

# サッカーリーグのデータ解析システム構築とデータ分析

理工学部 数理情報学科  
T050008 梅原 健太  
指導教員 佐野 彰

## 概要

今日のスポーツ界では、試合データを解析して、戦略やチーム構成に利用するということが一般的に行われている。本研究では、様々なスポーツ団体の中から、日本のサッカーリーグであるJリーグに注目した。そして、サッカーリーグデータの解析のためにデータベースを構築し、また、解析の便宜さを図るためにJavaによるインターフェイスを作成してデータ解析を行った。このデータ解析を通じて、Jリーグの各チームの特徴を求めていく。

まず、Jリーグ公式サイトから2007年度のJ1リーグのデータを元にデータベースを構築した。データベースを構築する上で、データベース管理システムのMySQLを利用した。MySQLを利用することで、データの検索、削除などを素早く正確に行うことができる。また、MySQLを効率よく利用するために、Javaによるインターフェイスを作成した。検索条件を細かく指定できるため、容易に目的のチームや選手を探し出すことができるようになった。

次に、作成したインターフェイスを利用してデータの解析を行った。まず、予備的な解析として、年間順位と各データとの相関係数を計算した。その結果、ゴール数や失点数などの試合の勝負に直接影響するような数値と年間順位の間には強い相関があるということが確認できた。次に、年間順位と相関の強いデータ項目において、その数値が同じにもかかわらず順位がまったく異なる2チームに注目した。そして、この2チームのどのような違いが順位を分けたかということに注目してデータ解析を行った。とくに、各チームのシュート決定率の内容に注目した。決定率の内容をより詳細にするためにチーム内の各選手ごとのデータを調べた。その結果、上位のチームにはシュート数と得点数を引っ張る一人の選手が存在し、さらに複数ポジションの複数選手が点を取っているということがわかった。逆に、下位のチームは攻撃専門の2選手のみがおおよそそのチーム得点を取っていることがわかった。このことは、相手チームの守備という視点からは、上位チームに対しては、複数の選手を抑えなければならないのに対し、下位チームに対してはある少数の選手を抑えることで点がとられにくくなることを意味する。さらに、2チームの攻撃スタイルの違いにも注目し、その中から、シュート数に関する数値であるパス数を調べた。パス数に関しては、両チームとも差がなかったが、パスの内容が大きく異なっていた。上位チームに関しては、シュートにつながるパス（スルーパス数、クロス数）が下位に比べて比較的多かったのである。

本研究で解析対象として、2つのチームは共に同じフォーメーションをとるチームであった。それにもかかわらず、データ分析の結果、この2チームには、決定率の内容や攻撃スタイルに差が表れており、これが順位の違いとして表れていると考えられる。

2008 年度卒業論文

サッカーリーグのデータ解析システム構築とデータ分析

龍谷大学工学部  
数理情報学科

T050008 梅原 健太

指導教員 佐野 彰

## 目次

1	はじめに	1
2	データベース	2
2.1	データベースとは	
2.2	リレーショナル・データベースとは	
2.3	SQL と MySQL	
3	サッカーリーグデータのSQL化	3
3.2	サッカーリーグデータ	
3.1	データの抽出	
3.3	SQL データベース化	
4	Javaによるデータ解析システムの構築	7
4.1	データの表示 (解析システムの構築)	
4.2	相関関係とは	
4.3	相関係数表の表示	
5	データ分析	10
5.1	分析方法	
5.2	予備的な解析 (年間順位と強さの関係性)	
5.3	決定率の同じ2チームの比較	
6	おわりに	21
	謝辞	22
	参考文献	23