

# Relatório 1

Renan de Oliveira da Cruz

2022-05-12

Abaixo temos bibliotecas que serão utilizadas nesse notebook.

```
library(readr)
library(ggplot2)
library(readr)
library(stringr)
library(MASS)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following object is masked from 'package:MASS':
##
##      select

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

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

```
library(tidyr)
```

Leitura dos dados

```
basic_info <- read_csv("basic_info.csv", show_col_types = FALSE)
```

```
## New names:
## * '' -> '...1'
```

```
detailed_info <- read_csv("detailed_info.csv", show_col_types = FALSE)
```

```
## New names:
## * '' -> '...1'
```

```
dados <- basic_info %>% left_join(detailed_info, by = "ID")
glimpse(dados)
```

```
## Rows: 19,825
## Columns: 86
## $ ...1.x <dbl> 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 1~
## $ ID <dbl> 236988, 225863, 241721, 224371, 200104, 238~
## $ Name <chr> "Eddie Nketiah", "Olivier Boscagli", "Rafae~
## $ Age <dbl> 22, 23, 22, 24, 28, 21, 23, 21, 16, 22, 25,~
## $ Nationality <chr> "England", "France", "Portugal", "England",~
## $ Overall <dbl> 72, 77, 82, 79, 89, 78, 81, 81, 77, 78, 77,~
## $ Potential <dbl> 79, 82, 90, 82, 89, 83, 87, 88, 89, 86, 79,~
## $ Club <chr> "Arsenal", "PSV", "AC Milan", "West Ham Uni~
## $ Contract <chr> "2016 ~ 2022", "2019 ~ 2025", "2019 ~ 2024"~
## $ Value <chr> "€4.8M", "€14.5M", "€68.5M", "€24M", "€104M~
## $ Wage <chr> "€45K", "€15K", "€52K", "€63K", "€240K", "€~
## $ 'Total stat' <dbl> 1698, 1961, 1959, 1966, 2141, 1955, 1730, 2~
## $ ...1.y <dbl> 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 1~
## $ LS <dbl> 74, 66, 82, 77, 88, 67, 59, 77, 69, 80, 56,~
## $ ST <dbl> 74, 66, 82, 77, 88, 67, 59, 77, 69, 80, 56,~
## $ RS <dbl> 74, 66, 82, 77, 88, 67, 59, 77, 69, 80, 56,~
## $ LW <dbl> 70, 69, 82, 78, 87, 70, 55, 76, 76, 76, 57,~
## $ LF <dbl> 72, 68, 82, 78, 87, 68, 57, 77, 74, 77, 57,~
## $ CF <dbl> 72, 68, 82, 78, 87, 68, 57, 77, 74, 77, 57,~
## $ RF <dbl> 72, 68, 82, 78, 87, 68, 57, 77, 74, 77, 57,~
## $ RW <dbl> 70, 69, 82, 78, 87, 70, 55, 76, 76, 76, 57,~
## $ LAM <dbl> 71, 73, 82, 79, 89, 70, 61, 80, 79, 77, 61,~
## $ CAM <dbl> 71, 73, 82, 79, 89, 70, 61, 80, 79, 77, 61,~
## $ RAM <dbl> 71, 73, 82, 79, 89, 70, 61, 80, 79, 77, 61,~
## $ LM <dbl> 70, 74, 83, 80, 89, 73, 61, 79, 78, 78, 62,~
## $ LCM <dbl> 62, 76, 74, 74, 83, 70, 68, 83, 79, 71, 67,~
## $ CM <dbl> 62, 76, 74, 74, 83, 70, 68, 83, 79, 71, 67,~
## $ RCM <dbl> 62, 76, 74, 74, 83, 70, 68, 83, 79, 71, 67,~
## $ RM <dbl> 70, 74, 83, 80, 89, 73, 61, 79, 78, 78, 62,~
## $ LWB <dbl> 50, 78, 60, 65, 71, 78, 72, 81, 73, 64, 72,~
## $ LDM <dbl> 47, 79, 56, 62, 67, 74, 79, 84, 74, 61, 76,~
## $ CDM <dbl> 47, 79, 56, 62, 67, 74, 79, 84, 74, 61, 76,~
## $ RDM <dbl> 47, 79, 56, 62, 67, 74, 79, 84, 74, 61, 76,~
## $ RWB <dbl> 50, 78, 60, 65, 71, 78, 72, 81, 73, 64, 72,~
## $ LB <dbl> 48, 78, 56, 62, 67, 78, 75, 81, 71, 62, 73,~
## $ LCB <dbl> 42, 79, 48, 55, 57, 75, 83, 82, 66, 58, 79,~
## $ CB <dbl> 42, 79, 48, 55, 57, 75, 83, 82, 66, 58, 79,~
## $ RCB <dbl> 42, 79, 48, 55, 57, 75, 83, 82, 66, 58, 79,~
## $ RB <dbl> 48, 78, 56, 62, 67, 78, 75, 81, 71, 62, 73,~
## $ GK <dbl> 18, 19, 21, 19, 22, 18, 19, 19, 19, 21, 22,~
## $ Height <dbl> 180, 181, 188, 175, 183, 169, 190, 187, 173~
## $ Weight <dbl> 73, 68, 81, 70, 78, 67, 87, 81, 68, 81, 75,~
## $ DOB <chr> "May 30, 1999", "Nov 18, 1997", "Jun 10, 19~
## $ 'Preferred foot' <chr> "Right", "Left", "Right", "Left", "Right", ~
## $ 'Weak foot' <dbl> 3, 4, 4, 3, 5, 3, 2, 3, 3, 3, 3, 3, 3, 4, 5~
## $ 'Skill move' <dbl> 3, 3, 4, 4, 4, 3, 2, 3, 3, 3, 2, 2, 4, 4, 4~
## $ 'International reputation' <dbl> 1, 1, 1, 1, 4, 1, 2, 2, 1, 1, 1, 1, 2, 1, 3~
## $ 'Work rate' <chr> "High/Medium", "Medium/Medium", "Medium/Low~
```

```
## $ 'Body type' <chr> "Lean (170-185)", "Lean (170-185)", "Lean (~
## $ 'Real face' <chr> "Yes", "Yes", "No", "Yes", "Yes", "No", "No~
## $ 'Release clause' <chr> "€10.1M", "€21.4M", "€130.2M", "€47.4M", "€~
## $ Crossing <dbl> 43, 73, 69, 79, 83, 75, 38, 68, 64, 67, 50,~
## $ Finishing <dbl> 73, 33, 80, 80, 88, 50, 31, 67, 64, 81, 30,~
## $ 'Heading Accuracy' <dbl> 66, 74, 63, 64, 68, 58, 85, 73, 42, 73, 72,~
## $ 'Short Passing' <dbl> 66, 78, 76, 74, 84, 75, 75, 84, 81, 73, 75,~
## $ Volleys <dbl> 69, 29, 74, 58, 79, 59, 28, 50, 49, 67, 27,~
## $ Dribbling <dbl> 74, 73, 89, 82, 87, 75, 52, 78, 82, 73, 64,~
## $ Curve <dbl> 54, 52, 80, 71, 85, 36, 47, 76, 74, 67, 33,~
## $ 'FK Accuracy' <dbl> 38, 56, 60, 70, 74, 42, 27, 53, 47, 61, 30,~
## $ 'Long Passing' <dbl> 45, 78, 62, 67, 74, 56, 72, 83, 76, 52, 66,~
## $ 'Ball Control' <dbl> 73, 77, 85, 79, 84, 74, 65, 80, 81, 75, 69,~
## $ Acceleration <dbl> 85, 71, 90, 78, 85, 84, 62, 70, 78, 88, 55,~
## $ 'Sprint Speed' <dbl> 82, 75, 92, 76, 90, 85, 73, 73, 75, 86, 58,~
## $ Agility <dbl> 80, 72, 82, 79, 86, 82, 51, 75, 86, 77, 54,~
## $ Reactions <dbl> 74, 77, 83, 77, 91, 73, 81, 81, 76, 80, 78,~
## $ Balance <dbl> 72, 70, 80, 83, 78, 89, 37, 71, 89, 54, 61,~
## $ 'Shot Power' <dbl> 72, 73, 80, 72, 88, 65, 51, 78, 56, 83, 48,~
## $ Jumping <dbl> 73, 82, 62, 64, 60, 85, 70, 74, 70, 73, 76,~
## $ Stamina <dbl> 66, 76, 75, 80, 88, 80, 67, 86, 77, 87, 68,~
## $ Strength <dbl> 59, 75, 80, 62, 64, 67, 87, 82, 56, 86, 78,~
## $ 'Long Shots' <dbl> 60, 64, 72, 76, 89, 62, 30, 72, 59, 71, 27,~
## $ Aggression <dbl> 52, 71, 60, 57, 62, 77, 81, 82, 82, 78, 75,~
## $ Interceptions <dbl> 19, 80, 24, 37, 39, 74, 80, 84, 65, 41, 78,~
## $ Positioning <dbl> 74, 62, 80, 79, 91, 60, 44, 75, 73, 77, 40,~
## $ Vision <dbl> 58, 74, 74, 74, 84, 52, 59, 82, 81, 74, 57,~
## $ Penalties <dbl> 68, 36, 63, 70, 73, 46, 40, 55, 51, 76, 38,~
## $ Composure <dbl> 68, 77, 81, 75, 89, 66, 74, 82, 74, 82, 74,~
## $ 'Defensive Awareness' <dbl> 22, 76, 22, 49, 47, 70, 83, 80, 59, 36, 77,~
## $ 'Standing Tackle' <dbl> 19, 78, 24, 40, 34, 80, 84, 83, 66, 39, 81,~
## $ 'Sliding Tackle' <dbl> 15, 77, 21, 36, 33, 78, 80, 78, 62, 33, 78,~
## $ 'GK Diving' <dbl> 12, 8, 10, 14, 11, 9, 12, 9, 8, 14, 14, 8, ~
## $ 'GK Handling' <dbl> 10, 7, 12, 8, 13, 9, 8, 8, 10, 10, 11, 11, ~
## $ 'GK Kicking' <dbl> 11, 14, 15, 14, 13, 12, 13, 12, 11, 12, 10,~
## $ 'GK Positioning' <dbl> 9, 8, 11, 10, 6, 8, 10, 12, 7, 10, 14, 5, 1~
## $ 'GK Reflexes' <dbl> 5, 12, 9, 7, 10, 8, 7, 7, 13, 13, 13, 9, 5,~
## $ Traits <chr> "Chip Shot (AI)", "Dives Into Tackles (AI),~
```

Ao todo temos 19825 linhas e 86. Vamos agora trabalhar para a diminuição do numero de covariáveis.

Vamos tomar uma amostra de tamanho 1000 desses dados.

```
dados <- dados %>% drop_na() %>% filter(`Release clause` != "</label>")
set.seed(22)
dados <- dados[sample(nrow(dados), 1000), ]
```

A principio vamos remover as colunas "...1.x", "ID", "Name", "...1.y" e "DOB" (date of birth).

```
dadoss <- dados %>% dplyr::select(-`...1.x`, -`ID`, -`Name`, -`...1.y`, -`DOB`)
```

```
dadoss$Nationality %>% unique() %>% length()
```

```
## [1] 91
```

```
dadoss$Club %>% unique() %>% length()
```

```
## [1] 525
```

```
dadoss$Contract %>% unique()
```

```
## [1] "2018 ~ 2022" "2022 ~ 2023" "2017 ~ 2022" "2021 ~ 2023" "2022 ~ 2022"
## [6] "2021 ~ 2022" "2022 ~ 2026" "2018 ~ 2023" "2021 ~ 2024" "2020 ~ 2022"
## [11] "2020 ~ 2023" "2021 ~ 2025" "2021 ~ 2021" "2019 ~ 2022" "2019 ~ 2024"
## [16] "2019 ~ 2025" "2019 ~ 2023" "2019 ~ 2021" "2020 ~ 2024" "2022 ~ 2024"
## [21] "2020 ~ 2021" "2017 ~ 2023" "2019 ~ 2026" "2011 ~ 2025" "2014 ~ 2022"
## [26] "2018 ~ 2024" "2018 ~ 2025" "2021 ~ 2026" "2011 ~ 2024" "2014 ~ 2021"
## [31] "2020 ~ 2025" "2016 ~ 2023" "2022 ~ 2025" "2014 ~ 2026" "2011 ~ 2023"
## [36] "2017 ~ 2025" "2012 ~ 2025" "2014 ~ 2023" "2018 ~ 2021" "2018 ~ 2026"
## [41] "2021 ~ 2027" "2012 ~ 2023" "2020 ~ 2026" "2013 ~ 2024" "2015 ~ 2022"
## [46] "2011 ~ 2022" "2015 ~ 2023" "2017 ~ 2024" "2015 ~ 2025" "2016 ~ 2021"
## [51] "2016 ~ 2022" "2002 ~ 2022" "2018 ~ 2027" "2014 ~ 2025" "2019 ~ 2020"
## [56] "2016 ~ 2024" "2012 ~ 2022" "2015 ~ 2024" "2020 ~ 2027" "2017 ~ 2026"
## [61] "2013 ~ 2025" "2010 ~ 2023" "2009 ~ 2022"
```

```
dadoss2 <- dadoss %>%
  dplyr::select(-Nationality, -Club, -Contract, -`Total stat`) %>%
  mutate(Value = parse_number(Value),
         Wage = parse_number(Wage))
```

```
dadoss2 <- dadoss2 %>% mutate(`Preferred foot` = as.factor(`Preferred foot`))
dadoss2 %>% group_by(`Preferred foot`) %>% summarise(n = n())
```

```
## # A tibble: 2 x 2
##   `Preferred foot`      n
##   <fct>             <int>
## 1 Left              258
## 2 Right             742
```

```
dadoss2 %>% group_by(`Work rate`) %>% summarise(n = n()) %>% print()
```

```
## # A tibble: 9 x 2
##   `Work rate`      n
##   <chr>          <int>
## 1 High/High      60
## 2 High/Low       43
## 3 High/Medium   168
## 4 Low/High       22
## 5 Low/Low        2
## 6 Low/Medium     18
## 7 Medium/High    87
## 8 Medium/Low     52
## 9 Medium/Medium 548
```

```
dadoss2 <- dadoss2 %>% dplyr::select(-`Work rate`)
```

```
dadoss2 %>% group_by(`Body type`) %>% summarise(n = n()) %>% print()
```

```
## # A tibble: 10 x 2
##   'Body type'      n
##   <chr>          <int>
## 1 Lean (170-)      30
## 2 Lean (170-185)  217
## 3 Lean (185+)     111
## 4 Normal (170-)   34
## 5 Normal (170-185) 368
## 6 Normal (185+)   191
## 7 Stocky (170-)   9
## 8 Stocky (170-185) 25
## 9 Stocky (185+)   11
## 10 Unique         4
```

```
dadoss3 <- dadoss2 %>%
  mutate(body_type = as.factor(word(`Body type`, 1))) %>%
  dplyr::select(-`Body type`)
dadoss3 %>% group_by(body_type) %>% summarise(n = n()) %>% print()
```

```
## # A tibble: 4 x 2
##   body_type      n
##   <fct>        <int>
## 1 Lean        358
## 2 Normal      593
## 3 Stocky       45
## 4 Unique        4
```

```
dadoss3 <- dadoss3 %>% dplyr::select(-`Real face`)
```

```
dadoss4 <- dadoss3 %>%
  mutate(
    `Release clause` = ifelse(
      `Release clause` == "</label>", NA, `Release clause`
    ),
    `Release clause` = parse_number(`Release clause`)
  )
```

```
dadoss4 <- dadoss4 %>% dplyr::select(-Traits)
```

## Análise de correlação

```
data_num <- dadoss4 %>% dplyr::select(where(is.numeric))
data_num <- data_num %>% drop_na()
glimpse(data_num)
```

```

## Rows: 1,000
## Columns: 72
## $ Age <dbl> 29, 32, 32, 16, 24, 17, 17, 20, 23, 25, 20,~
## $ Overall <dbl> 75, 80, 74, 55, 69, 54, 60, 52, 66, 69, 62,~
## $ Potential <dbl> 75, 80, 74, 73, 72, 74, 71, 65, 71, 72, 69,~
## $ Value <dbl> 5.5, 16.5, 2.1, 325.0, 1.9, 300.0, 475.0, 1~
## $ Wage <dbl> 30, 58, 8, 500, 500, 500, 500, 500, 500, 2,~
## $ LS <dbl> 72, 79, 33, 51, 67, 52, 53, 36, 65, 53, 49,~
## $ ST <dbl> 72, 79, 33, 51, 67, 52, 53, 36, 65, 53, 49,~
## $ RS <dbl> 72, 79, 33, 51, 67, 52, 53, 36, 65, 53, 49,~
## $ LW <dbl> 74, 78, 34, 54, 69, 54, 53, 33, 66, 55, 53,~
## $ LF <dbl> 74, 78, 34, 52, 68, 53, 52, 32, 66, 53, 49,~
## $ CF <dbl> 74, 78, 34, 52, 68, 53, 52, 32, 66, 53, 49,~
## $ RF <dbl> 74, 78, 34, 52, 68, 53, 52, 32, 66, 53, 49,~
## $ RW <dbl> 74, 78, 34, 54, 69, 54, 53, 33, 66, 55, 53,~
## $ LAM <dbl> 75, 80, 39, 54, 71, 56, 54, 33, 68, 57, 54,~
## $ CAM <dbl> 75, 80, 39, 54, 71, 56, 54, 33, 68, 57, 54,~
## $ RAM <dbl> 75, 80, 39, 54, 71, 56, 54, 33, 68, 57, 54,~
## $ LM <dbl> 75, 80, 38, 56, 71, 56, 56, 36, 68, 60, 57,~
## $ LCM <dbl> 73, 80, 38, 50, 67, 53, 55, 36, 65, 60, 56,~
## $ CM <dbl> 73, 80, 38, 50, 67, 53, 55, 36, 65, 60, 56,~
## $ RCM <dbl> 73, 80, 38, 50, 67, 53, 55, 36, 65, 60, 56,~
## $ RM <dbl> 75, 80, 38, 56, 71, 56, 56, 36, 68, 60, 57,~
## $ LWB <dbl> 61, 80, 34, 48, 57, 51, 59, 47, 60, 69, 62,~
## $ LDM <dbl> 61, 80, 34, 43, 54, 49, 60, 46, 59, 67, 60,~
## $ CDM <dbl> 61, 80, 34, 43, 54, 49, 60, 46, 59, 67, 60,~
## $ RDM <dbl> 61, 80, 34, 43, 54, 49, 60, 46, 59, 67, 60,~
## $ RWB <dbl> 61, 80, 34, 48, 57, 51, 59, 47, 60, 69, 62,~
## $ LB <dbl> 58, 80, 33, 45, 54, 51, 60, 49, 58, 69, 62,~
## $ LCB <dbl> 53, 77, 29, 39, 47, 45, 62, 54, 54, 70, 60,~
## $ CB <dbl> 53, 77, 29, 39, 47, 45, 62, 54, 54, 70, 60,~
## $ RCB <dbl> 53, 77, 29, 39, 47, 45, 62, 54, 54, 70, 60,~
## $ RB <dbl> 58, 80, 33, 45, 54, 51, 60, 49, 58, 69, 62,~
## $ GK <dbl> 19, 21, 74, 17, 18, 17, 18, 14, 15, 18, 17,~
## $ Height <dbl> 176, 181, 180, 186, 169, 170, 188, 183, 170~
## $ Weight <dbl> 69, 74, 76, 70, 65, 60, 75, 75, 74, 75, 71,~
## $ 'Weak foot' <dbl> 4, 3, 4, 3, 3, 3, 2, 2, 4, 3, 3, 2, 3, 3, 3~
## $ 'Skill move' <dbl> 4, 3, 1, 2, 4, 2, 2, 2, 3, 2, 2, 3, 2, 3, 2~
## $ 'International reputation' <dbl> 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1~
## $ 'Release clause' <dbl> 9.4, 33.8, 3.4, 691.0, 4.2, 825.0, 998.0, 3~
## $ Crossing <dbl> 65, 79, 20, 55, 65, 41, 42, 28, 62, 72, 58,~
## $ Finishing <dbl> 67, 76, 12, 39, 68, 46, 35, 21, 61, 22, 26,~
## $ 'Heading Accuracy' <dbl> 48, 69, 13, 33, 50, 35, 59, 49, 59, 66, 55,~
## $ 'Short Passing' <dbl> 78, 80, 42, 49, 65, 55, 58, 30, 68, 64, 61,~
## $ Volleys <dbl> 69, 69, 19, 39, 65, 42, 37, 28, 54, 31, 28,~
## $ Dribbling <dbl> 79, 77, 17, 60, 75, 54, 64, 29, 70, 61, 54,~
## $ Curve <dbl> 76, 78, 21, 48, 60, 43, 46, 23, 80, 31, 40,~
## $ 'FK Accuracy' <dbl> 69, 84, 20, 39, 65, 35, 38, 21, 69, 34, 26,~
## $ 'Long Passing' <dbl> 75, 81, 38, 47, 65, 53, 53, 27, 60, 48, 59,~
## $ 'Ball Control' <dbl> 78, 80, 25, 56, 74, 52, 63, 31, 67, 62, 55,~
## $ Acceleration <dbl> 74, 73, 63, 68, 71, 75, 61, 57, 70, 44, 70,~
## $ 'Sprint Speed' <dbl> 72, 75, 62, 66, 70, 73, 63, 55, 75, 67, 70,~
## $ Agility <dbl> 77, 73, 69, 58, 75, 82, 66, 40, 66, 57, 63,~
## $ Reactions <dbl> 74, 80, 77, 48, 65, 50, 60, 44, 61, 66, 60,~

```

```
## $ Balance <dbl> 77, 71, 67, 60, 70, 83, 54, 59, 61, 58, 74, ~
## $ 'Shot Power' <dbl> 71, 85, 51, 56, 65, 48, 60, 32, 54, 45, 31, ~
## $ Jumping <dbl> 73, 68, 78, 48, 60, 51, 58, 74, 56, 67, 59, ~
## $ Stamina <dbl> 55, 87, 40, 54, 69, 53, 53, 55, 58, 69, 60, ~
## $ Strength <dbl> 67, 73, 54, 54, 51, 34, 58, 61, 59, 75, 56, ~
## $ 'Long Shots' <dbl> 71, 80, 13, 37, 68, 51, 32, 16, 75, 29, 32, ~
## $ Aggression <dbl> 45, 74, 21, 32, 51, 42, 70, 45, 45, 73, 49, ~
## $ Interceptions <dbl> 42, 79, 16, 23, 25, 42, 56, 52, 58, 70, 61, ~
## $ Positioning <dbl> 70, 76, 11, 48, 60, 52, 35, 24, 61, 49, 44, ~
## $ Vision <dbl> 77, 78, 60, 54, 73, 50, 40, 31, 66, 56, 41, ~
## $ Penalties <dbl> 69, 61, 12, 58, 70, 49, 41, 36, 68, 44, 37, ~
## $ Composure <dbl> 71, 77, 64, 46, 65, 52, 63, 44, 51, 67, 58, ~
## $ 'Defensive Awareness' <dbl> 45, 76, 21, 33, 30, 38, 59, 52, 51, 63, 60, ~
## $ 'Standing Tackle' <dbl> 35, 76, 14, 29, 40, 45, 60, 58, 42, 70, 59, ~
## $ 'Sliding Tackle' <dbl> 43, 75, 13, 36, 42, 43, 64, 52, 34, 70, 55, ~
## $ 'GK Diving' <dbl> 8, 11, 74, 6, 12, 14, 8, 14, 9, 6, 7, 13, 1~
## $ 'GK Handling' <dbl> 8, 9, 73, 12, 9, 7, 8, 8, 7, 8, 11, 6, 8, 6~
## $ 'GK Kicking' <dbl> 9, 10, 68, 14, 14, 14, 10, 8, 8, 12, 7, 13, ~
## $ 'GK Positioning' <dbl> 15, 13, 72, 9, 12, 13, 11, 7, 7, 12, 8, 8, ~
## $ 'GK Reflexes' <dbl> 8, 13, 74, 15, 6, 7, 15, 5, 6, 13, 13, 10, ~
```

```
corr <- cor(data_num)
#corr
```

Desses numericos uma série de atributos não apresentam descrição e serão desconsiderados Podemos ver que alguns atributos estão fortemente correlacionados com outros.

```
cols_rm <- c("LS", "ST", "RS", "LW", "LF", "CF", "RF", "RW", "LAM", "CAM",
            "RAM", "LM", "LCM", "CM", "RCM", "RM", "LWB", "LDM", "CDM",
            "RDM", "RWB", "LB", "LCB", "CB", "RCB", "RB", "GK",
            "GK Diving", "GK Handling", "GK Kicking", "GK Positioning",
            "GK Reflexes", "Weak foot", "Skill move",
            "International reputation")
dadosrm <- dadossc4[, -which(colnames(dadossc4) %in% cols_rm)]
```

```
glimpse(dadosrm)
```

```
## Rows: 1,000
## Columns: 39
## $ Age <dbl> 29, 32, 32, 16, 24, 17, 17, 20, 23, 25, 20, 20, ~
## $ Overall <dbl> 75, 80, 74, 55, 69, 54, 60, 52, 66, 69, 62, 59, ~
## $ Potential <dbl> 75, 80, 74, 73, 72, 74, 71, 65, 71, 72, 69, 70, ~
## $ Value <dbl> 5.5, 16.5, 2.1, 325.0, 1.9, 300.0, 475.0, 180.0, ~
## $ Wage <dbl> 30, 58, 8, 500, 500, 500, 500, 500, 500, 2, 500, ~
## $ Height <dbl> 176, 181, 180, 186, 169, 170, 188, 183, 170, 186~
## $ Weight <dbl> 69, 74, 76, 70, 65, 60, 75, 75, 74, 75, 71, 63, ~
## $ 'Preferred foot' <fct> Right, Right, Right, Left, Right, Right, Right, ~
## $ 'Release clause' <dbl> 9.4, 33.8, 3.4, 691.0, 4.2, 825.0, 998.0, 338.0, ~
## $ Crossing <dbl> 65, 79, 20, 55, 65, 41, 42, 28, 62, 72, 58, 61, ~
## $ Finishing <dbl> 67, 76, 12, 39, 68, 46, 35, 21, 61, 22, 26, 46, ~
## $ 'Heading Accuracy' <dbl> 48, 69, 13, 33, 50, 35, 59, 49, 59, 66, 55, 21, ~
## $ 'Short Passing' <dbl> 78, 80, 42, 49, 65, 55, 58, 30, 68, 64, 61, 40, ~
## $ Volleys <dbl> 69, 69, 19, 39, 65, 42, 37, 28, 54, 31, 28, 31, ~
```

```
## $ Dribbling <dbl> 79, 77, 17, 60, 75, 54, 64, 29, 70, 61, 54, 70, ~
## $ Curve <dbl> 76, 78, 21, 48, 60, 43, 46, 23, 80, 31, 40, 56, ~
## $ 'FK Accuracy' <dbl> 69, 84, 20, 39, 65, 35, 38, 21, 69, 34, 26, 32, ~
## $ 'Long Passing' <dbl> 75, 81, 38, 47, 65, 53, 53, 27, 60, 48, 59, 32, ~
## $ 'Ball Control' <dbl> 78, 80, 25, 56, 74, 52, 63, 31, 67, 62, 55, 62, ~
## $ Acceleration <dbl> 74, 73, 63, 68, 71, 75, 61, 57, 70, 44, 70, 89, ~
## $ 'Sprint Speed' <dbl> 72, 75, 62, 66, 70, 73, 63, 55, 75, 67, 70, 90, ~
## $ Agility <dbl> 77, 73, 69, 58, 75, 82, 66, 40, 66, 57, 63, 84, ~
## $ Reactions <dbl> 74, 80, 77, 48, 65, 50, 60, 44, 61, 66, 60, 39, ~
## $ Balance <dbl> 77, 71, 67, 60, 70, 83, 54, 59, 61, 58, 74, 77, ~
## $ 'Shot Power' <dbl> 71, 85, 51, 56, 65, 48, 60, 32, 54, 45, 31, 38, ~
## $ Jumping <dbl> 73, 68, 78, 48, 60, 51, 58, 74, 56, 67, 59, 38, ~
## $ Stamina <dbl> 55, 87, 40, 54, 69, 53, 53, 55, 58, 69, 60, 31, ~
## $ Strength <dbl> 67, 73, 54, 54, 51, 34, 58, 61, 59, 75, 56, 29, ~
## $ 'Long Shots' <dbl> 71, 80, 13, 37, 68, 51, 32, 16, 75, 29, 32, 44, ~
## $ Aggression <dbl> 45, 74, 21, 32, 51, 42, 70, 45, 45, 73, 49, 26, ~
## $ Interceptions <dbl> 42, 79, 16, 23, 25, 42, 56, 52, 58, 70, 61, 18, ~
## $ Positioning <dbl> 70, 76, 11, 48, 60, 52, 35, 24, 61, 49, 44, 56, ~
## $ Vision <dbl> 77, 78, 60, 54, 73, 50, 40, 31, 66, 56, 41, 61, ~
## $ Penalties <dbl> 69, 61, 12, 58, 70, 49, 41, 36, 68, 44, 37, 31, ~
## $ Composure <dbl> 71, 77, 64, 46, 65, 52, 63, 44, 51, 67, 58, 62, ~
## $ 'Defensive Awareness' <dbl> 45, 76, 21, 33, 30, 38, 59, 52, 51, 63, 60, 22, ~
## $ 'Standing Tackle' <dbl> 35, 76, 14, 29, 40, 45, 60, 58, 42, 70, 59, 26, ~
## $ 'Sliding Tackle' <dbl> 43, 75, 13, 36, 42, 43, 64, 52, 34, 70, 55, 19, ~
## $ body_type <fct> Lean, Normal, Lean, Lean, Lean, Normal, Lean, Le~
```

## Análise de correlação

Vamos retirar alguns atributos baseados na correção do seguinte modo.

1. Calculo a correlação 2 a 2 de todos os atributos;
2. Tomo as correlações que são maiores do que 0.8;
3. Dessas, vejo qual é o atributo que mais aparece, ou seja, o atributos que mais apresenta correlação maior do 0.8 com seus pares.

```
data_num2 <- dadosrm %>%
  dplyr::select(-Potential) %>%
  dplyr::select(where(is.numeric)) %>%
  drop_na()
all_rm_vars <- c()
for(j in 1:50) {
  corr <- cor(data_num2)
  corr[lower.tri(corr) == FALSE] <- NA
  cordf <- data.frame()
  for (i in seq(1, ncol(corr))) {
    corr_coli <- corr[, i]
    names(corr_coli) <- NULL
    df_aux <- data.frame(
      Variavel1 = rownames(corr),
      Variavel2 = colnames(corr)[i],
      Correlacao = corr_coli
```



```

    )
    cordf <- rbind(cordf, df_aux)
  }
  cordf <- cordf %>% filter(!is.na(Correlacao))
  corr__ <- cordf %>% filter(Correlacao > 0.9)
  if(0 %in% dim(corr__)) break
  var <- corr__ %>%
    group_by(Variavel2) %>%
    summarise(n = n()) %>%
    arrange(-n) %>%
    head(1) %>%
    pull(Variavel2)
  data_num2 <- data_num2[, -which(colnames(data_num2) == var)]
  all_rm_vars <- c(all_rm_vars, var)
}
all_rm_vars

```

```

## [1] "Interceptions"      "Defensive Awareness" "Acceleration"
## [4] "Dribbling"           "Finishing"           "Short Passing"
## [7] "Standing Tackle"

```

Segundo o critério que definimos devemos retirar as colunas acima

```

dados_m <- dadosrm[, -which(colnames(dadosrm) %in% all_rm_vars)]
glimpse(dados_m)

```

```

## Rows: 1,000
## Columns: 32
## $ Age <dbl> 29, 32, 32, 16, 24, 17, 17, 20, 23, 25, 20, 20, 18, ~
## $ Overall <dbl> 75, 80, 74, 55, 69, 54, 60, 52, 66, 69, 62, 59, 59, ~
## $ Potential <dbl> 75, 80, 74, 73, 72, 74, 71, 65, 71, 72, 69, 70, 72, ~
## $ Value <dbl> 5.5, 16.5, 2.1, 325.0, 1.9, 300.0, 475.0, 180.0, 1.~
## $ Wage <dbl> 30, 58, 8, 500, 500, 500, 500, 500, 500, 2, 500, 1, ~
## $ Height <dbl> 176, 181, 180, 186, 169, 170, 188, 183, 170, 186, 1~
## $ Weight <dbl> 69, 74, 76, 70, 65, 60, 75, 75, 74, 75, 71, 63, 82, ~
## $ 'Preferred foot' <fct> Right, Right, Right, Left, Right, Right, Right, Rig~
## $ 'Release clause' <dbl> 9.4, 33.8, 3.4, 691.0, 4.2, 825.0, 998.0, 338.0, 2.~
## $ Crossing <dbl> 65, 79, 20, 55, 65, 41, 42, 28, 62, 72, 58, 61, 44, ~
## $ 'Heading Accuracy' <dbl> 48, 69, 13, 33, 50, 35, 59, 49, 59, 66, 55, 21, 55, ~
## $ Volleys <dbl> 69, 69, 19, 39, 65, 42, 37, 28, 54, 31, 28, 31, 55, ~
## $ Curve <dbl> 76, 78, 21, 48, 60, 43, 46, 23, 80, 31, 40, 56, 40, ~
## $ 'FK Accuracy' <dbl> 69, 84, 20, 39, 65, 35, 38, 21, 69, 34, 26, 32, 42, ~
## $ 'Long Passing' <dbl> 75, 81, 38, 47, 65, 53, 53, 27, 60, 48, 59, 32, 50, ~
## $ 'Ball Control' <dbl> 78, 80, 25, 56, 74, 52, 63, 31, 67, 62, 55, 62, 55, ~
## $ 'Sprint Speed' <dbl> 72, 75, 62, 66, 70, 73, 63, 55, 75, 67, 70, 90, 60, ~
## $ Agility <dbl> 77, 73, 69, 58, 75, 82, 66, 40, 66, 57, 63, 84, 48, ~
## $ Reactions <dbl> 74, 80, 77, 48, 65, 50, 60, 44, 61, 66, 60, 39, 55, ~
## $ Balance <dbl> 77, 71, 67, 60, 70, 83, 54, 59, 61, 58, 74, 77, 55, ~
## $ 'Shot Power' <dbl> 71, 85, 51, 56, 65, 48, 60, 32, 54, 45, 31, 38, 36, ~
## $ Jumping <dbl> 73, 68, 78, 48, 60, 51, 58, 74, 56, 67, 59, 38, 70, ~
## $ Stamina <dbl> 55, 87, 40, 54, 69, 53, 53, 55, 58, 69, 60, 31, 60, ~
## $ Strength <dbl> 67, 73, 54, 54, 51, 34, 58, 61, 59, 75, 56, 29, 62, ~

```

```
## $ 'Long Shots'      <dbl> 71, 80, 13, 37, 68, 51, 32, 16, 75, 29, 32, 44, 34,~
## $ Aggression        <dbl> 45, 74, 21, 32, 51, 42, 70, 45, 45, 73, 49, 26, 55,~
## $ Positioning       <dbl> 70, 76, 11, 48, 60, 52, 35, 24, 61, 49, 44, 56, 30,~
## $ Vision            <dbl> 77, 78, 60, 54, 73, 50, 40, 31, 66, 56, 41, 61, 40,~
## $ Penalties         <dbl> 69, 61, 12, 58, 70, 49, 41, 36, 68, 44, 37, 31, 52,~
## $ Composure         <dbl> 71, 77, 64, 46, 65, 52, 63, 44, 51, 67, 58, 62, 55,~
## $ 'Sliding Tackle'  <dbl> 43, 75, 13, 36, 42, 43, 64, 52, 34, 70, 55, 19, 58,~
## $ body_type         <fct> Lean, Normal, Lean, Lean, Lean, Normal, Lean, Lean,~
```

Vamos agora utilizar diferentes métodos de seleção de variáveis.

```
model <- lm(Potential ~ ., data = dados_m)
```

```
stepAIC(model, direction = "both", trace = -1)
```

```
## Start:  AIC=1928.48
## Potential ~ Age + Overall + Value + Wage + Height + Weight +
##   'Preferred foot' + 'Release clause' + Crossing + 'Heading Accuracy' +
##   Volleys + Curve + 'FK Accuracy' + 'Long Passing' + 'Ball Control' +
##   'Sprint Speed' + Agility + Reactions + Balance + 'Shot Power' +
##   Jumping + Stamina + Strength + 'Long Shots' + Aggression +
##   Positioning + Vision + Penalties + Composure + 'Sliding Tackle' +
##   body_type
##
##           Df Sum of Sq    RSS   AIC
## - Strength      1      0.1 6426.9 1926.5
## - 'Shot Power'   1      0.1 6426.9 1926.5
## - Reactions      1      0.4 6427.1 1926.5
## - 'Preferred foot' 1      0.4 6427.2 1926.5
## - 'Heading Accuracy' 1      0.5 6427.3 1926.6
## - Aggression     1      1.4 6428.2 1926.7
## - Height         1      1.4 6428.2 1926.7
## - 'FK Accuracy'  1      1.8 6428.6 1926.8
## - Positioning    1      3.4 6430.2 1927.0
## - Balance        1      3.9 6430.7 1927.1
## - 'Sliding Tackle' 1      4.0 6430.8 1927.1
## - Composure      1      4.8 6431.6 1927.2
## - Agility        1      6.4 6433.2 1927.5
## - Jumping        1      7.1 6433.9 1927.6
## - Curve          1      7.3 6434.1 1927.6
## - Weight         1      8.0 6434.8 1927.7
## - Penalties      1     10.0 6436.8 1928.0
## - Volleys        1     10.3 6437.1 1928.1
## - Vision         1     12.7 6439.5 1928.4
## <none>           6426.8 1928.5
## - 'Long Shots'   1     20.3 6447.1 1929.6
## - Crossing       1     21.5 6448.3 1929.8
## - 'Ball Control' 1     22.8 6449.6 1930.0
## - Value          1     23.0 6449.8 1930.0
## - body_type      3     49.9 6476.7 1930.2
## - Wage           1     27.0 6453.8 1930.7
## - 'Long Passing' 1     30.7 6457.5 1931.2
## - 'Sprint Speed' 1     59.5 6486.3 1935.7
```

```

## - 'Release clause'      1      66.3  6493.1 1936.7
## - Stamina               1     115.0  6541.8 1944.2
## - Overall               1    6603.0 13029.8 2633.2
## - Age                   1   10747.4 17174.2 2909.4
##
## Step: AIC=1926.49
## Potential ~ Age + Overall + Value + Wage + Height + Weight +
##   'Preferred foot' + 'Release clause' + Crossing + 'Heading Accuracy' +
##   Volleys + Curve + 'FK Accuracy' + 'Long Passing' + 'Ball Control' +
##   'Sprint Speed' + Agility + Reactions + Balance + 'Shot Power' +
##   Jumping + Stamina + 'Long Shots' + Aggression + Positioning +
##   Vision + Penalties + Composure + 'Sliding Tackle' + body_type
##
##              Df Sum of Sq      RSS      AIC
## - 'Shot Power'      1      0.1  6427.0 1924.5
## - Reactions          1      0.3  6427.2 1924.5
## - 'Preferred foot'   1      0.4  6427.3 1924.5
## - 'Heading Accuracy' 1      0.5  6427.4 1924.6
## - Aggression         1      1.3  6428.2 1924.7
## - Height             1      1.4  6428.2 1924.7
## - 'FK Accuracy'      1      1.8  6428.7 1924.8
## - Positioning        1      3.4  6430.3 1925.0
## - Balance            1      3.9  6430.8 1925.1
## - 'Sliding Tackle'   1      4.0  6430.9 1925.1
## - Composure          1      4.8  6431.6 1925.2
## - Agility            1      6.5  6433.3 1925.5
## - Jumping            1      7.2  6434.0 1925.6
## - Curve              1      7.4  6434.2 1925.6
## - Weight             1      8.8  6435.7 1925.9
## - Penalties          1     10.1  6437.0 1926.0
## - Volleys            1     10.4  6437.3 1926.1
## - Vision             1     12.6  6439.5 1926.5
## <none>                6426.9 1926.5
## - 'Long Shots'       1     20.3  6447.2 1927.6
## - Crossing           1     21.4  6448.3 1927.8
## - 'Ball Control'     1     22.8  6449.7 1928.0
## - Value              1     23.0  6449.9 1928.1
## - body_type          3     50.1  6476.9 1928.2
## + Strength           1      0.1  6426.8 1928.5
## - Wage               1     27.1  6454.0 1928.7
## - 'Long Passing'     1     30.7  6457.5 1929.2
## - 'Sprint Speed'     1     60.6  6487.4 1933.9
## - 'Release clause'   1     66.5  6493.4 1934.8
## - Stamina            1    122.1  6548.9 1943.3
## - Overall            1   6604.1 13031.0 2631.3
## - Age                1  11037.3 17464.1 2924.2
##
## Step: AIC=1924.5
## Potential ~ Age + Overall + Value + Wage + Height + Weight +
##   'Preferred foot' + 'Release clause' + Crossing + 'Heading Accuracy' +
##   Volleys + Curve + 'FK Accuracy' + 'Long Passing' + 'Ball Control' +
##   'Sprint Speed' + Agility + Reactions + Balance + Jumping +
##   Stamina + 'Long Shots' + Aggression + Positioning + Vision +
##   Penalties + Composure + 'Sliding Tackle' + body_type

```

```

##
##           Df Sum of Sq      RSS      AIC
## - Reactions      1         0.4  6427.3  1922.6
## - 'Preferred foot' 1         0.4  6427.4  1922.6
## - 'Heading Accuracy' 1         0.5  6427.4  1922.6
## - Aggression      1         1.4  6428.3  1922.7
## - Height           1         1.5  6428.4  1922.7
## - 'FK Accuracy'    1         1.9  6428.8  1922.8
## - Positioning      1         3.5  6430.5  1923.0
## - Balance          1         3.9  6430.9  1923.1
## - 'Sliding Tackle' 1         4.3  6431.3  1923.2
## - Composure        1         4.8  6431.7  1923.2
## - Agility          1         6.4  6433.3  1923.5
## - Jumping          1         7.4  6434.4  1923.7
## - Curve            1         7.6  6434.6  1923.7
## - Weight           1         8.8  6435.8  1923.9
## - Penalties        1        10.2  6437.2  1924.1
## - Volleys          1        11.1  6438.0  1924.2
## - Vision           1        12.5  6439.5  1924.5
## <none>                                6427.0  1924.5
## - Crossing         1        21.5  6448.4  1925.8
## - 'Long Shots'     1        22.3  6449.3  1926.0
## - 'Ball Control'   1        22.8  6449.7  1926.0
## - Value            1        23.0  6449.9  1926.1
## - body_type        3        50.2  6477.1  1926.3
## + 'Shot Power'     1         0.1  6426.9  1926.5
## + Strength         1         0.0  6426.9  1926.5
## - Wage            1        27.1  6454.0  1926.7
## - 'Long Passing'   1        31.0  6457.9  1927.3
## - 'Sprint Speed'   1        60.7  6487.7  1931.9
## - 'Release clause' 1        66.9  6493.9  1932.9
## - Stamina          1       122.8  6549.7  1941.4
## - Overall          1      6814.4 13241.3 2645.3
## - Age             1     11039.7 17466.7 2922.3
##
## Step:  AIC=1922.56
## Potential ~ Age + Overall + Value + Wage + Height + Weight +
##   'Preferred foot' + 'Release clause' + Crossing + 'Heading Accuracy' +
##   Volleys + Curve + 'FK Accuracy' + 'Long Passing' + 'Ball Control' +
##   'Sprint Speed' + Agility + Balance + Jumping + Stamina +
##   'Long Shots' + Aggression + Positioning + Vision + Penalties +
##   Composure + 'Sliding Tackle' + body_type
##
##           Df Sum of Sq      RSS      AIC
## - 'Preferred foot' 1         0.4  6427.7  1920.6
## - 'Heading Accuracy' 1         0.4  6427.8  1920.6
## - Aggression      1         1.4  6428.7  1920.8
## - Height           1         1.5  6428.8  1920.8
## - 'FK Accuracy'    1         1.8  6429.1  1920.8
## - Positioning      1         3.4  6430.7  1921.1
## - Balance          1         4.0  6431.4  1921.2
## - 'Sliding Tackle' 1         4.1  6431.4  1921.2
## - Composure        1         4.9  6432.2  1921.3
## - Agility          1         6.2  6433.5  1921.5

```

```

## - Curve          1          7.5  6434.8 1921.7
## - Jumping        1          7.8  6435.2 1921.8
## - Weight         1          8.7  6436.0 1921.9
## - Penalties      1         10.1  6437.4 1922.1
## - Volleys        1         11.3  6438.6 1922.3
## - Vision         1         12.3  6439.6 1922.5
## <none>           6427.3 1922.6
## - Crossing       1         21.9  6449.2 1924.0
## - 'Long Shots'   1         22.1  6449.4 1924.0
## - 'Ball Control' 1         22.7  6450.0 1924.1
## - Value          1         22.9  6450.2 1924.1
## - body_type      3         50.2  6477.6 1924.3
## + Reactions      1          0.4  6427.0 1924.5
## + 'Shot Power'   1          0.1  6427.2 1924.5
## + Strength       1          0.0  6427.3 1924.5
## - Wage           1         27.1  6454.4 1924.8
## - 'Long Passing' 1         30.7  6458.0 1925.3
## - 'Sprint Speed' 1         60.8  6488.1 1930.0
## - 'Release clause' 1        67.2  6494.6 1931.0
## - Stamina        1        122.4  6549.7 1939.4
## - Overall        1       10300.4 16727.8 2877.1
## - Age            1       11098.9 17526.2 2923.7
##
## Step:  AIC=1920.62
## Potential ~ Age + Overall + Value + Wage + Height + Weight +
##   'Release clause' + Crossing + 'Heading Accuracy' + Volleys +
##   Curve + 'FK Accuracy' + 'Long Passing' + 'Ball Control' +
##   'Sprint Speed' + Agility + Balance + Jumping + Stamina +
##   'Long Shots' + Aggression + Positioning + Vision + Penalties +
##   Composure + 'Sliding Tackle' + body_type
##
##           Df Sum of Sq    RSS    AIC
## - 'Heading Accuracy' 1         0.4  6428.1 1918.7
## - Aggression          1         1.3  6429.0 1918.8
## - Height              1         1.5  6429.3 1918.9
## - 'FK Accuracy'       1         1.7  6429.4 1918.9
## - Positioning         1         3.3  6431.0 1919.1
## - 'Sliding Tackle'    1         4.0  6431.7 1919.2
## - Balance             1         4.1  6431.8 1919.2
## - Composure           1         4.9  6432.6 1919.4
## - Agility             1         6.1  6433.8 1919.6
## - Curve              1         7.5  6435.2 1919.8
## - Jumping            1         7.7  6435.4 1919.8
## - Weight              1         8.6  6436.3 1920.0
## - Penalties          1        10.0  6437.7 1920.2
## - Volleys            1        11.1  6438.8 1920.3
## - Vision              1        12.4  6440.1 1920.5
## <none>                6427.7 1920.6
## - Crossing           1        21.5  6449.2 1922.0
## - 'Long Shots'       1        22.3  6450.0 1922.1
## - 'Ball Control'     1        22.9  6450.6 1922.2
## - Value              1        23.0  6450.7 1922.2
## - body_type          3        50.4  6478.1 1922.4
## + 'Preferred foot'   1          0.4  6427.3 1922.6

```

```

## + Reactions          1      0.3  6427.4 1922.6
## + 'Shot Power'       1      0.1  6427.6 1922.6
## + Strength           1      0.1  6427.7 1922.6
## - Wage               1     26.9  6454.6 1922.8
## - 'Long Passing'     1     30.4  6458.1 1923.3
## - 'Sprint Speed'     1     60.4  6488.2 1928.0
## - 'Release clause'   1     67.1  6494.9 1929.0
## - Stamina            1    122.6  6550.3 1937.5
## - Overall            1  10300.2 16727.9 2875.1
## - Age                1  11114.0 17541.7 2922.6
##
## Step:  AIC=1918.68
## Potential ~ Age + Overall + Value + Wage + Height + Weight +
##   'Release clause' + Crossing + Volleys + Curve + 'FK Accuracy' +
##   'Long Passing' + 'Ball Control' + 'Sprint Speed' + Agility +
##   Balance + Jumping + Stamina + 'Long Shots' + Aggression +
##   Positioning + Vision + Penalties + Composure + 'Sliding Tackle' +
##   body_type
##
##           Df Sum of Sq    RSS    AIC
## - Height          1      1.6  6429.7 1916.9
## - Aggression       1      1.7  6429.8 1917.0
## - 'FK Accuracy'    1      1.9  6430.0 1917.0
## - Positioning      1      3.1  6431.2 1917.2
## - 'Sliding Tackle' 1      3.6  6431.7 1917.2
## - Balance          1      3.7  6431.9 1917.3
## - Composure        1      4.9  6433.1 1917.5
## - Agility          1      5.7  6433.9 1917.6
## - Curve            1      7.4  6435.6 1917.8
## - Weight           1      8.5  6436.6 1918.0
## - Jumping          1     10.3  6438.5 1918.3
## - Penalties        1     11.9  6440.0 1918.5
## - Volleys          1     12.5  6440.6 1918.6
## <none>              6428.1 1918.7
## - Vision           1     14.7  6442.8 1919.0
## - Crossing          1     21.8  6450.0 1920.1
## - 'Long Shots'     1     22.3  6450.5 1920.2
## - Value            1     22.7  6450.8 1920.2
## - body_type        3     50.3  6478.5 1920.5
## + 'Heading Accuracy' 1      0.4  6427.7 1920.6
## + 'Preferred foot'   1      0.4  6427.8 1920.6
## + Reactions         1      0.3  6427.8 1920.6
## + 'Shot Power'     1      0.1  6428.1 1920.7
## + Strength          1      0.0  6428.1 1920.7
## - Wage             1     27.0  6455.1 1920.9
## - 'Ball Control'    1     27.3  6455.5 1920.9
## - 'Long Passing'    1     30.2  6458.4 1921.4
## - 'Sprint Speed'    1     60.0  6488.2 1926.0
## - 'Release clause'  1     67.1  6495.3 1927.1
## - Stamina           1    122.3  6550.4 1935.5
## - Overall           1  10304.3 16732.4 2873.3
## - Age              1  11149.1 17577.3 2922.6
##
## Step:  AIC=1916.93

```

```

## Potential ~ Age + Overall + Value + Wage + Weight + 'Release clause' +
## Crossing + Volleys + Curve + 'FK Accuracy' + 'Long Passing' +
## 'Ball Control' + 'Sprint Speed' + Agility + Balance + Jumping +
## Stamina + 'Long Shots' + Aggression + Positioning + Vision +
## Penalties + Composure + 'Sliding Tackle' + body_type
##
##      Df Sum of Sq    RSS    AIC
## - Aggression      1      1.8  6431.5 1915.2
## - 'FK Accuracy'    1      2.2  6432.0 1915.3
## - Balance          1      2.4  6432.2 1915.3
## - Positioning      1      3.5  6433.3 1915.5
## - 'Sliding Tackle' 1      3.6  6433.3 1915.5
## - Agility          1      5.2  6434.9 1915.7
## - Composure        1      5.5  6435.2 1915.8
## - Weight           1      7.0  6436.7 1916.0
## - Curve            1      7.9  6437.7 1916.2
## - Jumping          1     11.6  6441.4 1916.7
## - Penalties        1     11.7  6441.5 1916.8
## - Volleys          1     12.5  6442.3 1916.9
## <none>              6429.7 1916.9
## - Vision           1     14.7  6444.4 1917.2
## - Crossing         1     21.6  6451.3 1918.3
## - 'Long Shots'     1     22.0  6451.8 1918.3
## - Value            1     22.6  6452.4 1918.5
## + Height           1      1.6  6428.1 1918.7
## + 'Heading Accuracy' 1      0.5  6429.3 1918.9
## + 'Preferred foot'  1      0.4  6429.3 1918.9
## + Reactions        1      0.3  6429.4 1918.9
## + 'Shot Power'     1      0.1  6429.6 1918.9
## - Wage             1     25.7  6455.4 1918.9
## + Strength         1      0.0  6429.7 1918.9
## - body_type        3     52.2  6481.9 1919.0
## - 'Long Passing'   1     28.8  6458.6 1919.4
## - 'Ball Control'   1     30.0  6459.8 1919.6
## - 'Sprint Speed'   1     61.9  6491.6 1924.5
## - 'Release clause' 1     67.2  6497.0 1925.3
## - Stamina          1    120.8  6550.6 1933.5
## - Overall          1  10303.8 16733.5 2871.4
## - Age              1  11347.0 17776.8 2931.9
##
## Step:  AIC=1915.21
## Potential ~ Age + Overall + Value + Wage + Weight + 'Release clause' +
## Crossing + Volleys + Curve + 'FK Accuracy' + 'Long Passing' +
## 'Ball Control' + 'Sprint Speed' + Agility + Balance + Jumping +
## Stamina + 'Long Shots' + Positioning + Vision + Penalties +
## Composure + 'Sliding Tackle' + body_type
##
##      Df Sum of Sq    RSS    AIC
## - 'Sliding Tackle' 1      2.0  6433.5 1913.5
## - Balance          1      2.1  6433.6 1913.5
## - 'FK Accuracy'    1      2.3  6433.8 1913.6
## - Positioning      1      2.9  6434.4 1913.7
## - Agility          1      5.2  6436.7 1914.0
## - Composure        1      6.6  6438.1 1914.2

```

```

## - Weight          1          6.9  6438.4 1914.3
## - Curve           1          8.0  6439.5 1914.5
## - Penalties       1         11.5  6443.0 1915.0
## <none>                                6431.5 1915.2
## - Volleys         1         12.9  6444.4 1915.2
## - Jumping          1         13.2  6444.7 1915.3
## - Vision           1         16.1  6447.6 1915.7
## - 'Long Shots'    1         21.8  6453.3 1916.6
## - Crossing         1         23.0  6454.5 1916.8
## - Value            1         23.4  6454.9 1916.8
## + Aggression       1          1.8  6429.7 1916.9
## + Height           1          1.7  6429.8 1917.0
## - Wage             1         24.5  6456.0 1917.0
## + 'Heading Accuracy' 1          0.9  6430.6 1917.1
## + Reactions         1          0.4  6431.1 1917.2
## + 'Preferred foot'  1          0.3  6431.2 1917.2
## + 'Shot Power'     1          0.2  6431.3 1917.2
## + Strength         1          0.1  6431.4 1917.2
## - body_type        3         51.8  6483.3 1917.2
## - 'Long Passing'   1         30.2  6461.7 1917.9
## - 'Ball Control'   1         31.7  6463.2 1918.1
## - 'Sprint Speed'   1         63.9  6495.4 1923.1
## - 'Release clause' 1         66.0  6497.5 1923.4
## - Stamina          1        119.1  6550.6 1931.6
## - Overall           1       10303.7 16735.2 2869.5
## - Age              1       11392.9 17824.4 2932.6
##
## Step:  AIC=1913.52
## Potential ~ Age + Overall + Value + Wage + Weight + 'Release clause' +
## Crossing + Volleys + Curve + 'FK Accuracy' + 'Long Passing' +
## 'Ball Control' + 'Sprint Speed' + Agility + Balance + Jumping +
## Stamina + 'Long Shots' + Positioning + Vision + Penalties +
## Composure + body_type
##
##           Df Sum of Sq    RSS    AIC
## - Balance      1      2.1  6435.7 1911.8
## - Positioning   1      2.5  6436.1 1911.9
## - 'FK Accuracy' 1      2.7  6436.2 1911.9
## - Composure     1      5.6  6439.1 1912.4
## - Agility       1      6.6  6440.2 1912.5
## - Weight        1      6.7  6440.2 1912.6
## - Curve         1      8.1  6441.6 1912.8
## - Jumping       1     12.2  6445.7 1913.4
## <none>                                6433.5 1913.5
## - Penalties     1     14.0  6447.5 1913.7
## - Vision         1     14.3  6447.8 1913.7
## - Volleys       1     15.2  6448.7 1913.9
## - 'Long Shots'  1     20.9  6454.4 1914.8
## + 'Sliding Tackle' 1      2.0  6431.5 1915.2
## + Height        1      1.6  6431.9 1915.3
## - Value         1     24.3  6457.8 1915.3
## - Wage          1     24.5  6458.0 1915.3
## + 'Shot Power'  1      0.4  6433.1 1915.5
## + 'Preferred foot' 1      0.3  6433.3 1915.5

```



```

## + Reactions          1      0.2  6433.3 1915.5
## + Aggression          1      0.2  6433.3 1915.5
## + 'Heading Accuracy'  1      0.2  6433.4 1915.5
## + Strength            1      0.1  6433.4 1915.5
## - Crossing            1     27.1  6460.6 1915.7
## - body_type           3     53.2  6486.7 1915.8
## - 'Long Passing'      1     28.9  6462.5 1916.0
## - 'Ball Control'      1     29.7  6463.2 1916.1
## - 'Sprint Speed'      1     63.8  6497.4 1921.4
## - 'Release clause'    1     66.8  6500.3 1921.8
## - Stamina             1    155.4  6588.9 1935.4
## - Overall             1   10349.1 16782.6 2870.3
## - Age                 1   11390.9 17824.5 2930.6
##
## Step:  AIC=1911.85
## Potential ~ Age + Overall + Value + Wage + Weight + 'Release clause' +
## Crossing + Volleys + Curve + 'FK Accuracy' + 'Long Passing' +
## 'Ball Control' + 'Sprint Speed' + Agility + Jumping + Stamina +
## 'Long Shots' + Positioning + Vision + Penalties + Composure +
## body_type
##
##              Df Sum of Sq      RSS      AIC
## - 'FK Accuracy'      1      2.6  6438.3 1910.3
## - Positioning         1      2.6  6438.3 1910.3
## - Composure           1      6.0  6441.6 1910.8
## - Curve               1      7.9  6443.6 1911.1
## - Agility             1     10.6  6446.2 1911.5
## - Weight              1     12.2  6447.9 1911.8
## <none>                6435.7 1911.8
## - Penalties          1     13.7  6449.3 1912.0
## - Vision              1     13.9  6449.5 1912.0
## - Jumping             1     14.8  6450.4 1912.2
## - Volleys             1     14.9  6450.6 1912.2
## - 'Long Shots'       1     20.9  6456.6 1913.1
## + Balance             1      2.1  6433.5 1913.5
## + 'Sliding Tackle'    1      2.1  6433.6 1913.5
## - Wage                1     24.5  6460.1 1913.7
## - Value               1     24.6  6460.3 1913.7
## + 'Shot Power'        1      0.4  6435.2 1913.8
## + Height              1      0.3  6435.3 1913.8
## + 'Preferred foot'    1      0.3  6435.4 1913.8
## + Reactions           1      0.3  6435.4 1913.8
## + Aggression          1      0.1  6435.5 1913.8
## + 'Heading Accuracy'  1      0.0  6435.6 1913.8
## + Strength            1      0.0  6435.7 1913.8
## - Crossing            1     26.0  6461.7 1913.9
## - body_type           3     52.9  6488.5 1914.0
## - 'Long Passing'      1     27.8  6463.5 1914.2
## - 'Ball Control'      1     34.3  6470.0 1915.2
## - 'Sprint Speed'      1     64.1  6499.7 1919.8
## - 'Release clause'    1     66.8  6502.5 1920.2
## - Stamina             1    153.8  6589.5 1933.5
## - Overall             1   10583.4 17019.0 2882.3
## - Age                 1   11388.8 17824.5 2928.6

```

```

##
## Step: AIC=1910.26
## Potential ~ Age + Overall + Value + Wage + Weight + 'Release clause' +
## Crossing + Volleys + Curve + 'Long Passing' + 'Ball Control' +
## 'Sprint Speed' + Agility + Jumping + Stamina + 'Long Shots' +
## Positioning + Vision + Penalties + Composure + body_type
##
##           Df Sum of Sq    RSS    AIC
## - Positioning      1      2.0  6440.2 1908.6
## - Composure        1      5.9  6444.1 1909.2
## - Curve            1      5.9  6444.2 1909.2
## - Agility          1     10.6  6448.8 1909.9
## - Weight           1     11.4  6449.7 1910.0
## - Penalties        1     11.7  6450.0 1910.1
## <none>                                6438.3 1910.3
## - Vision           1     14.5  6452.7 1910.5
## - Volleys          1     14.6  6452.9 1910.5
## - Jumping          1     15.5  6453.8 1910.7
## + 'FK Accuracy'    1      2.6  6435.7 1911.8
## + 'Sliding Tackle' 1      2.5  6435.8 1911.9
## - Wage             1     23.7  6462.0 1911.9
## + Balance          1      2.1  6436.2 1911.9
## - Value            1     25.2  6463.4 1912.2
## + 'Shot Power'     1      0.6  6437.7 1912.2
## + Height           1      0.5  6437.7 1912.2
## + Reactions        1      0.2  6438.1 1912.2
## + 'Preferred foot' 1      0.1  6438.1 1912.2
## + Aggression       1      0.1  6438.2 1912.2
## + 'Heading Accuracy' 1      0.0  6438.2 1912.2
## + Strength         1      0.0  6438.3 1912.3
## - 'Long Passing'   1     26.4  6464.7 1912.4
## - body_type        3     53.6  6491.9 1912.5
## - 'Long Shots'     1     27.9  6466.1 1912.6
## - Crossing         1     28.8  6467.1 1912.7
## - 'Ball Control'   1     35.8  6474.1 1913.8
## - 'Sprint Speed'   1     63.5  6501.8 1918.1
## - 'Release clause' 1     65.5  6503.8 1918.4
## - Stamina          1    153.3  6591.5 1931.8
## - Overall          1   10635.0 17073.2 2883.5
## - Age              1   11458.9 17897.1 2930.6
##
## Step: AIC=1908.56
## Potential ~ Age + Overall + Value + Wage + Weight + 'Release clause' +
## Crossing + Volleys + Curve + 'Long Passing' + 'Ball Control' +
## 'Sprint Speed' + Agility + Jumping + Stamina + 'Long Shots' +
## Vision + Penalties + Composure + body_type
##
##           Df Sum of Sq    RSS    AIC
## - Composure        1      5.5  6445.8 1907.4
## - Curve            1      6.1  6446.3 1907.5
## - Agility          1      9.6  6449.9 1908.1
## - Penalties        1     10.7  6450.9 1908.2
## - Weight           1     11.8  6452.0 1908.4
## <none>                                6440.2 1908.6

```

```

## - Volleys          1      13.5  6453.7 1908.7
## - Jumping          1      16.4  6456.7 1909.1
## - Vision           1      17.3  6457.5 1909.2
## - Wage             1      23.4  6463.6 1910.2
## + Balance          1       2.2  6438.0 1910.2
## + 'Sliding Tackle' 1       2.0  6438.2 1910.2
## + Positioning      1       2.0  6438.3 1910.3
## + 'FK Accuracy'    1       1.9  6438.3 1910.3
## - Value            1      24.6  6464.9 1910.4
## + 'Shot Power'     1       0.7  6439.5 1910.5
## + Height           1       0.7  6439.6 1910.5
## + 'Preferred foot' 1       0.2  6440.1 1910.5
## + Reactions        1       0.1  6440.1 1910.5
## + Aggression       1       0.0  6440.2 1910.6
## + Strength         1       0.0  6440.2 1910.6
## + 'Heading Accuracy' 1       0.0  6440.2 1910.6
## - body_type        3      54.1  6494.3 1910.9
## - 'Long Passing'    1      30.2  6470.5 1911.2
## - 'Ball Control'    1      34.0  6474.2 1911.8
## - Crossing         1      34.2  6474.4 1911.9
## - 'Long Shots'     1      36.9  6477.2 1912.3
## - 'Release clause' 1      65.4  6505.6 1916.7
## - 'Sprint Speed'   1      67.1  6507.3 1916.9
## - Stamina          1     165.5  6605.7 1931.9
## - Overall          1    10863.0 17303.2 2894.9
## - Age              1    11457.7 17897.9 2928.7
##
## Step:  AIC=1907.42
## Potential ~ Age + Overall + Value + Wage + Weight + 'Release clause' +
##      Crossing + Volleys + Curve + 'Long Passing' + 'Ball Control' +
##      'Sprint Speed' + Agility + Jumping + Stamina + 'Long Shots' +
##      Vision + Penalties + body_type
##
##
##      Df Sum of Sq    RSS    AIC
## - Curve          1      7.8  6453.5 1906.6
## - Agility         1     10.7  6456.4 1907.1
## - Weight          1     10.8  6456.6 1907.1
## - Penalties       1     11.7  6457.5 1907.2
## <none>                        6445.8 1907.4
## - Volleys        1     14.6  6460.3 1907.7
## - Vision          1     16.5  6462.2 1908.0
## - Jumping         1     17.1  6462.8 1908.1
## + Composure       1      5.5  6440.2 1908.6
## - Wage            1     22.2  6468.0 1908.9
## + Balance         1      2.6  6443.2 1909.0
## + 'FK Accuracy'   1      1.9  6443.9 1909.1
## + Positioning     1      1.6  6444.1 1909.2
## + 'Sliding Tackle' 1      1.0  6444.7 1909.3
## + Height          1      0.9  6444.8 1909.3
## + 'Shot Power'    1      0.7  6445.0 1909.3
## + Aggression      1      0.4  6445.3 1909.3
## + Reactions       1      0.2  6445.5 1909.4
## + 'Preferred foot' 1      0.1  6445.6 1909.4
## + 'Heading Accuracy' 1      0.1  6445.7 1909.4

```

```

## + Strength          1          0.0 6445.7 1909.4
## - Value             1          25.9 6471.7 1909.4
## - body_type         3          55.1 6500.8 1909.9
## - 'Long Passing'    1          36.2 6482.0 1911.0
## - Crossing          1          37.7 6483.5 1911.2
## - 'Long Shots'      1          38.6 6484.3 1911.4
## - 'Ball Control'    1          42.8 6488.5 1912.0
## - 'Release clause'  1          64.5 6510.2 1915.4
## - 'Sprint Speed'    1          69.5 6515.2 1916.1
## - Stamina           1         163.0 6608.7 1930.4
## - Age               1       11491.6 17937.3 2928.9
## - Overall           1       12558.5 19004.3 2986.7
##
## Step:  AIC=1906.63
## Potential ~ Age + Overall + Value + Wage + Weight + 'Release clause' +
## Crossing + Volleys + 'Long Passing' + 'Ball Control' + 'Sprint Speed' +
## Agility + Jumping + Stamina + 'Long Shots' + Vision + Penalties +
## body_type
##
##              Df Sum of Sq    RSS    AIC
## - Agility      1      11.3  6464.8 1906.4
## - Weight       1      11.4  6464.9 1906.4
## - Penalties    1      11.7  6465.2 1906.4
## <none>                     6453.5 1906.6
## - Vision       1      13.9  6467.4 1906.8
## - Jumping      1      17.7  6471.2 1907.4
## + Curve        1       7.8  6445.8 1907.4
## + Composure    1       7.2  6446.3 1907.5
## - Volleys      1      22.0  6475.6 1908.0
## - Wage         1      22.1  6475.6 1908.0
## + Balance      1       2.5  6451.0 1908.2
## + Positioning  1       1.8  6451.8 1908.3
## + Height       1       1.2  6452.3 1908.4
## + 'Shot Power'  1       1.1  6452.4 1908.5
## + 'Sliding Tackle' 1       0.8  6452.7 1908.5
## + Aggression   1       0.6  6452.9 1908.5
## + 'Preferred foot' 1       0.2  6453.3 1908.6
## + Reactions    1       0.2  6453.3 1908.6
## + 'FK Accuracy' 1       0.2  6453.3 1908.6
## + 'Heading Accuracy' 1       0.0  6453.5 1908.6
## + Strength     1       0.0  6453.5 1908.6
## - Value        1      28.6  6482.2 1909.0
## - body_type    3      55.7  6509.3 1909.2
## - Crossing     1      30.3  6483.8 1909.3
## - 'Long Shots' 1      33.3  6486.9 1909.8
## - 'Long Passing' 1      40.0  6493.5 1910.8
## - 'Ball Control' 1      43.2  6496.7 1911.3
## - 'Release clause' 1      64.1  6517.7 1914.5
## - 'Sprint Speed' 1      68.7  6522.2 1915.2
## - Stamina      1     168.8  6622.4 1930.5
## - Age          1    11543.0 17996.5 2930.2
## - Overall      1    12596.6 19050.1 2987.1
##
## Step:  AIC=1906.38

```

```

## Potential ~ Age + Overall + Value + Wage + Weight + 'Release clause' +
## Crossing + Volleys + 'Long Passing' + 'Ball Control' + 'Sprint Speed' +
## Jumping + Stamina + 'Long Shots' + Vision + Penalties + body_type
##
##           Df Sum of Sq    RSS    AIC
## - Penalties      1      9.4  6474.2 1905.8
## - Vision          1     10.3  6475.1 1906.0
## <none>                        6464.8 1906.4
## + Agility         1     11.3  6453.5 1906.6
## + Composure       1      8.4  6456.4 1907.1
## + Curve           1      8.4  6456.4 1907.1
## - Jumping         1     18.3  6483.2 1907.2
## + Balance         1      6.9  6457.9 1907.3
## - Wage            1     19.5  6484.3 1907.4
## - Weight          1     23.4  6488.2 1908.0
## + 'Sliding Tackle' 1      2.3  6462.5 1908.0
## - Volleys         1     23.7  6488.5 1908.0
## + Positioning     1      0.8  6464.0 1908.2
## + 'Shot Power'    1      0.7  6464.2 1908.3
## + 'Heading Accuracy' 1      0.4  6464.4 1908.3
## + Aggression      1      0.2  6464.6 1908.3
## + 'FK Accuracy'   1      0.2  6464.6 1908.3
## + Strength        1      0.1  6464.7 1908.4
## + Reactions       1      0.1  6464.7 1908.4
## + Height          1      0.1  6464.8 1908.4
## + 'Preferred foot' 1      0.1  6464.8 1908.4
## - Value           1     28.0  6492.8 1908.7
## - Crossing         1     28.3  6493.1 1908.7
## - 'Long Shots'    1     30.1  6494.9 1909.0
## - body_type       3     57.8  6522.6 1909.3
## - 'Long Passing'  1     35.8  6500.7 1909.9
## - 'Ball Control'  1     50.5  6515.3 1912.2
## - 'Sprint Speed'  1     57.4  6522.2 1913.2
## - 'Release clause' 1     61.3  6526.1 1913.8
## - Stamina         1    167.6  6632.4 1930.0
## - Age             1   11548.6 18013.4 2929.1
## - Overall         1   12622.9 19087.8 2987.1
##
## Step:  AIC=1905.83
## Potential ~ Age + Overall + Value + Wage + Weight + 'Release clause' +
## Crossing + Volleys + 'Long Passing' + 'Ball Control' + 'Sprint Speed' +
## Jumping + Stamina + 'Long Shots' + Vision + body_type
##
##           Df Sum of Sq    RSS    AIC
## - Vision          1     10.4  6484.7 1905.4
## <none>                        6474.2 1905.8
## + Composure       1      9.4  6464.8 1906.4
## + Penalties       1      9.4  6464.8 1906.4
## + Agility         1      9.0  6465.2 1906.4
## + Curve           1      8.3  6465.9 1906.5
## - Jumping         1     19.4  6493.6 1906.8
## + Balance         1      5.9  6468.3 1906.9
## - Wage            1     21.9  6496.1 1907.2
## + 'Sliding Tackle' 1      3.9  6470.3 1907.2

```

```

## - Weight          1      23.3  6497.5 1907.4
## - 'Long Shots'    1      24.1  6498.3 1907.5
## + 'Shot Power'    1       0.9  6473.3 1907.7
## - Value           1      25.5  6499.7 1907.8
## + Positioning     1       0.3  6473.9 1907.8
## + Strength        1       0.2  6474.0 1907.8
## + 'Preferred foot' 1       0.1  6474.2 1907.8
## + Height          1       0.1  6474.2 1907.8
## + Aggression      1       0.0  6474.2 1907.8
## + 'FK Accuracy'   1       0.0  6474.2 1907.8
## + Reactions       1       0.0  6474.2 1907.8
## + 'Heading Accuracy' 1       0.0  6474.2 1907.8
## - Crossing        1      29.3  6503.6 1908.3
## - 'Long Passing'   1      33.2  6507.4 1908.9
## - body_type       3      60.2  6534.4 1909.1
## - Volleys         1      47.7  6521.9 1911.2
## - 'Sprint Speed'   1      58.1  6532.4 1912.8
## - 'Release clause' 1      61.5  6535.7 1913.3
## - 'Ball Control'   1      63.0  6537.2 1913.5
## - Stamina         1     167.2  6641.4 1929.3
## - Age             1    11545.5 18019.7 2927.5
## - Overall         1    12624.8 19099.0 2985.6
##
## Step:  AIC=1905.44
## Potential ~ Age + Overall + Value + Wage + Weight + 'Release clause' +
##           Crossing + Volleys + 'Long Passing' + 'Ball Control' + 'Sprint Speed' +
##           Jumping + Stamina + 'Long Shots' + body_type
##
##           Df Sum of Sq    RSS    AIC
## <none>                    6484.7 1905.4
## + Vision                 1     10.4  6474.2 1905.8
## + Penalties              1      9.5  6475.1 1906.0
## + Composure              1      8.0  6476.7 1906.2
## + Agility                1      5.9  6478.8 1906.5
## + Curve                  1      5.8  6478.9 1906.5
## - Weight                 1     20.6  6505.3 1906.6
## + Balance                1      4.6  6480.1 1906.7
## - Jumping                1     21.7  6506.4 1906.8
## - Wage                   1     22.1  6506.7 1906.8
## - Value                  1     23.8  6508.5 1907.1
## - 'Long Passing'         1     24.4  6509.1 1907.2
## + Positioning            1      1.5  6483.1 1907.2
## + 'Heading Accuracy'     1      1.3  6483.4 1907.2
## + 'Sliding Tackle'       1      1.1  6483.6 1907.3
## + Aggression             1      0.8  6483.9 1907.3
## + 'Shot Power'          1      0.3  6484.4 1907.4
## + Height                 1      0.2  6484.5 1907.4
## + 'Preferred foot'       1      0.1  6484.6 1907.4
## + Strength               1      0.0  6484.6 1907.4
## + Reactions              1      0.0  6484.6 1907.4
## + 'FK Accuracy'         1      0.0  6484.7 1907.4
## - body_type              3     56.2  6540.9 1908.1
## - Crossing               1     33.0  6517.7 1908.5
## - Volleys                1     45.3  6530.0 1910.4

```

```

## - 'Long Shots'      1      45.5  6530.1 1910.4
## - 'Sprint Speed'    1      58.2  6542.9 1912.4
## - 'Release clause'  1      60.7  6545.4 1912.8
## - 'Ball Control'    1      69.0  6553.6 1914.0
## - Stamina           1     157.6  6642.2 1927.5
## - Age               1    11539.2 18023.9 2925.7
## - Overall           1    13635.4 20120.1 3035.7

##
## Call:
## lm(formula = Potential ~ Age + Overall + Value + Wage + Weight +
##     'Release clause' + Crossing + Volleys + 'Long Passing' +
##     'Ball Control' + 'Sprint Speed' + Jumping + Stamina + 'Long Shots' +
##     body_type, data = dados_m)
##
## Coefficients:
##      (Intercept)           Age           Overall           Value
##      41.4776392      -0.9793704       0.8747629      -0.0005714
##           Wage           Weight 'Release clause'           Crossing
##      0.0006004      -0.0267471      -0.0009362      -0.0217757
##          Volleys 'Long Passing' 'Ball Control' 'Sprint Speed'
##      0.0269119       0.0185471       0.0374856      -0.0250990
##          Jumping           Stamina 'Long Shots' body_typeNormal
##      0.0145216      -0.0407000      -0.0267433      -0.3196466
## body_typeStocky body_typeUnique
##      0.0179388       2.7108693

model2 <- lm(formula = Potential ~ Age + Overall + Value + Wage + Weight +
  `Release clause` + Crossing + Volleys + `Long Passing` +
  `Ball Control` + `Sprint Speed` + Jumping + Stamina + `Long Shots` +
  body_type, data = dados_m)

# Distancia de Cook

dcookAnalysis <- function(model, y_values) {
  cook_vec <- cooks.distance(model)
  dfcook <- data.frame(
    Index = seq(1, length(cook_vec)),
    Cook = cook_vec
  )
  corte <- c(4/nrow(dfcook))
  chart <- ggplot(dfcook, aes(x = Index, y = Cook)) +
    geom_point() +
    geom_segment(aes(x = Index, xend = Index, y = 0, yend = Cook)) +
    geom_hline(yintercept = corte, color = "red") +
    ggtitle("Cook's Distance")
  dfcook1 <- dfcook %>% filter(Cook > corte)
  out <- dfcook1 %>% mutate(y = y_values[dfcook1$Index]) %>% arrange(-Cook)
  return(list(threshold = corte, chartCook = chart, table = out))
}

dffitAnalysis <- function(model, y_values) {
  dffit_vec <- dffits(model)

```

```

dfdffit <- data.frame(
  Index = seq(1, length(dffit_vec)),
  Dffits = abs(dffit_vec)
)
p <- length(model$coefficients)
n <- nrow(dfdffit)
corte <- 2 * sqrt(p/n)
chart <- ggplot(dfdffit, aes(x = Index, y = Dffits)) +
  geom_point() +
  geom_segment(aes(x = Index, xend = Index, y = 0, yend = Dffits)) +
  geom_hline(yintercept = corte, color = "red") +
  ggtitle("Dffits") + ylab("| Dffits |")
dfdffit1 <- dfdffit %>% filter(Dffits > corte)
out <- dfdffit1 %>% mutate(y = y_values[dfdffit1$Index]) %>% arrange(-Dffits)
return(list(threshold = corte, chartDffits = chart, table = out))
}

```