Project: Analytics for sports
League : Bundesliga 2
Under The Guidance of
Prof.Narayan Rangaraj

Team members :
   1. *Tejas Gunaghe (213194002)*
   2. *Anuj Meena (213190006)*
   3. *Aditya Dutta (213190001)*

- **Predicted Winners of Bundesliga2 :**
  Based on the probability based prediction model, higher winning chances of following teams is found:

1.	Werder Bremen
2.	Schalke 04
3.	Hamburg

- **Model link:**
  docs.google.com/spreadsheets/d/1SgOV_0SHNPdidA24DK4q2pNQyobJazBGPSobLYb9nA4/edit#gid=2143961598

- **Data considered :** *link*
  Data from 1988 - till 2022 is considered for predicting the winner in 2022

  - Key Features Considered:
    1. Total Goals for and against each other
    2. Number of matches won against each other
    3. Performance in current league
    4. Luck probability -30%(randomly)

  Refer below image for sample data

Team1	Against Team	Goals For	Goals Against	Win_Team1	Draw	Loss_Team1	Total Matche
Darmstadt	St. Pauli	21	8	8	1	2	11
Darmstadt	Werder Bremen	18	16	3	5	2	10
Darmstadt	Nürnberg	25	23	5	4	5	14
Darmstadt	Schalke 04	12	18	2	1	5	8
Darmstadt	Heidenheim	24	28	6	6	8	20

- **Prediction Model Details :**
  - Methodology : Probability based model
  - Model Accuracy: 51.07%

Sheets	Information
Current Performance :	Calculates the winning probability of a team based on current leagues performance till week25.
Individual Probability Calculations:	Calculation of winning probability from historic data is calculated based on number of goals and number of times won against a particular team
Match Summary:	This sheet list down all the matches in current league and based on current performance and historic performance the winning probability of each team is calculated match wise
Prediction Summary:	Shows final summary of Winners based on the probability calculation

- **Sample Calculations:**

Example:
The 1st match in the current league was played between Schalke 04 Vs Hamburg on Friday 24/07/2021, Hamburg won the match by 3-1.

Let's analyze this scenario and try to predict the winner of this match using the winning probability:
Winning probability is calculated based on:

1. **Current league performance : (Performance)**
   a. Schalke 04 played 25 matches out of which they won 12, hence winning probability of this team is $12/25 = 0.56 \sim 56\%$
   b. Hamburg played 25 matches out of which they won 10, hence winning probability of this team is $10/25 = 0.4 \sim 40\%$

2. **Historic Goals against each other:**
   From 1988 till 2021 Schalke 04 has made 149 goals against Hamburg, and Hamburg made 158 goals against Schalke 04 , hence total goals made by both = 149+158 = 307 Goals, hence
   a. P(Winner = Schalke 04) = $149/307 = 0.49 \sim 49\%$
   b. P(Winner = Hamburg) = $158/307 = 0.51 \sim 51\%$

3. **Historic Matches won against each other:**
   From 1988 till 2021 Schalke 04 and Hamburg played 102 mathes out of which Schalke 04 won 39 matches , Hamburg won 38 matches and Draw were 25
   a. P(Winner = Schalke 04) = $39/102 = 0.38$
   b. P(Winner = Hamburg) = $38/102 = 0.37$
   c. P(Draw) = 0.25

4. **Luck Probability :upto 30%**
   To add the randomness in the model a randomly selected probability is calculated using the Rand() excel function, this probability is completely based on luck and is added to the earlier probability.
   a. P(Winner = Schalke 04) = 0.13
   b. P(Winner = Hamburg) = 0.21

Hence Overall probability : (from 1 to 4 points)
Considering all the above probabilities are independent of each other

$$P(Winner = Schalke\ 04) = 0.56 * 0.49 * 0.38 * 0.13 = 0.0115796896\ \textbf{(A)}$$
$$P(Winner = Hamburg) = 0.4 * 0.51 * 0.37 * 0.21 = 0.01610576739\ \textbf{(B)}$$

As **(B) > (A)**, the probability of winning is more for Hamburg and is predicted to be the winner of this match.

Similarly, all the winners and draws are calculated for each match. Finally, considering 3 points for winner, 1 point for Draw and 0 point for loser, overall points are calculated for each team. Top 3 teams having maximum points are the winners as per our prediction.