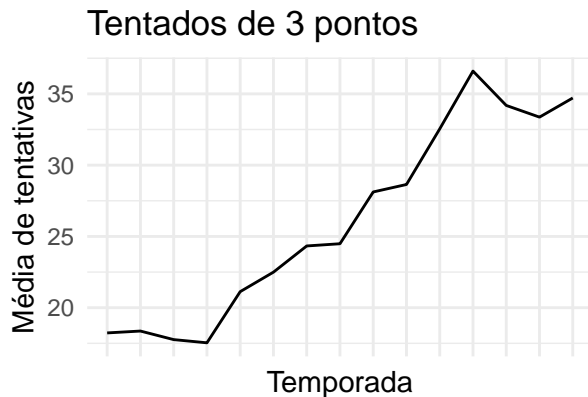
#Gráfico de linha com a variação das bolas de 3 acertadas durante as temporadas

```
acertadas3 <- dados_playoffs %>% group_by(Temporada) %>% summarise(chute = mean(`3PM`)) %>%
  ggplot(aes(x = Temporada, y = chute, group = 1)) +
  geom_line() +
  labs(title = "Acertados de 3 pontos",
       x = "Temporada",
       y = "Média de acertos") +
  theme_bw() +
  theme(axis.text.x = element_blank(),
        axis.ticks.x = element_blank())
```

#Gráfico de linha com a variação da porcentagem de bolas de 3 durante as temporadas

```
porcentagem3 <- dados_playoffs %>% group_by(Temporada) %>% summarise(chute = mean(`3PP`)) %>%
  ggplot(aes(x = Temporada, y = chute, group = 1,xlab="")) +
  geom_line() +
```

```
labs(title = "Porcentagem de 3 pontos",
     x = "Temporada",
     y = "Porcentagem") +
theme_bw() +
theme(axis.text.x = element_blank(),
      axis.ticks.x = element_blank())
(tentadas3 + acertadas3)/(porcentagem3)
```



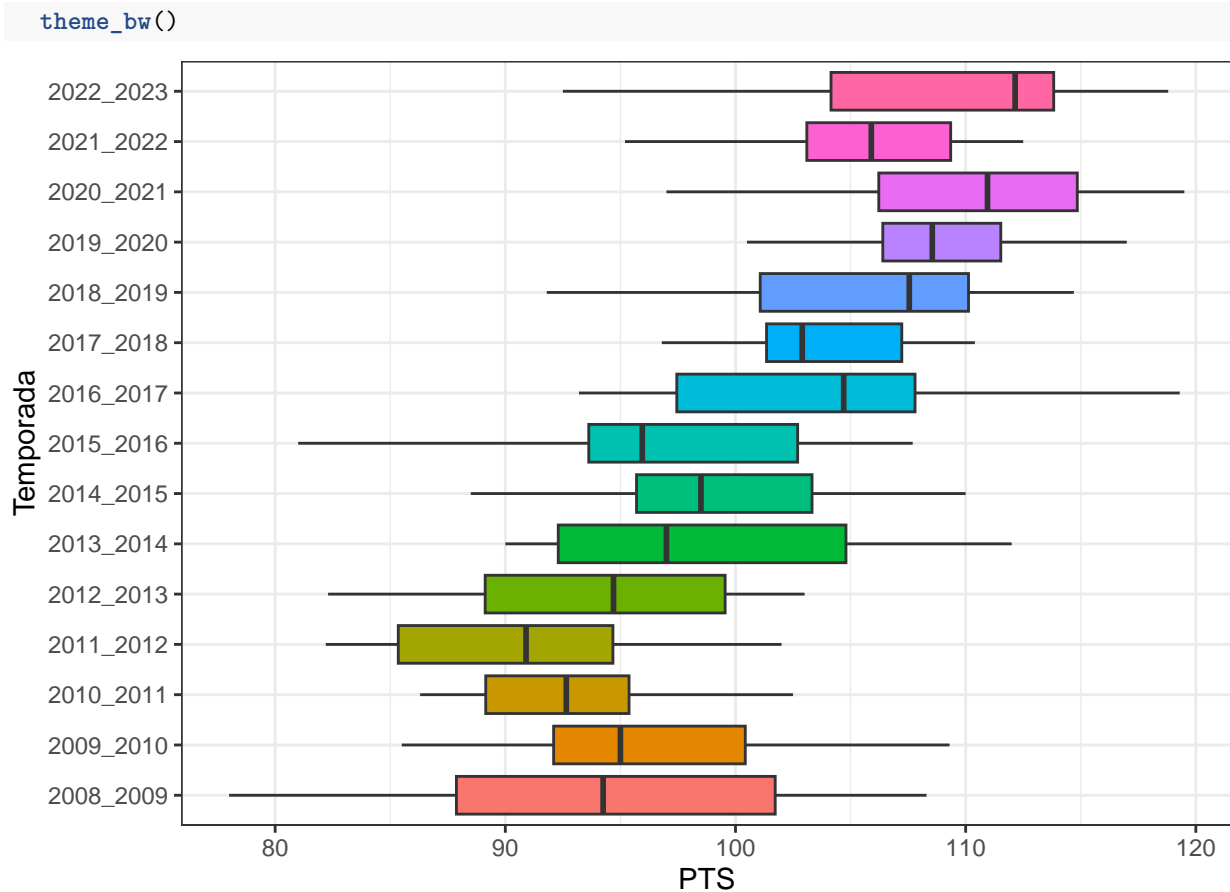
```
#####

#####

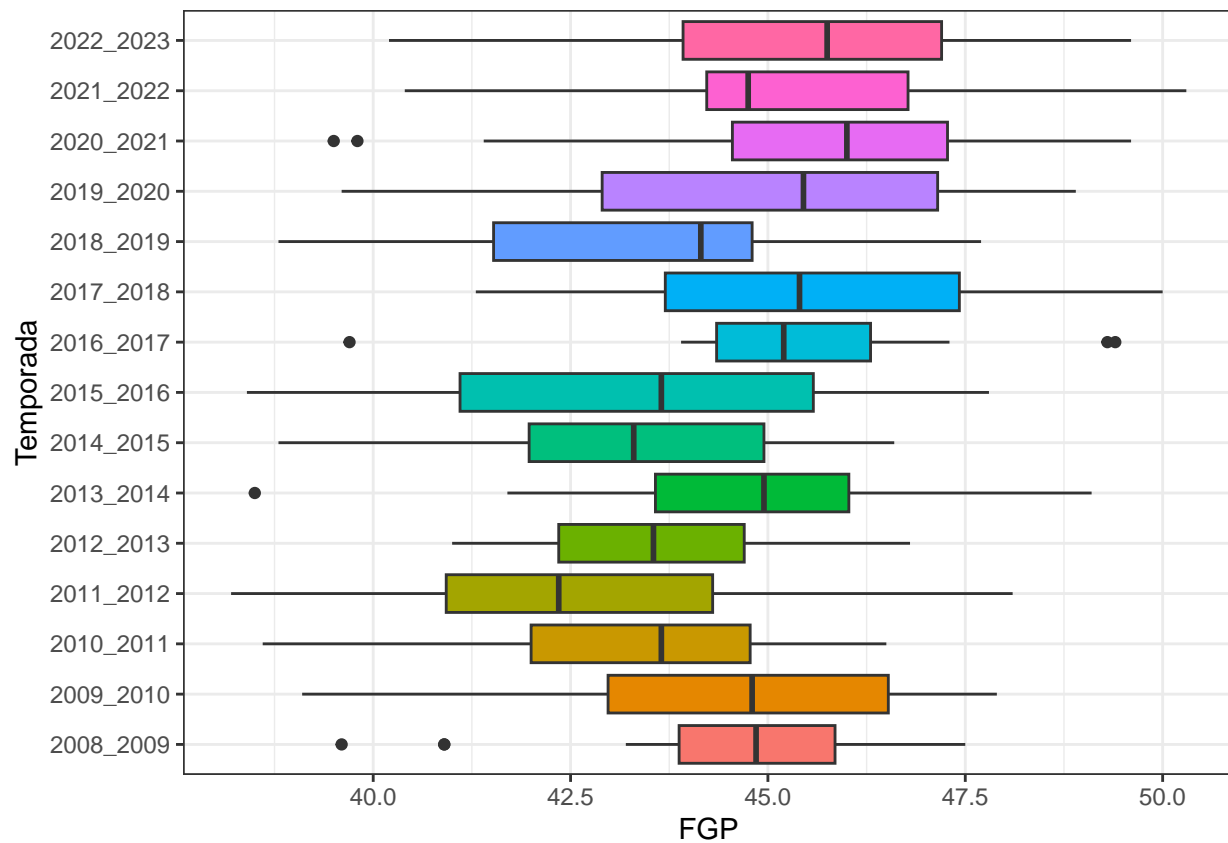
#Gráfico de linha com a Média dae pontos por temporada
pontos_playoffs <- dados_playoffs %>% group_by(Temporada) %>% summarise(Media_Pontos = mean(PTS)) %>%
  ggplot(aes(x = Temporada, y = Media_Pontos, group = 1)) +
  geom_line() +
  labs(title = "Média de pontos nos playoffs",
       x = "Temporada",
       y = "Média das Porcentagens de Vitórias") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus

##### Boxplots ao decorrer das temporadas da variável resposta com a variável: #####

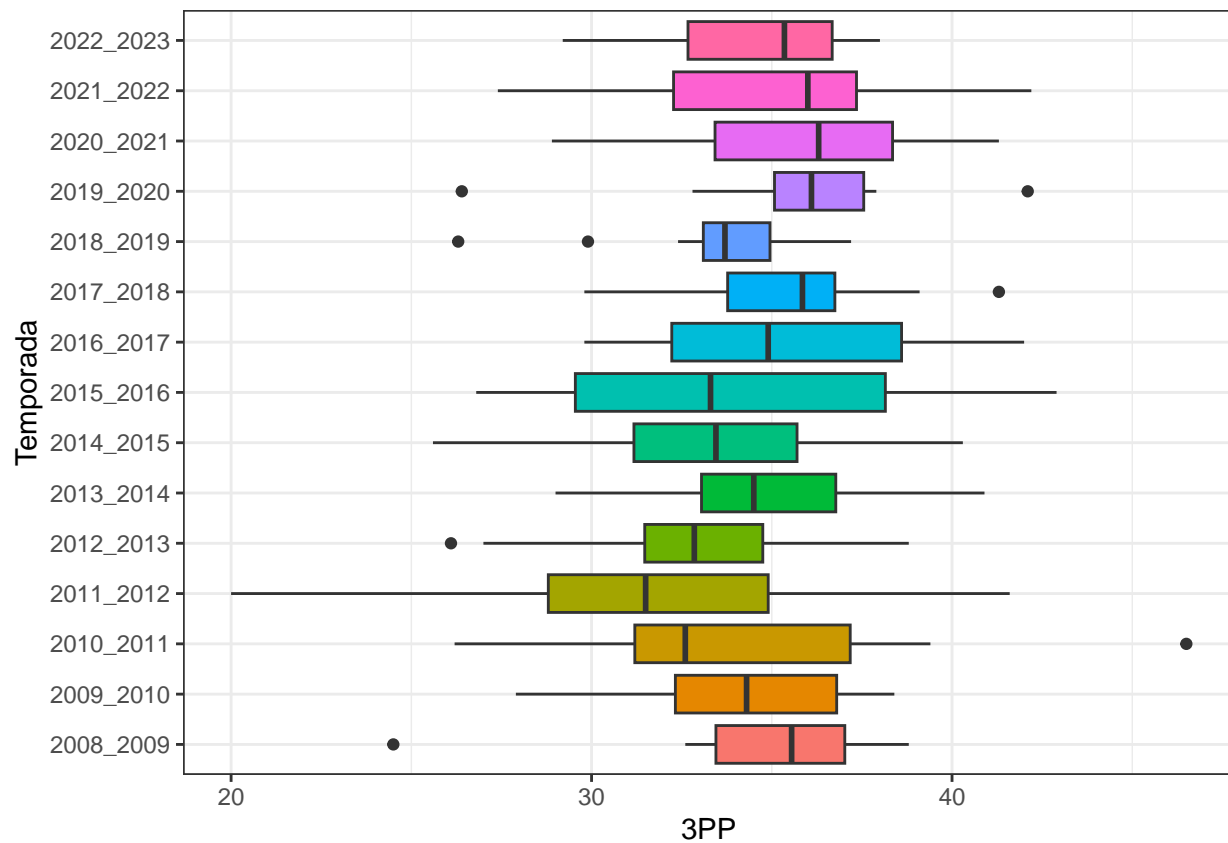
#Pontos
dados_playoffs %>%
  ggplot() +
  geom_boxplot(aes(x = PTS, y = Temporada, fill = Temporada), show.legend = FALSE) +
```



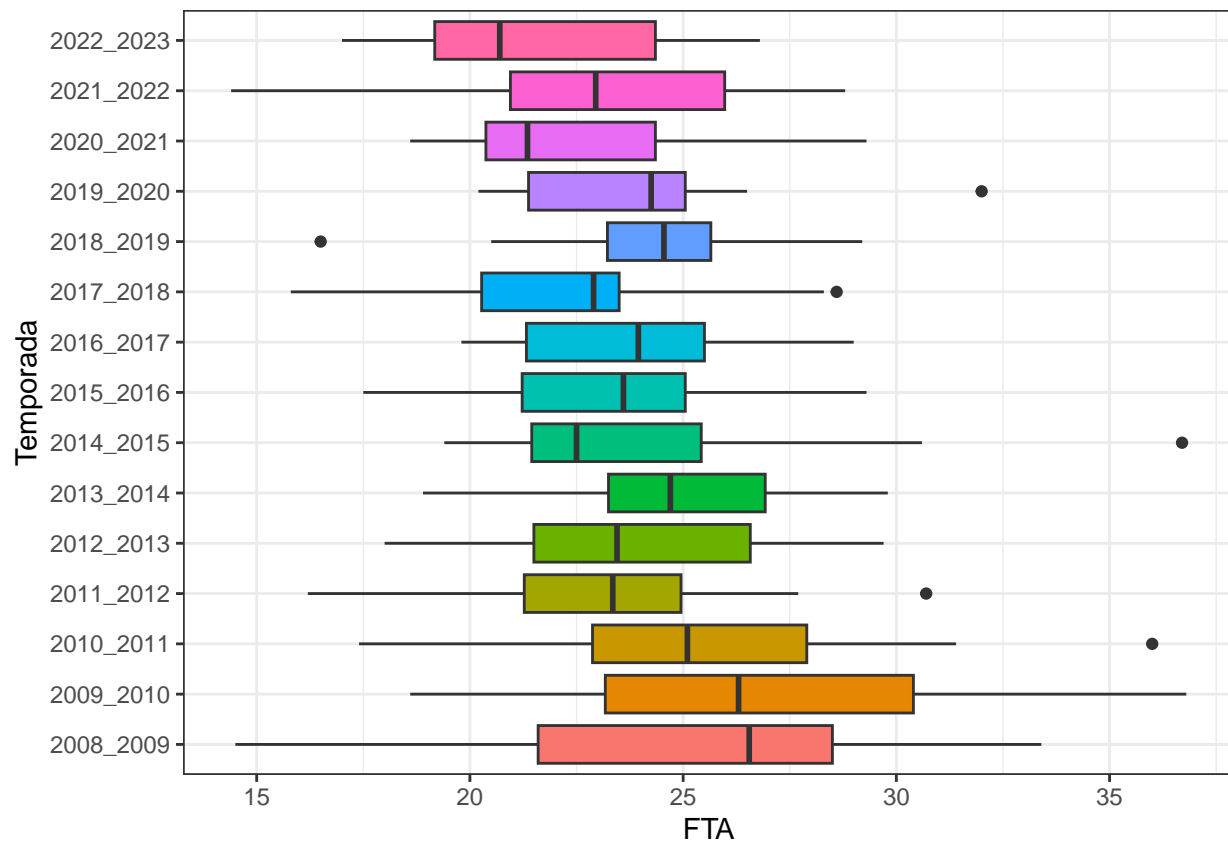
```
#Porcentagem de arremessos livres
dados_playoffs %>%
  ggplot() +
  geom_boxplot(aes(x = FGP, y = Temporada, fill = Temporada), show.legend = FALSE) +
  theme_bw()
```



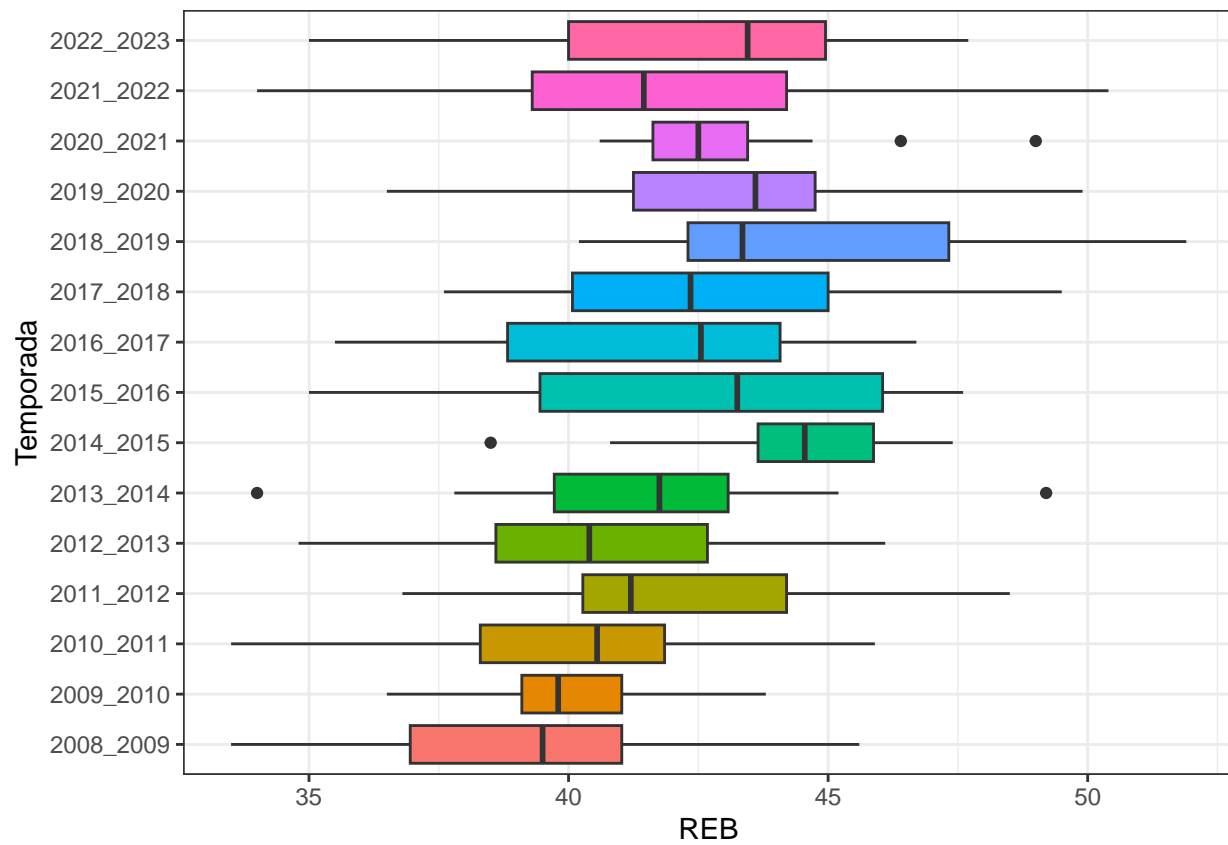
```
#Porcentagem de bolas de três convertidas
dados_playoffs %>%
  ggplot() +
  geom_boxplot(aes(x = `3PP`, y = Temporada, fill = Temporada), show.legend = FALSE) +
  theme_bw()
```



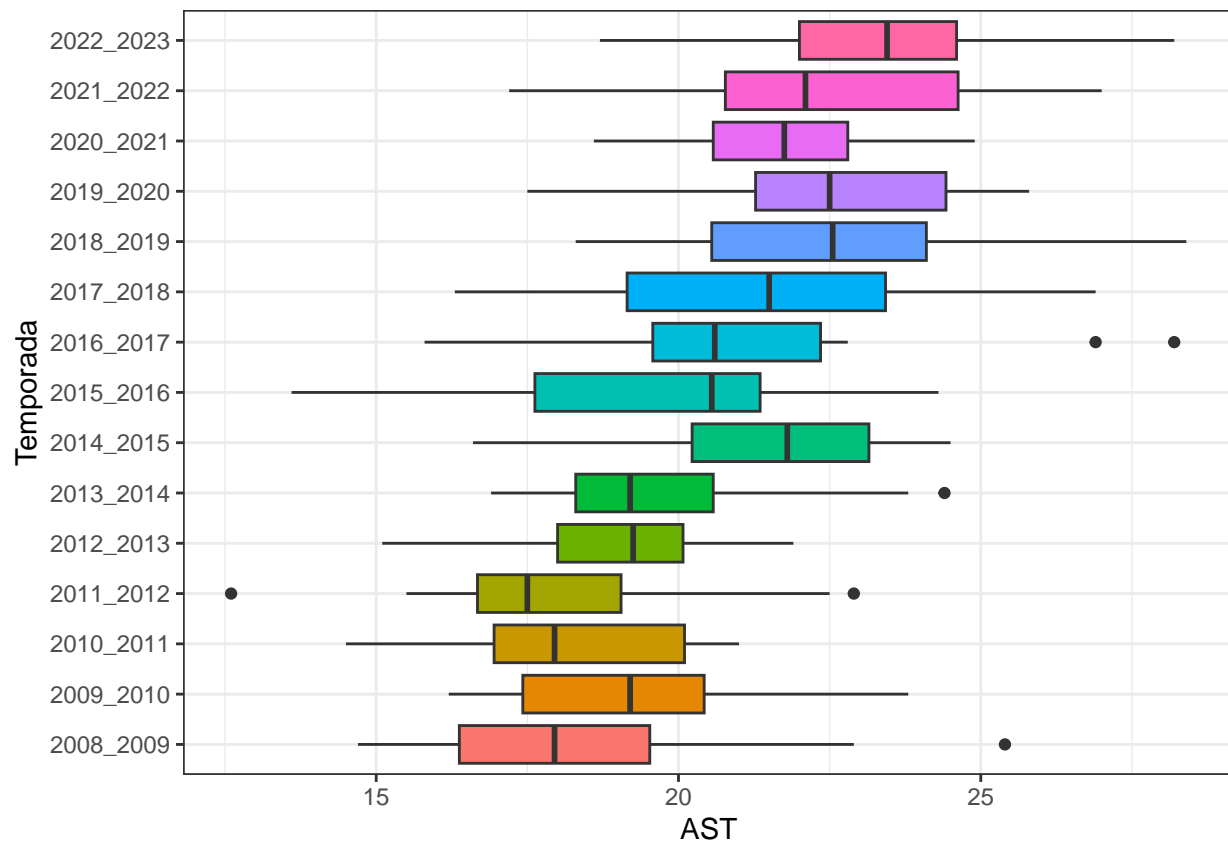
```
#Tentativas de lance livre
dados_playoffs %>%
  ggplot() +
  geom_boxplot(aes(x = FTA, y = Temporada, fill = Temporada), show.legend = FALSE) +
  theme_bw()
```

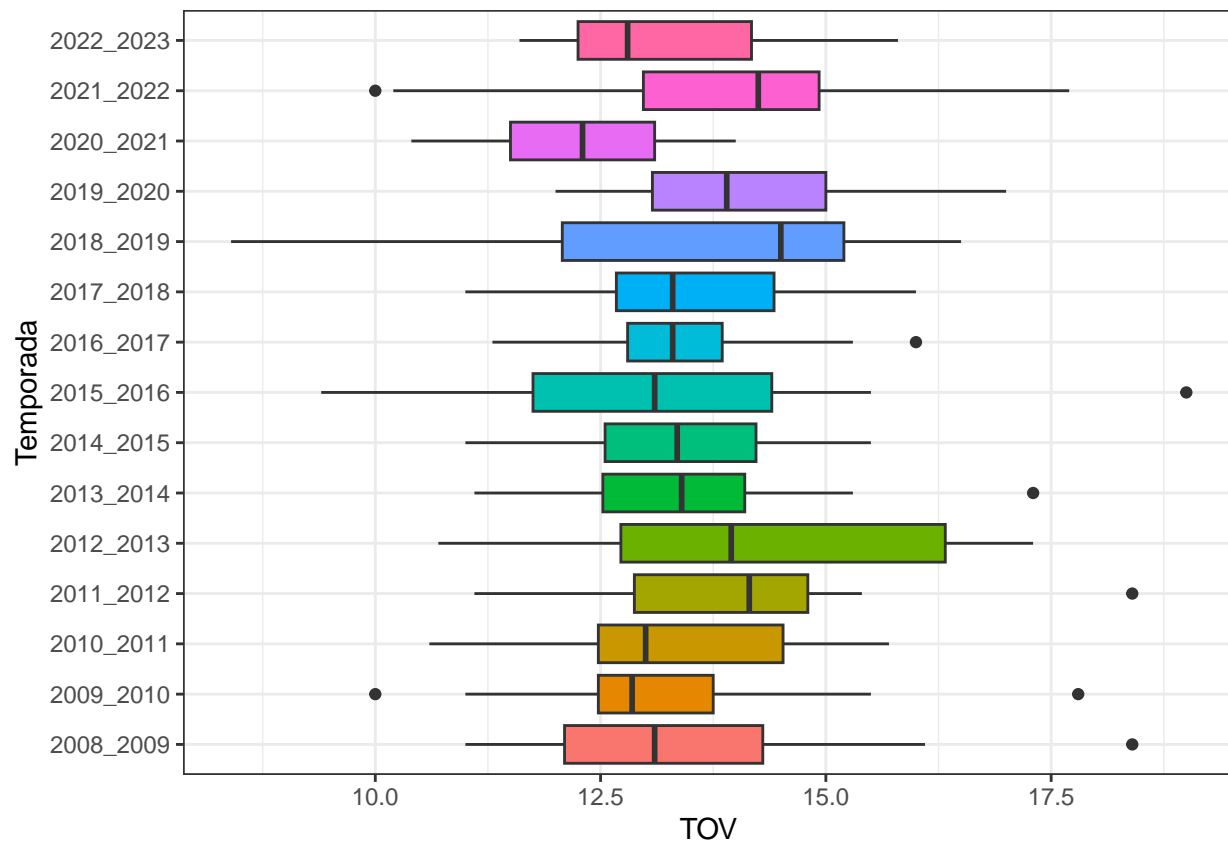
```
#Rebotes
dados_playoffs %>%
  ggplot() +
  geom_boxplot(aes(x = REB, y = Temporada, fill = Temporada), show.legend = FALSE) +
  theme_bw()
```



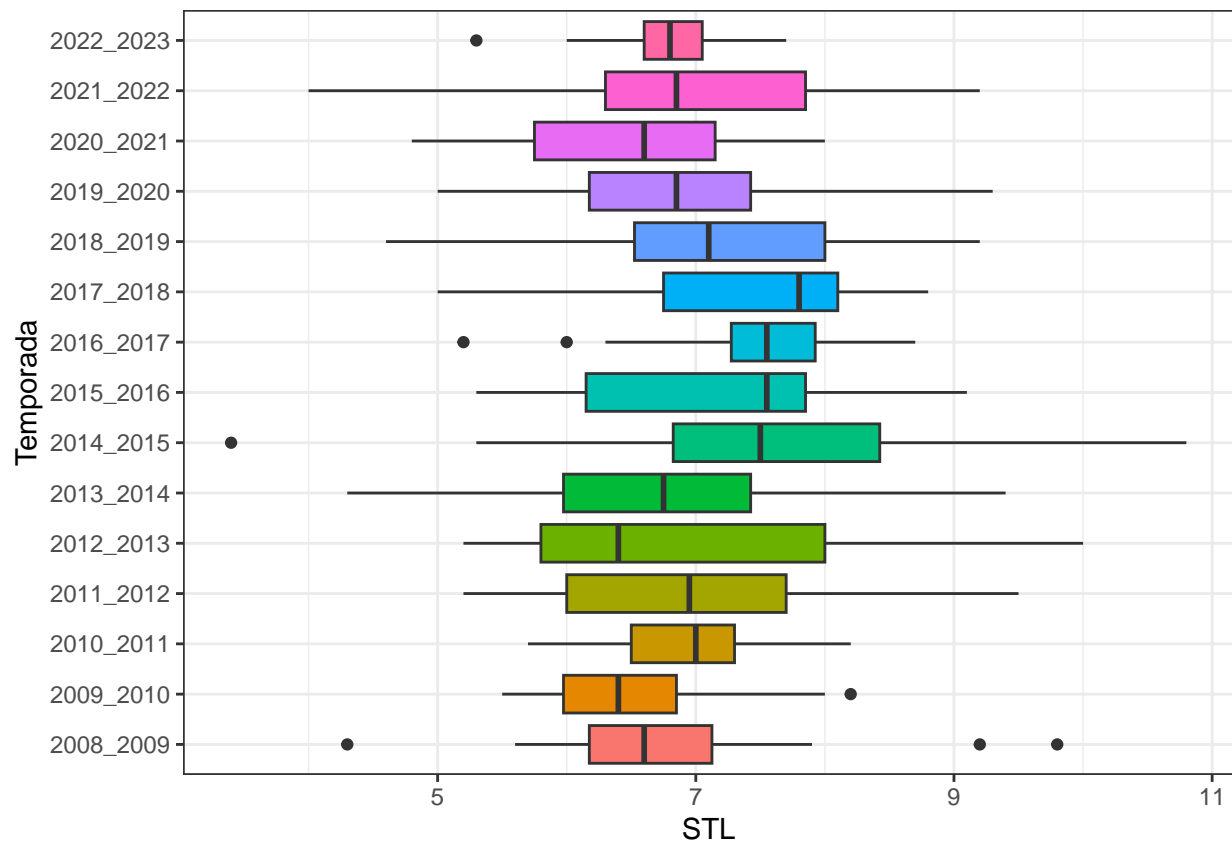
```
#Assistência
dados_playoffs %>%
  ggplot() +
  geom_boxplot(aes(x = AST, y = Temporada, fill = Temporada), show.legend = FALSE) +
  theme_bw()
```



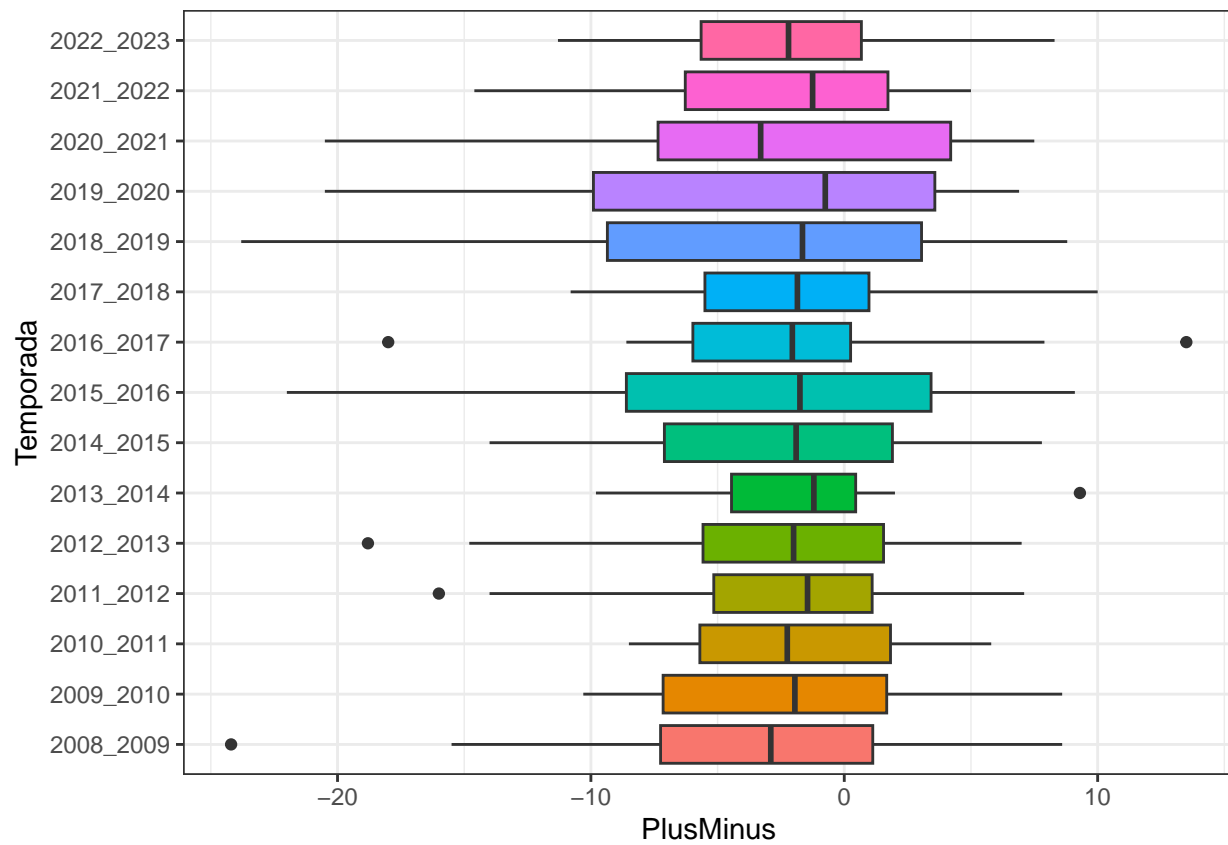
```
#Turnovers
dados_playoffs %>%
  ggplot() +
  geom_boxplot(aes(x = TOV, y = Temporada, fill = Temporada), show.legend = FALSE) +
  theme_bw()
```



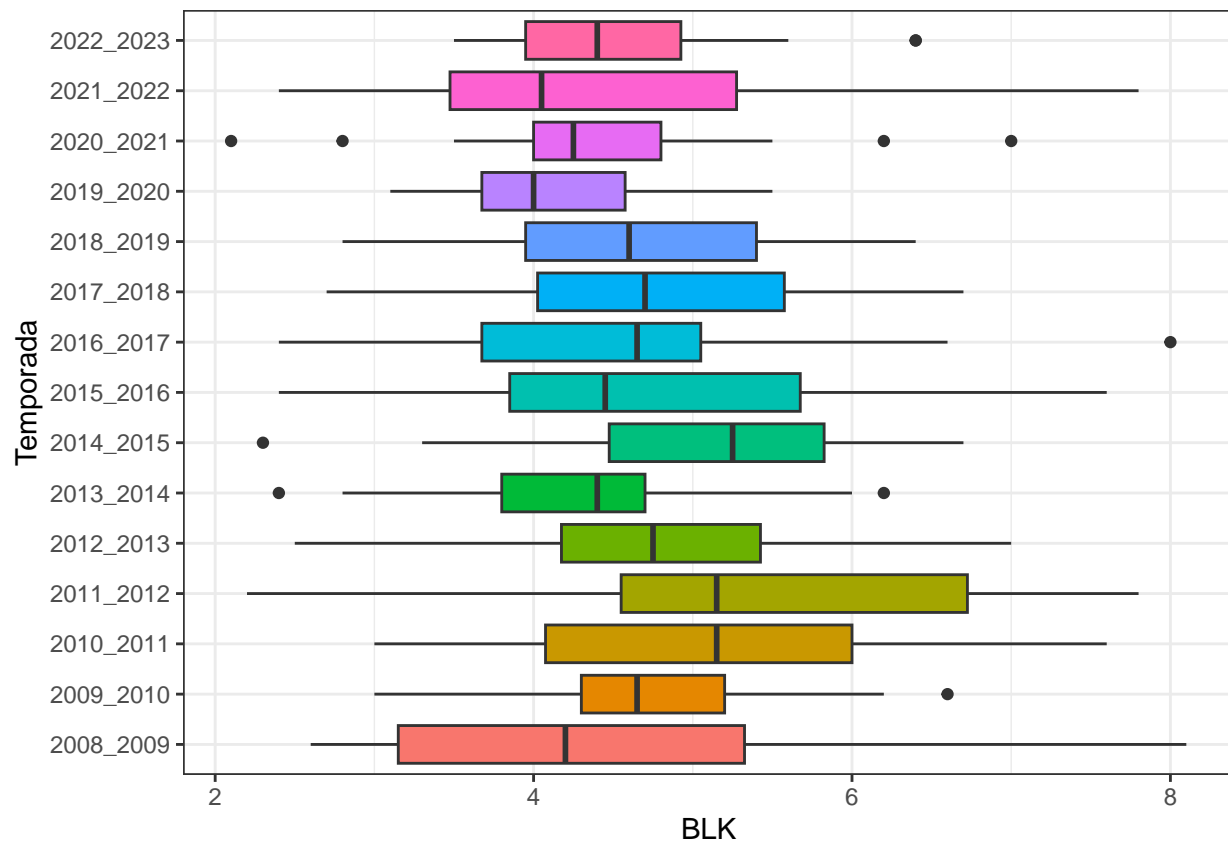
```
#Roubada de bola
dados_playoffs %>%
  ggplot() +
  geom_boxplot(aes(x = STL, y = Temporada, fill = Temporada), show.legend = FALSE) +
  theme_bw()
```



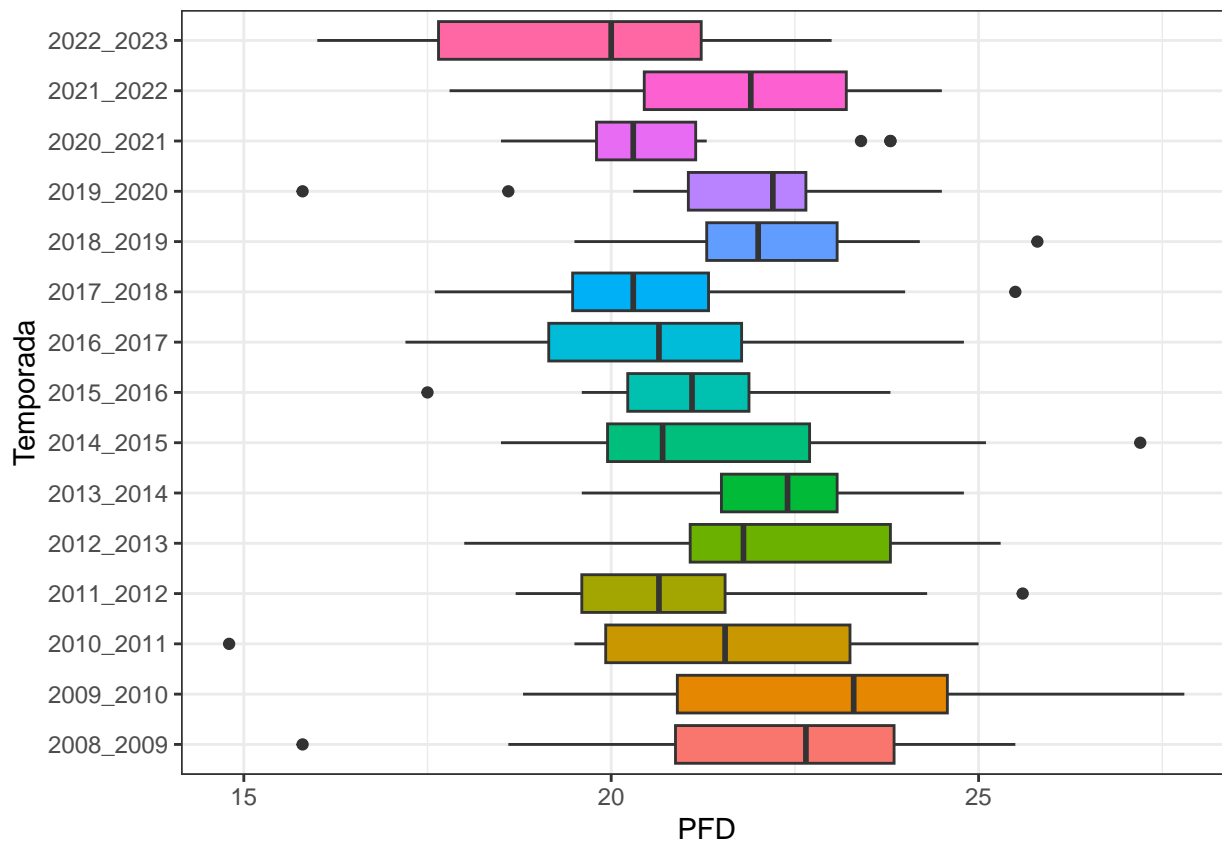
```
#Plus_minus
dados_playoffs %>%
  ggplot() +
  geom_boxplot(aes(x = PlusMinus, y = Temporada, fill = Temporada), show.legend = FALSE) +
  theme_bw()
```



```
#Tocos
dados_playoffs %>%
  ggplot() +
  geom_boxplot(aes(x = BLK, y = Temporada, fill = Temporada), show.legend = FALSE) +
  theme_bw()
```



```
#Faltas
dados_playoffs %>%
  ggplot() +
  geom_boxplot(aes(x = PFD, y = Temporada, fill = Temporada), show.legend = FALSE) +
  theme_bw()
```



Análise de Correlação

```
library(GGally)
```

```
## Registered S3 method overwritten by 'GGally':
```

```
##   method from
```

```
##   +.gg      ggplot2
```

```
#Correlação entre todas as variáveis que farão parte do modelo
```

```
#A visualização não está muito boa
```

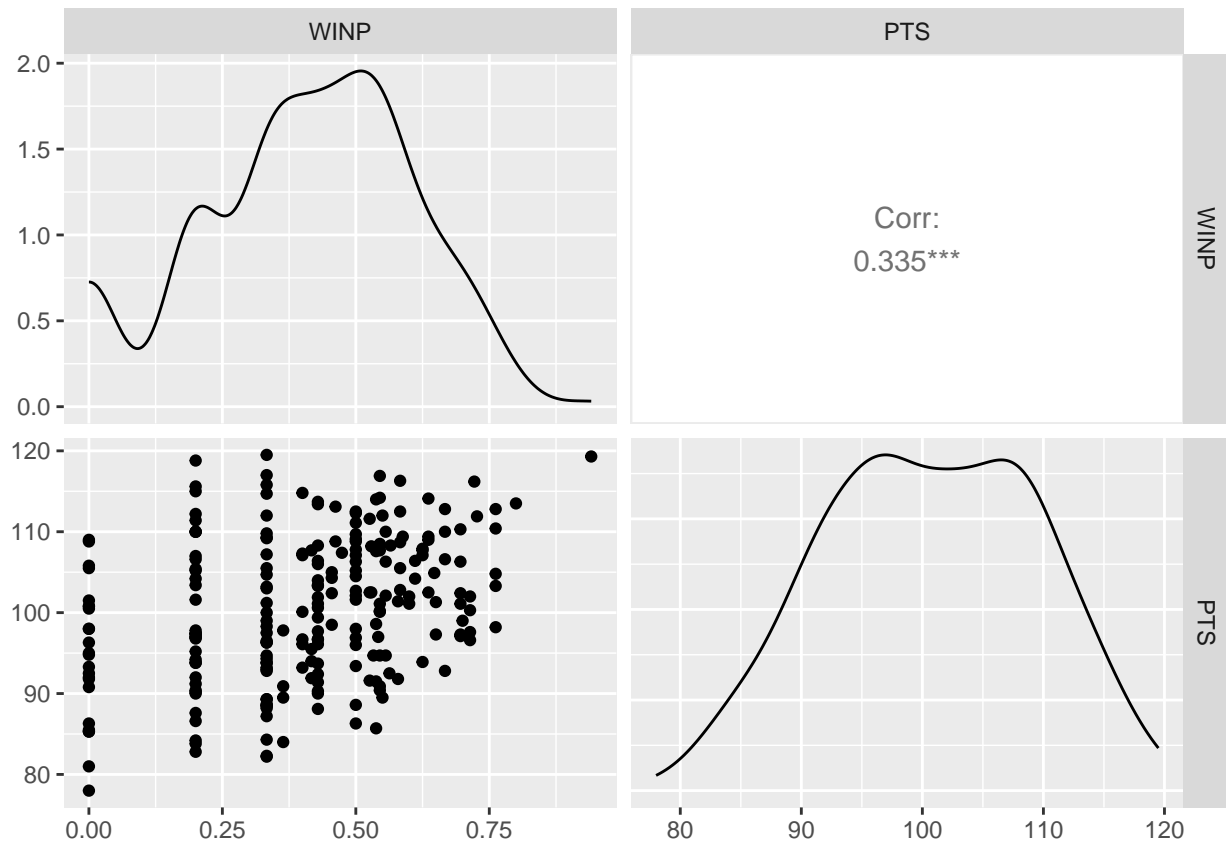
```
#Rodar apenas se quiser ver, mas não compensa
```

```
#dados_playoffs %>% dplyr::select(-c(Posicao,TEAM, GP, W, L, MIN, Temporada)) %>% ggpairs() #Demora pa
```

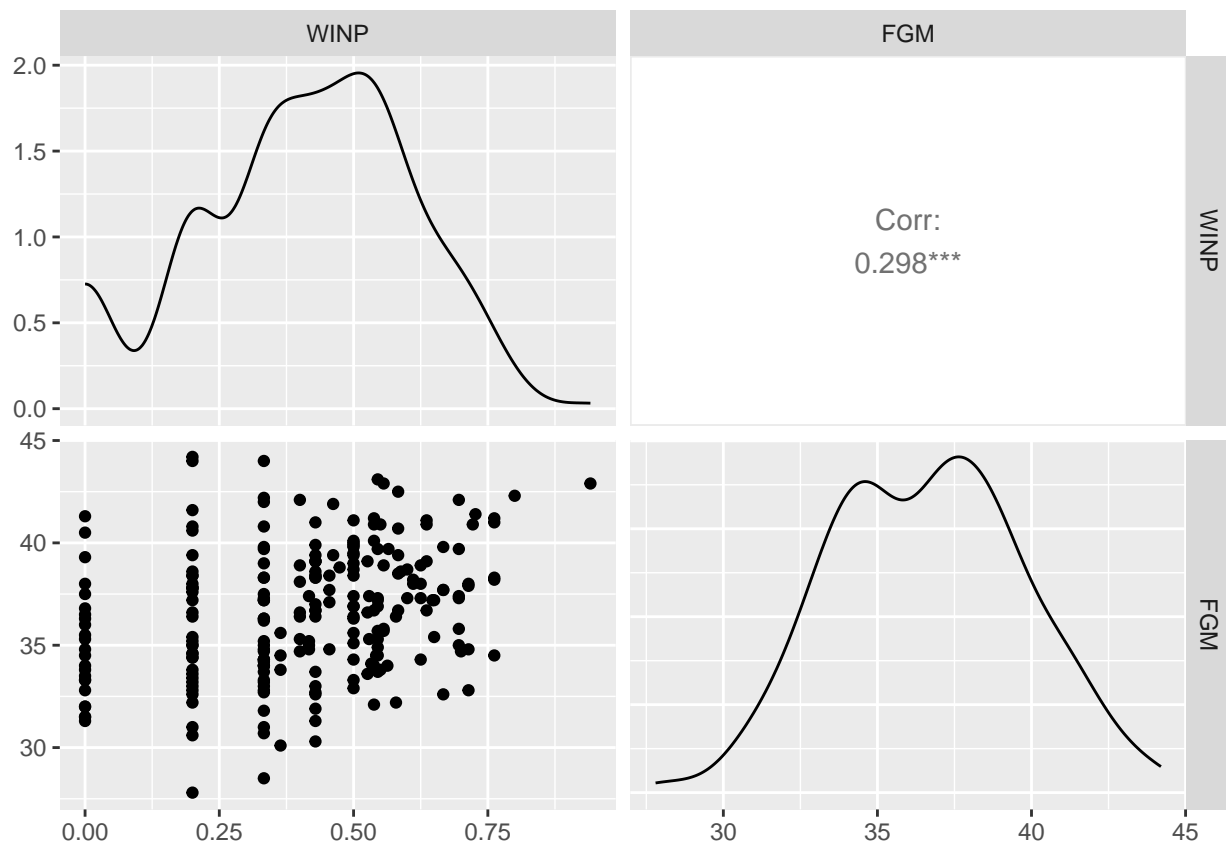
```
#### Análise correlação entre a variável resposta e a seguintes variável:
```

```
#Pontos
```

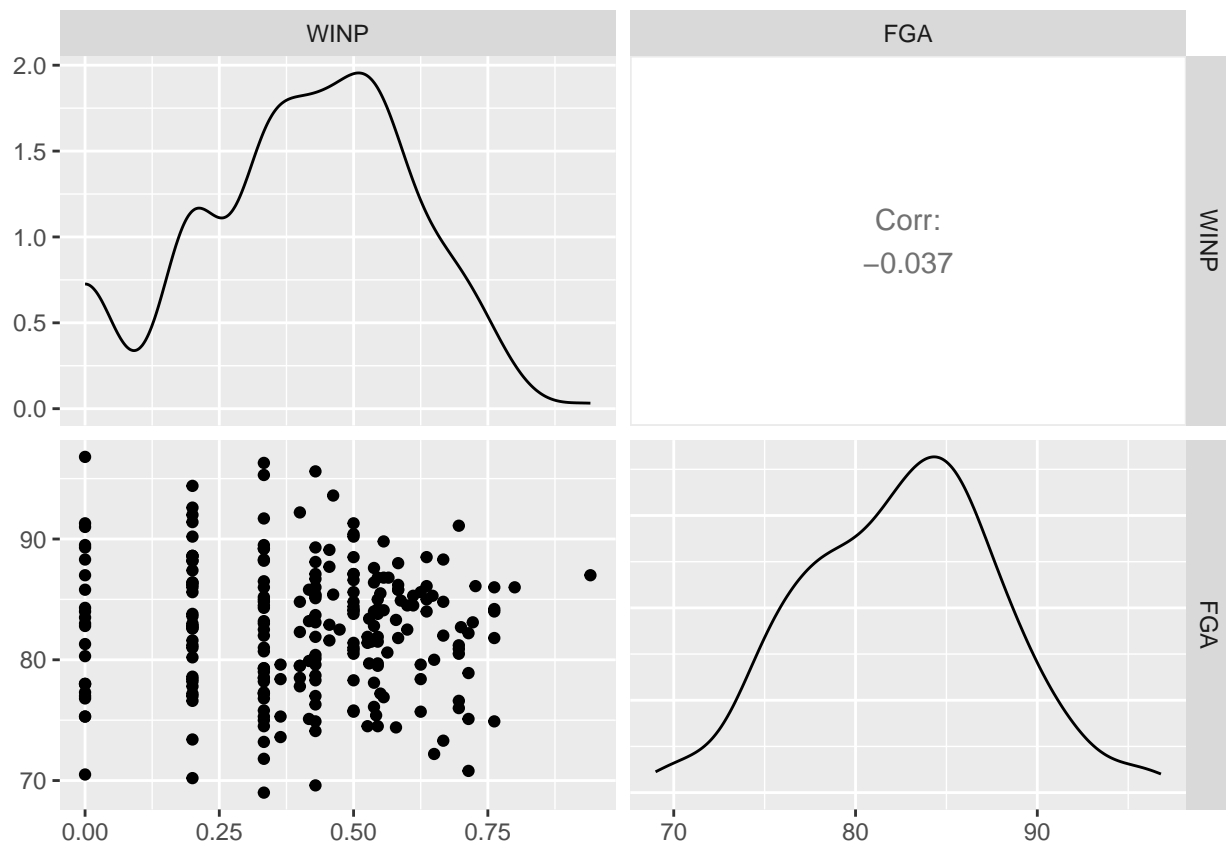
```
dados_playoffs %>% dplyr::select(c(WINP, PTS)) %>% ggpairs()
```

```
#Arremessos feitos
dados_playoffs %>% dplyr::select(c(WINP, FGM)) %>% ggpairs()
```

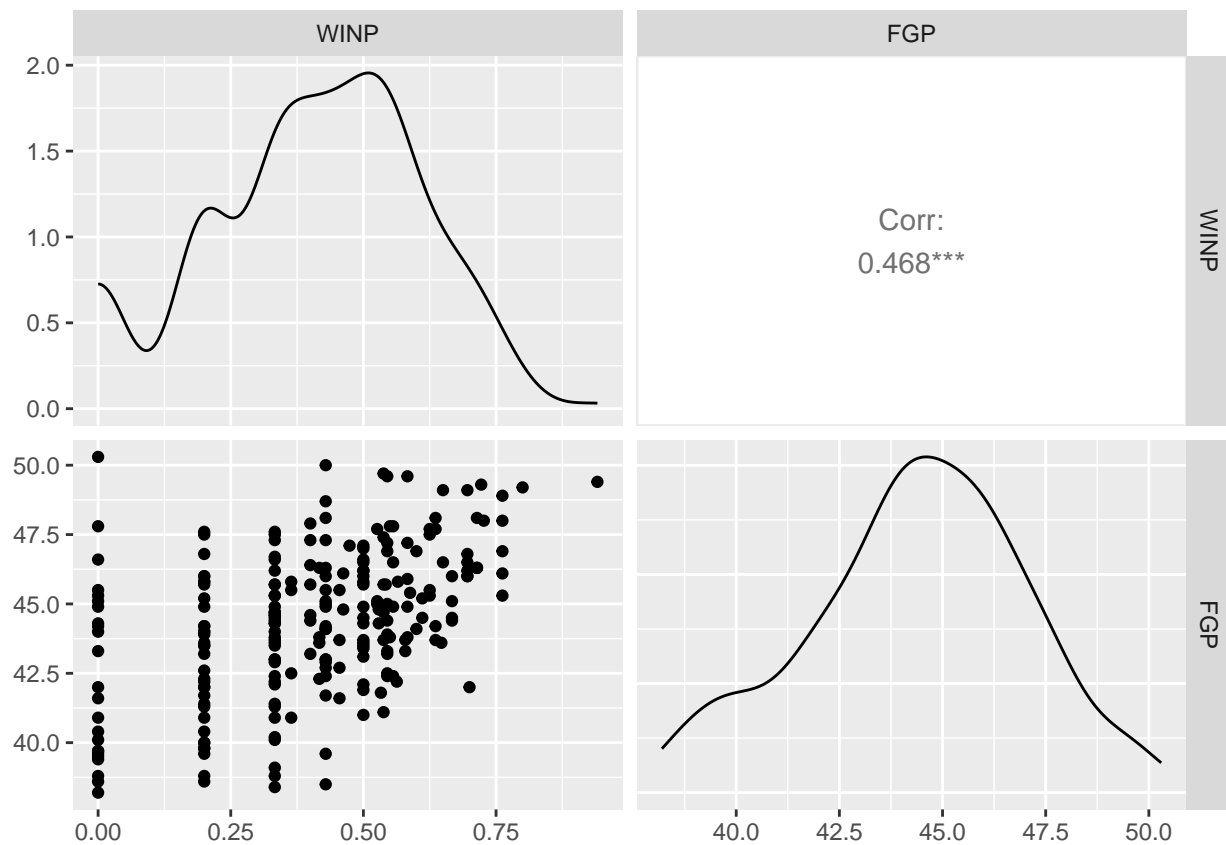


```
#Tentativas de Arremessos
dados_playoffs %>% dplyr::select(c(WINP, FGA)) %>% ggpairs()
```

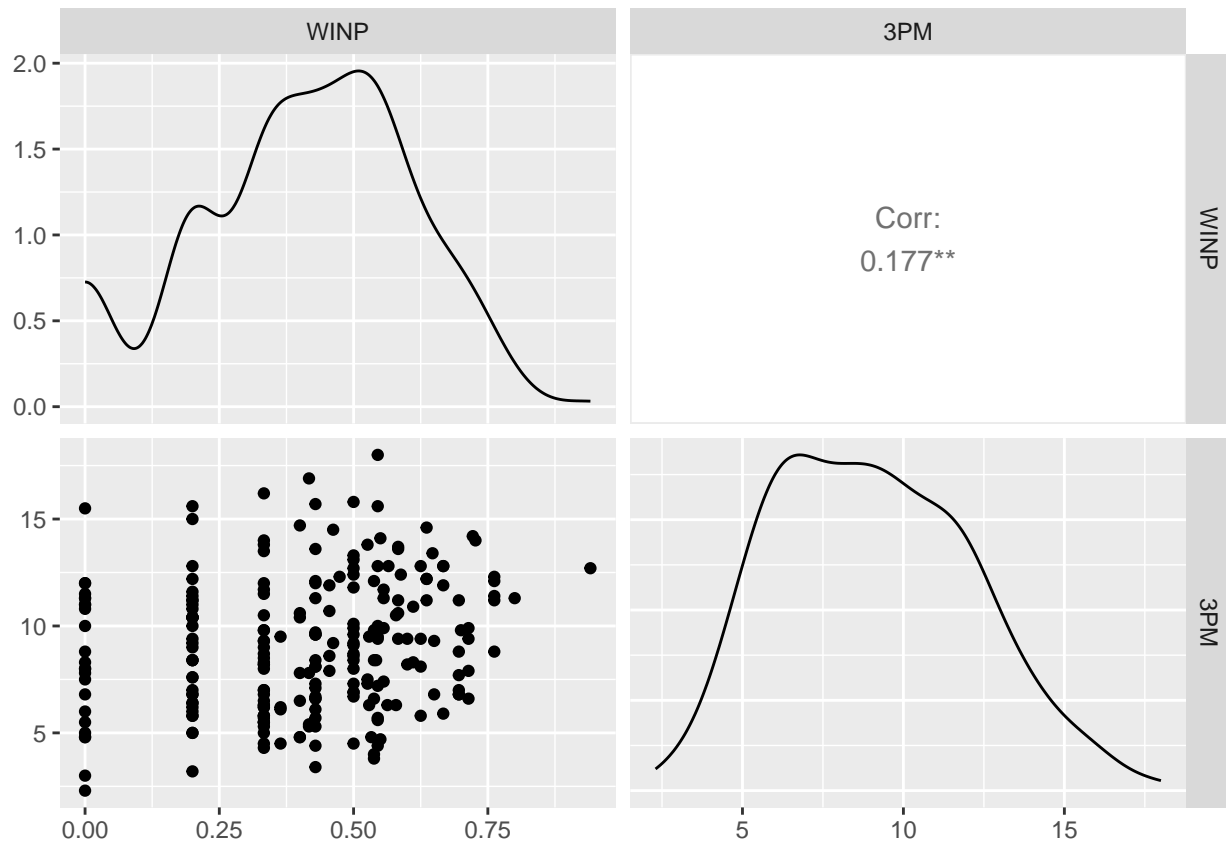


##Porcentagem de Arremessos

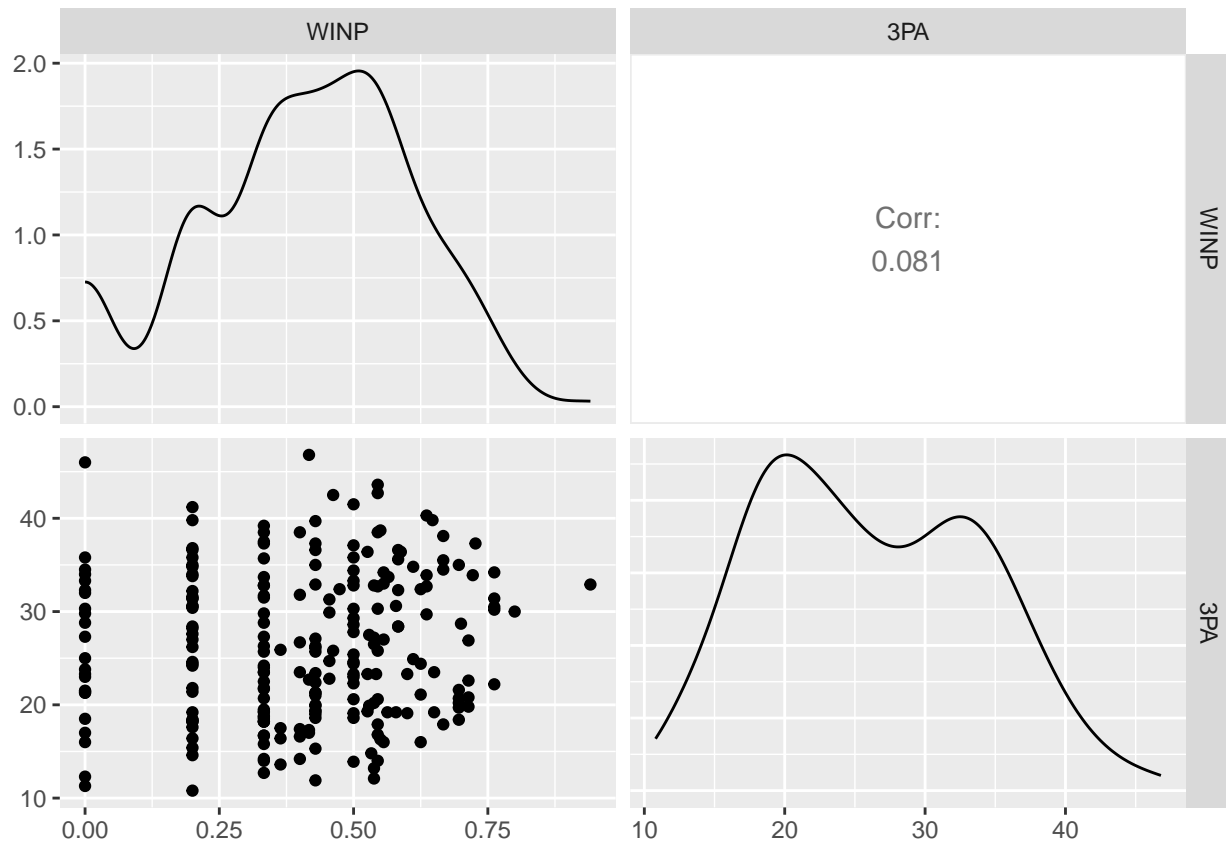
```
dados_playoffs %>% dplyr::select(c(WNP, FGA)) %>% ggpairs()
```



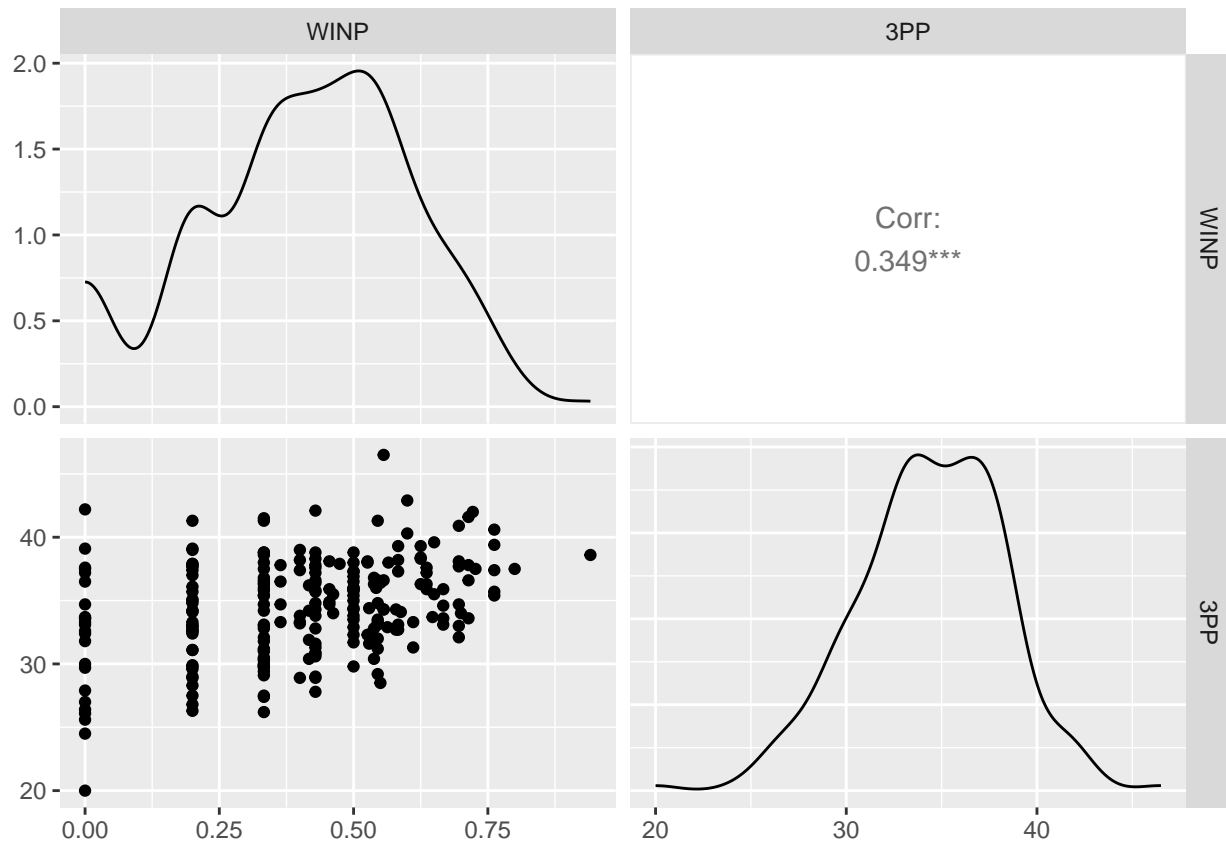
```
#3 pontos feitos
dados_playoffs %>% dplyr::select(c(WINP, `3PM`)) %>% ggpairs()
```



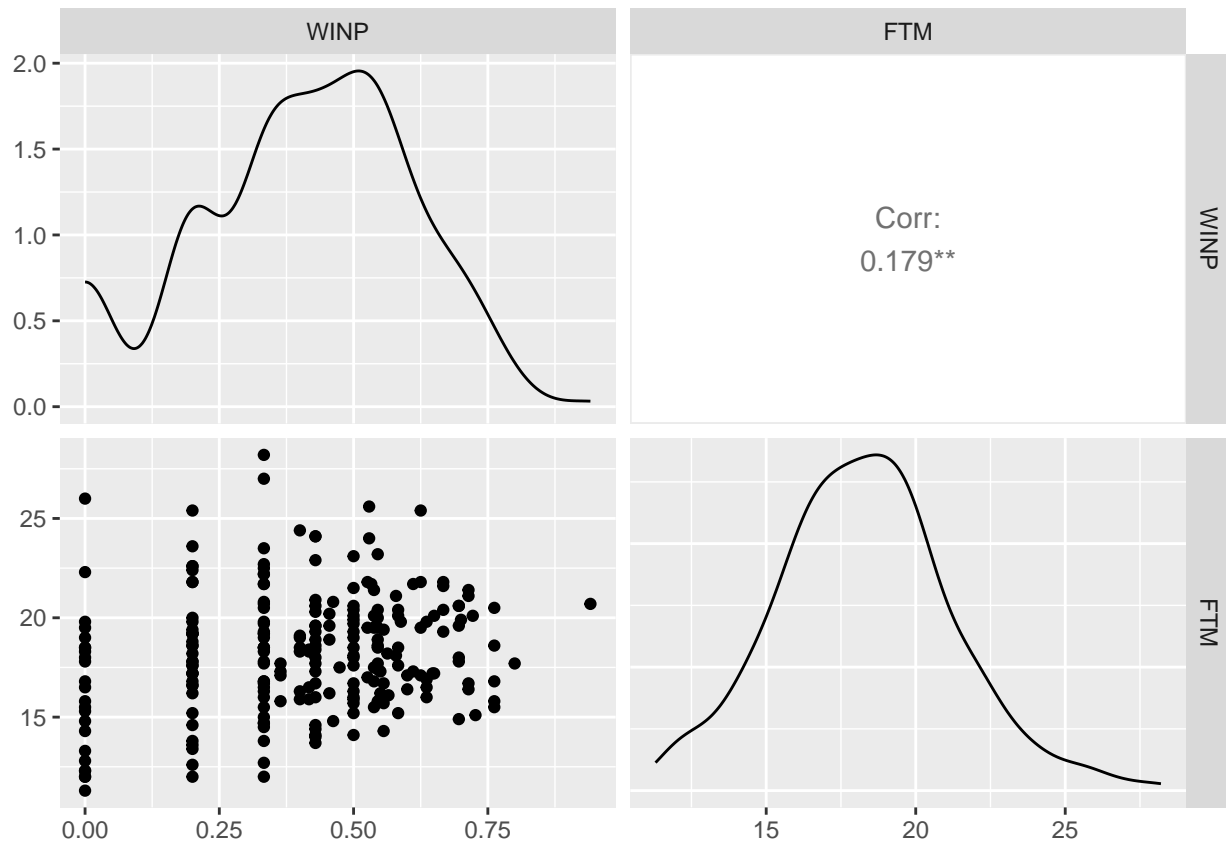
```
#3 pontos tentados
dados_playoffs %>% dplyr::select(c(WINP, `3PA`)) %>% ggpairs()
```



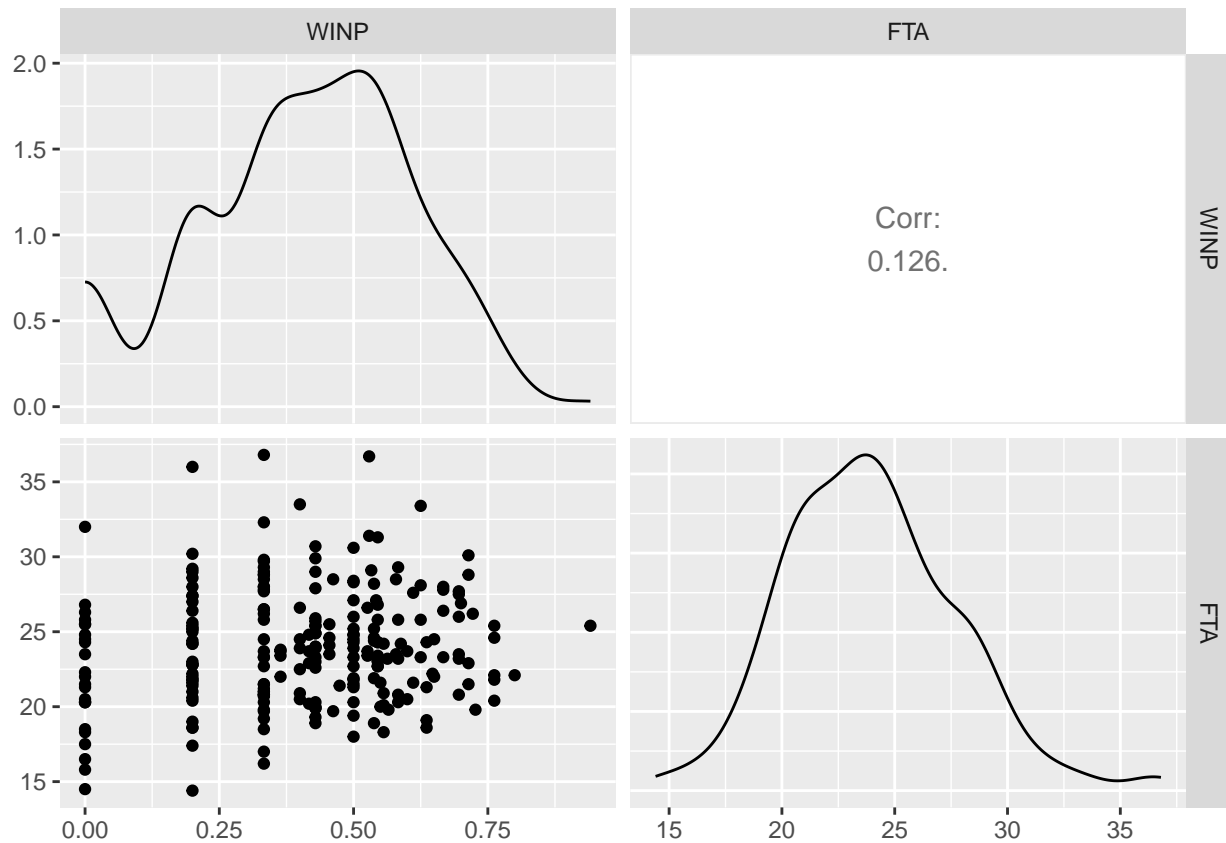
```
#Porcentagem de 3 pontos acertados
dados_playoffs %>% dplyr::select(c(WINP, `3PP`)) %>% ggpairs()
```



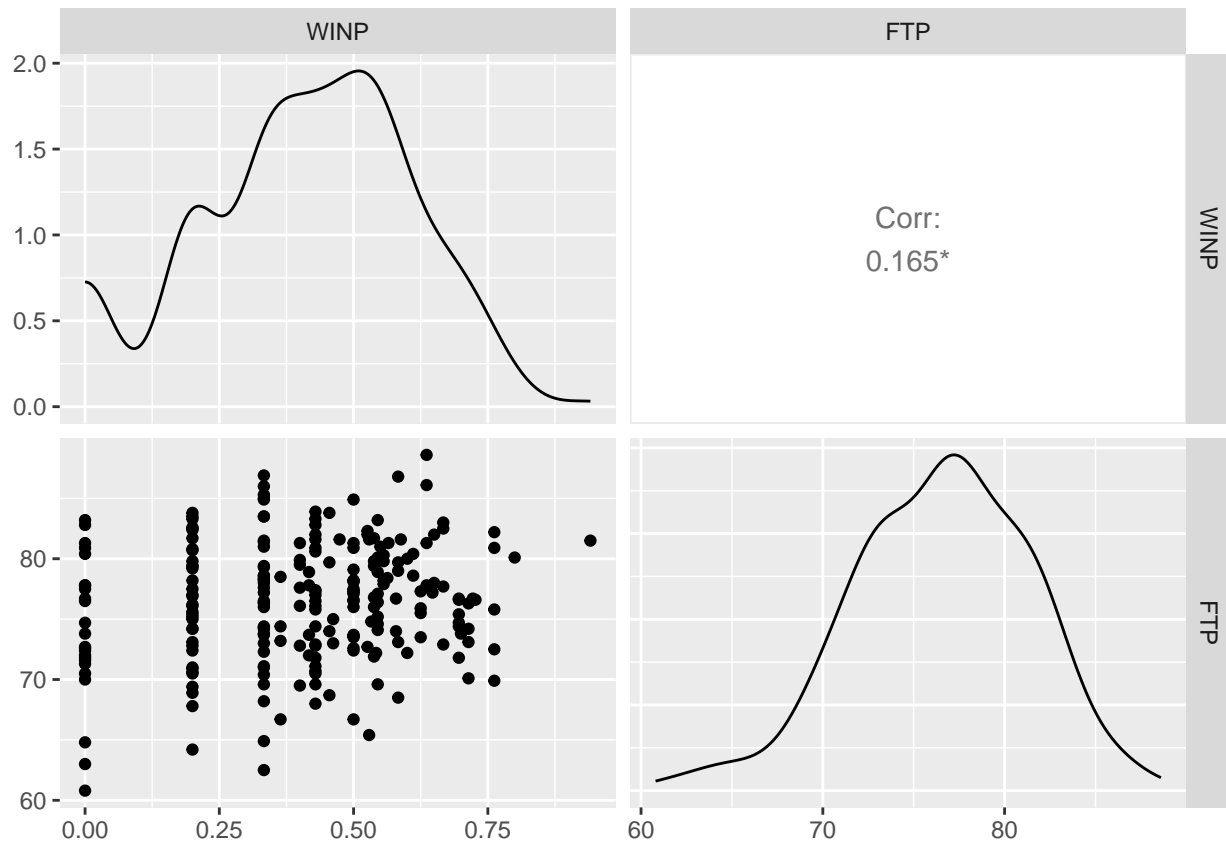
```
#Lances livres feitos
dados_playoffs %>% dplyr::select(c(WINP, FTM)) %>% ggpairs()
```



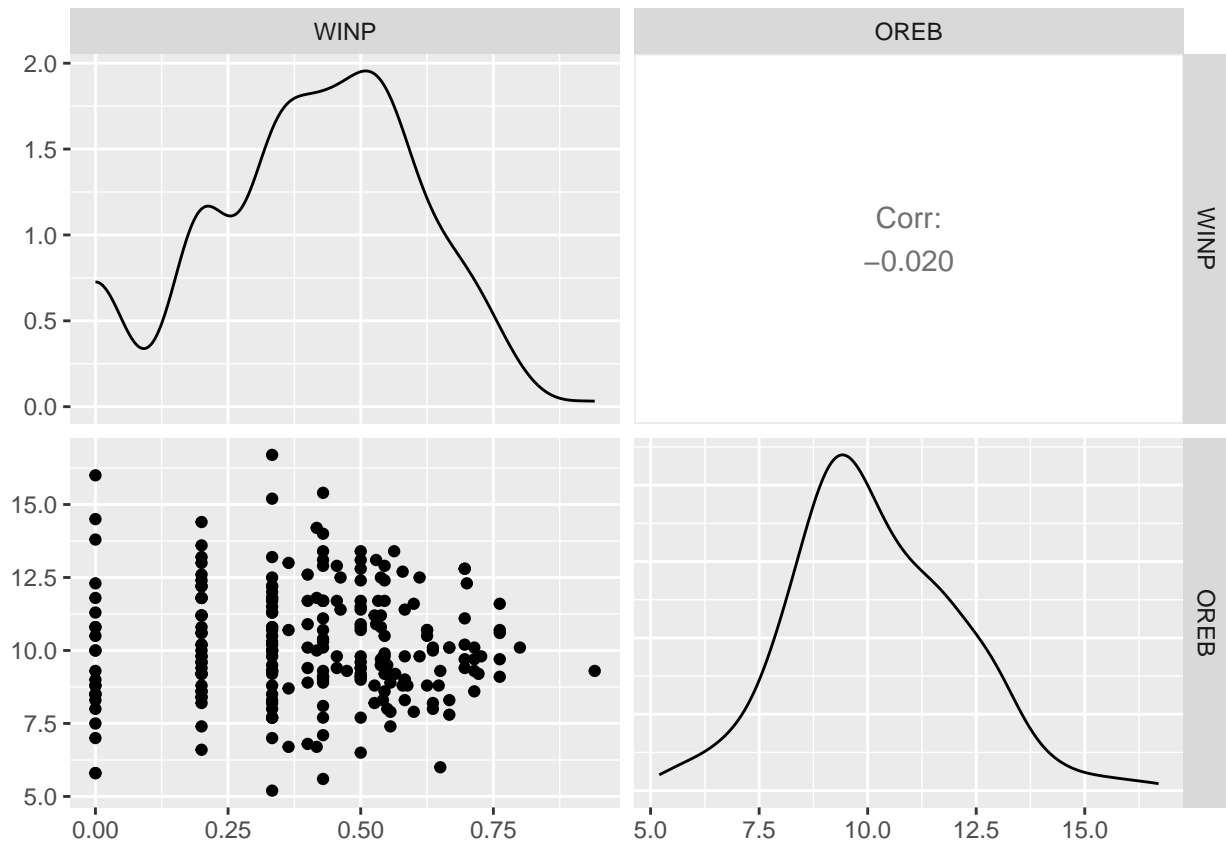
```
#Tentativas de lance livre
dados_playoffs %>% dplyr::select(c(WINP, FTA)) %>% ggpairs()
```

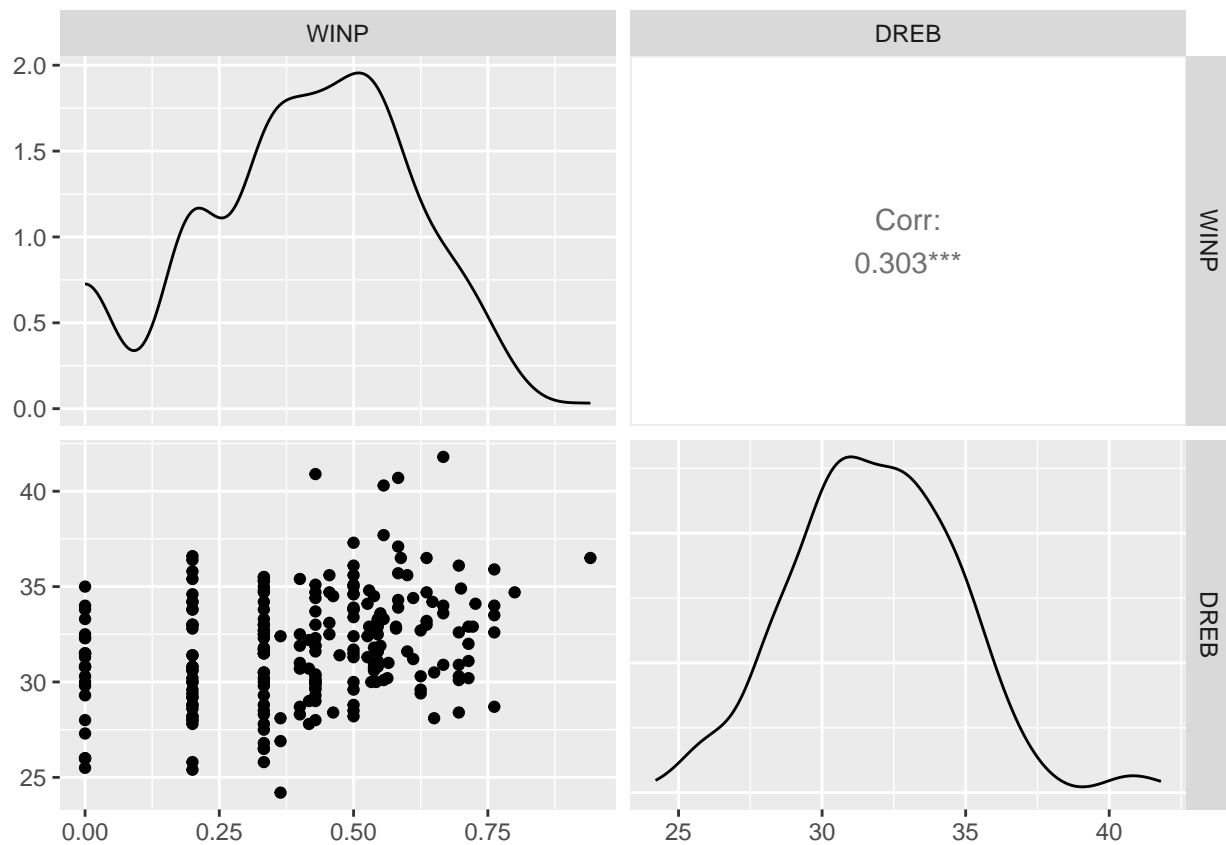
```
##Porcentagem de lance livre
dados_playoffs %>% dplyr::select(c(WINP, FTA)) %>% ggpairs()
```



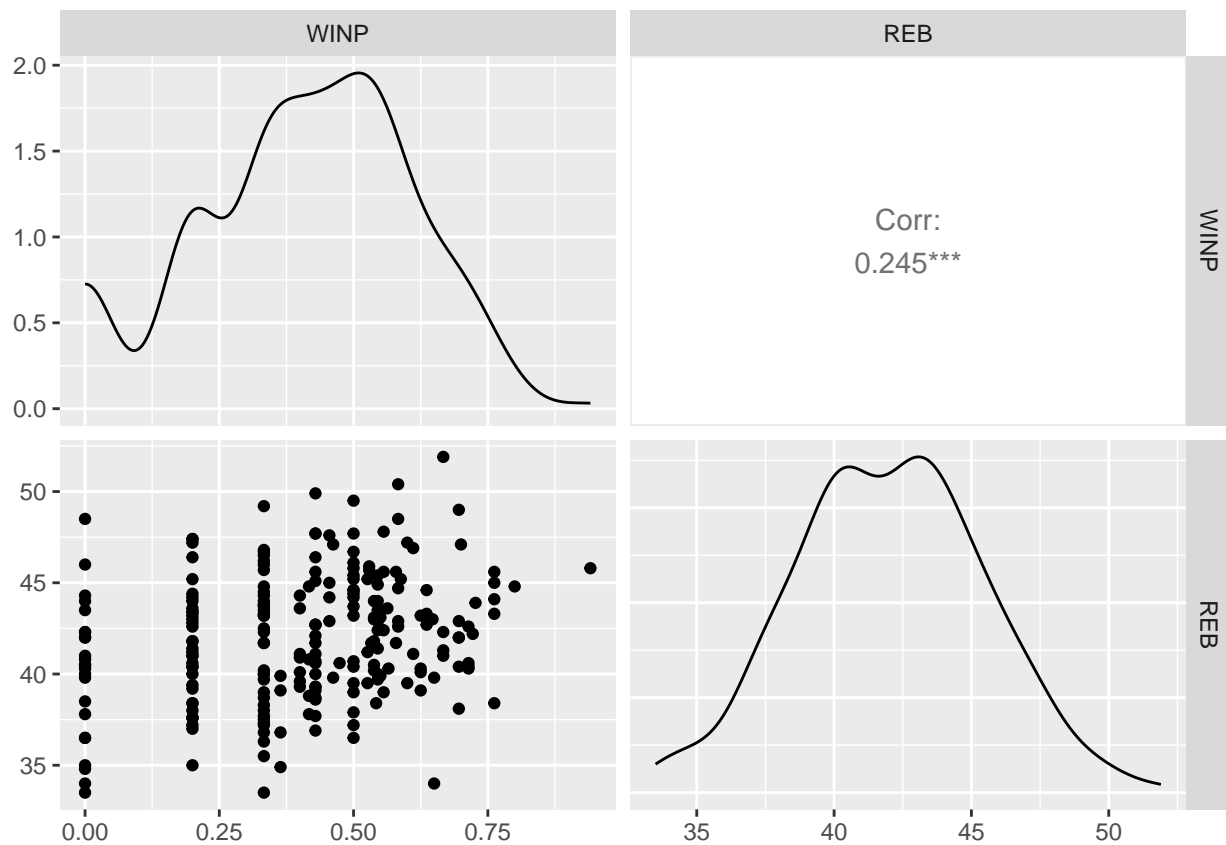
```
#Rebote ofensivo
dados_playoffs %>% dplyr::select(c(WINP, OREB)) %>% ggpairs()
```



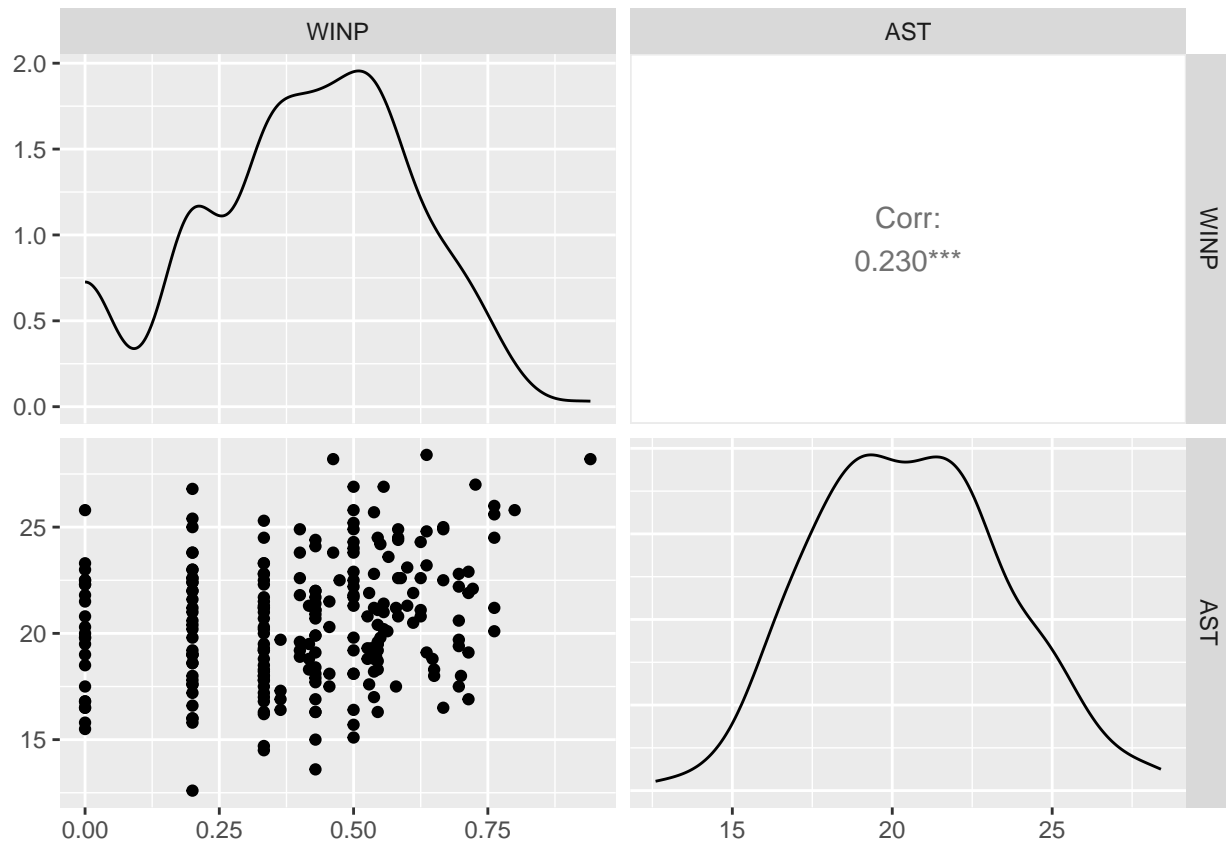
```
#Rebote defensivo
dados_playoffs %>% dplyr::select(c(WINP, DREB)) %>% ggpairs()
```



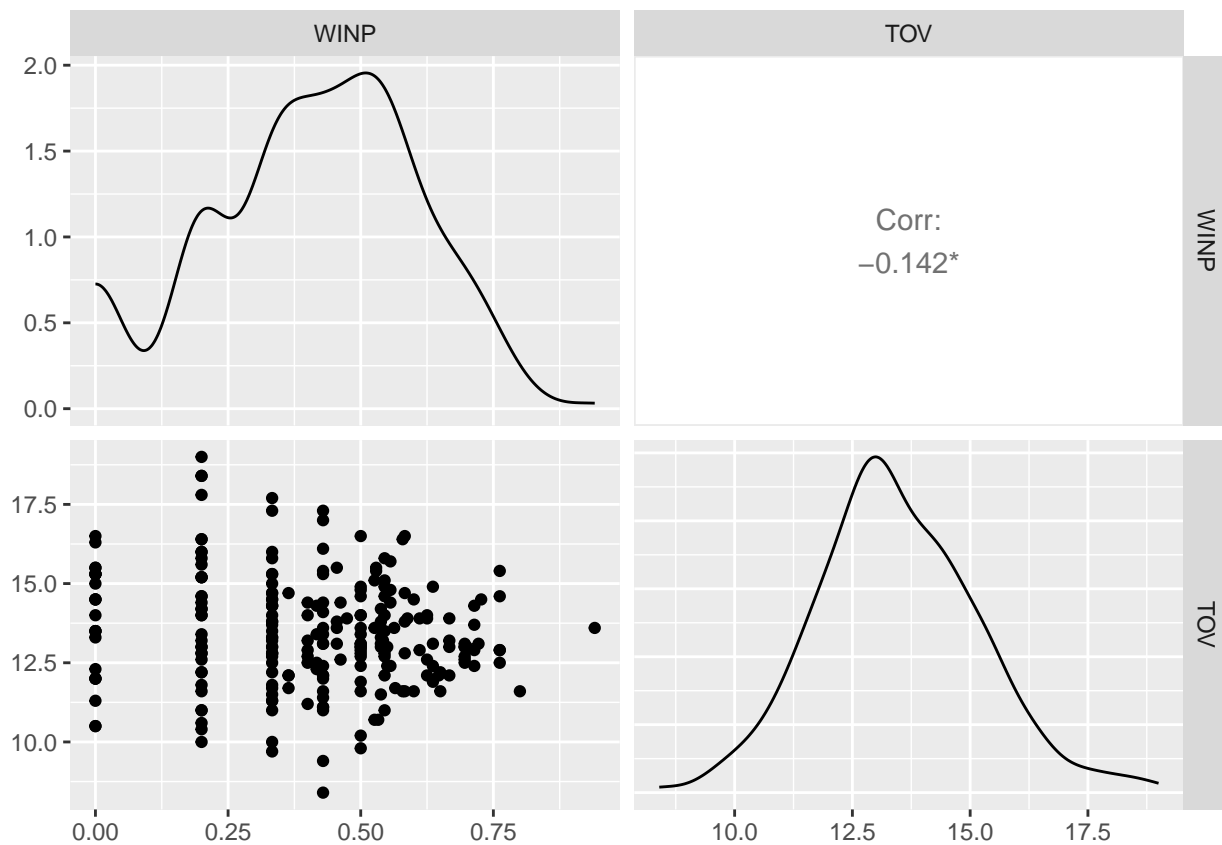
```
#Rebotes
dados_playoffs %>% dplyr::select(c(WINP, REB)) %>% ggpairs()
```



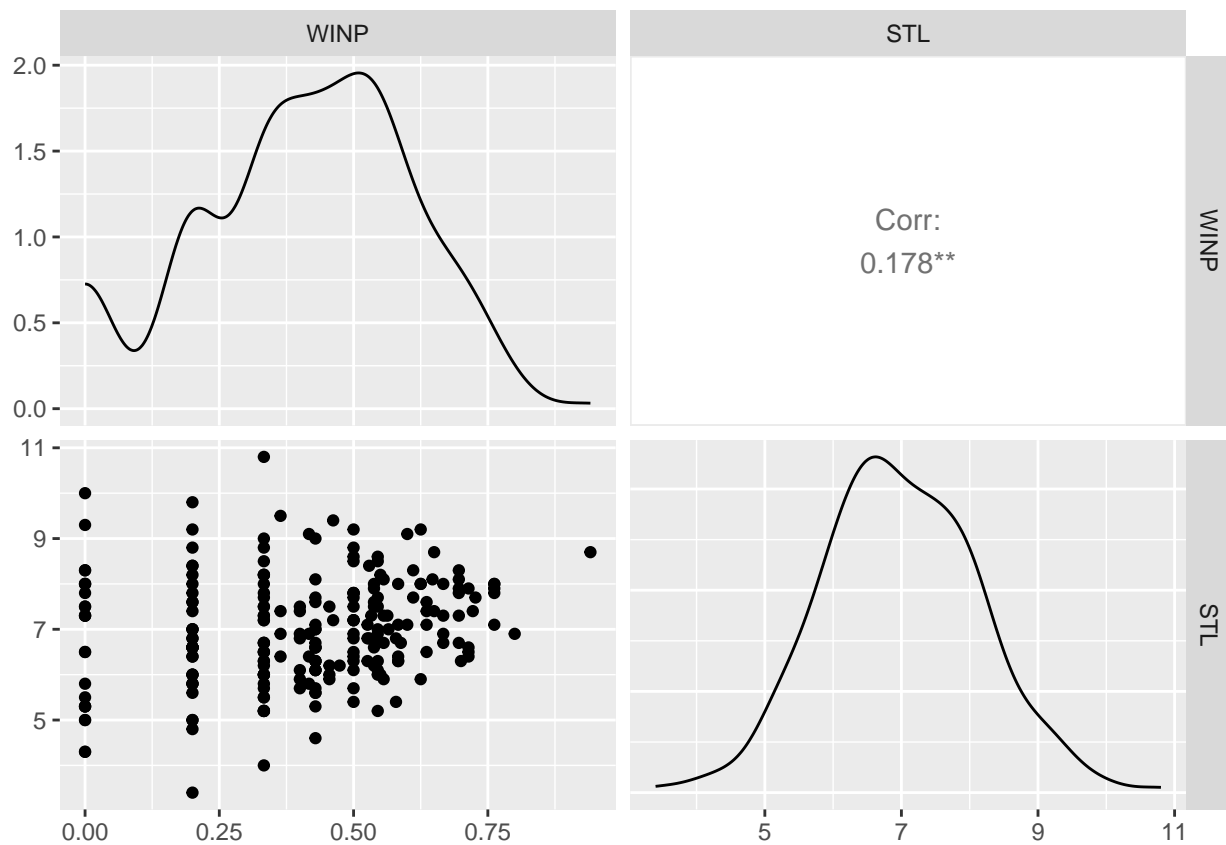
```
#Assistência
dados_playoffs %>% dplyr::select(c(WINP, AST)) %>% ggpairs()
```



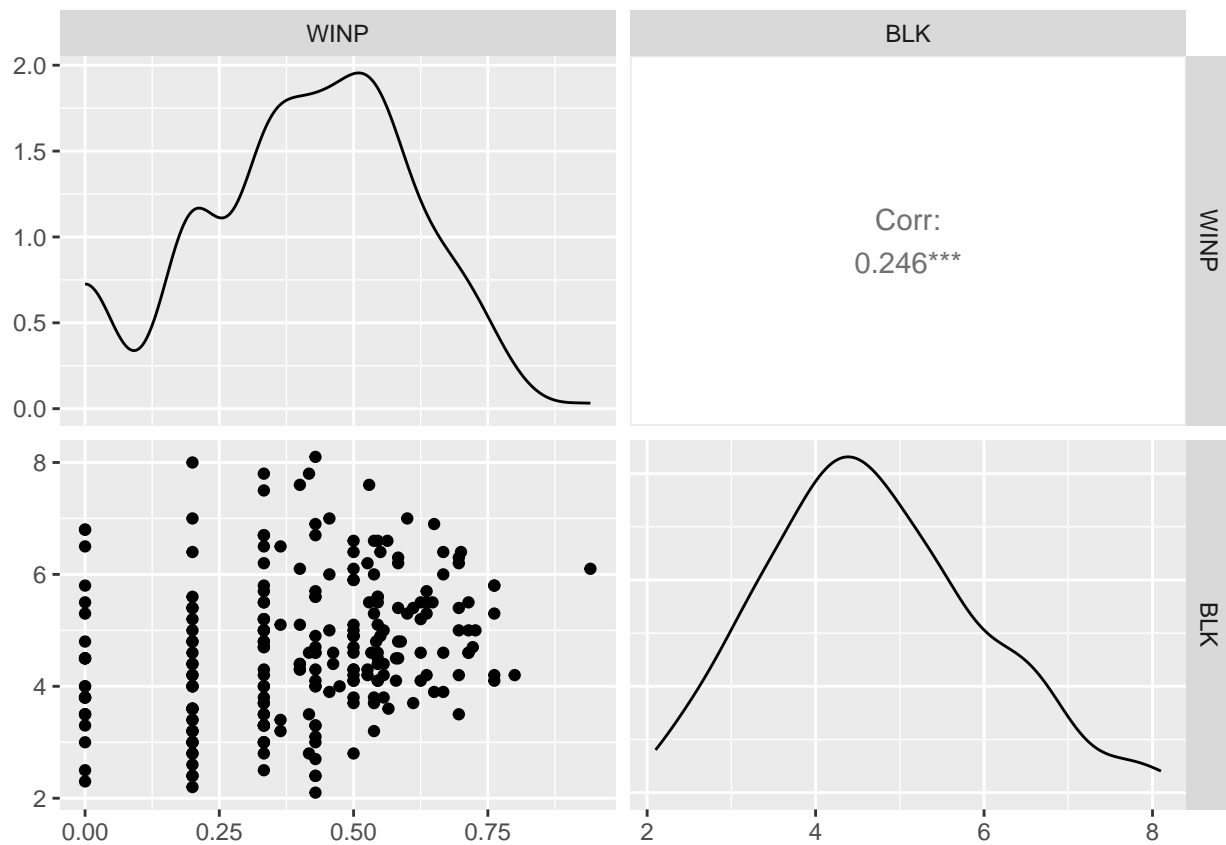
```
#Turnovers
dados_playoffs %>% dplyr::select(c(WINP, TOV)) %>% ggpairs()
```



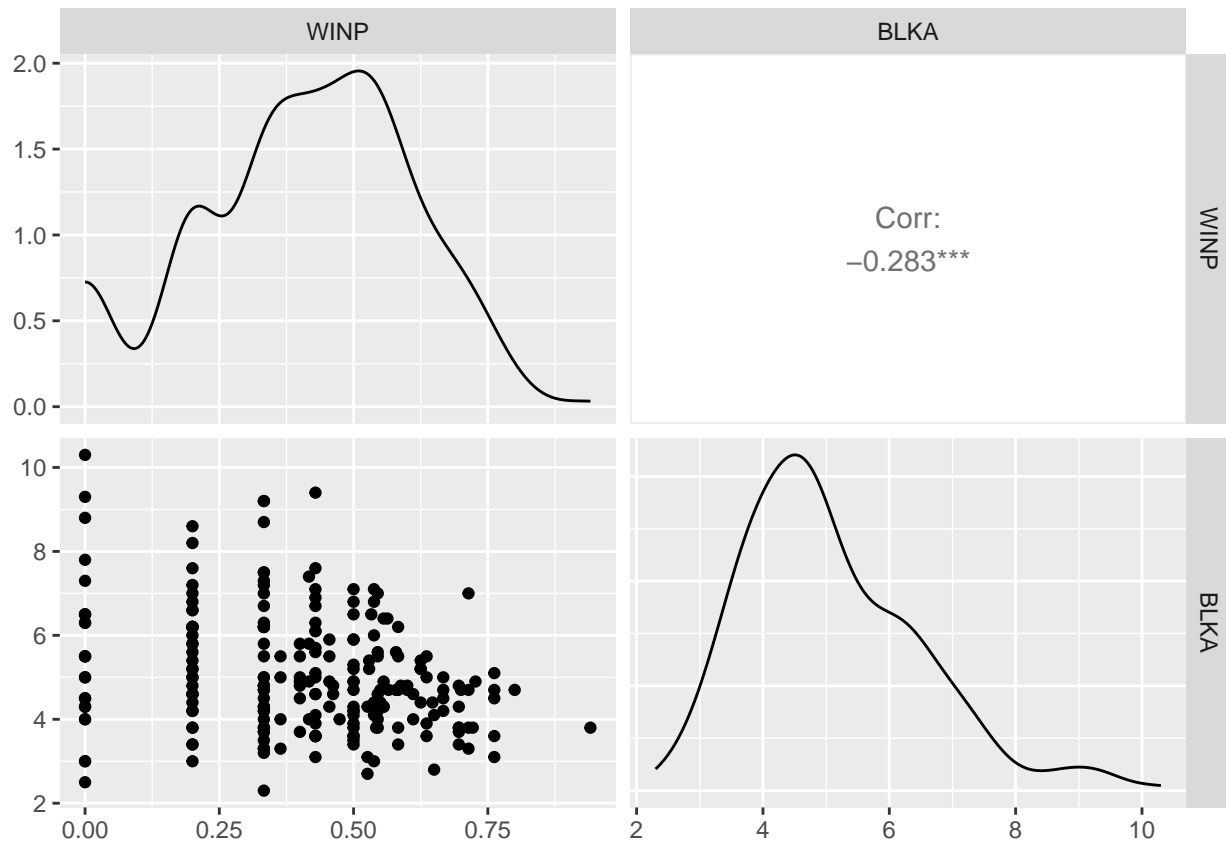
```
#Rouba de bola
dados_playoffs %>% dplyr::select(c(WINP, STL)) %>% ggpairs()
```



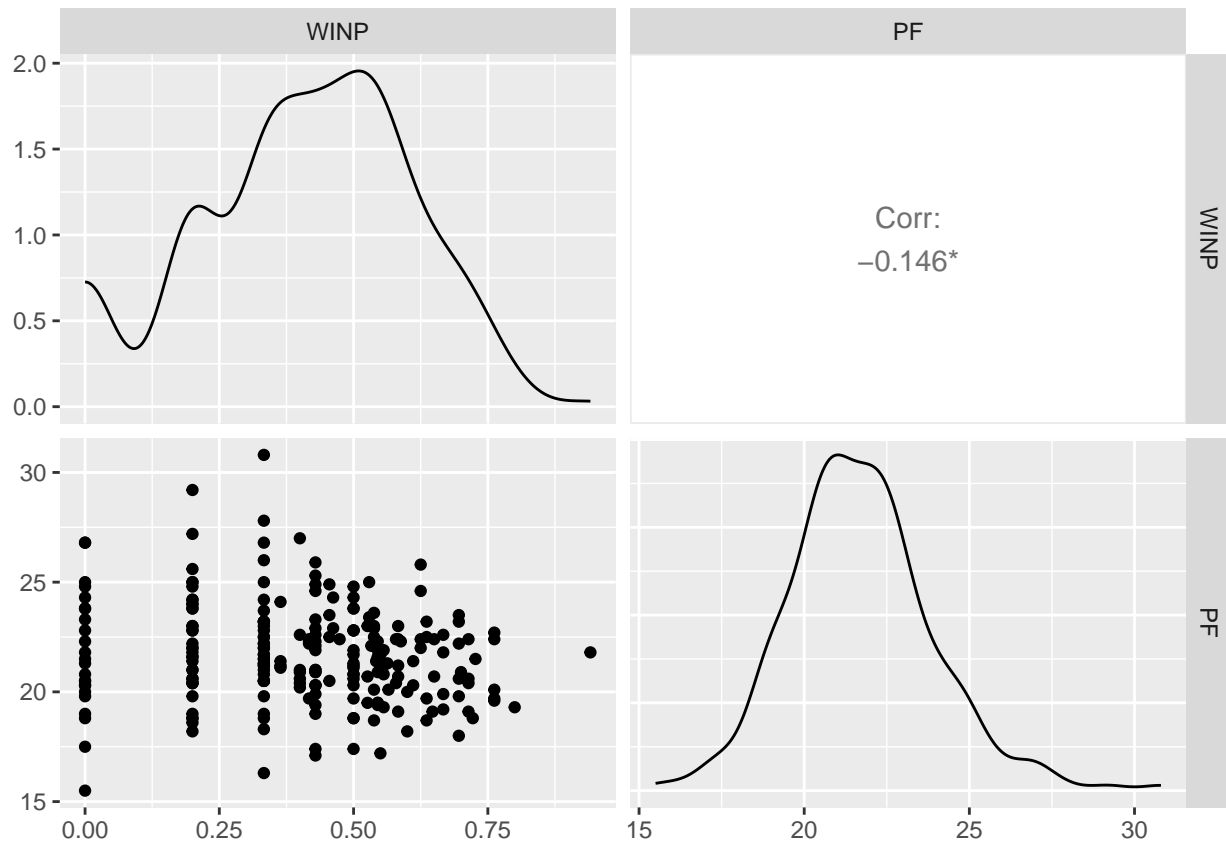
```
#Toco
dados_playoffs %>% dplyr::select(c(WINP, BLK)) %>% ggpairs()
```

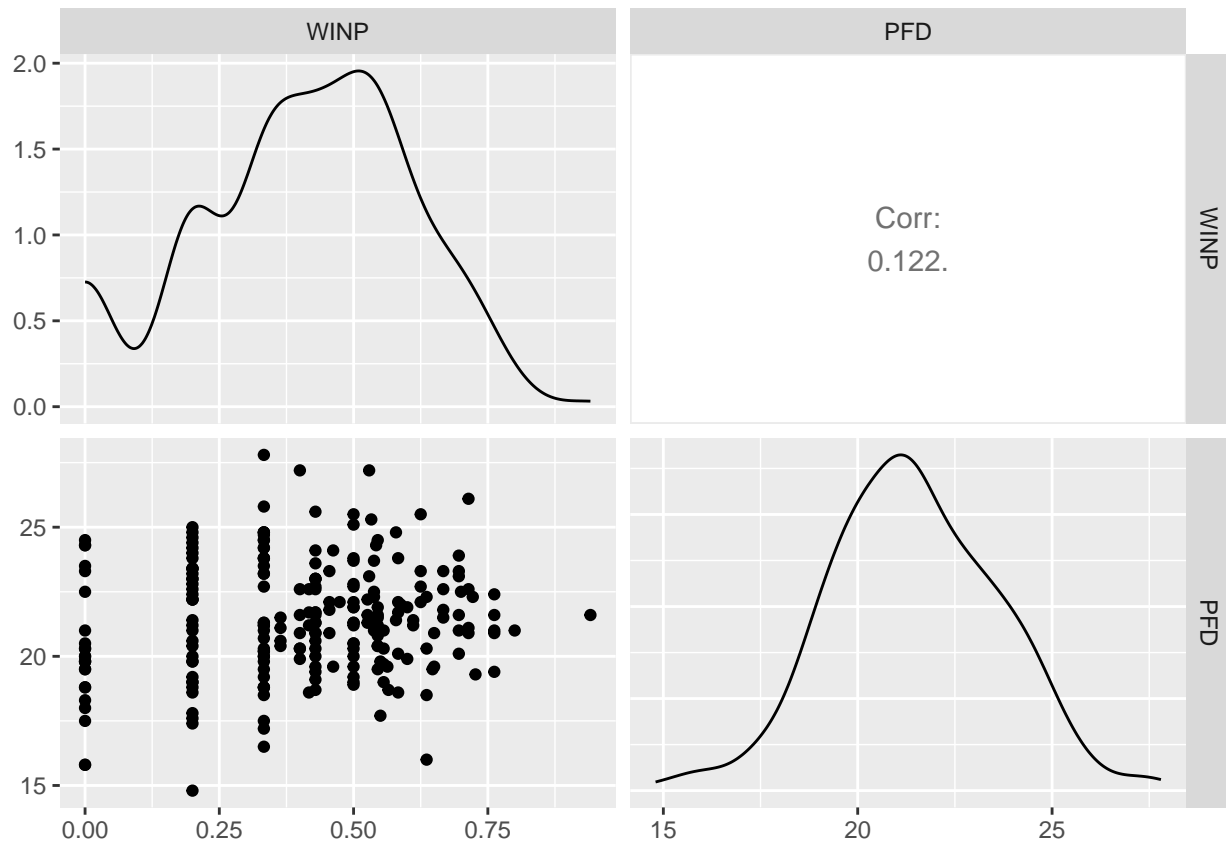
```
#Tocos tomados
dados_playoffs %>% dplyr::select(c(WINP, BLKA)) %>% ggpairs()
```



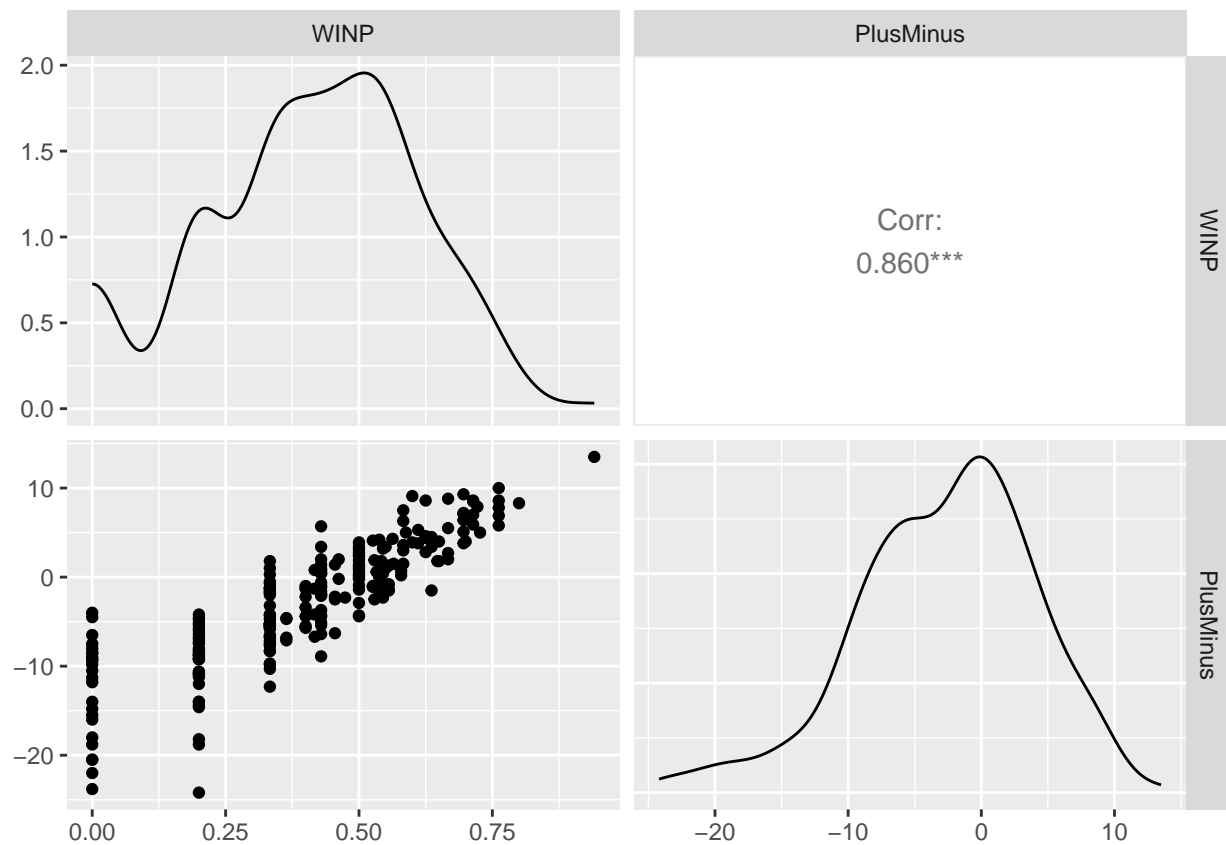
```
#Faltas
dados_playoffs %>% dplyr::select(c(WINP, PF)) %>% ggpairs()
```



```
#Faltas cedidas
dados_playoffs %>% dplyr::select(c(WINP, PFD)) %>% ggpairs()
```



```
#Plus_Minus
dados_playoffs %>% dplyr::select(c(WINP, PlusMinus)) %>% ggpairs() #Correlação muito alta
```



```
cor(dados_playoffs$WINP,dados_playoffs$PlusMinus) #0.860124
```

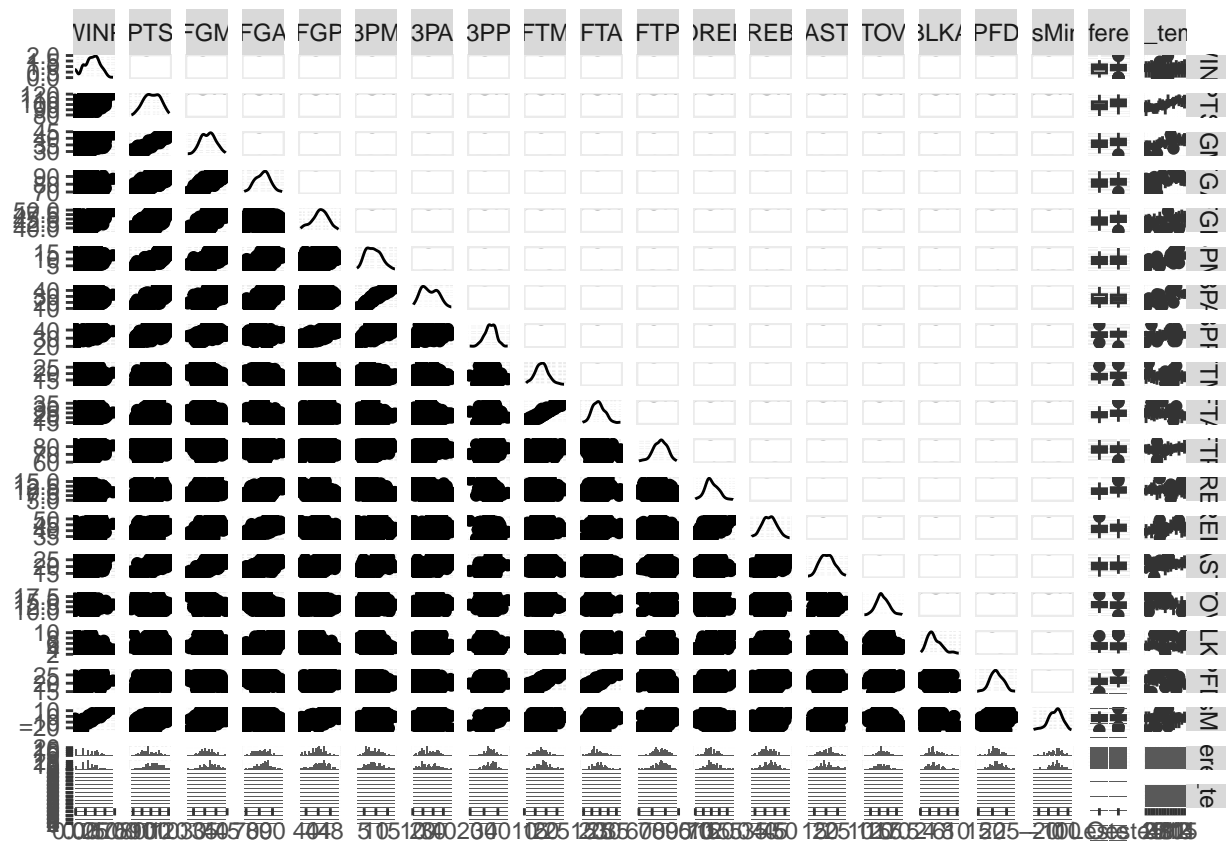
```
## [1] 0.860124
```

```
#Correlação de características defensivas
```

```
dados_playoffs %>% dplyr::select(c(WINP, DREB, STL, BLK, PF)) %>% ggpairs() #Sem correlações de interesse
```



```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
# FG_P e Plus_Minus correlação de 0,564
# WINP e Plus_Minus Correlação muito alta
# PFD e FTA correlação de 0,866
# PFD e FTM correlação de 0,784
#AST e 3PM correlação 0.542
#AST e PTS correlação 0.683
#AST e FGA correlação 0.539
#AST e FGM correlação 0.732
#REB e OREB correlação 0.540
#REB e FGA correlação de 0.619
# Alta correlação entre FTA e FTM de 0.920
# 3P_P e FG_P correlação 0.5616
#PTS e FGM correlação 0.907
#PTS e FGA correlação 0.656
```

```
#FGM e FGA correlação 0.718
#PTS e FG_P correlação 0.589
#FG_P e FGM correlação 0.655
#PTS e 3PM correlação 0.745
#3PM e FGM correlação 0.593
#3PM e FGA correlação 0.514
# 3PA e 3PM correlação de 0.951
#PTS e 3PA correlação 0.683
#3PA e FGM correlação 0.530
#3PA e FGA correlação 0.586
```

Aparições em playoffs

```
aparicoes <- dados_playoffs %>% group_by(TEAM) %>% summarise(Aparicoes = n()) %>% arrange(desc(Aparicoes))
aparicoes
```

```
## # A tibble: 33 x 2
##   TEAM                Aparicoes
##   <fct>              <int>
## 1 Boston Celtics      14
## 2 Atlanta Hawks       12
## 3 Miami Heat          12
## 4 Portland Trail Blazers 11
## 5 San Antonio Spurs    11
## 6 Dallas Mavericks     10
## 7 Denver Nuggets      10
## 8 Memphis Grizzlies    10
## 9 Milwaukee Bucks      10
## 10 Oklahoma City Thunder 10
## # i 23 more rows
```

```
aparicoes_leste <- dados_playoffs %>% filter(Conferencia == "Leste") %>% group_by(TEAM) %>% summarise(Aparicoes = n())
aparicoes_leste
```

```
## # A tibble: 16 x 2
##   TEAM                Aparicoes
##   <fct>              <int>
## 1 Boston Celtics      14
## 2 Atlanta Hawks       12
## 3 Miami Heat          12
## 4 Milwaukee Bucks     10
## 5 Chicago Bulls        9
## 6 Indiana Pacers        9
## 7 Philadelphia 76ers    9
## 8 Brooklyn Nets        8
## 9 Toronto Raptors      8
## 10 Cleveland Cavaliers  7
## 11 Orlando Magic        6
## 12 New York Knicks      5
## 13 Washington Wizards   5
## 14 Detroit Pistons      3
## 15 Charlotte Bobcats    2
## 16 Charlotte Hornets    1
```



```
aparicoes_oeste <- dados_playoffs %>% filter(Conferencia == "Oeste") %>% group_by(TEAM) %>% summarise(Aparicoes_oeste)
```

```
## # A tibble: 17 x 2
##   TEAM                Aparicoes
##   <fct>              <int>
## 1 Portland Trail Blazers      11
## 2 San Antonio Spurs          11
## 3 Dallas Mavericks           10
## 4 Denver Nuggets             10
## 5 Memphis Grizzlies           10
## 6 Oklahoma City Thunder       10
## 7 Golden State Warriors       9
## 8 Houston Rockets             9
## 9 Utah Jazz                   9
## 10 Los Angeles Lakers          8
## 11 LA Clippers                 6
## 12 Los Angeles Clippers        4
## 13 Phoenix Suns                4
## 14 Minnesota Timberwolves       3
## 15 New Orleans Pelicans         3
## 16 New Orleans Hornets          2
## 17 Sacramento Kings            1
```

Análise das conferências

#Média da porcentagem de vitória por conferência

```
dados_playoffs %>% group_by(Conferencia) %>% summarise(Media_conf = mean(WINP))
```

```
## # A tibble: 2 x 2
##   Conferencia Media_conf
##   <chr>          <dbl>
## 1 Leste          0.389
## 2 Oeste          0.417
```

#Teste t para saber se as médias tem diferença significativa

```
t.test(WINP ~ Conferencia, data = dados_playoffs)
```

```
##
## Welch Two Sample t-test
##
## data: WINP by Conferencia
## t = -1.0578, df = 234.18, p-value = 0.2912
## alternative hypothesis: true difference in means between group Leste and group Oeste is not equal to 0
## 95 percent confidence interval:
## -0.07993514 0.02408514
## sample estimates:
## mean in group Leste mean in group Oeste
## 0.3891583 0.4170833
```

#Summary de cada conferencia de todas as variaveis

```
dados_playoffs %>% filter(Conferencia == "Leste") %>% summary()
```

```
##      Posicao      TEAM      GP      W
## Min.   : 1.000 Boston Celtics :14 Min.   : 4.00 Min.   : 0.00
## 1st Qu.: 4.000 Atlanta Hawks  :12 1st Qu.: 5.00 1st Qu.: 1.00
```

```

## Median : 8.000 Miami Heat :12 Median : 8.00 Median : 3.50
## Mean : 8.133 Milwaukee Bucks:10 Mean :10.43 Mean : 5.15
## 3rd Qu.:12.000 Chicago Bulls : 9 3rd Qu.:14.00 3rd Qu.: 7.25
## Max. :16.000 Indiana Pacers : 9 Max. :24.00 Max. :16.00
## (Other) :54
## L WINP MIN PTS
## Min. : 4.000 Min. :0.0000 Min. :48.00 Min. : 78.00
## 1st Qu.: 4.000 1st Qu.:0.2000 1st Qu.:48.00 1st Qu.: 91.80
## Median : 4.000 Median :0.4290 Median :48.00 Median : 97.05
## Mean : 5.283 Mean :0.3892 Mean :48.36 Mean : 98.17
## 3rd Qu.: 6.250 3rd Qu.:0.5560 3rd Qu.:48.50 3rd Qu.:105.28
## Max. :11.000 Max. :0.7620 Max. :53.00 Max. :118.80
##
## FGM FGA FGP 3PM
## Min. :28.50 Min. :69.00 Min. :38.40 Min. : 3.000
## 1st Qu.:33.70 1st Qu.:77.30 1st Qu.:42.08 1st Qu.: 6.600
## Median :35.35 Median :81.80 Median :44.10 Median : 8.600
## Mean :35.81 Mean :81.66 Mean :43.87 Mean : 8.929
## 3rd Qu.:38.08 3rd Qu.:85.65 3rd Qu.:45.70 3rd Qu.:11.200
## Max. :44.00 Max. :96.80 Max. :50.30 Max. :15.600
##
## 3PA 3PP FTM FTA
## Min. :10.80 Min. :24.50 Min. :12.00 Min. :14.40
## 1st Qu.:19.30 1st Qu.:31.98 1st Qu.:15.80 1st Qu.:20.80
## Median :24.60 Median :34.15 Median :17.55 Median :22.75
## Mean :26.03 Mean :34.16 Mean :17.64 Mean :23.07
## 3rd Qu.:32.92 3rd Qu.:36.83 3rd Qu.:19.52 3rd Qu.:25.43
## Max. :46.00 Max. :46.50 Max. :26.00 Max. :32.00
##
## FTP OREB DREB REB
## Min. :63.00 Min. : 5.200 Min. :24.20 Min. :33.50
## 1st Qu.:72.67 1st Qu.: 8.600 1st Qu.:29.88 1st Qu.:39.10
## Median :76.65 Median : 9.650 Median :31.50 Median :41.20
## Mean :76.48 Mean : 9.807 Mean :31.73 Mean :41.54
## 3rd Qu.:79.90 3rd Qu.:11.125 3rd Qu.:33.80 3rd Qu.:43.85
## Max. :88.60 Max. :14.500 Max. :41.80 Max. :51.90
##
## AST TOV STL BLK
## Min. :12.60 Min. : 9.40 Min. : 4.300 Min. :2.200
## 1st Qu.:18.27 1st Qu.:12.47 1st Qu.: 6.175 1st Qu.:3.800
## Median :20.35 Median :13.40 Median : 6.850 Median :4.500
## Mean :20.36 Mean :13.58 Mean : 6.871 Mean :4.643
## 3rd Qu.:22.50 3rd Qu.:14.62 3rd Qu.: 7.600 3rd Qu.:5.500
## Max. :26.90 Max. :18.40 Max. :10.800 Max. :8.100
##
## BLKA PF PFD PlusMinus
## Min. :2.500 Min. :15.50 Min. :14.80 Min. : -23.800
## 1st Qu.:4.175 1st Qu.:20.00 1st Qu.:19.77 1st Qu.: -7.550
## Median :4.700 Median :21.25 Median :21.00 Median : -1.550
## Mean :5.018 Mean :21.33 Mean :20.99 Mean : -3.033
## 3rd Qu.:6.000 3rd Qu.:22.40 3rd Qu.:22.50 3rd Qu.: 1.800
## Max. :8.800 Max. :29.20 Max. :26.10 Max. : 8.800
##
## Temporada Conferencia Numero_temporada

```

```
## Length:120      Length:120      1      : 8
## Class :character Class :character 2      : 8
## Mode  :character Mode  :character 3      : 8
##                                         4      : 8
##                                         5      : 8
##                                         6      : 8
##                                         (Other):72
```

```
dados_playoffs %>% filter(Conferencia == "Oeste") %>% summary()
```

```
##      Posicao      TEAM      GP      W
## Min.   : 1.000 Portland Trail Blazers:11 Min.   : 4.00 Min.   : 0.000
## 1st Qu.: 5.000 San Antonio Spurs      :11 1st Qu.: 6.00 1st Qu.: 2.000
## Median : 8.000 Dallas Mavericks      :10 Median : 8.00 Median : 3.500
## Mean   : 7.867 Denver Nuggets      :10 Mean   :10.48 Mean   : 5.308
## 3rd Qu.:10.250 Memphis Grizzlies      :10 3rd Qu.:14.00 3rd Qu.: 7.250
## Max.   :16.000 Oklahoma City Thunder :10 Max.   :24.00 Max.   :16.000
##      (Other)      :58
##      L      WINP      MIN      PTS
## Min.   : 1.000 Min.   :0.0000 Min.   :48.00 Min.   : 81.0
## 1st Qu.: 4.000 1st Qu.:0.3330 1st Qu.:48.00 1st Qu.: 97.4
## Median : 4.000 Median :0.4290 Median :48.00 Median :103.3
## Mean   : 5.175 Mean   :0.4171 Mean   :48.31 Mean   :103.0
## 3rd Qu.: 6.000 3rd Qu.:0.5380 3rd Qu.:48.50 3rd Qu.:109.0
## Max.   :10.000 Max.   :0.9410 Max.   :50.90 Max.   :119.5
##
##      FGM      FGA      FGP      3PM
## Min.   :27.80 Min.   :70.20 Min.   :38.20 Min.   : 2.300
## 1st Qu.:35.15 1st Qu.:79.92 1st Qu.:43.58 1st Qu.: 6.675
## Median :37.50 Median :83.90 Median :45.10 Median : 9.100
## Mean   :37.45 Mean   :83.45 Mean   :44.88 Mean   : 9.206
## 3rd Qu.:39.42 3rd Qu.:86.40 3rd Qu.:46.62 3rd Qu.:11.525
## Max.   :44.20 Max.   :95.60 Max.   :49.70 Max.   :18.000
##
##      3PA      3PP      FTM      FTA
## Min.   :11.30 Min.   :20.00 Min.   :11.30 Min.   :15.80
## 1st Qu.:19.45 1st Qu.:32.25 1st Qu.:16.80 1st Qu.:22.27
## Median :25.80 Median :35.25 Median :18.90 Median :24.50
## Mean   :26.31 Mean   :34.62 Mean   :18.95 Mean   :24.83
## 3rd Qu.:32.40 3rd Qu.:37.50 3rd Qu.:20.43 3rd Qu.:27.02
## Max.   :46.80 Max.   :42.90 Max.   :28.20 Max.   :36.80
##
##      FTP      OREB      DREB      REB
## Min.   :60.80 Min.   : 6.500 Min.   :25.80 Min.   :33.50
## 1st Qu.:73.50 1st Qu.: 9.275 1st Qu.:30.00 1st Qu.:40.20
## Median :76.85 Median :10.250 Median :32.25 Median :42.85
## Mean   :76.45 Mean   :10.564 Mean   :31.92 Mean   :42.48
## 3rd Qu.:80.03 3rd Qu.:11.700 3rd Qu.:33.85 3rd Qu.:44.80
## Max.   :86.90 Max.   :16.700 Max.   :40.90 Max.   :49.90
##
##      AST      TOV      STL      BLK
## Min.   :14.70 Min.   : 8.40 Min.   :3.400 Min.   :2.100
## 1st Qu.:18.38 1st Qu.:12.10 1st Qu.:6.275 1st Qu.:3.800
## Median :20.55 Median :13.10 Median :7.000 Median :4.600
## Mean   :20.75 Mean   :13.32 Mean   :7.093 Mean   :4.714
```

```
## 3rd Qu.:22.65 3rd Qu.:14.50 3rd Qu.:8.000 3rd Qu.:5.500
## Max. :28.40 Max. :19.00 Max. :9.800 Max. :8.000
##
## BLKA PF PFD PlusMinus
## Min. : 2.300 Min. :17.10 Min. :17.50 Min. : -24.200
## 1st Qu.: 4.000 1st Qu.:20.50 1st Qu.:20.38 1st Qu.: -6.075
## Median : 4.900 Median :22.00 Median :21.60 Median : -1.800
## Mean : 5.139 Mean :22.10 Mean :21.87 Mean : -2.453
## 3rd Qu.: 5.950 3rd Qu.:23.32 3rd Qu.:23.23 3rd Qu.: 1.500
## Max. :10.300 Max. :30.80 Max. :27.80 Max. : 13.500
##
## Temporada Conferencia Numero_temporada
## Length:120 Length:120 1 : 8
## Class :character Class :character 2 : 8
## Mode :character Mode :character 3 : 8
## 4 : 8
## 5 : 8
## 6 : 8
## (Other):72
```

#Summary de cada conferencia da variavel resposta

```
dados_playoffs %>% filter(Conferencia == "Leste") %>% dplyr::select(WINP) %>% summary()
```

```
## WINP
## Min. :0.0000
## 1st Qu.:0.2000
## Median :0.4290
## Mean :0.3892
## 3rd Qu.:0.5560
## Max. :0.7620
```

```
var(dados_playoffs %>% filter(Conferencia == "Leste") %>% dplyr::select(WINP))
```

```
## WINP
## WINP 0.04715713
```

```
dados_playoffs %>% filter(Conferencia == "Oeste") %>% dplyr::select(WINP) %>% summary()
```

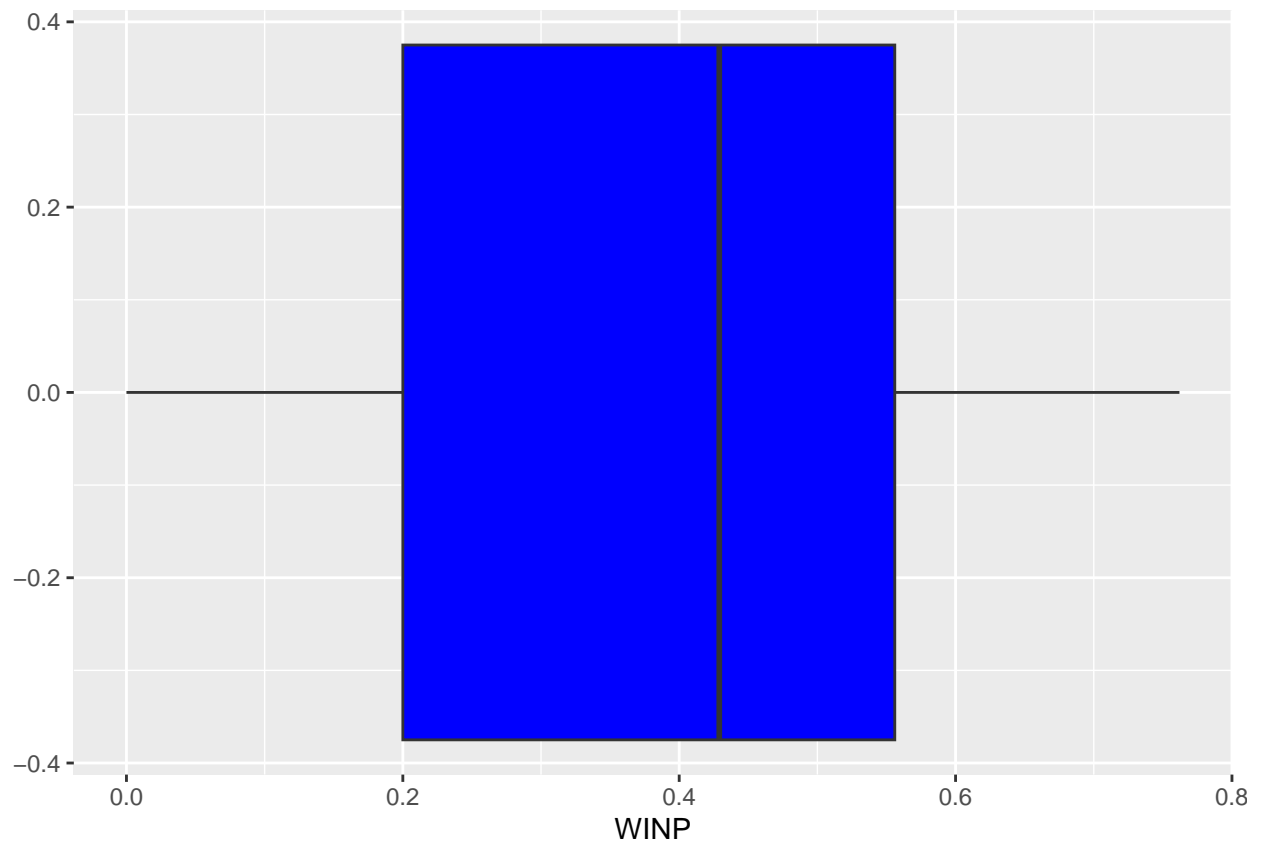
```
## WINP
## Min. :0.0000
## 1st Qu.:0.3330
## Median :0.4290
## Mean :0.4171
## 3rd Qu.:0.5380
## Max. :0.9410
```

```
var(dados_playoffs %>% filter(Conferencia == "Oeste") %>% dplyr::select(WINP))
```

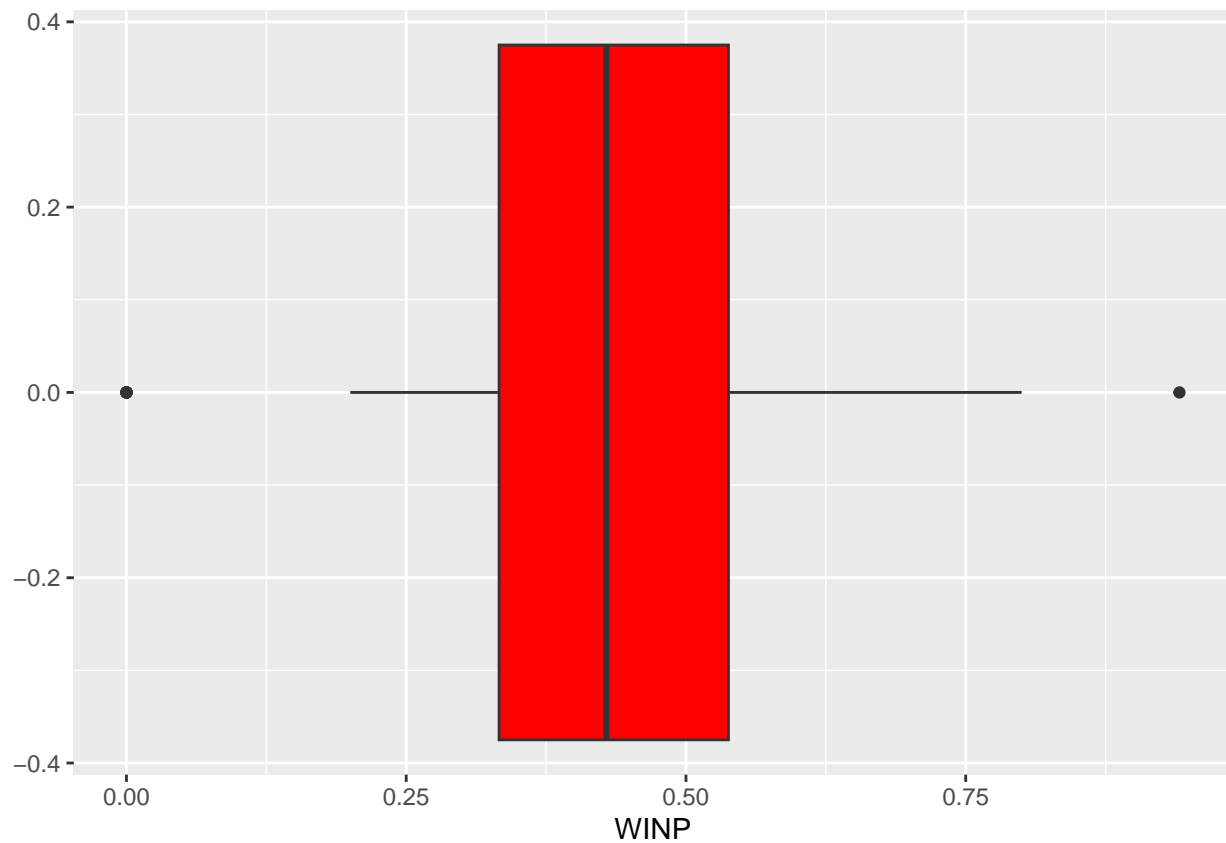
```
## WINP
## WINP 0.03647256
```

Boxplot da porcentagem de vitorias por conferencia

```
dados_playoffs %>% filter(Conferencia == "Leste") %>% dplyr::select(WINP) %>%
  ggplot() +
  geom_boxplot(aes(x = WINP), fill = "Blue")
```

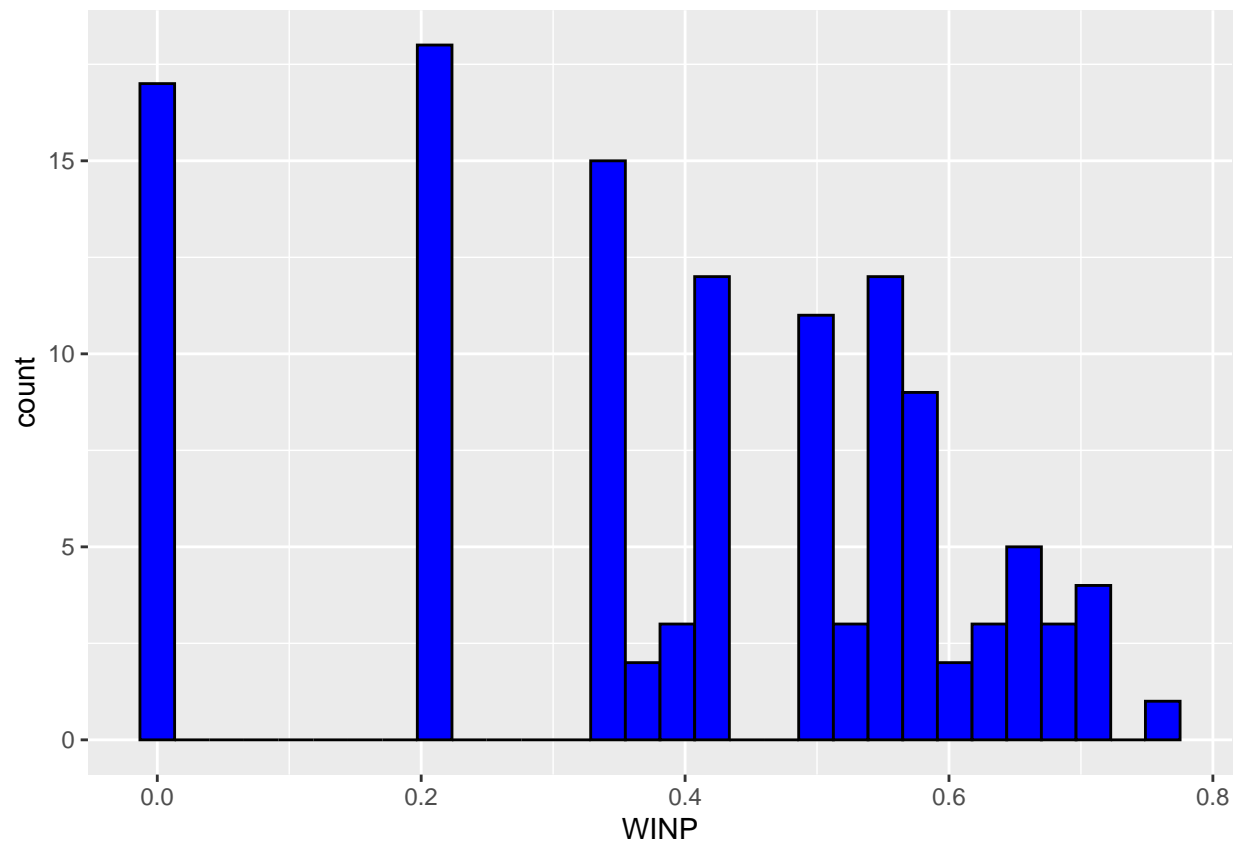


```
dados_playoffs %>% filter(Conferencia == "Oeste") %>% dplyr::select(WINP) %>%  
  ggplot() +  
  geom_boxplot(aes(x = WINP), fill = "Red")
```



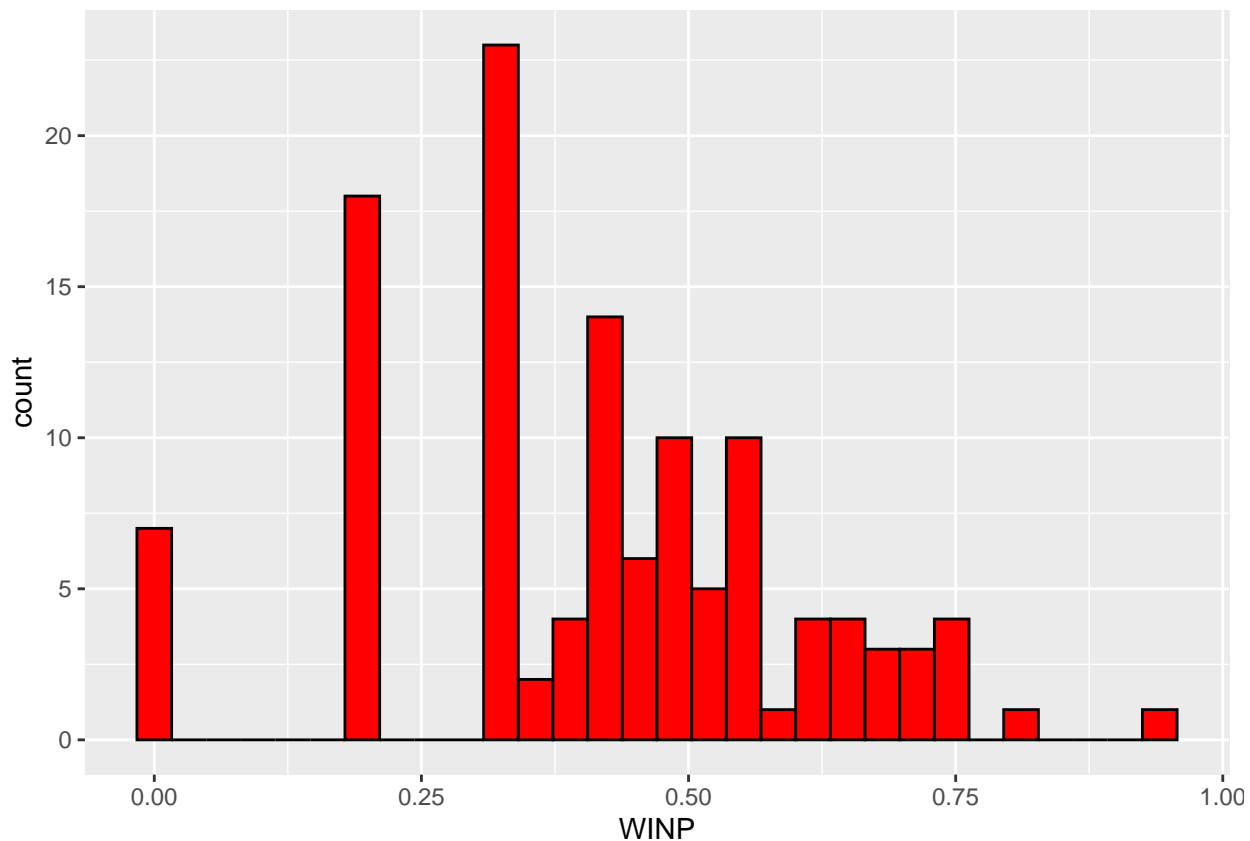
```
# Histograma da porcentagem de vitorias por conferencia
dados_playoffs %>% filter(Conferencia == "Leste") %>% dplyr::select(WINP) %>%
  ggplot() +
  geom_histogram(aes(x = WINP), fill = "Blue", color = "black")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
dados_playoffs %>% filter(Conferencia == "Oeste") %>% dplyr::select(WINP) %>%  
  ggplot() +  
  geom_histogram(aes(x = WINP), fill = "Red", color = "black")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Leste

#Gráfico de pontos com a Porcentagem de vitórias ao decorrer das temporadas

```
dados_playoffs %>% filter(Conferencia == "Leste") %>% ggplot() +
```

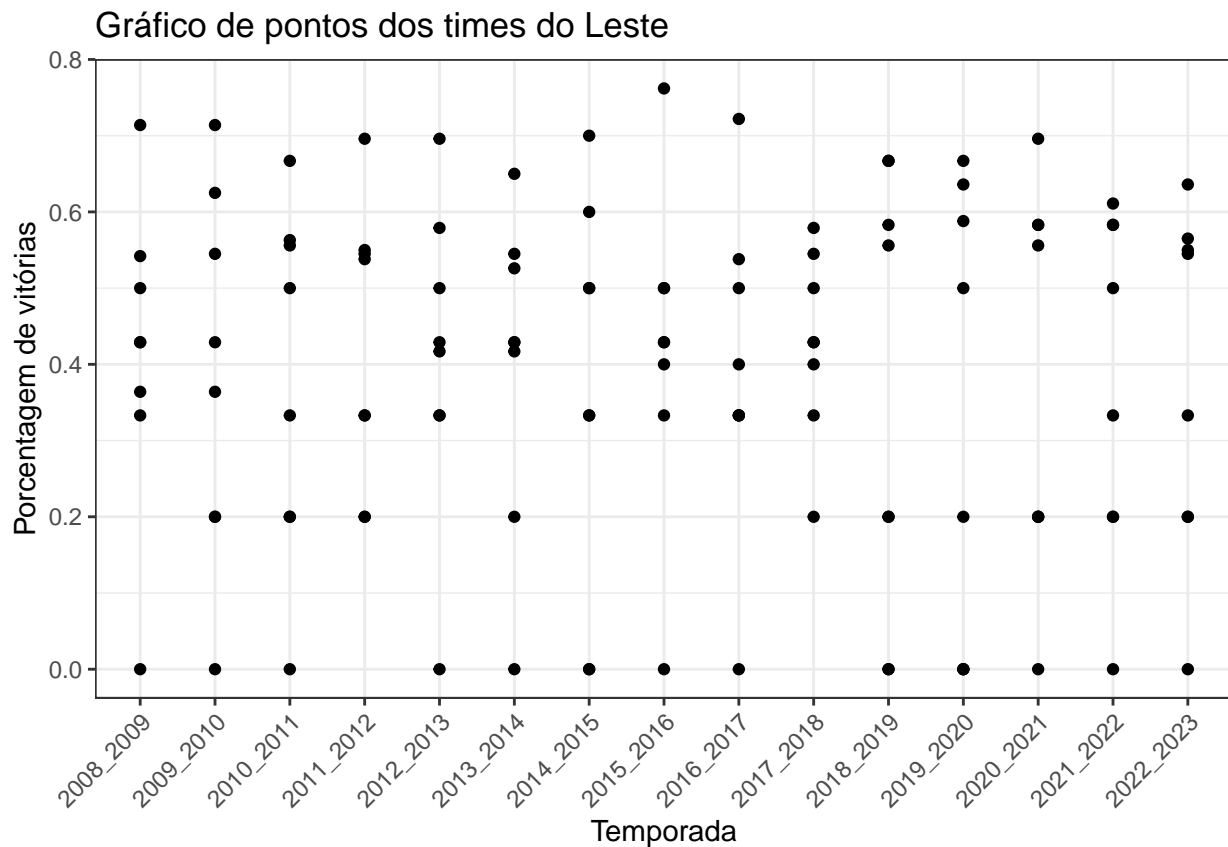
```
  geom_point(aes(x = Temporada, y = WINP)) +
```

```
  theme_bw() +
```

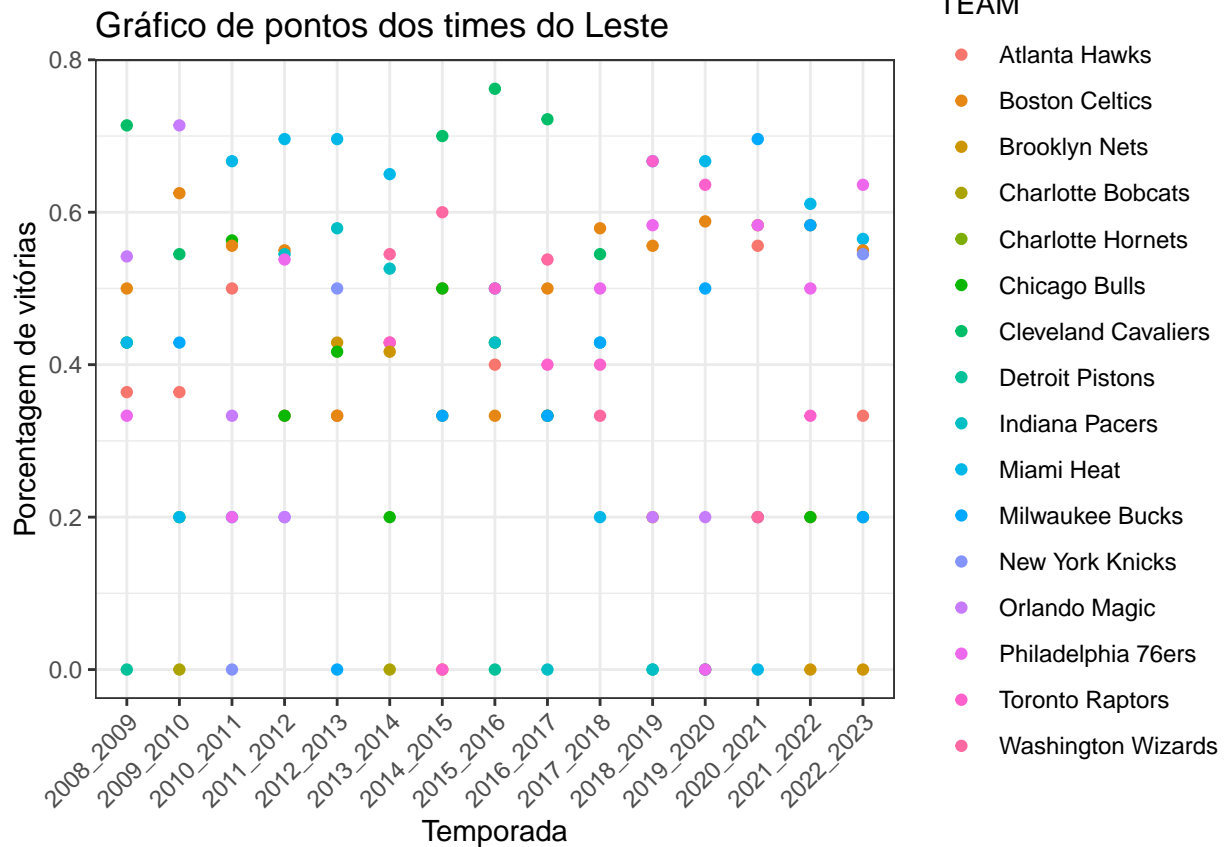
```
  labs(title = "Gráfico de pontos dos times do Leste",
```

```
        y = "Porcentagem de vitórias") +
```

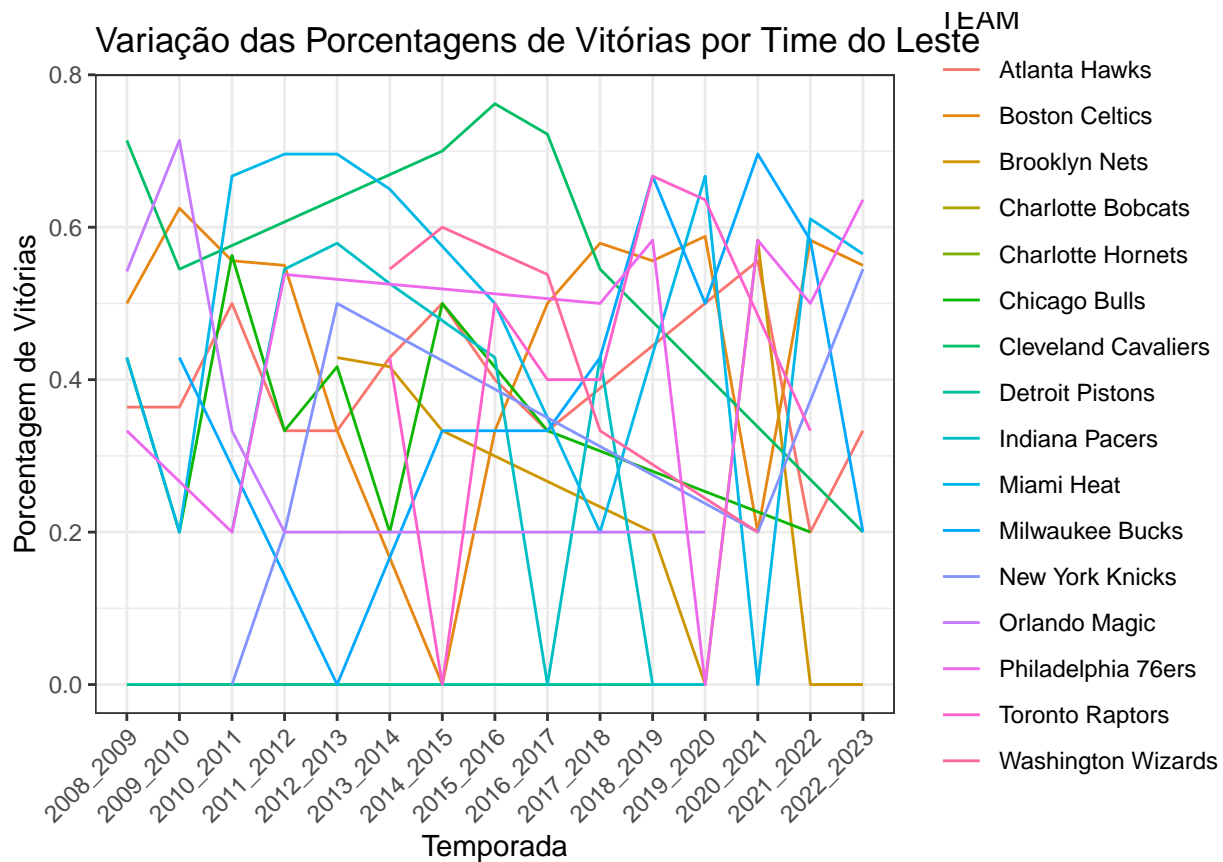
```
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```

```
#Gráfico de pontos com a Porcentagem de vitórias ao decorrer das temporadas por time
dados_playoffs %>% filter(Conferencia == "Leste") %>%
  ggplot() +
  geom_point(aes(x = Temporada, y = WINP, color = TEAM)) +
  theme_bw() +
  labs(title = "Gráfico de pontos dos times do Leste",
       y = "Porcentagem de vitórias") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```

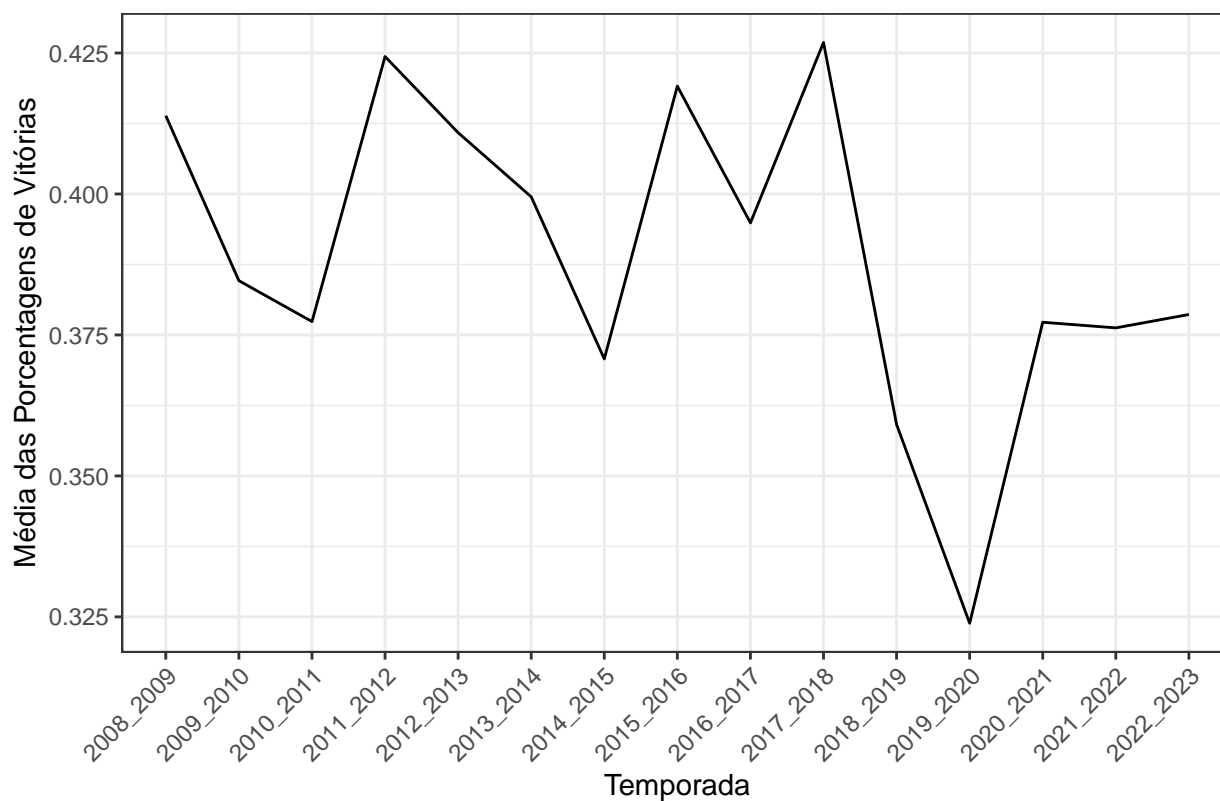


```
#Gráfico de linha com Variação das Porcentagens de Vitórias por Time do Leste
dados_playoffs %>% filter(Conferencia == "Leste") %>% ggplot(aes(x = Temporada, y = WINP, color = TEAM,
  geom_line() +
  labs(title = "Variação das Porcentagens de Vitórias por Time do Leste",
    x = "Temporada",
    y = "Porcentagem de Vitórias") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))) # Rotaciona os rótulos em 45 graus
```



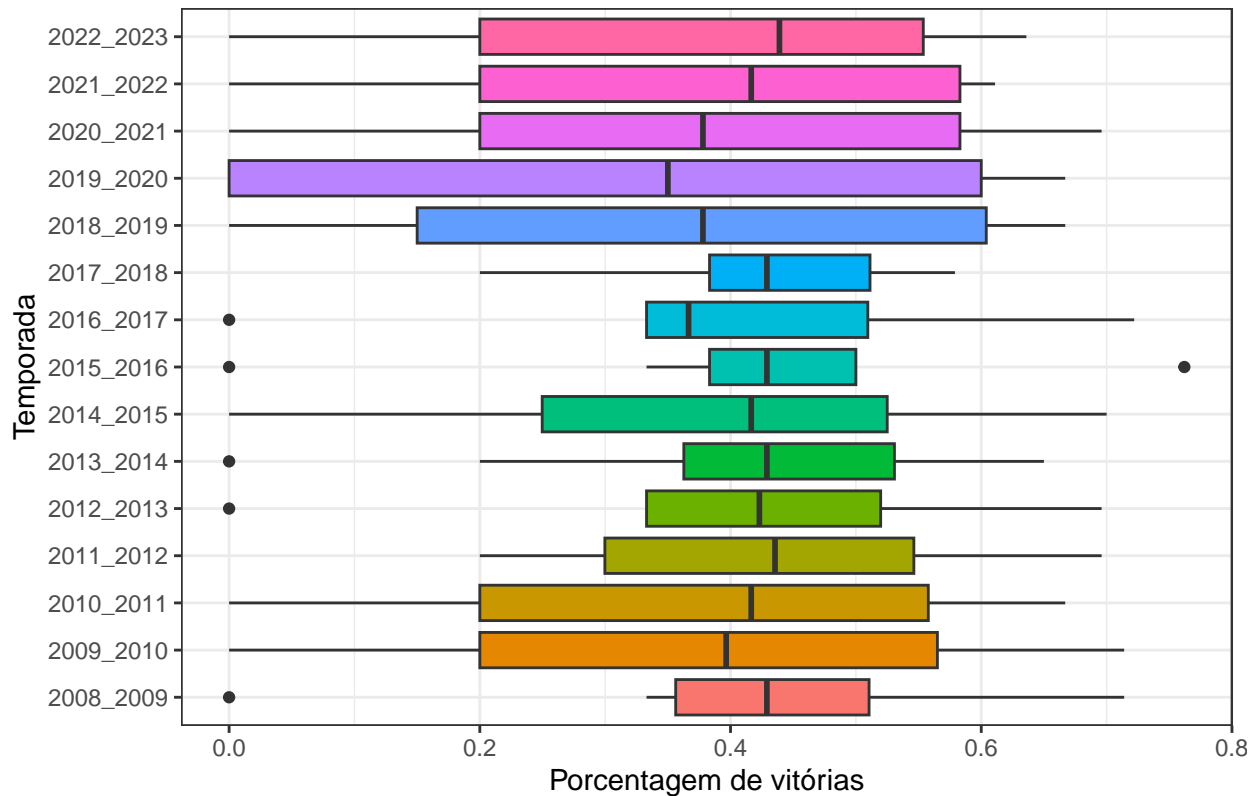
```
#Gráfico de linha com Média das Porcentagens de Vitórias por Temporada do Leste
dados_playoffs %>% filter(Conferencia == "Leste") %>% dplyr::select(TEAM, Temporada, WINP, Conferencia) %>%
  ggplot(aes(x = Temporada, y = Media, group = 1)) +
  geom_line() +
  labs(title = "Média das Porcentagens de Vitórias por Temporada do Leste",
        x = "Temporada",
        y = "Média das Porcentagens de Vitórias") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```

Média das Porcentagens de Vitórias por Temporada do Leste



```
#Boxplot das porcentagens de vitória nas temporadas do Leste
dados_playoffs %>% filter(Conferencia == "Leste") %>%
  ggplot() +
  geom_boxplot(aes(x = WINP, y = Temporada, fill = Temporada), show.legend = FALSE) +
  labs(title = "Boxplot das porcentagens de vitória nas temporadas do Leste",
        x = "Porcentagem de vitórias") +
  theme_bw()
```

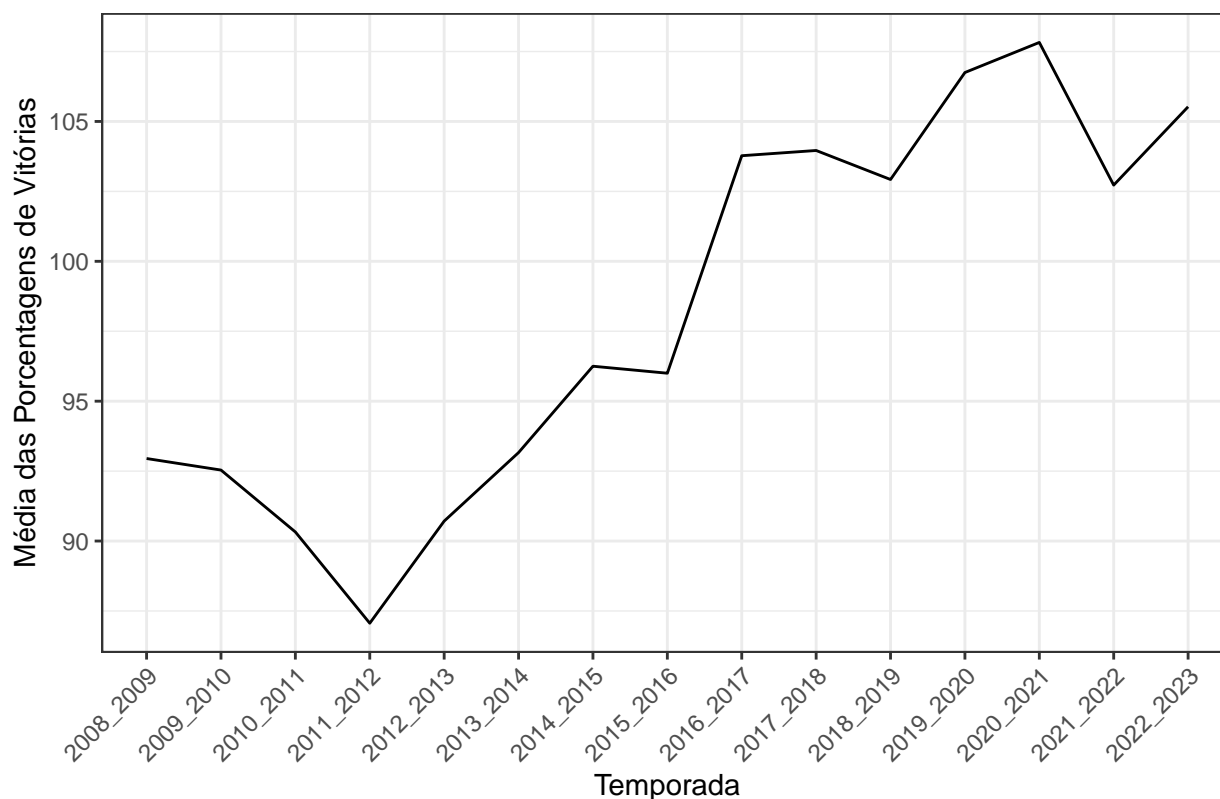
Boxplot das porcentagens de vitória nas temporadas do Leste



```
#Gráfico de linha com Média de pontos por Temporada do Leste
dados_playoffs %>% filter(Conferencia == "Leste") %>% group_by(Temporada) %>% summarise(Media_Pontos = mean(Pontos))

ggplot(aes(x = Temporada, y = Media_Pontos, group = 1)) +
  geom_line() +
  labs(title = "Média de pontos por Temporada do Leste",
       x = "Temporada",
       y = "Média das Porcentagens de Vitórias") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```

Média de pontos por Temporada do Leste



```
#Quantidade de vezes que o time do Leste ficou em cada posição
dados_playoffs %>% filter(Conferencia == "Leste") %>% count(Posicao)
```

```
## # A tibble: 16 x 2
##   Posicao     n
##   <int> <int>
## 1       1     6
## 2       2     9
## 3       3    12
## 4       4     8
## 5       5    10
## 6       6     8
## 7       7     1
## 8       8    10
## 9       9     4
## 10      10    11
## 11      11     7
## 12      12     6
## 13      13     8
## 14      14     7
## 15      15     6
## 16      16     7
```

```
##### Oeste #####
```

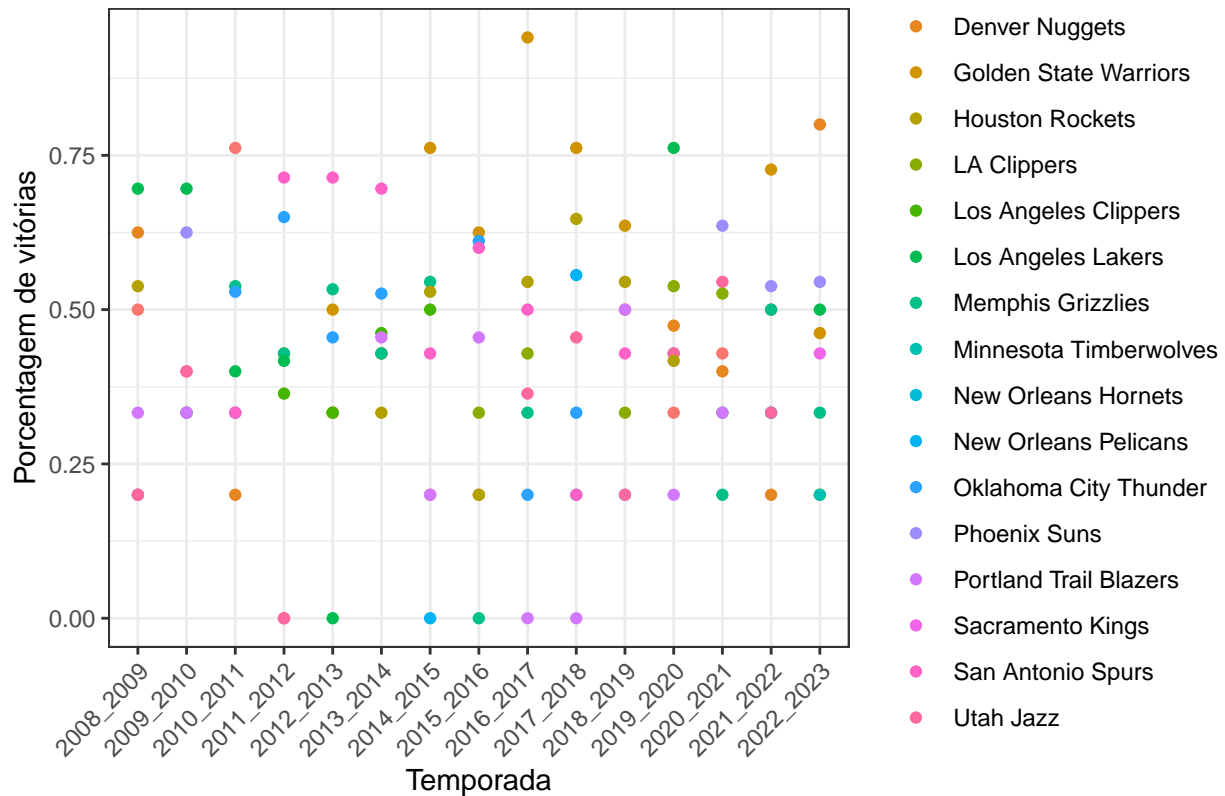
```
#Gráfico de pontos com a Porcentagem de vitórias ao decorrer das temporadas
dados_playoffs %>% filter(Conferencia == "Oeste") %>% ggplot() +
  geom_point(aes(x = Temporada, y = WINP)) +
```

```
theme_bw() +
labs(title = "Gráfico de pontos dos times do Oeste",
     y = "Porcentagem de vitórias") +
theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```



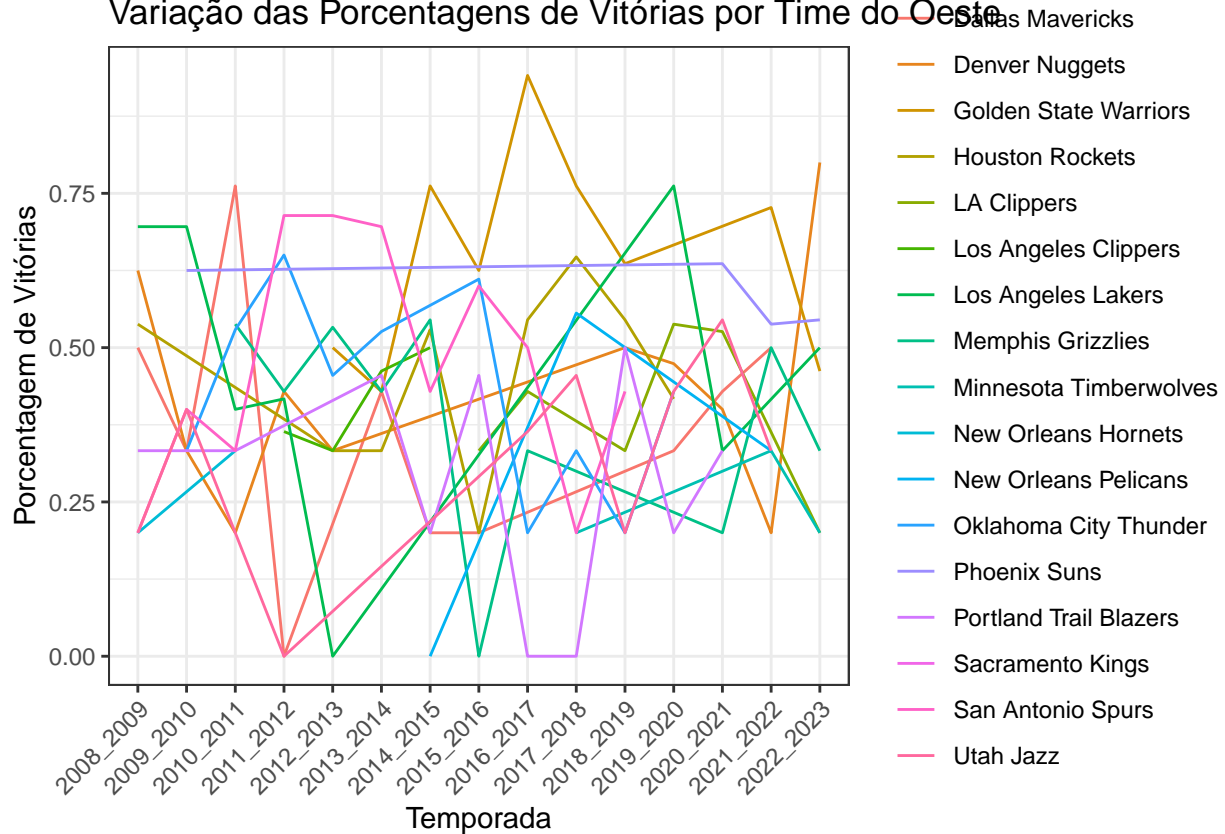
```
#Gráfico de pontos com a Porcentagem de vitórias ao decorrer das temporadas por time
dados_playoffs %>% filter(Conferencia == "Oeste") %>%
ggplot() +
geom_point(aes(x = Temporada, y = WINP, color = TEAM)) +
theme_bw() +
labs(title = "Gráfico de pontos dos times do Oeste",
     y = "Porcentagem de vitórias") +
theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```

Gráfico de pontos dos times do Oeste



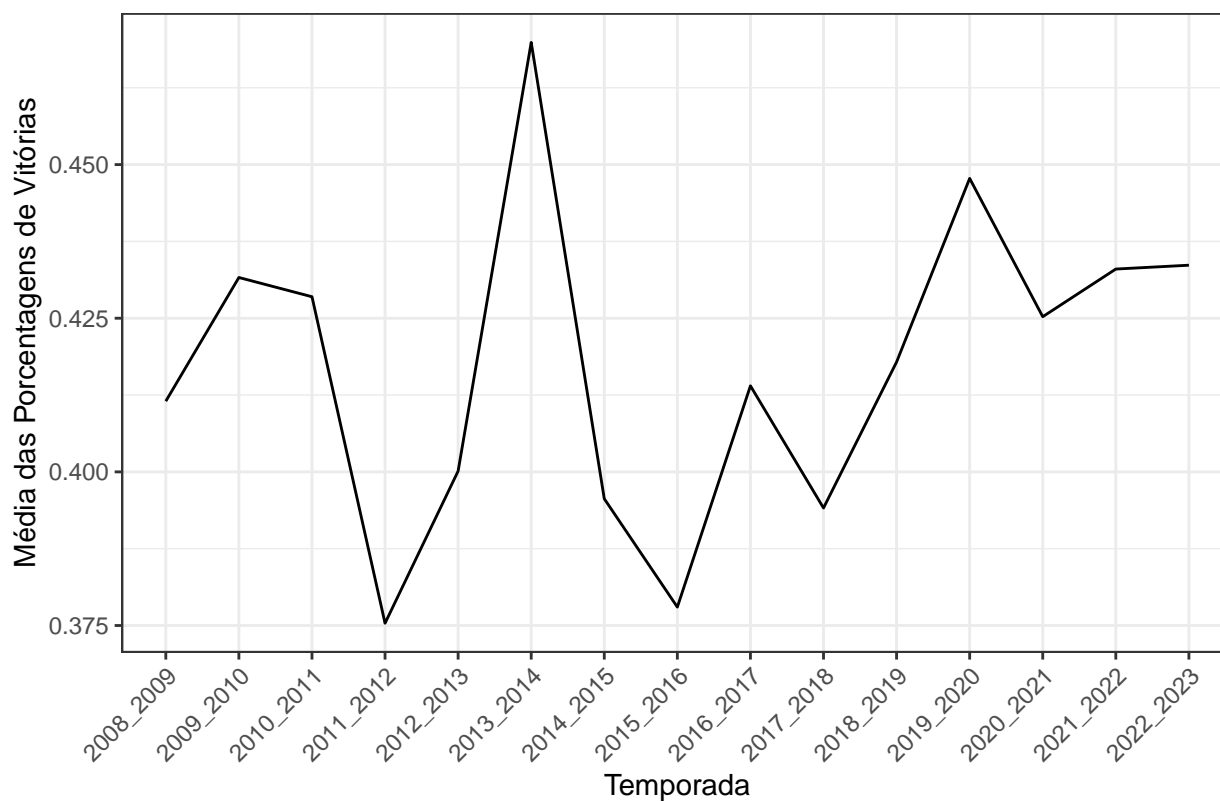
```
#Gráfico de linha com Variação das Porcentagens de Vitórias por Time do Oeste
dados_playoffs %>% filter(Conferencia == "Oeste") %>% ggplot(aes(x = Temporada, y = WINP, color = TEAM,
  geom_line() +
  labs(title = "Variação das Porcentagens de Vitórias por Time do Oeste",
    x = "Temporada",
    y = "Porcentagem de Vitórias") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```


Variação das Porcentagens de Vitórias por Time do Oeste



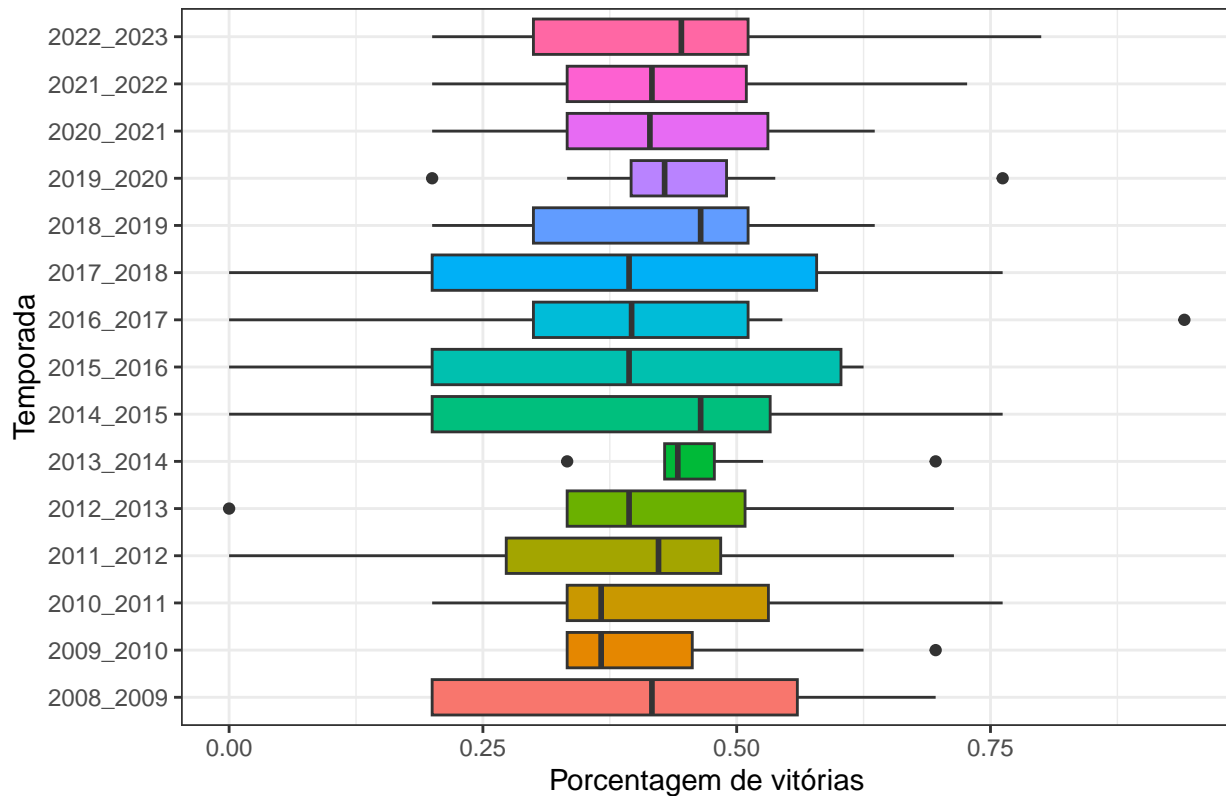
```
#Gráfico de linhas com Média das Porcentagens de Vitórias por Temporada do Oeste
dados_playoffs %>% filter(Conferencia == "Oeste") %>% dplyr::select(TEAM, Temporada, WINP, Conferencia) %>%
  ggplot(aes(x = Temporada, y = Media, group = 1)) +
  geom_line() +
  labs(title = "Média das Porcentagens de Vitórias por Temporada do Oeste",
       x = "Temporada",
       y = "Média das Porcentagens de Vitórias") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```

Média das Porcentagens de Vitórias por Temporada do Oeste



```
#Boxplot das porcentagens de vitória nas temporadas do Oeste
dados_playoffs %>% filter(Conferencia == "Oeste") %>%
  ggplot() +
  geom_boxplot(aes(x = WINP, y = Temporada, fill = Temporada), show.legend = FALSE) +
  labs(title = "Boxplot das porcentagens de vitória nas temporadas do Oeste",
       x = "Porcentagem de vitórias") +
  theme_bw()
```

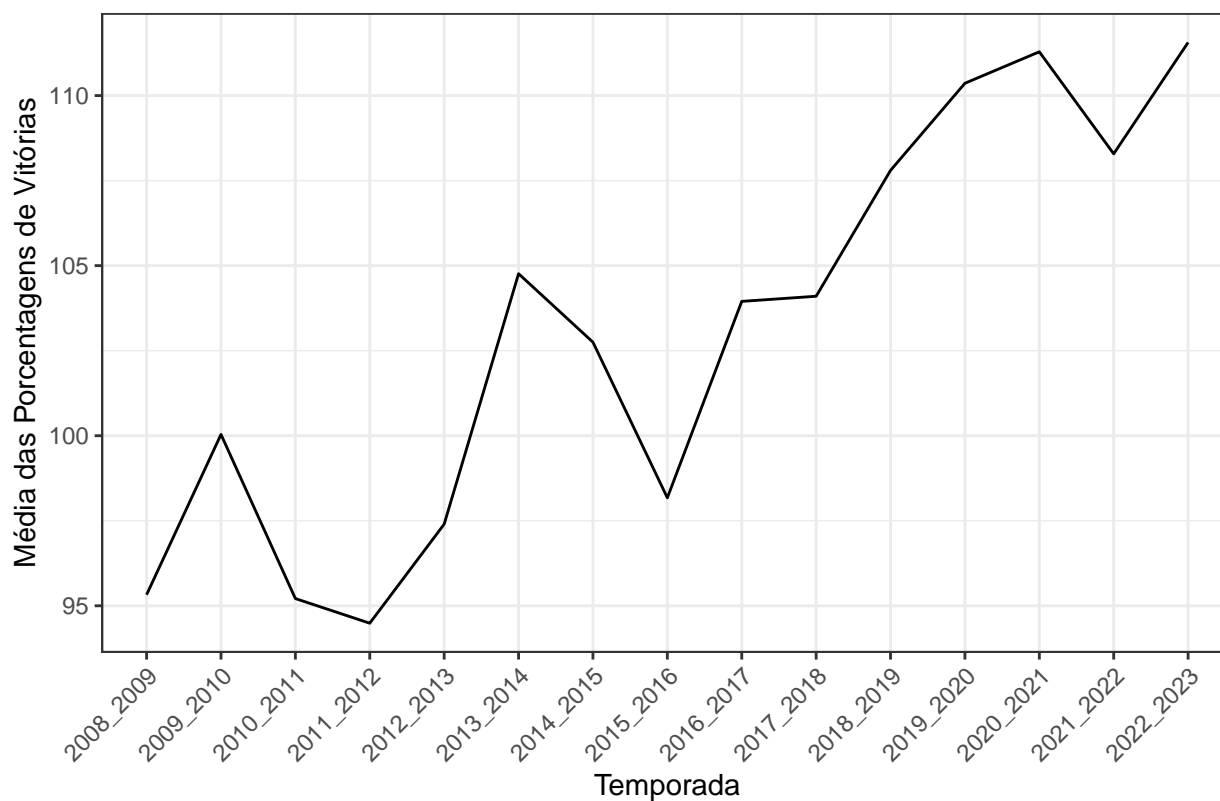
Boxplot das porcentagens de vitória nas temporadas do Oeste



```
#Gráfico de linha com a Média das pontas por temporada
dados_playoffs %>% filter(Conferencia == "Oeste") %>% group_by(Temporada) %>% summarise(Media_Pontos = mean(Pontos))

ggplot(aes(x = Temporada, y = Media_Pontos, group = 1)) +
  geom_line() +
  labs(title = "Média de pontos por Temporada do Oeste",
        x = "Temporada",
        y = "Média das Porcentagens de Vitórias") +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotaciona os rótulos em 45 graus
```

Média de pontos por Temporada do Oeste



```
#Quantidade de vezes que o time do oeste ficou em cada posição
dados_playoffs %>% filter(Conferencia == "Oeste") %>% count(Posicao)
```

```
## # A tibble: 16 x 2
##   Posicao     n
##   <int> <int>
## 1      1     10
## 2      2      5
## 3      3      6
## 4      4      5
## 5      5      8
## 6      6      8
## 7      7     14
## 8      8      8
## 9      9     12
## 10     10     14
## 11     11      6
## 12     12      6
## 13     13      9
## 14     14      3
## 15     15      5
## 16     16      1
```