Leicester City: An Improbable Story

By: Jonathan Lee (jwl274), Crystal Liu (crl95), Nikolaus Salvatore (nds34)

Our project took a look at the English Premier League and the historical run that one team, Leicester City, has had this season. We developed three interactive visualizations that helped to summarize and provide context to the history in the making that is happening within English Football. Leicester City who were only recently promoted to the top league of English Football have went from being promoted, fighting relegation and ending 14th in the league, and being champions all in the span of three seasons.

Our first visualization focused on an overview of the current and historical Premier League table. In the English Premier League, a win earns a team three points, a draw one, and a loss zero. The team with the highest number of points at the season's conclusion is crowned the victor; ties are broken by goal differential.

The first visualization draws from JSON's sourced from: https://github.com/jokecamp/FootballData/tree/master/EPL%201992%20-%202015/tables. These JSONs provide the final rankings of the English Premier League of the season. Each JSON object has the rank, team, number of games played, wins, draws, losses, points, goals-for, goals-against, and goal differential. We downloaded the last 10 seasons of JSON data for the sake of our visualization. We also manually constructed a JSON for the current '15-'16 season up to date with the current standings. This was sourced from: http://www.premierleague.com/en-gb/matchday/league-table.html.

For the first visualization, we elected to not use the number of games played, goals-for,goals-against, and goal differential attributes of the dataset. This is because games played for every season will consistently be the same at conclusion, and goal differential is only relevant in the case of a tie, which does not occur frequently. Instead, we used points, wins, goals, draws, rank, and team name which best succinctly help to paint a picture of each individual season, the victors, and how their season compared to their previous one.

In the visualization, text names are clearly mapped to the team name data and the team's crest is also mapped to the name provided for each rank.

Points for each team are mapped to rectangle blocks indicating and thus offering easy comparison between teams. On hover, a breakdown of the wins, draws, and losses are indicated as well as the points awarded from each category. The losses category indicates the number of total additional points had the team won every game they had lost. Lastly, the movement in position relative to the previous season is represented by either an up, down, or both arrows to indicate moving up, moving down, or no movement respectively compared to the previous season.

In order to position the 20 teams in the table, an ordinal scale was used with the domain [1..20] and rangeBands between padding and the height of the table. In order to determine the length

of the rectangular blocks for points as well as win,draw,loss breakdowns, a linear scale was used with the domain being 0 to 114 (the total number of points possible in a season).

The table visualization first helps to provide a comprehensive overview of the current season's standings and table. Using the slider, users can see the last ten winners of the English Premier League and more importantly track the success of particular teams that tend to remain top contenders. Investigating Leicester's path through the Premier League can be tracked by following its two previous seasons where they were first promoted and placed 14th and now are top champion contenders. What is so incredibly surprising is that in the past ten seasons, the victor has always been within the top three in the previous season. This season, the top two contenders were 14th and 5th respectively the year before. More surprising yet, Leicester City has only been in the Premier League for one season before being in the prime position to clinch the title. Lastly, what is surprising to a viewer is the apparent decline of historically very successful clubs such as Chelsea and Manchester United.

For the second visualization, we compiled data from three different datasets that reported game statistics throughout this current season that were specifically related to Leicester City. We manually created a JSON file where each object reports the date, match, odds, result, possession, and pass accuracy of that week.
The first dataset is taken from http://www.oddschecker.com/football/english/premier-league/winner/bet-history/leicester/all-history, which shows the betting odds for Leicester City everyday reported by several sources. For every week, we chose to only look at relevant odds by taking the highest offered odds by bookies the day of the match.

The second dataset is taken from http://www.espnfc.us/club/leicester-city/375/fixtures?leagueId=23, where we got the statistics on who Leicester City played against and the result of the match for each week.

The third dataset is taken from http://www.squawka.com/teams/leicester-city/stats#performance-score#english-barclays-premier-league#season-2015/2016#165#all-matches#1-33#by-match, which shows more detailed match statistics than the second dataset. We chose to only focus on possession of the soccer ball and pass accuracy for every week, two relevant metrics of performance. Lastly, we added images of the opposing team for every match to the JSON file.

Using this JSON file, we created an interactive line plot in order to show the drastic change in the team's performance over this season. We plotted the betting odds for each match against the matchweek and used the images of the opposing team as the points to give viewers a preview of the corresponding game's statistics. For the y-axis, we chose to transform the scale into a square root scale so that the points would be spaced out more evenly and viewers would be able to clearly see all the points with small betting odds. For the x-axis, we used a linear scale to also space out the weeks evenly. Below the line plot is are two dark blue horizontal bars that display the medium pass accuracy and possession percentages of the last ten season's title winners. This data was sourced from: http://www.bbc.com/sport/football/35553082

On hover, we added tooltips to every point that display that week's match date, betting odds, match, and its result. On mouse click, that match's pass accuracy and possession percentages will show up in light blue on top of the static dark blue horizontal bars in order to compare Leicester City's performance to those of past winners.

This visualization shows how Leicester City's position in the English Premier League improved dramatically over this season. The team's betting odds in the first few weeks were in the thousands, and then less than halfway through the season, the odds dropped down to double digits. The team continued to improve tremendously to the point that the odds are now less than one. Based on the statistics for each game, viewers can see that Leicester City significantly wins more matches than loses them.

Displaying the team's pass accuracy and possession per game helps frame and compare Leicester City's performance this season to that of previous title winners. Clearly, there is a definite disparity between the percentages of Leicester and the median title winners'. This suggests that either their style of play is significantly different but is still netting them success, or they have simply been extremely lucky. This helps to transition into the third visualization which helps to further the point that Leicester City may simply be bringing a new style of play to the Premier League that doesn't necessarily seek to maximize traditional metrics of performance.

For the third visualization, various game statistics were taken from the following csv file:
https://github.com/jargnar/premier-league-data/blob/master/2015-16/data.csv

The graph displays the results of each game from the most recent season with respect to the chosen metric as well as the currently selected teams. For these statistics, we felt the most relevant data included: the total number of shots taken, crosses, corners, goals, penalties, saves and goals conceded during each game. These statistics unlike others such as the number of fouls help paint the best picture of how well a team performed.

A drop-down menu on the right-hand side allows the user to select from a list of relevant game statistics and the accompanying form of checkboxes allows the user to select which team's data will be displayed. The data associated with each team is color coded with the colors indicated in the legend generated on the right-hand side of the plot. The visualization also has an option to enable cumulative viewing mode, where the game statistics are displayed cumulatively over the season. While showing the statistics of each game shows the relative consistency of performance of a single or few teams, the cumulative data allows for easier comparison between a larger set of teams at once. Both cumulative and non-cumulative data are shown using a linear scale with the chosen metric. In addition to the chosen game metric, a popup box appears when hovering over a point on the plot that shows additional game information such as which teams were competing and also the date and venue where the game took place.

The comparison visualization helps support the narrative that traditional manners of evaluating team performance aren't particularly relevant to Leicester City or this current season. For example, Leicester are #2 in goals scored and #3 in goals conceded. Likewise, taking a look

across different metrics, user can see that there is not one necessarily distinct criteria or trend that can capture the team's incredible success this season.

Our visualizations as a whole help to show first how astonishingly successful Leicester City have performed this season to be in the title-winning position, and then second to show that there isn't necessarily one particular reason that the team have performed so well. In actuality, the team have inspired a style of play that is in some ways contrary to how soccer has traditionally been understood (e.g lower possession and passing accuracy). While the visualizations themselves might not offer the exact formula for Leicesters' success, they do highlight Leicester's amazing run throughout the season as well as suggest that to outside observers, there might be more than just pure statistics that dictate the result of a game.