

# TD1

2023-02-03

## 1. Importer le jeu de données dans R

Le bloc de code suivant permet de charger dans R un fichier au format Excel.

```
library(readxl)
fifa22 <- read_excel("fifa22.xlsx")
```

Nous venons d'importer la base de données dans R en lui donnant le nom fifa22. C'est avec ce nom que nous nous référons à la base de données plus tard dans le reste du code.

## 2. Afficher les noms des colonnes du jeu de données actif

On utilise la fonction `colnames(fifa22)` pour afficher les noms des colonnes qui sont dans la base de données. Cela nous permet de vérifier que la base de donnée a été importée correctement.

```
colnames(fifa22)
```

```
## [1] "short_name"           "player_positions"
## [3] "overall"              "potential"
## [5] "value_eur"            "wage_eur"
## [7] "age"                  "dob"
## [9] "height_cm"            "weight_kg"
## [11] "club_team_id"         "club_name"
## [13] "league_name"          "league_level"
## [15] "club_position"        "club_jersey_number"
## [17] "club_loaned_from"     "club_joined"
## [19] "club_contract_valid_until" "nationality_id"
## [21] "nationality_name"     "nation_team_id"
## [23] "nation_position"      "nation_jersey_number"
## [25] "preferred_foot"       "weak_foot"
## [27] "skill_moves"          "international_reputation"
## [29] "work_rate"            "body_type"
## [31] "real_face"            "release_clause_eur"
## [33] "player_traits"        "pace"
## [35] "shooting"             "passing"
## [37] "dribbling"            "defending"
## [39] "physic"               "attacking_crossing"
## [41] "attacking_finishing"  "attacking_heading_accuracy"
## [43] "attacking_short_passing" "attacking_volleys"
## [45] "skill_dribbling"      "skill_curve"
## [47] "skill_fk_accuracy"    "skill_long_passing"
## [49] "skill_ball_control"   "movement_acceleration"
## [51] "movement_sprint_speed" "movement_agility"
## [53] "movement_reactions"   "movement_balance"
## [55] "power_shot_power"     "power_jumping"
## [57] "power_stamina"        "power_strength"
## [59] "power_long_shots"     "mentality_aggression"
## [61] "mentality_interceptions" "mentality_positioning"
## [63] "mentality_vision"     "mentality_penalties"
## [65] "mentality_composure"  "defending_marking_awareness"
## [67] "defending_standing_tackle" "defending_sliding_tackle"
## [69] "goalkeeping_diving"    "goalkeeping_handling"
## [71] "goalkeeping_kicking"   "goalkeeping_positioning"
## [73] "goalkeeping_reflexes"  "goalkeeping_speed"
```

## 3. Afficher les types de données de chaque colonne

Nous allons utiliser la fonction `str(fifa22)` pour afficher la structure de la base de donnée.

Cette fonction affiche aussi les types de données pour chaque type de colonne. Par exemple `chr` pour des données texte, `num` pour des données numériques. Il y a des opérations qui ne sont permises que pour certains types de données.

```
str(fifa22)
```

```

## tibble [19,239 × 74] (S3: tbl_df/tbl/data.frame)
## $ short_name          : chr [1:19239] "L, Messi" "R, Lewandowski" "Cristiano Ronaldo" "Neymar Jr" ...
## $ player_positions    : chr [1:19239] "RW, ST, CF" "ST" "ST, LW" "LW, CAM" ...
## $ overall             : num [1:19239] 93 92 91 91 91 91 91 90 90 90 ...
## $ potential           : num [1:19239] 93 92 91 91 91 93 95 90 92 90 ...
## $ value_eur           : num [1:19239] 7.80e+07 1.20e+08 4.50e+07 1.29e+08 1.26e+08 ...
## $ wage_eur           : num [1:19239] 320000 270000 270000 270000 350000 130000 230000 86000 250000 240000
## ...
## $ age                 : num [1:19239] 34 32 36 29 30 28 22 35 29 27 ...
## $ dob                 : POSIXct[1:19239], format: "1987-06-24" "1988-08-21" ...
## $ height_cm           : num [1:19239] 170 185 187 175 181 188 182 193 187 188 ...
## $ weight_kg           : num [1:19239] 72 81 83 68 70 87 73 93 85 89 ...
## $ club_team_id        : num [1:19239] 73 21 11 73 10 240 73 21 241 18 ...
## $ club_name           : chr [1:19239] "Paris Saint-Germain" "FC Bayern München" "Manchester United" "Paris
Saint-Germain" ...
## $ league_name         : chr [1:19239] "French Ligue 1" "German 1, Bundesliga" "English Premier League" "Fre
nch Ligue 1" ...
## $ league_level        : num [1:19239] 1 1 1 1 1 1 1 1 1 1 ...
## $ club_position       : chr [1:19239] "RW" "ST" "ST" "LW" ...
## $ club_jersey_number  : num [1:19239] 30 9 7 10 17 13 7 1 1 10 ...
## $ club_loaned_from    : chr [1:19239] NA NA NA NA ...
## $ club_joined         : POSIXct[1:19239], format: "2021-08-10" "2014-07-01" ...
## $ club_contract_valid_until : num [1:19239] 2023 2023 2023 2025 2025 ...
## $ nationality_id       : num [1:19239] 52 37 38 54 7 44 18 21 21 14 ...
## $ nationality_name     : chr [1:19239] "Argentina" "Poland" "Portugal" "Brazil" ...
## $ nation_team_id      : num [1:19239] 1369 1353 1354 NA 1325 ...
## $ nation_position     : chr [1:19239] "RW" "RS" "ST" NA ...
## $ nation_jersey_number : num [1:19239] 10 9 7 NA 7 NA 10 1 NA 9 ...
## $ preferred_foot      : chr [1:19239] "Left" "Right" "Right" "Right" ...
## $ weak_foot           : num [1:19239] 4 4 4 5 5 3 4 4 4 5 ...
## $ skill_moves         : num [1:19239] 4 4 5 5 4 1 5 1 1 3 ...
## $ international_reputation : num [1:19239] 5 5 5 5 4 5 4 5 4 4 ...
## $ work_rate           : chr [1:19239] "Medium/Low" "High/Medium" "High/Low" "High/Medium" ...
## $ body_type           : chr [1:19239] "Unique" "Unique" "Unique" "Unique" ...
## $ real_face           : chr [1:19239] "Yes" "Yes" "Yes" "Yes" ...
## $ release_clause_eur   : num [1:19239] 1.44e+08 1.97e+08 8.33e+07 2.39e+08 2.32e+08 ...
## $ player_traits       : chr [1:19239] "Finesse Shot, Long Shot Taker (AI), Playmaker (AI), Outside Foot Sho
t, One Club Player, Chip Shot (AI), Technic" | __truncated__ "Solid Player, Finesse Shot, Outside Foot Shot, Chip Shot
(AI)" "Power Free-Kick, Flair, Long Shot Taker (AI), Speed Dribbler (AI), Outside Foot Shot" "Injury Prone, Flair, Spe
ed Dribbler (AI), Playmaker (AI), Outside Foot Shot, Technical Dribbler (AI)" ...
## $ pace               : num [1:19239] 85 78 87 91 76 NA 97 NA NA 70 ...
## $ shooting           : num [1:19239] 92 92 94 83 86 NA 88 NA NA 91 ...
## $ passing            : num [1:19239] 91 79 80 86 93 NA 80 NA NA 83 ...
## $ dribbling          : num [1:19239] 95 86 88 94 88 NA 92 NA NA 83 ...
## $ defending           : num [1:19239] 34 44 34 37 64 NA 36 NA NA 47 ...
## $ physic             : num [1:19239] 65 82 75 63 78 NA 77 NA NA 83 ...
## $ attacking_crossing  : num [1:19239] 85 71 87 85 94 13 78 15 18 80 ...
## $ attacking_finishing : num [1:19239] 95 95 95 83 82 11 93 13 14 94 ...
## $ attacking_heading_accuracy : num [1:19239] 70 90 90 63 55 15 72 25 11 86 ...
## $ attacking_short_passing : num [1:19239] 91 85 80 86 94 43 85 60 61 85 ...
## $ attacking_volleys   : num [1:19239] 88 89 86 86 82 13 83 11 14 88 ...
## $ skill_dribbling     : num [1:19239] 96 85 88 95 88 12 93 30 21 83 ...
## $ skill_curve         : num [1:19239] 93 79 81 88 85 13 80 14 18 83 ...
## $ skill_fk_accuracy   : num [1:19239] 94 85 84 87 83 14 69 11 12 65 ...
## $ skill_long_passing  : num [1:19239] 91 70 77 81 93 40 71 68 63 86 ...
## $ skill_ball_control  : num [1:19239] 96 88 88 95 91 30 91 46 30 85 ...
## $ movement_acceleration : num [1:19239] 91 77 85 93 76 43 97 54 38 65 ...
## $ movement_sprint_speed : num [1:19239] 80 79 88 89 76 60 97 60 50 74 ...
## $ movement_agility   : num [1:19239] 91 77 86 96 79 67 92 51 39 71 ...
## $ movement_reactions  : num [1:19239] 94 93 94 89 91 88 93 87 86 92 ...
## $ movement_balance   : num [1:19239] 95 82 74 84 78 49 83 35 43 70 ...
## $ power_shot_power    : num [1:19239] 86 90 94 80 91 59 86 68 66 91 ...
## $ power_jumping       : num [1:19239] 68 85 95 64 63 78 78 77 79 79 ...
## $ power_stamina       : num [1:19239] 72 76 77 81 89 41 88 43 35 83 ...
## $ power_strength      : num [1:19239] 69 86 77 53 74 78 77 80 78 85 ...
## $ power_long_shots    : num [1:19239] 94 87 93 81 91 12 82 16 10 86 ...
## $ mentality_aggression : num [1:19239] 44 81 63 63 76 34 62 29 43 80 ...
## $ mentality_interceptions : num [1:19239] 40 49 29 37 66 19 38 30 22 44 ...
## $ mentality_positioning : num [1:19239] 93 95 95 86 88 11 92 12 11 94 ...

```

```
## $ mentality_vision      : num [1:19239] 95 81 76 90 94 65 82 70 70 87 ...
## $ mentality_penalties   : num [1:19239] 75 90 88 93 83 11 79 47 25 91 ...
## $ mentality_composure   : num [1:19239] 96 88 95 93 89 68 88 70 70 91 ...
## $ defending_marking_awareness: num [1:19239] 20 35 24 35 68 27 26 17 25 50 ...
## $ defending_standing_tackle : num [1:19239] 35 42 32 32 65 12 34 10 13 36 ...
## $ defending_sliding_tackle : num [1:19239] 24 19 24 29 53 18 32 11 10 38 ...
## $ goalkeeping_diving     : num [1:19239] 6 15 7 9 15 87 13 88 88 8 ...
## $ goalkeeping_handling   : num [1:19239] 11 6 11 9 13 92 5 88 85 10 ...
## $ goalkeeping_kicking    : num [1:19239] 15 12 15 15 5 78 7 91 88 11 ...
## $ goalkeeping_positioning : num [1:19239] 14 8 14 15 10 90 11 89 88 14 ...
## $ goalkeeping_reflexes    : num [1:19239] 8 10 11 11 13 90 6 88 90 11 ...
## $ goalkeeping_speed       : num [1:19239] NA NA NA NA NA 50 NA 56 43 NA ...
```

#### 4. Est-ce que R a correctement reconnu le type de chaque colonne ?

Pour la plupart des données, R semble avoir bien reconnu le type des données. Par contre toutes les données qui comportent du texte ont été lus comme chr alors que certaines données texte peuvent représenter des catégories. Par exemple la variable `work_rate` représente une catégorie (plusieurs joueurs peuvent être caractérisés pour un même `work_rate`, cette donnée n'est pas unique à chaque joueur.).