**1. Introduction**

**1.1 Problem Statement and Research Motivation**

Analyzing game performance using statistical metrics reveals important insights in competitive sports like football. Ball possession is frequently emphasized as a crucial component of gameplay by coaches, commentators, and players. This is because it is believed that higher possession rates are correlated with improved performance and more opportunities to score goals. However, recent studies, such as Collet (2013), suggest that possession alone may not guarantee success, and its correlation with goals scored remains inconclusive. The purpose of this study is to use actual results to support or contradict this commonly held belief by examining the correlation between ball possession and goals scored at the 2018 FIFA World Cup.

**1.2 The Dataset**

The dataset "DS051 and FIFA 2018 Statistics.csv," which contains 128 rows of match-related metrics from the 2018 FIFA World Cup, is used for the analysis. Important variables of interest:

* Goals scored (dependent variable): shows how well a team performed.
* Ball Possession (Independent Variable): Indicates the percentage of the game a team had the ball under control.

The structured format of the dataset makes it possible to perform thorough statistical analysis, such as correlation testing, to ascertain whether the chosen variables have any meaningful correlations.

**1.3 Research Question**

Is there a correlation between the number of goals scored and ball possession during the FIFA world cup 2018?? The purpose of this question is to either contradict or support the widely held belief in sports analytics that having possession of the ball for longer periods directly results in more goals.

To answer the research question, we will analyze the "DS051 and FIFA 2018 Statistics.csv" dataset using statistical and visualization methods. Scatterplots will assess the relationship between ball possession and goals scored, while histograms evaluate the data distribution. Spearman’s Rho correlation analysis will determine if a significant correlation exists, ensuring a thorough reliable investigation of the research question.

**1.4 Null Hypothesis and Alternative Hypothesis**

The study aims to test the relationship between ball possession and the number of goals scored during the FIFA World Cup 2018 by formulating the following hypotheses:

* **Null Hypothesis (H₀)**: there is no correlation between the number of goals scored and the ball possession.
* **Alternative Hypothesis (H₁)**: there is a correlation between the number of goals scored and the ball possession.

In order to make a data-driven conclusion about the correlation between ball possession and number of goals scored, this study will test these hypotheses at a 5% significance level.