- **Esercizio R.** Nella pagina web `http://www.memotef.uniroma1.it/node/6787` è presente il data set `germania` contenente dati relativi alle performance delle squadre tedesche nel loro campionato di calcio. In particolare sono presenti le variabili "Position" (posizione alla fine del campionato) "Team" (squadra) "GamesPlayed" (partite giocate) "Wins" (partite vinte)"Draws" (partite pareggiate) "Losses" (partite perse) "GoalsScored" (goal segnati) "GoalsAgainst" (goal subiti) "PointsScored" (punti)

  1. Scaricare il data set. Quante partite hanno giocato le prime 5 squadre del data set?
  2. Calcolare la media della variabile GamesPlayed. Che informazione relativa al data set ne desumi?
  3. Creare una nuova variabile y data dal rapporto tra i punti ottenuti e le partite giocate
  4. Stimare un modello di regressione avente variabile risposta la y appena creata e come esplicative il numero di goal fatti e subiti. Interpretare i valori dei coefficienti del modello
  5. Una squadra che pareggia 0 a 0 tutte le partite del suo campionato, in base al modello stimato quanti punti per partita ottiene?
  6. Le variabili esplicative sono entrambe significative?
  7. Calcolare i residui delle prime 3 osservazioni. Quanto vale la media dei residui
  8. disegare un grafico con i residui sull'asse delle y e i goal segnati sull'asse delle x. Che informazioni ne desumi?
  9. Introdurre all'interno del modello anche il quadrato della variabile goal subiti