

2022_transfers.csv

player_name: Player's name
age: Age of the player
fee: Transfer fee of the player
transfer_period: Transfer window for the transfer of the player
season: Upcoming season ahead of the transfer
Squad: Player's team or squad
player_code: Unique identifier for the player

squad_comparisons.csv

'Squad': Team name
'sq_Poss': Team's possession statistics
'sq_Gls': Team's total goals scored
'sq_Ast': Team's total assists
'sq_Gls/90': Team's average goals scored per 90 minutes
'sq_Ast/90': Team's average assists per 90 minutes
'opp_sq_Poss': Opponent team's possession statistics
'opp_sq_Gls': Opponent team's total goals scored
'opp_sq_Ast': Opponent team's total assists
'opp_sq_Gls/90': Opponent team's average goals scored per 90 minutes
'opp_sq_Ast/90': Opponent team's average assists per 90 minutes

squads_2022-2023.csv

'Squad': Team name
'sq_Poss': Team's possession statistics
'sq_Gls': Team's total goals scored
'sq_Ast': Team's total assists
'sq_Gls/90': Team's average goals scored per 90 minutes
'sq_Ast/90': Team's average assists per 90 minutes
'opp_sq_Poss': Opponent team's possession statistics
'opp_sq_Gls': Opponent team's total goals scored
'opp_sq_Ast': Opponent team's total assists
'opp_sq_Gls/90': Opponent team's average goals scored per 90 minutes
'opp_sq_Ast/90': Opponent team's average assists per 90 minutes

squads 2021-2022.csv

'Squad': Team name
'sq_Poss': Team's possession statistics
'sq_Gls': Team's total goals scored
'sq_Ast': Team's total assists
'sq_Gls/90': Team's average goals scored per 90 minutes
'sq_Ast/90': Team's average assists per 90 minutes
'opp_sq_Poss': Opponent team's possession statistics
'opp_sq_Gls': Opponent team's total goals scored
'opp_sq_Ast': Opponent team's total assists
'opp_sq_Gls/90': Opponent team's average goals scored per 90 minutes
'opp_sq_Ast/90': Opponent team's average assists per 90 minutes

transfermarkt_data.csv

'player_id': Unique identifier for the player
'first_name': Player's first name
'last_name': Player's last name
'name': Player's full name
'current_club_id': Unique identifier for the player's current club
'player_code': Player's code or identifier within the system
'country_of_birth': Player's country of birth
'date_of_birth': Player's date of birth
'sub_position': Player's more specific position
'foot': Player's dominant foot
'height_in_cm': Player's height in centimeters
'contract_expiration_date': Date when the player's contract expires
'current_club_domestic_competition_id': Unique identifier for the domestic competition of the player's current club
'current_club_name': Name of the player's current club
'club_id': Unique identifier for the club
'club_code': Code or identifier for the club
'last_season': Last season the player participated in
'date': Date of the data entry or data snapshot
'market_value_in_eur': Player's market value in euros
'player_club_domestic_competition_id': Unique identifier for the domestic competition of the player's club
'Born': Player's birth year (last two digits)
'Based': Player's team's league
'Squad': Player's team or squad

final_nongk_data.csv

'Player': Player's name
'Squad': Player's team or squad
'Age': Player's age
'Born': Player's birth year (last two digits)
'90s': Minutes played converted to 90-minute intervals
'Gls': Goals scored
'Sh': Total shots taken
'SoT': Shots on target
'SoT%': Percentage of shots on target
'Sh/90': Shots per 90 minutes
'SoT/90': Shots on target per 90 minutes
'G/Sh': Goals per shot
'G/SoT': Goals per shot on target
'Dist': Total distance covered
'FK': Free kicks taken
'PK': Penalty kicks scored
'PKatt': Penalty kick attempts
'xG': Expected goals
'npG': Non-penalty expected goals
'npG/Sh': Non-penalty expected goals per shot
'G-xG': Goals minus expected goals
'npG-xG': Non-penalty goals minus non-penalty expected goals
'Based': Player's team's league
'Tkl': Total tackles
'TklW': Tackles won
'Def 3rd': Tackles in the defensive third
'Mid 3rd': Tackles in the middle third
'Att 3rd': Tackles in the attacking third
'drTkl': Dribbled past tackles
'Att': Total attempted tackles
'Tkl%': Tackle success rate
'Lost': Total times possession was lost
'Blocks': Total blocks
'Blocks Sh': Blocked shots
'Blocks Pass': Blocked passes
'Int': Interceptions
'Tkl+Int': Total tackles plus interceptions
'Clr': Total clearances
'Err': Errors leading to goals
'Position': Player's position on the field
'Total Cmp%': Passing accuracy percentage
'TotDist': Total distance of passes
'PrgDist': Progressive distance of passes
'Short Cmp%': Short passing accuracy percentage

'Medium Cmp%': Medium passing accuracy percentage
'Long Cmp%': Long passing accuracy percentage
'Ast': Assists
'xAG': Expected assists
'xA': Expected assists
'KP': Key passes
'1/3': Passes into the final third
'PPA': Passes leading to a shot
'CrsPA': Crosses leading to a shot
'PrgP': Progressive passes
'PPM': Points per match
'onG': Goals scored when player is on the field
'onGA': Goals conceded when player is on the field
'+/-': Goal difference when player is on the field
'+/-90': Goal difference per 90 minutes when player is on the field
'On-Off': Goal difference when player is on versus off the field
'onxG': Expected goals when player is on the field
'onxGA': Expected goals conceded when player is on the field
'xG+/-': Difference between actual goals and expected goals when player is on the field
'xG+/-90': Difference between actual goals and expected goals per 90 minutes when player is on the field
'xG_On-Off': Difference in expected goals when player is on versus off the field
'player_code': Unique identifier for the player
'country_of_birth': Player's country of birth
'sub_position': more specific position
'foot': Dominant foot
'height_in_cm': Height in centimeters
'market_value_in_eur': Player's market value in euros
'fee': Transfer fee
'transfer_period': Period of transfer
'Target': Target variable or outcome (in a predictive modeling context)

NOTE: nongk_data_2023.csv and player_data_2023.csv are subsets of final_nongk_data.csv above.

final_gk_data.csv:

'Player': Player's name
'Squad': Player's team or squad
'Age': Player's age
'Born': Player's birth year (last two digits)
'90s': Minutes played converted to 90-minute intervals
'GA': Goals conceded
'GA90': Goals conceded per 90 minutes
'SoTA': Shots on target faced
'Saves': Total saves made
'Save%': Save percentage
'CS': Clean sheets (matches without conceding goals)
'CS%': Clean sheet percentage
'Based': Player's team's league
'Tkl': Total tackles
'TklW': Tackles won
'Def 3rd': Tackles in the defensive third
'Mid 3rd': Tackles in the middle third
'Att 3rd': Tackles in the attacking third
'drTkl': Dribbled past tackles
'Att': Total attempted tackles
'Tkl%': Tackle success rate
'Lost': Total times possession was lost
'Blocks': Total blocks
'Blocks Sh': Blocked shots
'Blocks Pass': Blocked passes
'Int': Interceptions
'Tkl+Int': Total tackles plus interceptions
'Clr': Total clearances
'Err': Errors leading to goals
'Position': Player's position on the field
'Total Cmp%': Total passing completion percentage
'TotDist': Total distance of passes
'PrgDist': Progressive distance of passes
'Short Cmp%': Short passing completion percentage
'Medium Cmp%': Medium passing completion percentage
'Long Cmp%': Long passing completion percentage
'Ast': Assists
'xAG': Expected assists
'xA': Expected assists
'A-xAG': Assists minus expected assists
'KP': Key passes
'1/3': Passes into the final third
'PPA': Passes leading to a shot
'CrsPA': Crosses leading to a shot

'PrgP': Progressive passes
'PPM': Points per match
'onG': Goals scored when player is on the field
'onGA': Goals conceded when player is on the field
'+/-': Goal difference (goals scored minus goals conceded) when player is on the field
'+/-90': Goal difference per 90 minutes when player is on the field
'On-Off': Goal difference when player is on versus off the field
'onxG': Expected goals when player is on the field
'onxGA': Expected goals conceded when player is on the field
'xG+/-': Difference between actual goals and expected goals when player is on the field
'xG+/-90': Difference between actual goals and expected goals per 90 minutes when player is on the field
'xG_On-Off': Difference in expected goals when player is on versus off the field
'player_code': Unique identifier for the player
'country_of_birth': Player's country of birth
'sub_position': more specific position
'foot': Dominant foot
'height_in_cm': Height in centimeters
'market_value_in_eur': Player's market value in euros
'fee': Transfer fee
'transfer_period': Period of transfer
'Target': Target variable or outcome (in a predictive modeling context)

NOTE: gk_data_2023.csv is a subset of final_gk_data.csv above.