

# 情報可視化ソフトウェア Hidden によるデータ分析

～浦和レッズの過去 5 年分の成績について～

私は地元が埼玉県の浦和というところで、地元のチームである浦和レッズの過去 5 年分の試合の成績や、そこからさらに細かな情報等を分析してみた。

試合の成績は浦和レッズ公式ホームページから引用した。

(引用元 URL <http://www.urawa-reds.co.jp/game/pastlist/>)

## ➤ データ作成

まずデータを csv 形式のファイルでまとめる際に、年(Year)、月(Month)、対戦相手(Opponent)、スタジアム(Stadium)、得点(Score)、勝敗(Judge)の 6 つの属性に分けてまとめた。また、具体的な数値で表せない部分は以下のようにクラス分けをして数値で表せるようにした。

## 対戦相手

(※基本的に 2018 年の J リーグ(J1)の順位順に並んでいる)

1	川崎フロンターレ	11	横浜 F・マリノス
2	サンフレッチェ広島	12	湘南ベルマーレ

---

3	鹿島アントラーズ	13	サガン鳥栖
4	北海道コンサドーレ札幌	14	名古屋グランパスエイト
5	FC東京	15	ジュビロ磐田
6	セレッソ大阪	16	柏レイソル
7	清水エスパルス	17	V・ファーレン長崎
8	ガンバ大阪	18	国内(天皇杯等)
9	ヴィッセル神戸	19	海外(ACL等)
10	ベガルタ仙台		

---

### スタジアム

---

0 ホーム(埼玉スタジアム)

1 それ以外

### 勝敗

---

0 負け

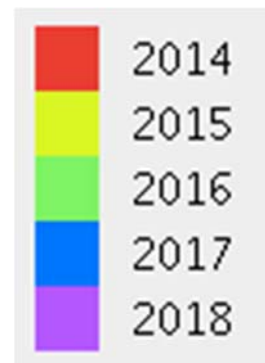
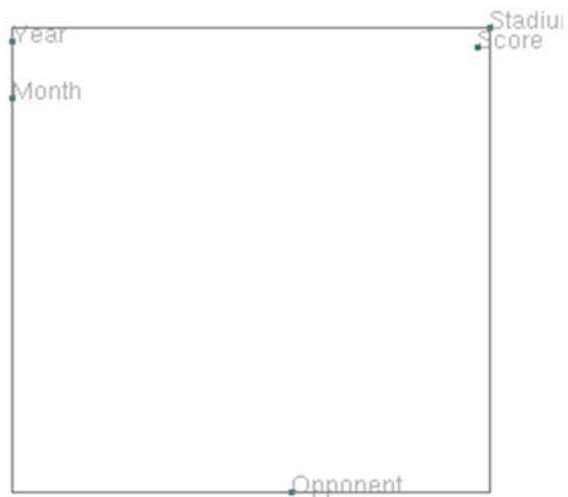
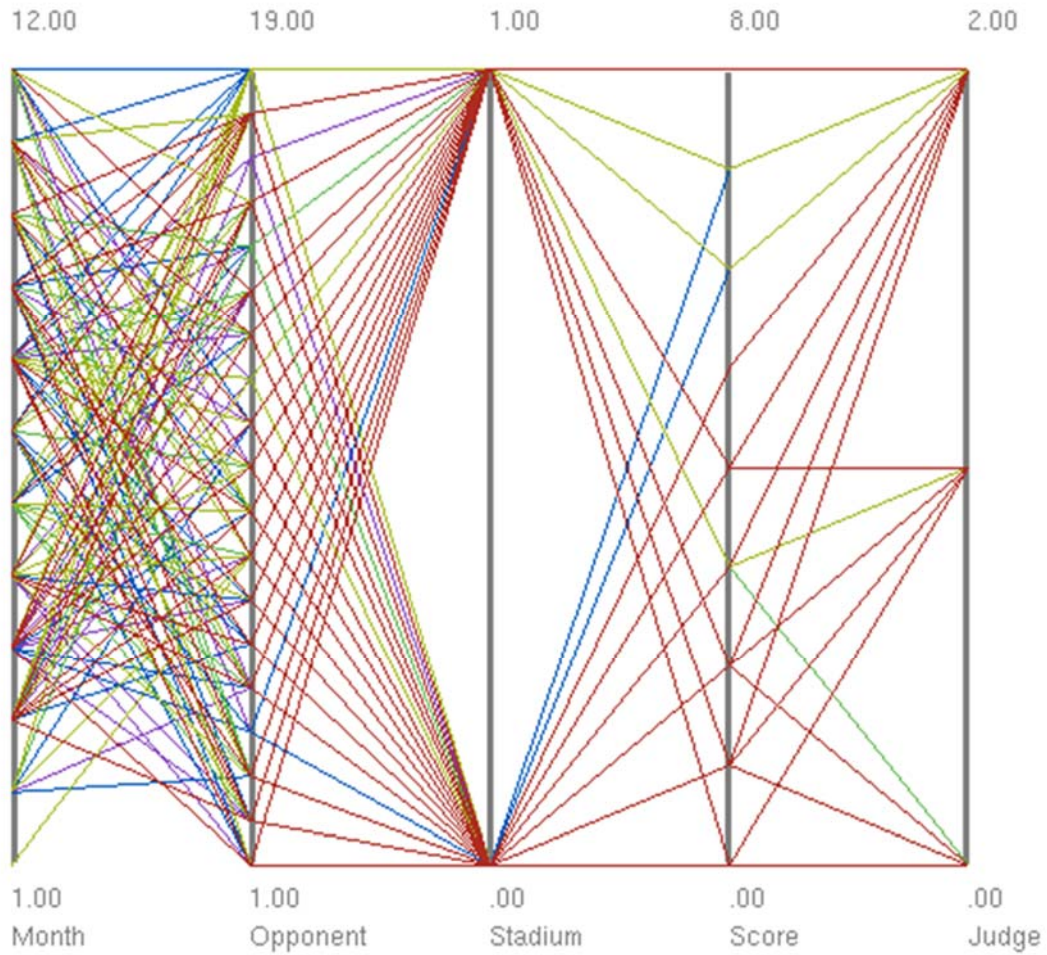
1 引き分け

2 勝ち

---

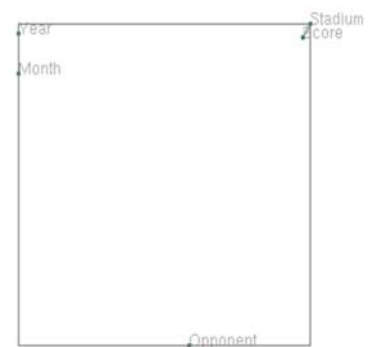
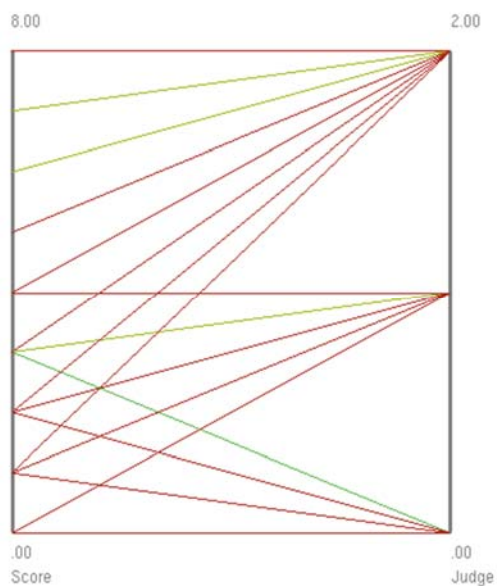
➤ データ分析

作成した csv 形式ファイルを Hidden で開くと以下のような結果になった。



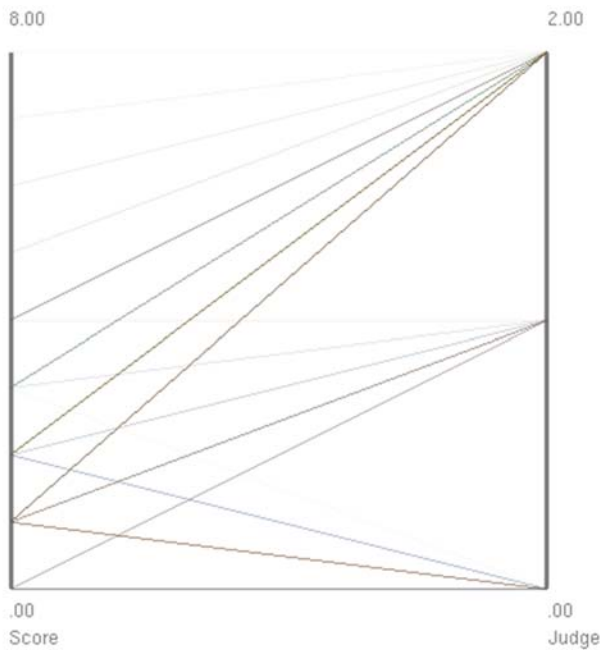
この状態から、色々動かしてデータの各要素における特徴や関連性を調べていくことにする。

- 右上のスライダーを動かす(相関の強さ)

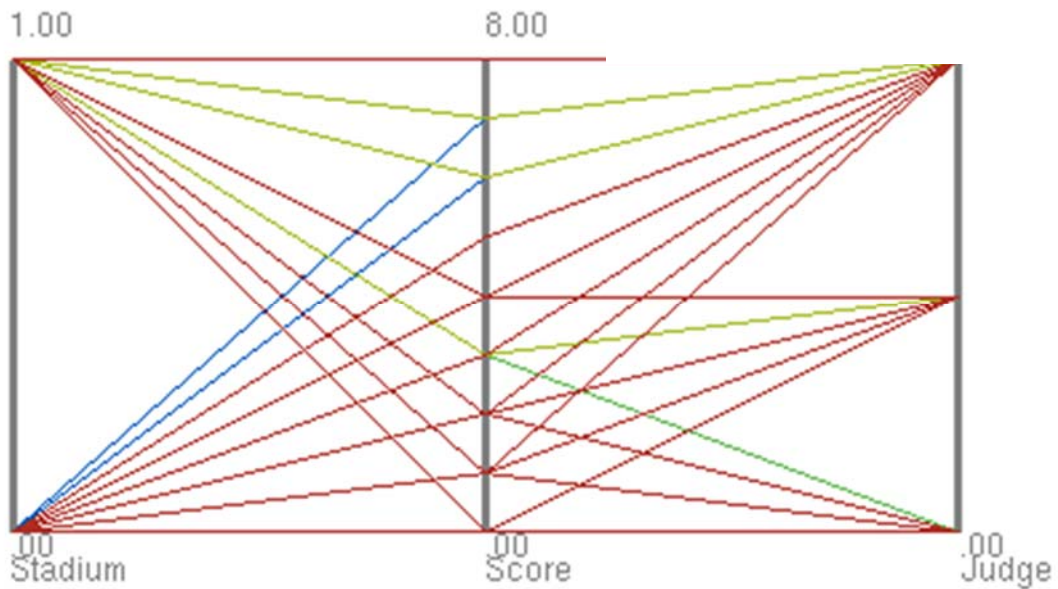


まず Score と Judge の相関が強いことがわかる。Score は 0~8 点であり、グラフから、0~3 点だと負け、0~4 点で引き分け、1~8 点で勝ちというの

が読み取れる。サッカーというスポーツの性質上、点がたくさん入る競技でもないのに、1点でも取れば勝てる場合も多いにある。また、3点取っても負ける場合もあることがわかり、点がある程度取れたから必ず勝てるという保証はないのだということがわかる。



また、Transparency を動かしてみると、1~4 点で勝ち、0~1 点で引き分け、0~2 点で負け、というパターンが特に多いことがわかる。よって、3 点以上取れている場合は勝ちの可能性が高いことがわかる。また、勝敗の数は線



次に、スタジアムと点数の関係について見ていく。自分の予想ではホームでの試合の方が自分のチームのサポーターも多く応援もその分大きくなるので勝つ可能性がアウェイの時と比べて高いと思っていたが、実際グラフを見てみると点のばらつきはどちらも大きく、スタジアムがホームか否かという条件は点や勝敗にあまり大きく影響しないのではないかと考えられる。

先ほどのデータを可視化して見ていった時に5年分を一個のデータファイルにまとめているので、だいぶ複雑に入り組んでいてわかりにくい部分も多かった。なので、以下は2018年を詳細に見て、成績を分析していくこととする。さらに、もっとそれぞれの試合を細かく見るため、Jリーグの試合に絞ってみていくこととする。また、新たにデータの要素を作り、シュート(Shoot)、シュート成功率(Shoot rate)、支配率(Dominate rate)、監督(Director)を増やして分析してみた。新しく追加した情報は以下のURLから引用した。ここで、監督は以下のように分類分けする。

(引用 URL <http://www.football-lab.jp/uraw/match/>)

監督

---

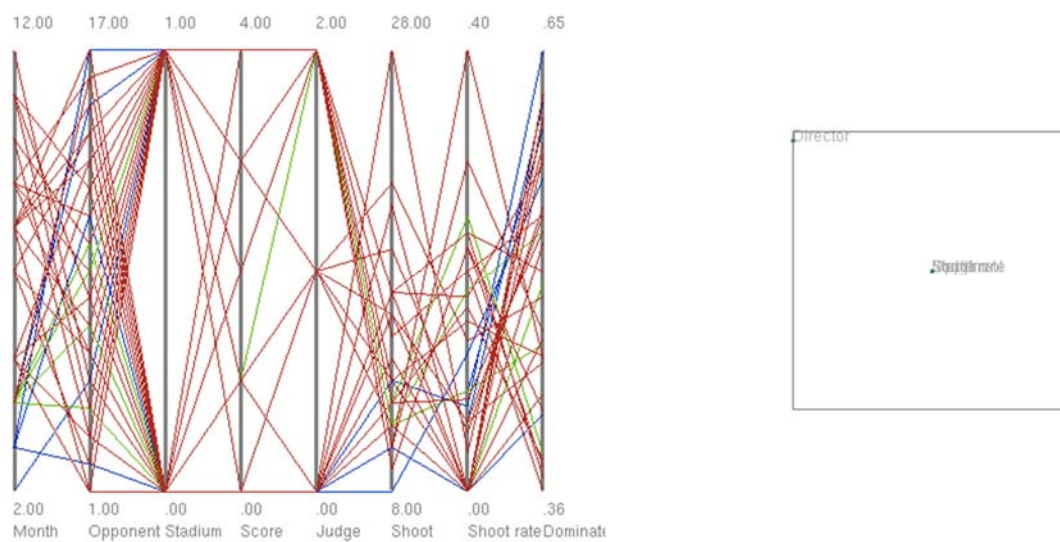
1            オズワルド・オリヴェイラ(現在の監督)

2            大槻 毅

3            堀 孝史

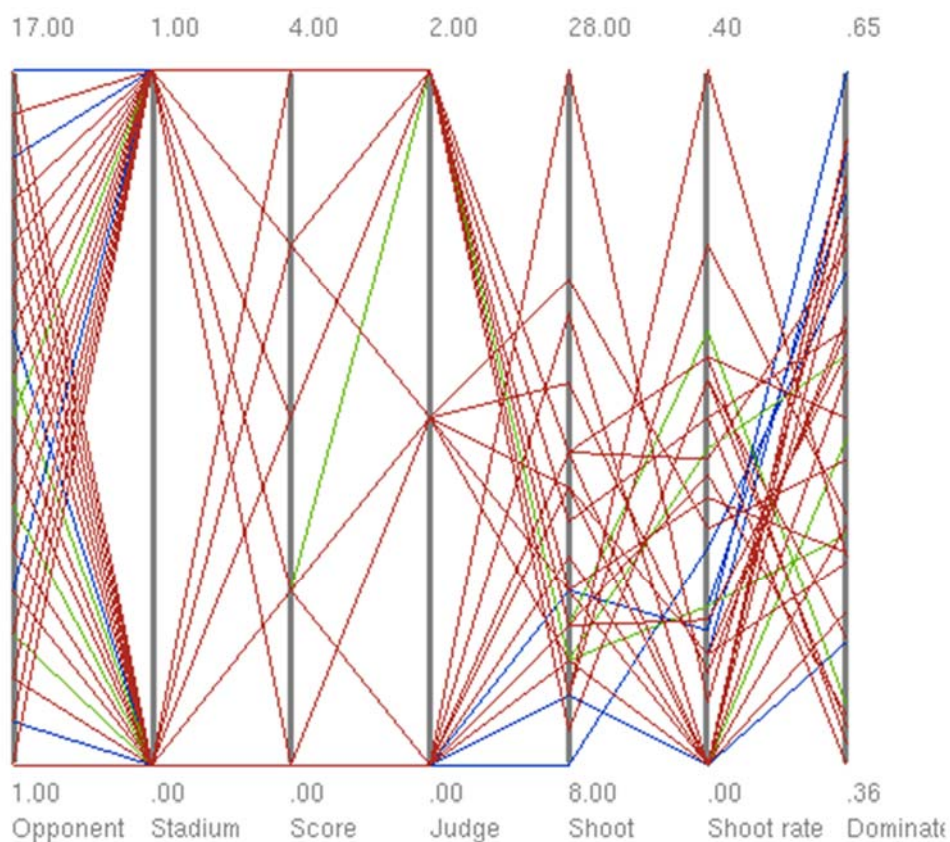
---

● 2018



線を監督ごとに色分けして表示した結果が上の図である。

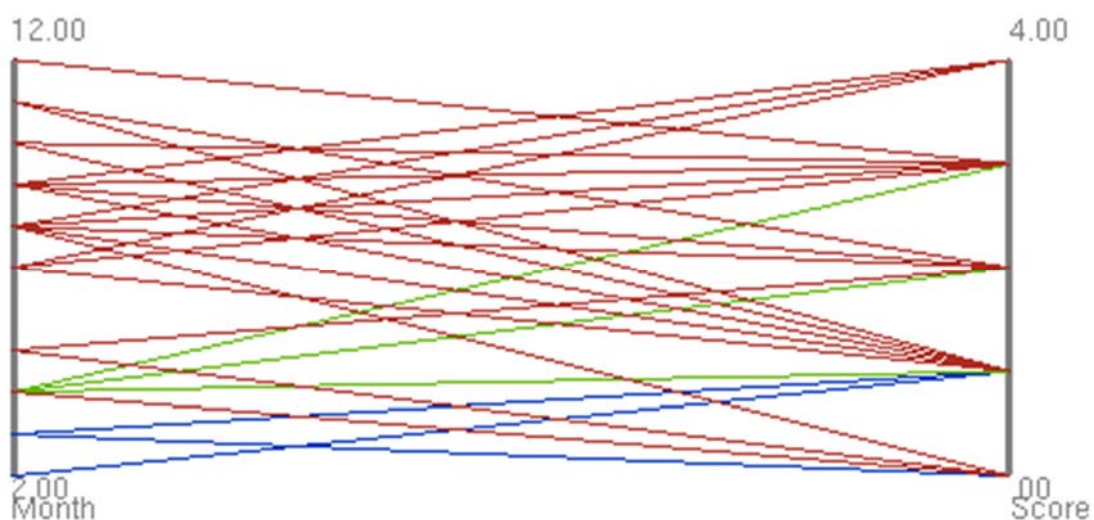




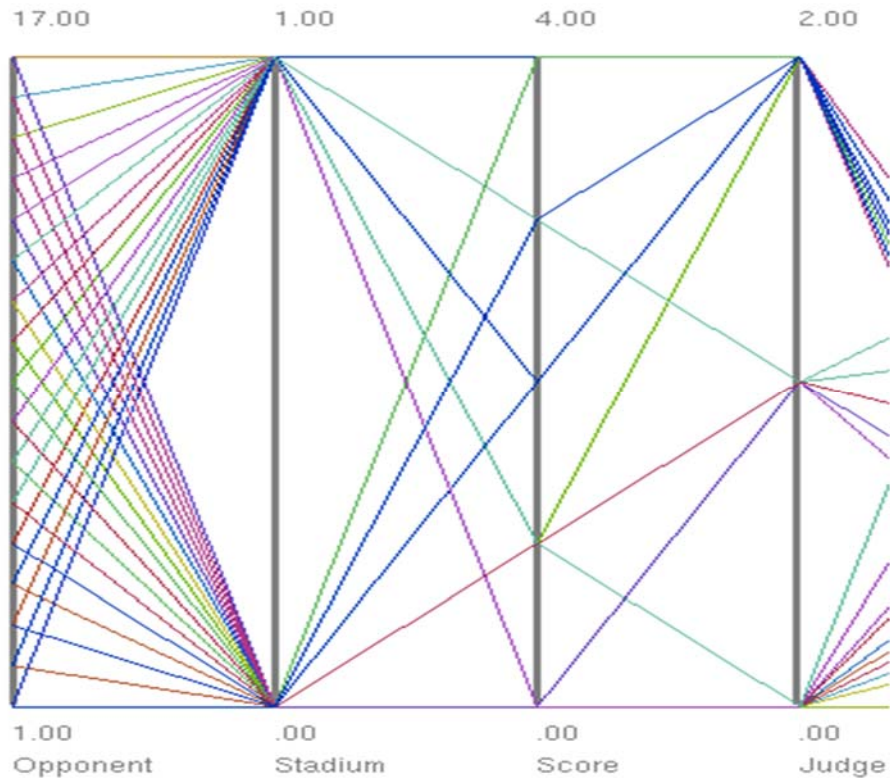
右上のスライダーを調節して詳しく見ていく。Shoot と Score に注目するとシュート本数が多すぎても、少なすぎても試合結果は負けていることが多いことがわかる。また、勝っている試合の時は、シュート本数は 10~20 本くらいだといえる。さらに、シュート本数が少ないのに勝った試合に注目すると、もちろん Shoot rate は高いのだが、Dominate rate はそれほど高いわけではないことがわかる。苦しい状況の中で確実に打ったシュートを決めて勝利につながった試合だといえよう。また、監督別に見ていくと、青線(Director3)の時は、ボールの支配率は高いが、シュート成功率が低く、結果負けてしまうと



いうゲームが多い。反対に緑線(Director2)は、ボール支配率は低いものの、シュート成功率は青より高く、試合結果も勝ちが多い。赤線(Director1)はボール支配率にあまり規則性がないものの、シュート成功率がかなり高い試合をしていることが多いことがわかる。



実際、Month と Score を見てみると、青から緑、赤と監督が変わるにつれて試合で取れる得点が上がってきていることがわかり、この年は後半になるにつれてチーム成績が良くなっていったことがわかる。



また、対戦相手ごとに色分けして表示してみると、Opponent の値が小さい(J1リーグでの順位が高いチーム)の時の方が強いので点を取りにくいのではないかと思ったが、意外とグラフを見ると、強いチームと戦う時に点が取れていないというより、自分たちのチーム(2018年は5位)より格下のチームと戦った時に意外と点が取れていないことが見て取れる。

➤ まとめ

あまりデータの相関関係が綺麗に出なかったこと、対戦相手と試合結果に関する関係性がうまく見ることができなかつたことが今回の反省点ではあるが、得

点の分布や傾向、シュート本数やボール支配率などの分析ができ、また監督の影響力が大きいということなど、データを可視化することで様々なことがわかった。今回はデータの要素において、数値で表せないものを無理やりクラス分けして数値に変換することが多かったので、今度また可視化するならもっと細かい数値のデータを拾えるような内容のもので実践したいと思う。