10-11

Diya Patel

2024-06-14

```r
# Load the dataset
dataset_path <- "/Users/diya/Downloads/Football_dataset.xlsx"
football_data <- read_excel(dataset_path)

# Display the first few rows of the dataset
head(football_data)
```

```
## # A tibble: 6 x 13
##   appearance_id  game_id player_id player_club_id player_current_club_id
##   <chr>            <dbl>     <dbl>          <dbl>                  <dbl>
## 1 2231978_38004  2231978     38004            853                    235
## 2 2233748_79232  2233748     79232           8841                   2698
## 3 2234413_42792  2234413     42792           6251                    465
## 4 2234418_73333  2234418     73333           1274                   6646
## 5 2234421_122011 2234421    122011            195                   3008
## 6 2234421_146889 2234421    146889            195                    190
## # i 8 more variables: date <dttm>, player_name <chr>, competition_id <chr>,
## #   yellow_cards <dbl>, red_cards <dbl>, goals <dbl>, assists <dbl>,
## #   minutes_played <dbl>
```

```r
# 10. Display current column names
print('Current Columns:')
```

```
## [1] "Current Columns:"
```

```r
colnames(football_data)
```

```
##  [1] "appearance_id"          "game_id"                "player_id"
##  [4] "player_club_id"         "player_current_club_id" "date"
##  [7] "player_name"            "competition_id"         "yellow_cards"
## [10] "red_cards"              "goals"                  "assists"
## [13] "minutes_played"
```

```r
# Rename some of the columns
football_data <- football_data %>%
  rename(
    player_full_name = player_name,
    goals_scored = goals,
    minutes = minutes_played
  )
```

```r
# Display new column names
print('Updated Columns:')
```

```
## [1] "Updated Columns:"
```

```r
colnames(football_data)
```

```
##  [1] "appearance_id"        "game_id"                "player_id"
##  [4] "player_club_id"       "player_current_club_id" "date"
##  [7] "player_full_name"     "competition_id"         "yellow_cards"
## [10] "red_cards"            "goals_scored"           "assists"
## [13] "minutes"
```

```r
# 11. Add new variable by multiplying an existing column by 2
football_data <- football_data %>%
  mutate(double_goals = goals_scored * 2)

# Display the first few rows of the updated dataset
head(football_data)
```

```
## # A tibble: 6 x 14
##   appearance_id  game_id player_id player_club_id player_current_club_id
##   <chr>            <dbl>     <dbl>          <dbl>                  <dbl>
## 1 2231978_38004  2231978     38004            853                    235
## 2 2233748_79232  2233748     79232           8841                   2698
## 3 2234413_42792  2234413     42792           6251                    465
## 4 2234418_73333  2234418     73333           1274                   6646
## 5 2234421_122011 2234421    122011            195                   3008
## 6 2234421_146889 2234421    146889            195                    190
## # i 9 more variables: date <dttm>, player_full_name <chr>,
## #   competition_id <chr>, yellow_cards <dbl>, red_cards <dbl>,
## #   goals_scored <dbl>, assists <dbl>, minutes <dbl>, double_goals <dbl>
```