

Data Collection Report

Kiran Gadhave - u1143683

Matthew Bradley - u1206298

February 13, 2019

- **Source:**

The data is taken from a Kaggle repository located at following url:

<https://www.kaggle.com/hugomathien/soccer/version/10>

The data is in form of an sqlite database which we downloaded and processed as described below.

- **Size:**

The original sqlite database has multiple tables which have varying sizes. The table below describes the dimensions of each:

Table	Dimensions (Rows \times Columns)
Country	11 \times 2
League	11 \times 3
Match	25979 \times 115
Player	11060 \times 7
Player Attributes	183978 \times 42
Team	299 \times 5
Team Attributes	1458 \times 2

- **Format:**

- **Preprocessing:**

- **Simulation:**