Project Proposal

EPL Competitive Balance Analysis

Members:
Rahul Purushottam Gaonkar (rpg283)
Achyut Tibrewalla (at3617)
Divyanshu Singh (ds5387)

Project goal

The English Premier League is the top level of the English football league system. Contested by 20 clubs, it operates on a system of promotion and relegation. The top 4 teams of the English Premier League get qualified for the UEFA Champions League. The bottom 3 teams get relegated to a lower division league (EFL Championship).

The main goal of the project is to show how the distribution of teams changed based on league positions over different seasons and the game statistic metrics contributing to it. The project will try to convey how the traditional top 4 (Manchester United, Arsenal, Liverpool and Chelsea) and bottom 3 teams (relegated teams) changed over different seasons thereby emphasizing the competitive balance and unpredictability of EPL.

Data Set

We are going to use multiple related dataset to describe the goal of the project. The list of datasets from Kaggle are given below:

1. **Link**: https://www.kaggle.com/limmen/premierleague-league-tables-188889-201617
   File Name: result.csv
   Description: Data set containing the league standings of all the EPL teams from 88/89 - 16/17
   Important Attributes:

Attribute Name	Attribute Description	Attribute Type	Attribute Semantic
year	Year of the EPL Season	Quantitative	Temporal
Team	Name of the soccer Team	Categorical	
A	Goals conceded by the team	Quantitative	Sequential
D	Number of matches drawn	Quantitative	Sequential
F	Goals scored by the team	Quantitative	Sequential
GD	Goal Difference of Each team	Quantitative	Divergent
L	Number of matches lost	Quantitative	Sequential

Pos	League Position of the team	Quantitative	Sequential
Pts	Total points of the team	Quantitative	Sequential
W	Number of matches won	Quantitative	Sequential

2. **Link:** https://www.kaggle.com/zaeemnalla/premier-league#stats.csv
   File Name: stats.csv
   Description: Dataset contains each team's stats like wins, losses, yellow & red cards, shots, clean sheets, etc. in every season from 06/07 - 17/18
   Important Attributes:

Attribute Name	Attribute Description	Attribute Type	Attribute Semantic
Team	Name of the soccer Team	Categorical	
wins	Number of matches won	Quantitative	Sequential
losses	Number of matches lost	Quantitative	Sequential
goals	Goals scored by the team	Quantitative	Sequential
att_hd_goal	Goals from headers	Quantitative	Sequential
att_pen_goal	Goals from penalties	Quantitative	Sequential
att_freekick_goal	Goals from freekicks	Quantitative	Sequential
att_ibox_goal	Goals from inside the box	Quantitative	Sequential
att_obox_goal	Goals from outside the box	Quantitative	Sequential
goal_fastbreak	Goals from the counter attack	Quantitative	Sequential
clean_sheet	Clean Sheets	Quantitative	Sequential
goals_conceded	Goals conceded by the team	Quantitative	Sequential
total_pass	Total passes done by the team	Quantitative	Sequential
total_long_balls	Total long balls by the team		Sequential
total_through_ball	Total through balls by the team	Quantitative	Sequential
total_cross	Total cross by the team	Quantitative	Sequential

3. **Link:** https://www.kaggle.com/thefc17/epl-results-19932018
   File Name: EPL Set.csv
   Description: Dataset contains full time and half-time stats of every premier league match from 93/94 – 17/18
   Important Attributes:

Attribute Name	Attribute Description	Attribute Type	Attribute Semantic
Date	Date of the match	Quantitative	Temporal
HomeTeam	Name of the Home Team	Categorical	
AwayTeam	Name of the Away Team	Categorical	
FTR	The full-time result, denoted as 'H' for home team win, 'A' for away team win, or 'D' for draw	Categorical	
Season	The season in which the match was played	Ordinal	

Analytical Questions and Proxy Tasks

1. Question: What is the distribution of teams over different seasons?
   Proxy Task 1: What is the distribution of teams who have ever been in top 4 over different seasons based on league positions?
   Proxy Value 1: Distribution -> League Position (Pos)
   teams -> Team
   Seasons -> year
   Proxy Task 2: What is the distribution of the teams who have ever been in the top 4 over different seasons based on league points?
   Proxy Value 2: Distribution -> League points (Pts)
   teams -> Team
   Seasons -> year

2. Question: Does the type of the result influence the league positions of the teams?
   Proxy Task: Does the proportion of win, draw and loss influence the league positions of the teams who have ever been in top 4 over different seasons?
   Proxy Value: type of result -> Wins(W), Draw(D), Loss(L)
   League Positions -> Pos
   teams -> Team
   seasons -> year

3. Question: Does the playing style of the teams affect the league positions?
   Proxy Task1: Does the proportion of goals scored, and goals conceded influence the league positions of the teams who have ever been in top 4 over different seasons?
   Proxy Value1: playing style -> goals scored(F), goals conceded(A)
   league positions -> Pos
   teams -> Team
   seasons -> year

   Proxy Task2: Does the proportion of total pass, total through balls, total long balls and total cross influence the league positions of the teams who have ever been in top 4 over different seasons?

Proxy Value2: playing style -> total pass(total_pass), total through balls (total_through_ball), total long balls(total_long_balls), total Cross(total_cross)
league positions -> Pos
teams -> Team
seasons -> year

Proxy Task3: Does the number of clean sheets influence the league positions of the teams who have ever been in top 4 over the different seasons?
Proxy Value3: playing style -> clean sheets
league positions -> Pos
teams -> Team
seasons -> year

Proxy Task4: Does the proportion of goal scored by headers, penalties, freekicks, inside & outside the box play influence the league positions of the teams who have ever been in top 4 over the different seasons?
Proxy Value4: playing style -> goals scored by headers(att_hd_goal), goal scored by penalties(att_pen_goal), goals scored by freekicks(att_freekick_goal), goals scored from play (inside & outside the box play)
league positions -> Pos
teams -> Team
seasons -> year

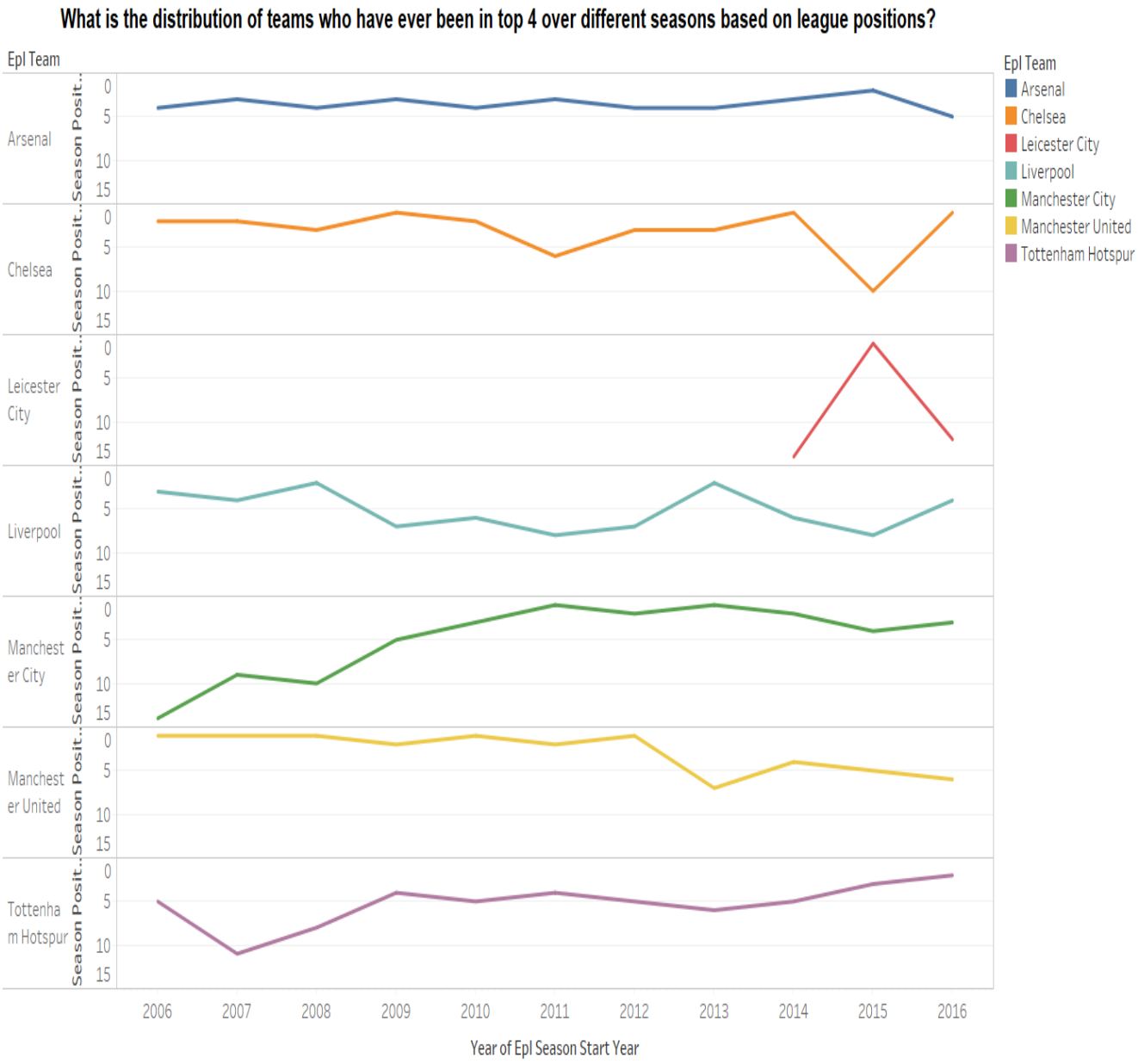
Proxy Task5: Does the number of goals scored through counter attacks influence the league positions of the teams who have ever been in top 4 over different seasons?
Proxy Value5: playing style -> goals scored through counter attacks (goal_fastbreak)
league position -> Pos
teams -> Team
seasons -> year

4. Question: Does the distribution of overall team form affect the league positions?
Proxy Task: Does the proportion of home and away form influence the league positions of the teams who have ever been in top 4 over different seasons?
Proxy Value: overall team form -> Home and Away Form (FTR)
league positions -> Pos
teams -> Team
seasons -> year

Story Design

Data Analysis

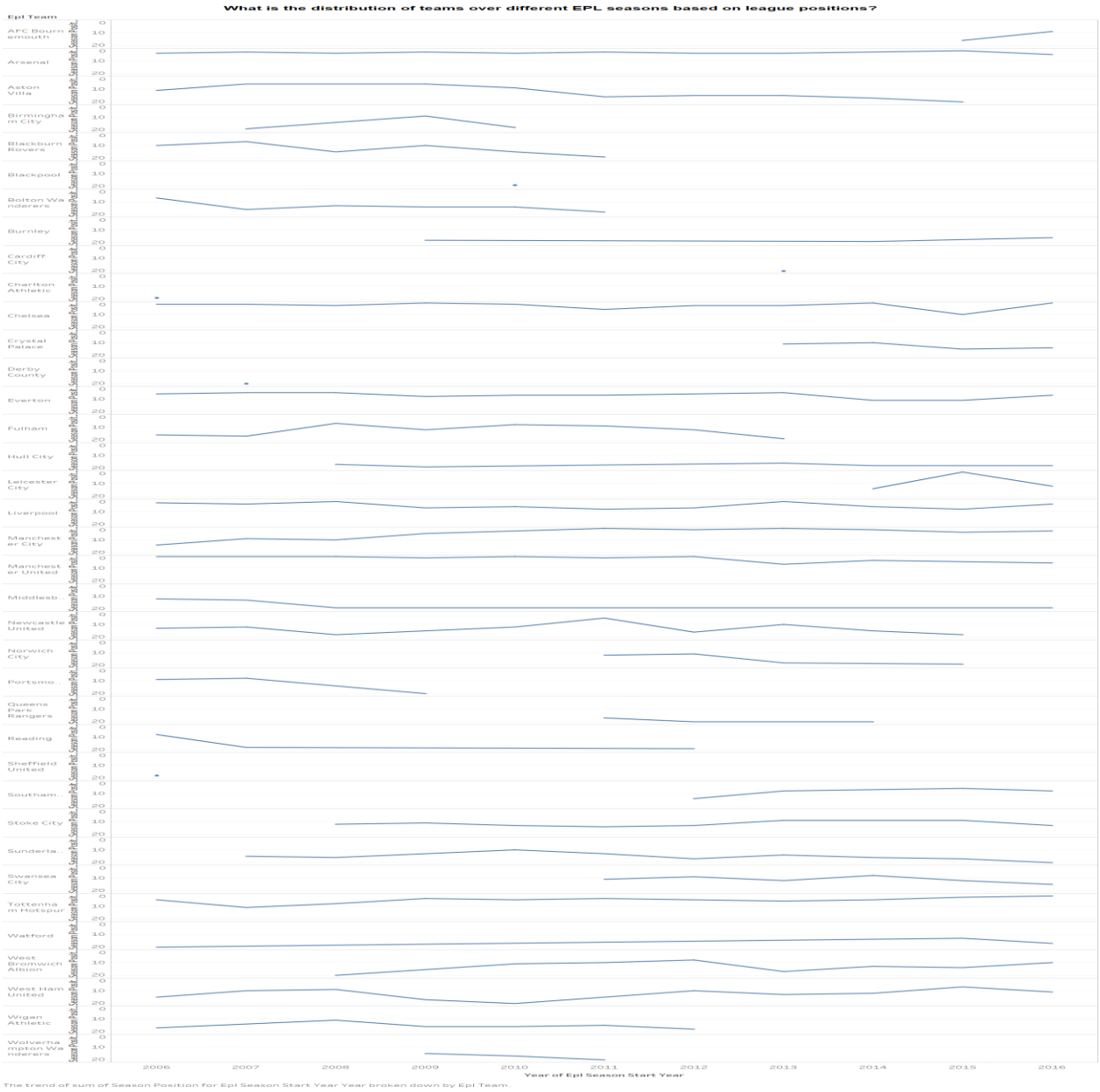
1. <u>Question:</u> What is the distribution of teams over different seasons?
   <u>Proxy Task 1</u>: What is the distribution of teams who have ever been in top 4 over different seasons based on league positions?
   <u>Visualization:</u>



The trend of sum of Season Position for Epl Season Start Year Year broken down by Epl Team. Color shows details about Epl Team. The view is filtered on Epl Team, which keeps 7 of 37 members.

Proxy Task 2: What is the distribution of teams over different seasons based on league positions?

Visualization:



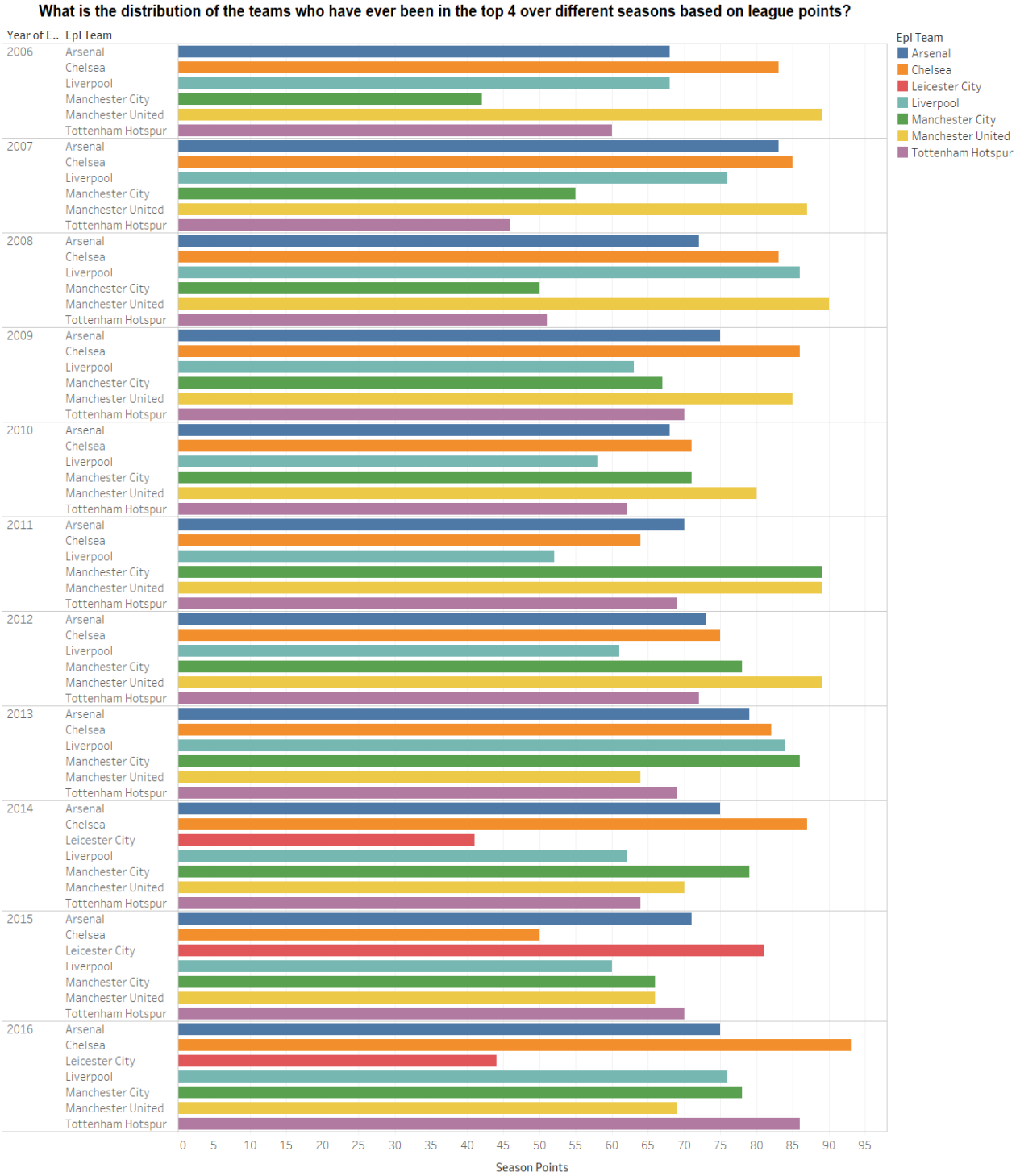
Visualization Description: The above line charts show how the league positions of teams change over different seasons starting from 2006/2007 – 2016/2017

Important Insights:

1. Arsenal have consistently been in the top four over different seasons
2. Leicester City got promoted to EPL in 2014/2015 season and won the league in 2015/2016 season
3. Manchester City have shown consistent form after 2009/2010 season and have been in top four since then

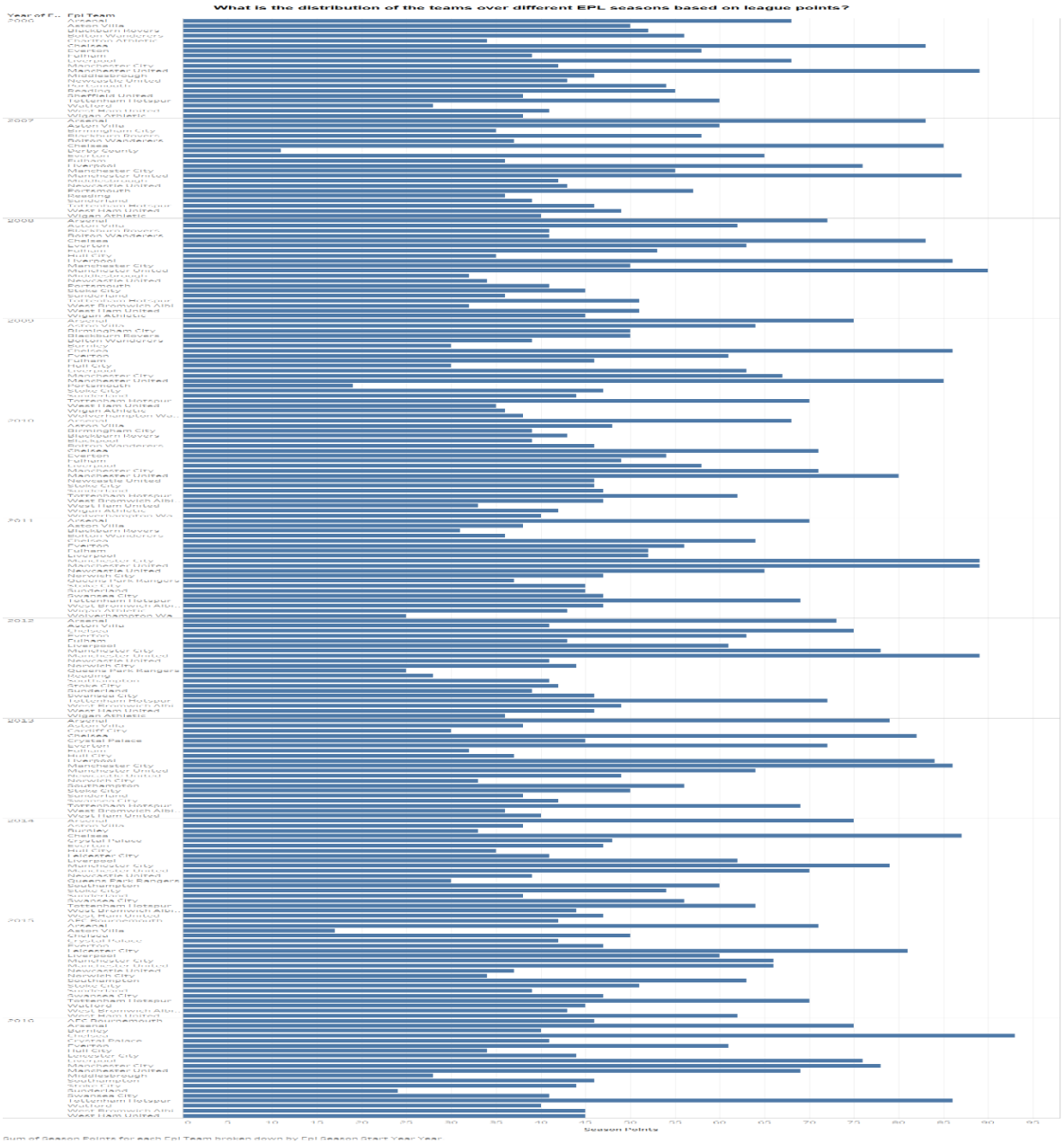
4. Manchester United have also shown a drop-in form after the 2012/2013 season

<u>Proxy Task 3</u>: What is the distribution of the teams who have ever been in the top 4 over different seasons based on league points?
<u>Visualization</u>:



Proxy Task 4: What is the distribution of the teams over different seasons based on league points?

Visualization:



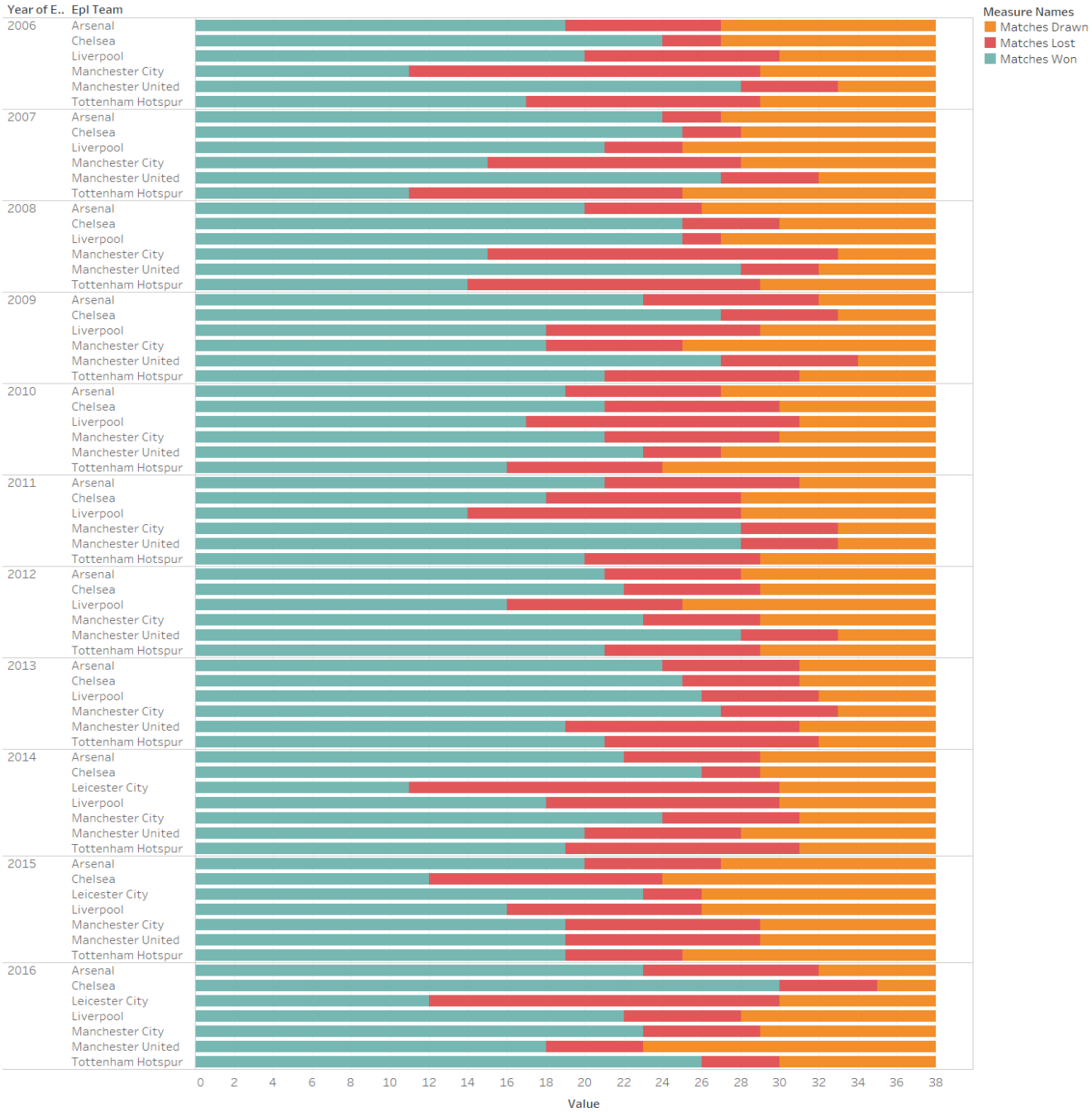
Visualization Description: The above bar charts show how the league points of teams change over different seasons starting from 2006/2007 – 2016/2017

Important Insights:

1. Manchester United's League Point tally was the best till 2012/2013 season after which it dropped.
2. Leicester City's league point tally significantly improved in the season 2015/2016 when they won the league.

2. <u>Question</u>: Does the type of the result influence the league positions of the teams?
   <u>Proxy Task 1</u>: Does the proportion of win, draw and loss influence the league positions of the teams who have ever been in top 4 over different seasons?
   <u>Visualization</u>:

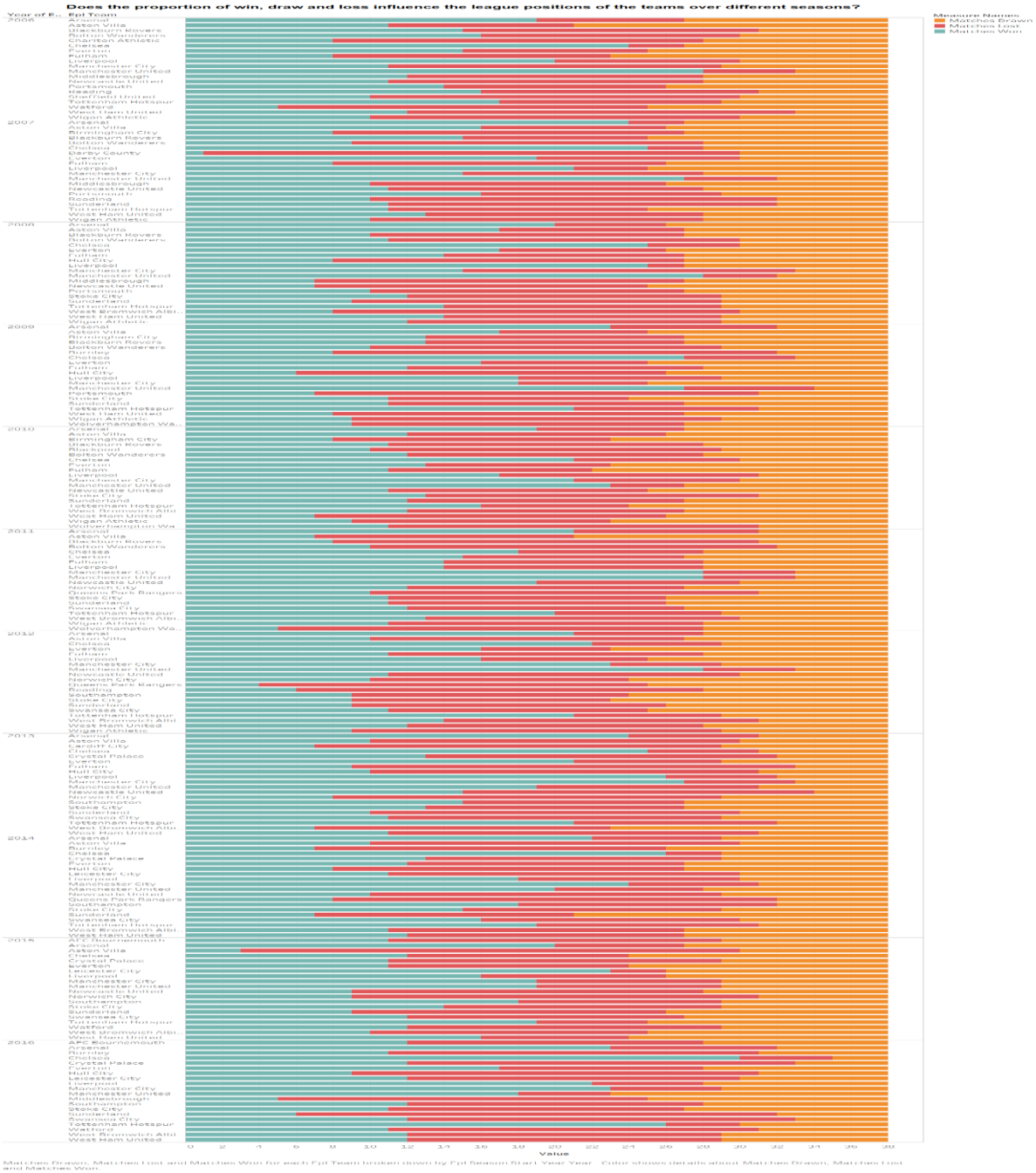
Does the proportion of win, draw and loss influence the league positions of the teams who have ever been in top 4 over different seasons?



Matches Drawn, Matches Lost and Matches Won for each Epl Team broken down by Epl Season Start Year Year. Color shows details about Matches Drawn, Matches Lost and Matches Won. The view is filtered on Epl Team, which keeps 7 of 37 members.

Proxy Task 2: Does the proportion of win, draw and loss influence the league positions of the teams over different seasons?

Visualization:

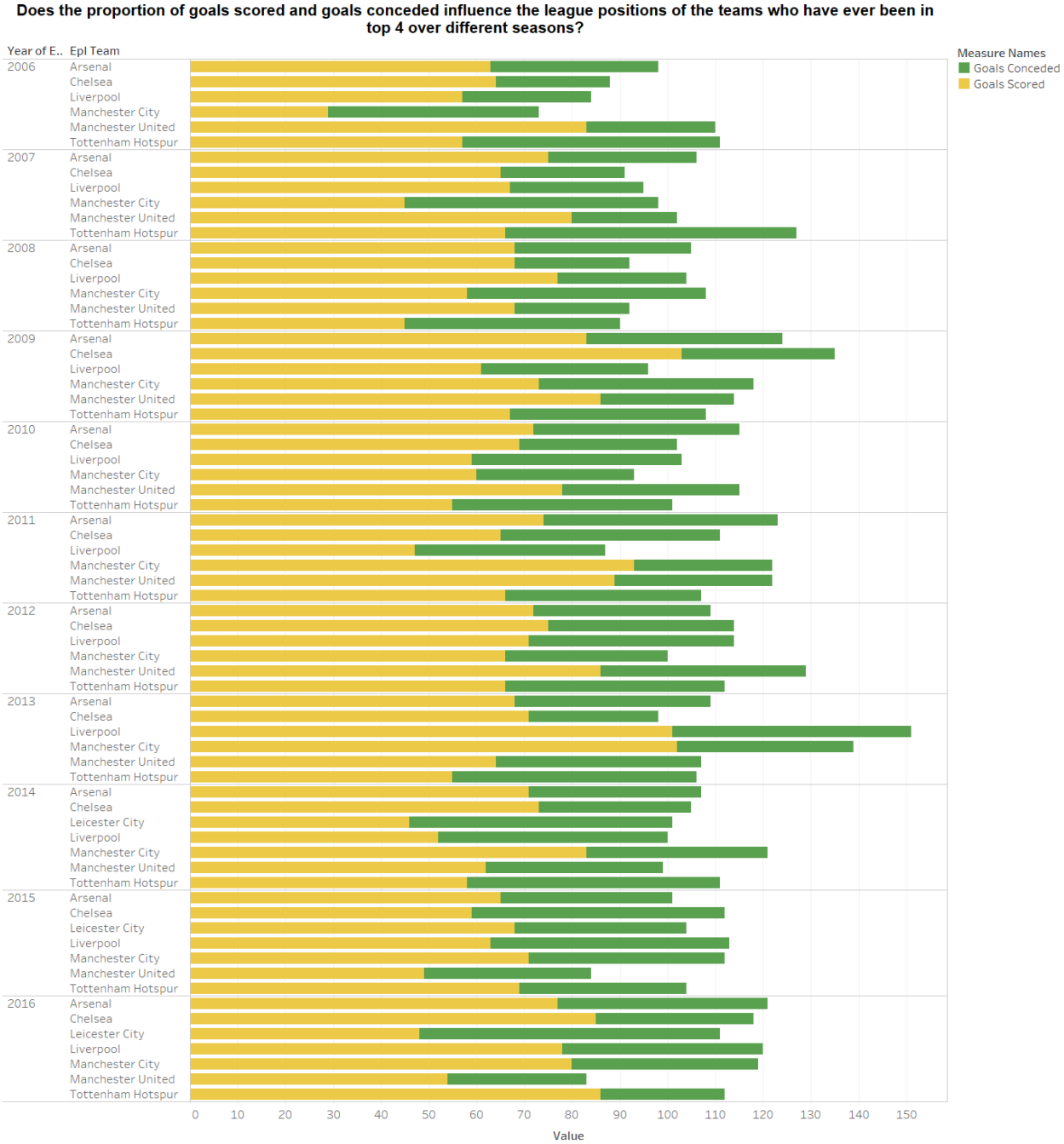


Visualization Description: The above stacked bar charts show how the proportion of win, draw and loss of teams change over different seasons starting from 2006/2007 – 2016/2017

Important Insight:

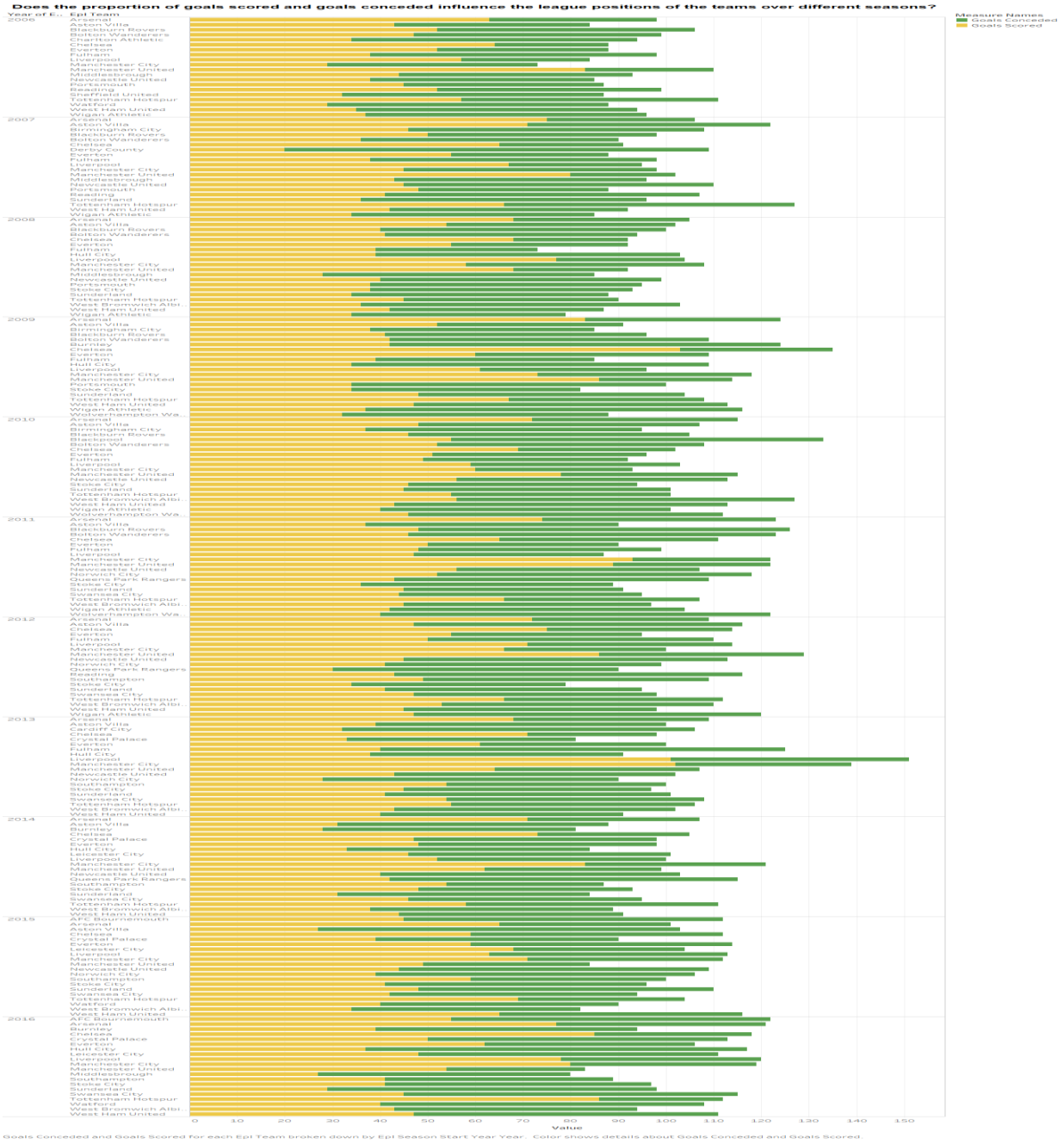
When we consider the teams in the top four, the number of wins influence the league position of the teams over different seasons more rather than the proportion of number of wins and number of draws

3. <u>Question:</u> Does the playing style of the teams affect the league positions?
   <u>Proxy Task 1:</u> Does the proportion of goals scored, and goals conceded influence the league positions of the teams who have ever been in top 4 over different seasons?
   <u>Visualization:</u>



Goals Conceded and Goals Scored for each Epl Team broken down by Epl Season Start Year Year. Color shows details about Goals Conceded and Goals Scored. The view is filtered on Epl Team, which keeps 7 of 37 members.

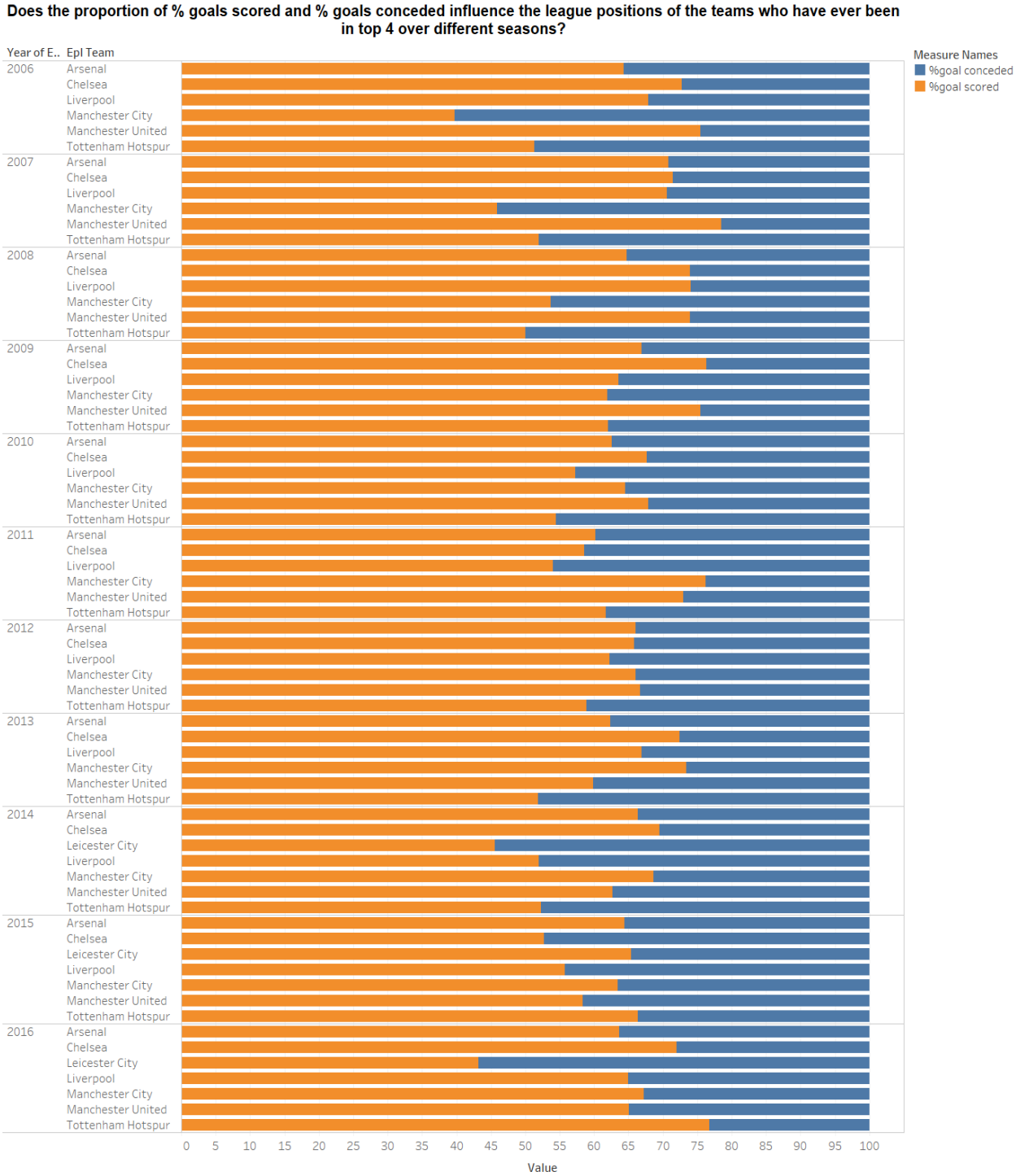
Proxy Task 2: Does the proportion of goals scored, and goals conceded influence the league positions of the teams over different seasons?
Visualization:



Visualization Description: The above stacked bar charts show how the proportion of goals scored and goals conceded of teams change over different seasons starting from 2006/2007 – 2016/2017
Important Insights:
The good proportion of goals scored to goals conceded i.e. high number of goals scored and less goals conceded influence the league position of the teams over different seasons. This is more evident from the next graph.

Proxy Task 3: Does the proportion of % goals scored and % goals conceded influence the league positions of the teams who have ever been in top 4 over different seasons?

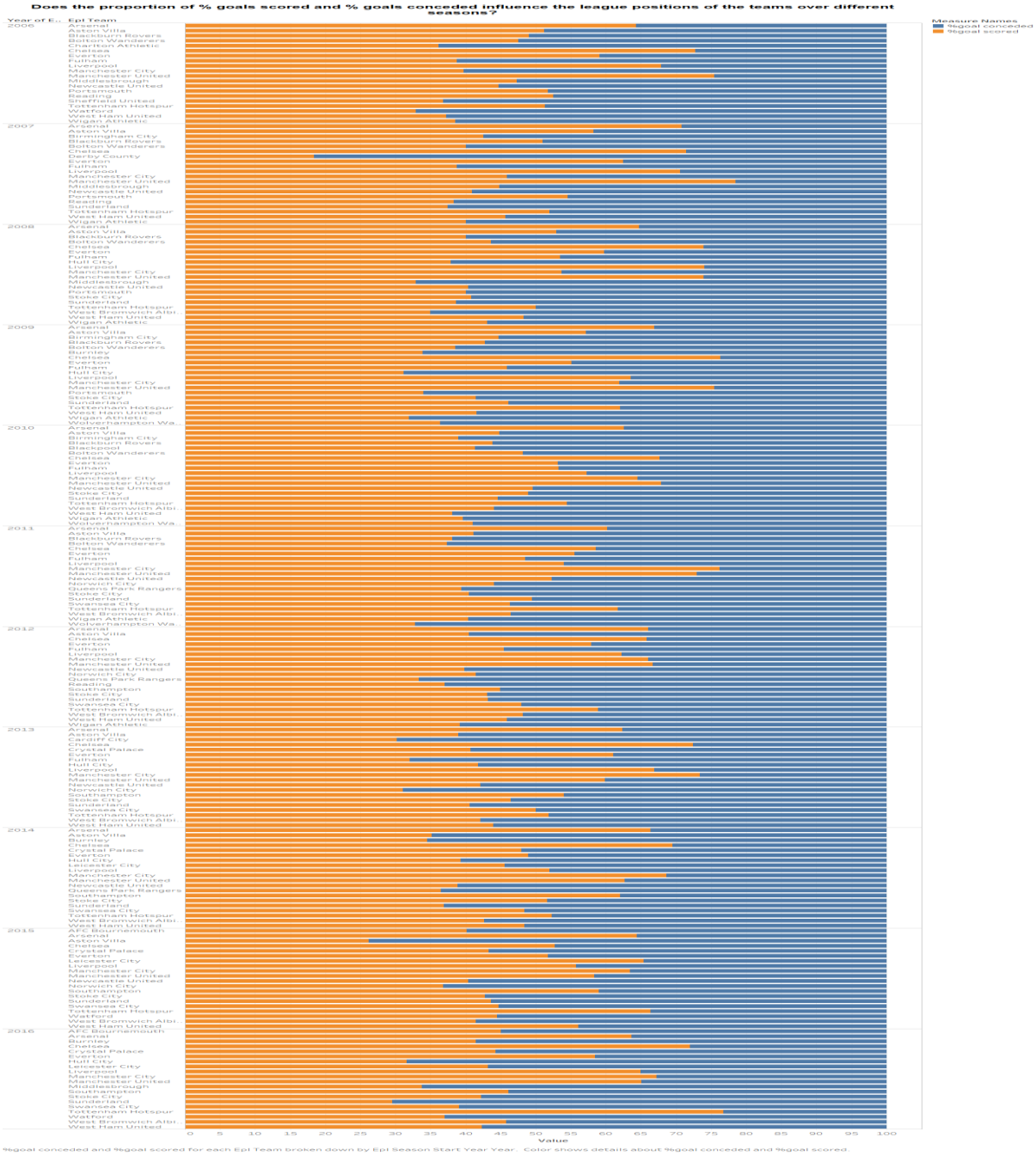
Visualization:



%goal conceded and %goal scored for each Epl Team broken down by Epl Season Start Year Year. Color shows details about %goal conceded and %goal scored. The view is filtered on Epl Team, which keeps 7 of 37 members.

Proxy Task 4: Does the proportion of % goals scored and % goals conceded influence the league positions of the teams over different seasons?

Visualization:



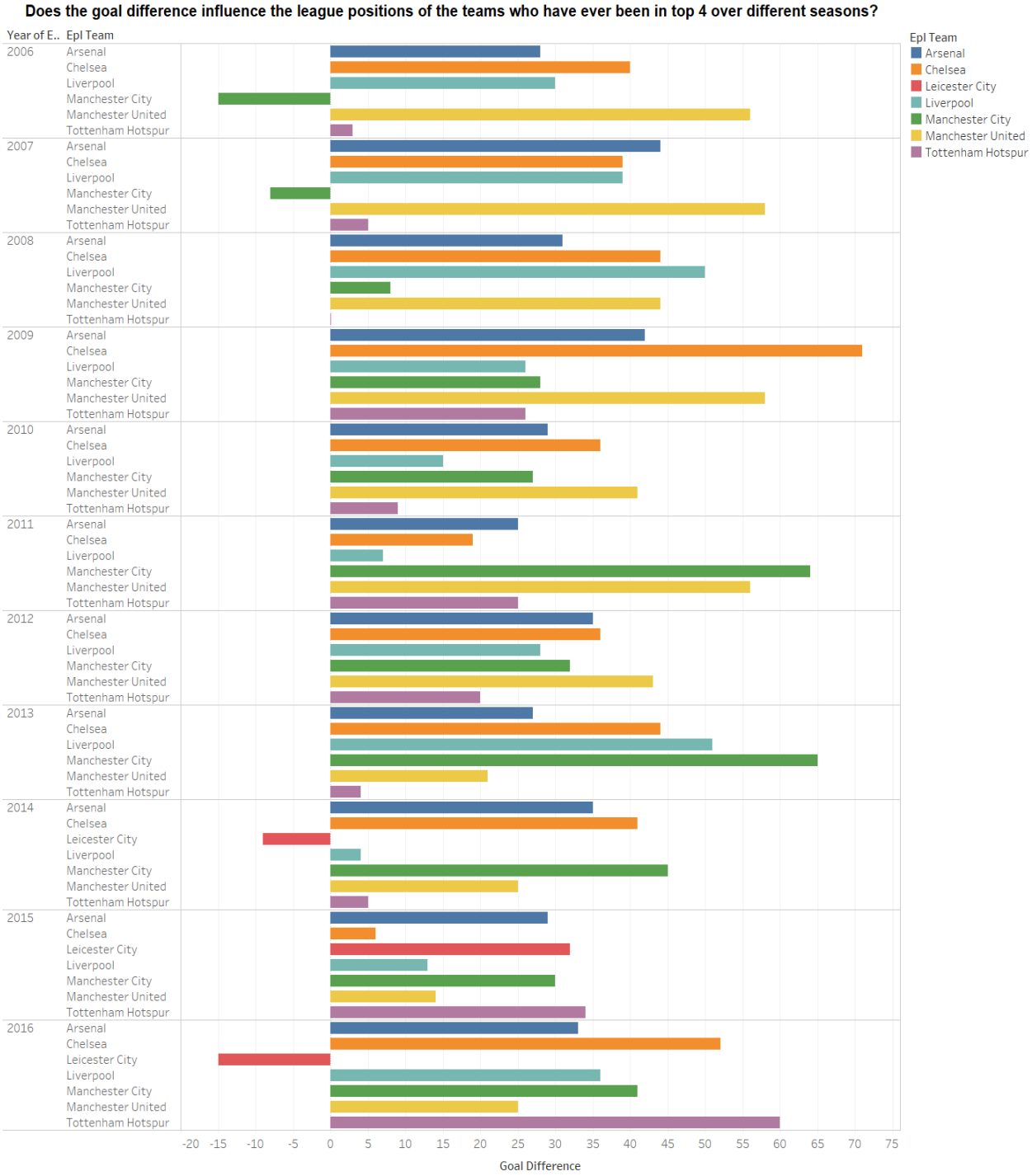
Visualization Description: The above stacked bar charts show how the proportion of % goals scored and % goals conceded of teams change over different seasons starting from 2006/2007 – 2016/2017

Important Insights:

The good proportion of goals scored to goals conceded i.e. high number of goals scored and less goals conceded influence the league position of the teams over different seasons.

Proxy Task 5: Does the goal difference influence the league positions of the teams who have ever been in top 4 over different seasons?

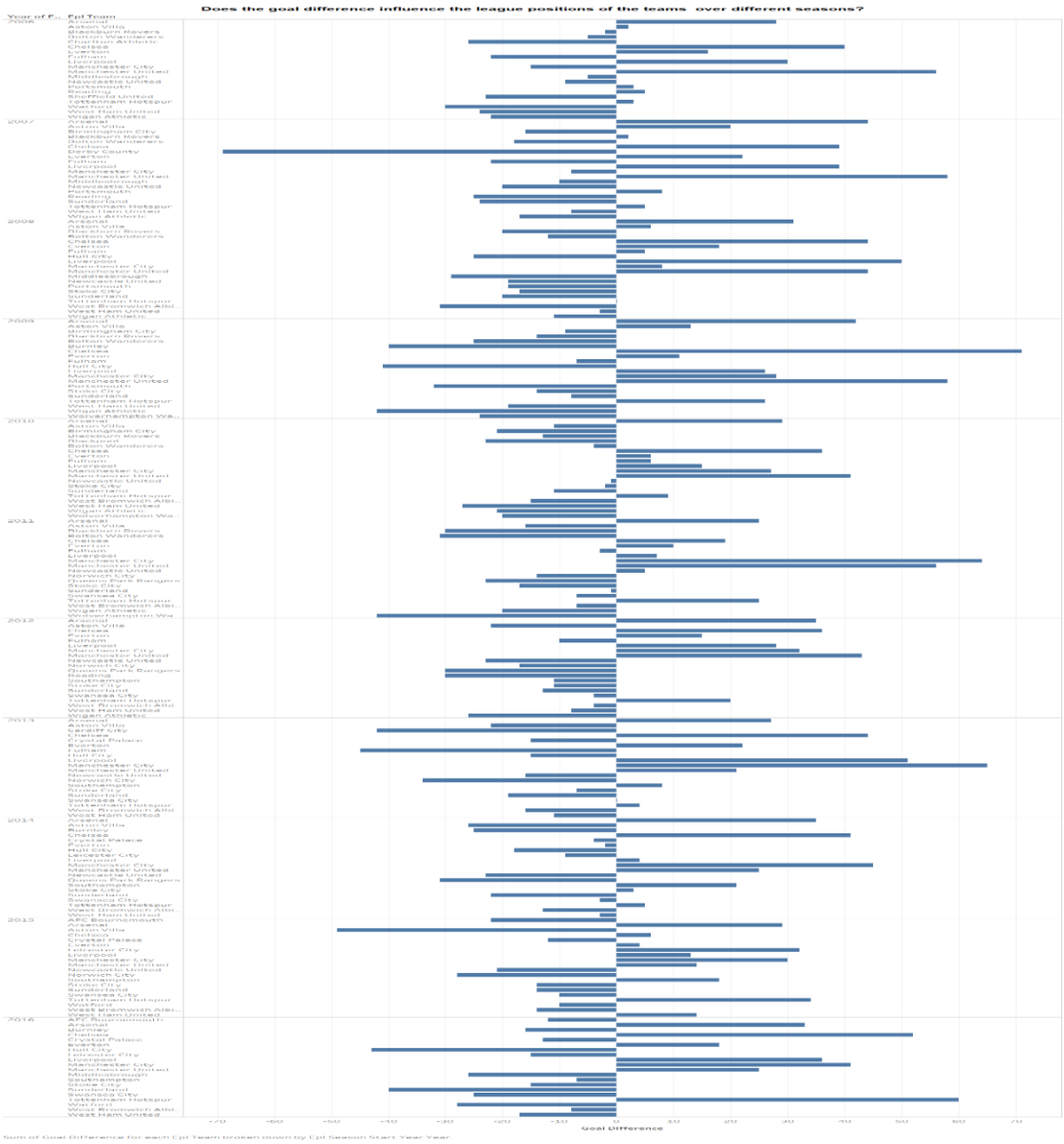
Visualization:



Sum of Goal Difference for each Epl Team broken down by Epl Season Start Year Year. Color shows details about Epl Team. The view is filtered on Epl Team, which keeps 7 of 37 members.

Proxy Task 6: Does the goal difference influence the league positions of the teams over different seasons?

Visualization:



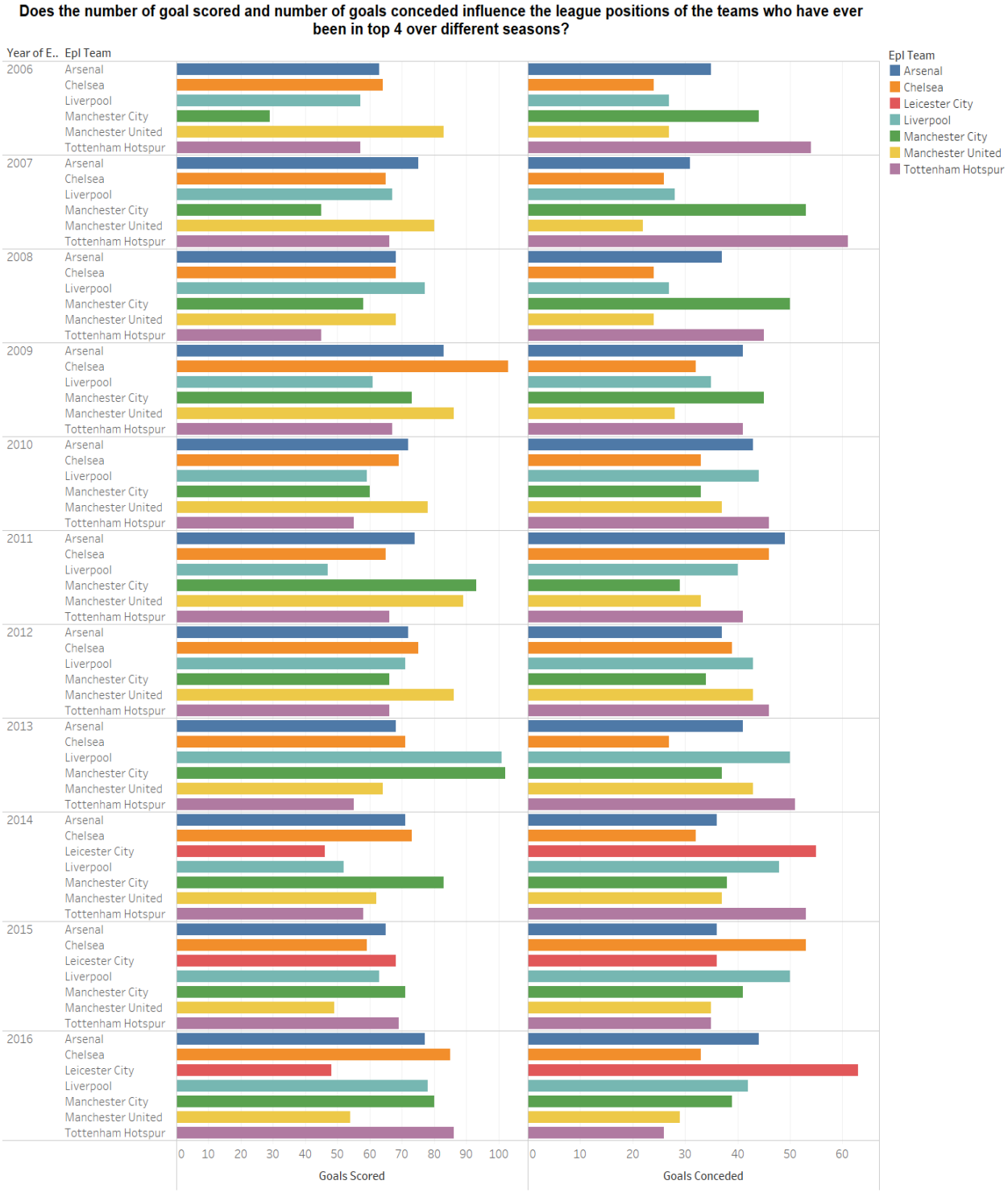
Visualization Description: The above bar charts shows how the goal difference of teams change over different seasons starting from 2006/2007 – 2016/2017

Important Insights:

1. A good goal difference influences the league position as you can see a range of goal difference values (from negative to positive values)
2. Manchester City had a negative goal difference in season 2006/2007 and 2007/2008 and improved on their goal difference since then which helped them improve their league position.
3. Leicester city had a negative goal difference since getting promoted in EPL except for the season 2015/2016 when the won the league (i.e. were present in the top four)

Proxy Task 7: Does the number of goals scored, and number of goals conceded influence the league positions of the teams who have ever been in top 4 over different seasons?

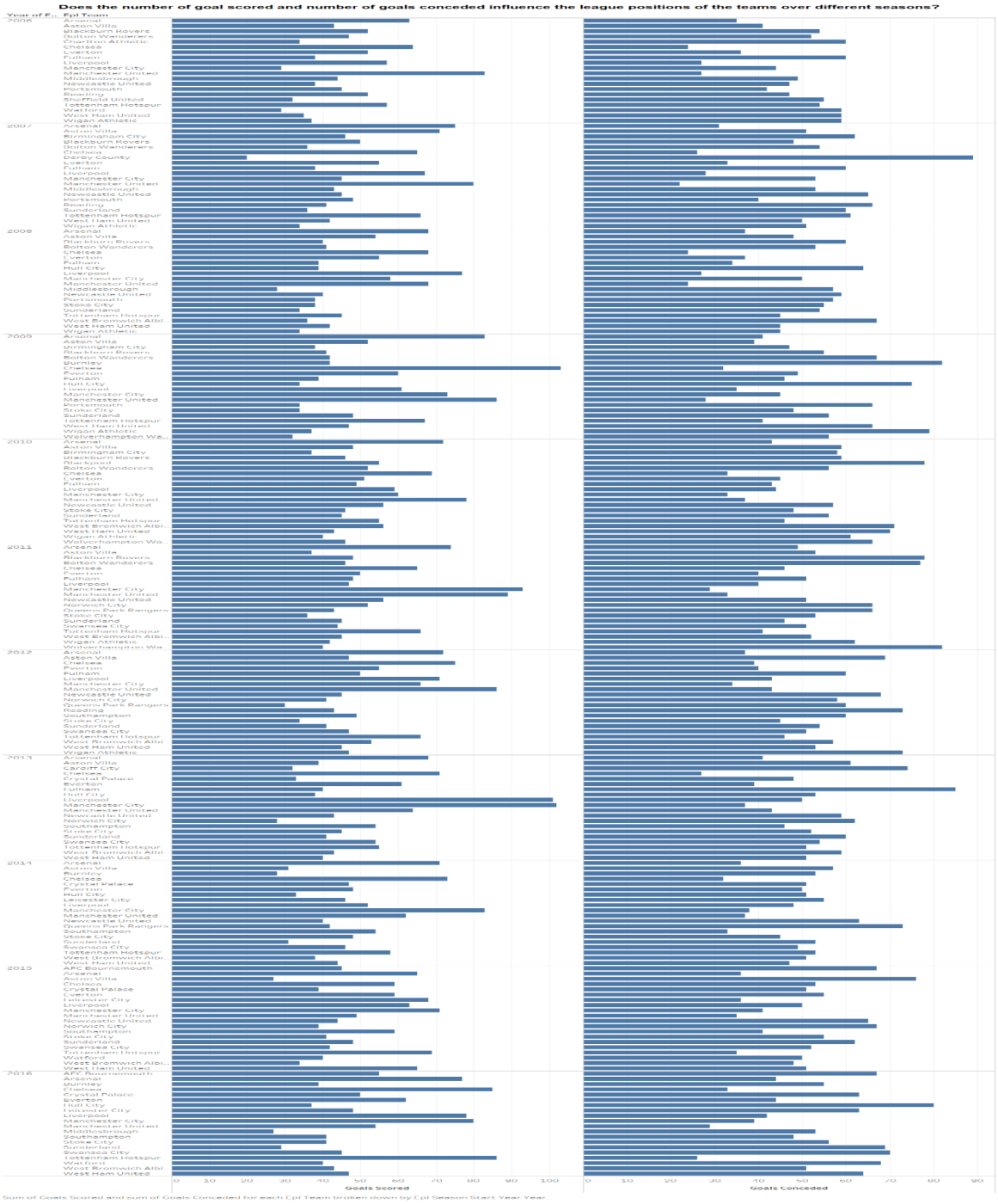
Visualization:



Sum of Goals Scored and sum of Goals Conceded for each Epl Team broken down by Epl Season Start Year Year. Color shows details about Epl Team. The view is filtered on Epl Team, which keeps 7 of 37 members.

Proxy Task 8: Does the number of goals scored, and number of goals conceded influence the league positions of the teams over different seasons?

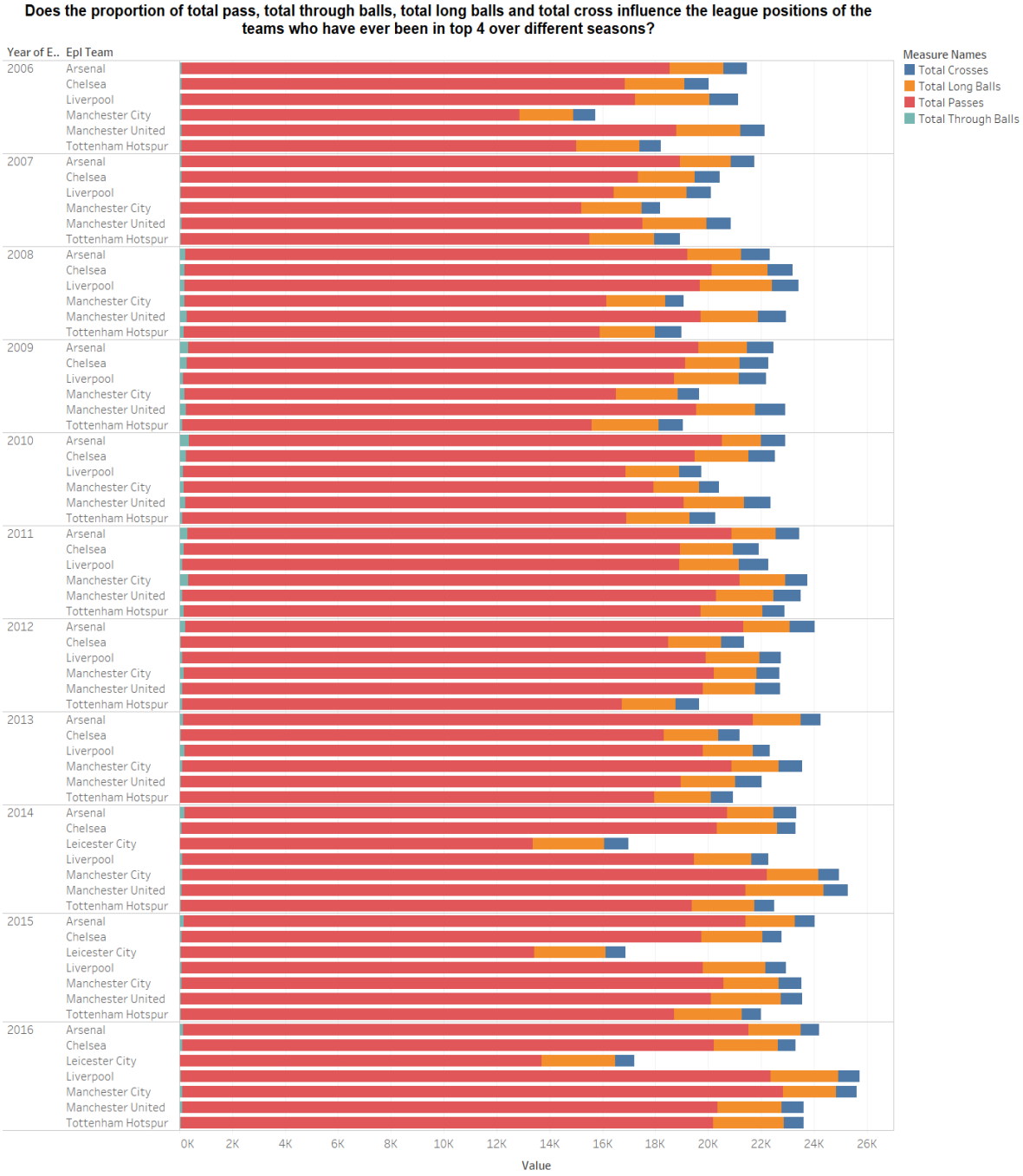
Visualization:



Visualization Description: The above bar charts show how the number of goals scored and number of goals conceded of teams change over different seasons starting from 2006/2007 – 2016/2017

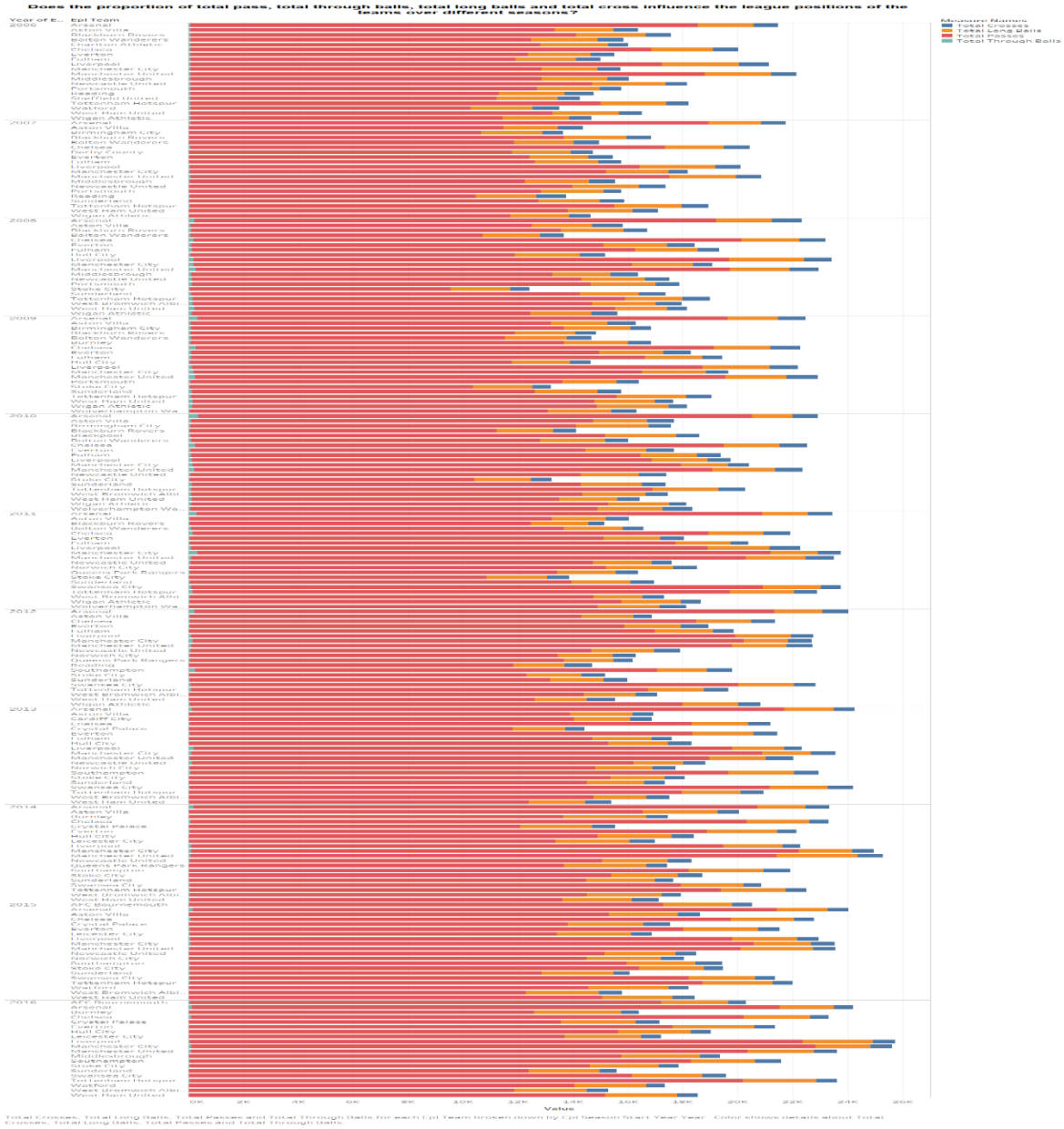
Proxy Task 9: Does the proportion of total pass, total through balls, total long balls and total cross influence the league positions of the teams who have ever been in top 4 over different seasons?

Visualization:



Total Crosses, Total Long Balls, Total Passes and Total Through Balls for each Epl Team broken down by Epl Season Start Year Year. Color shows details about Total Crosses, Total Long Balls, Total Passes and Total Through Balls. The view is filtered on Epl Team, which keeps 7 of 37 members.

Proxy Task 10: Does the proportion of total pass, total through balls, total long balls and total cross influence the league positions of the teams over different seasons?
Visualization:



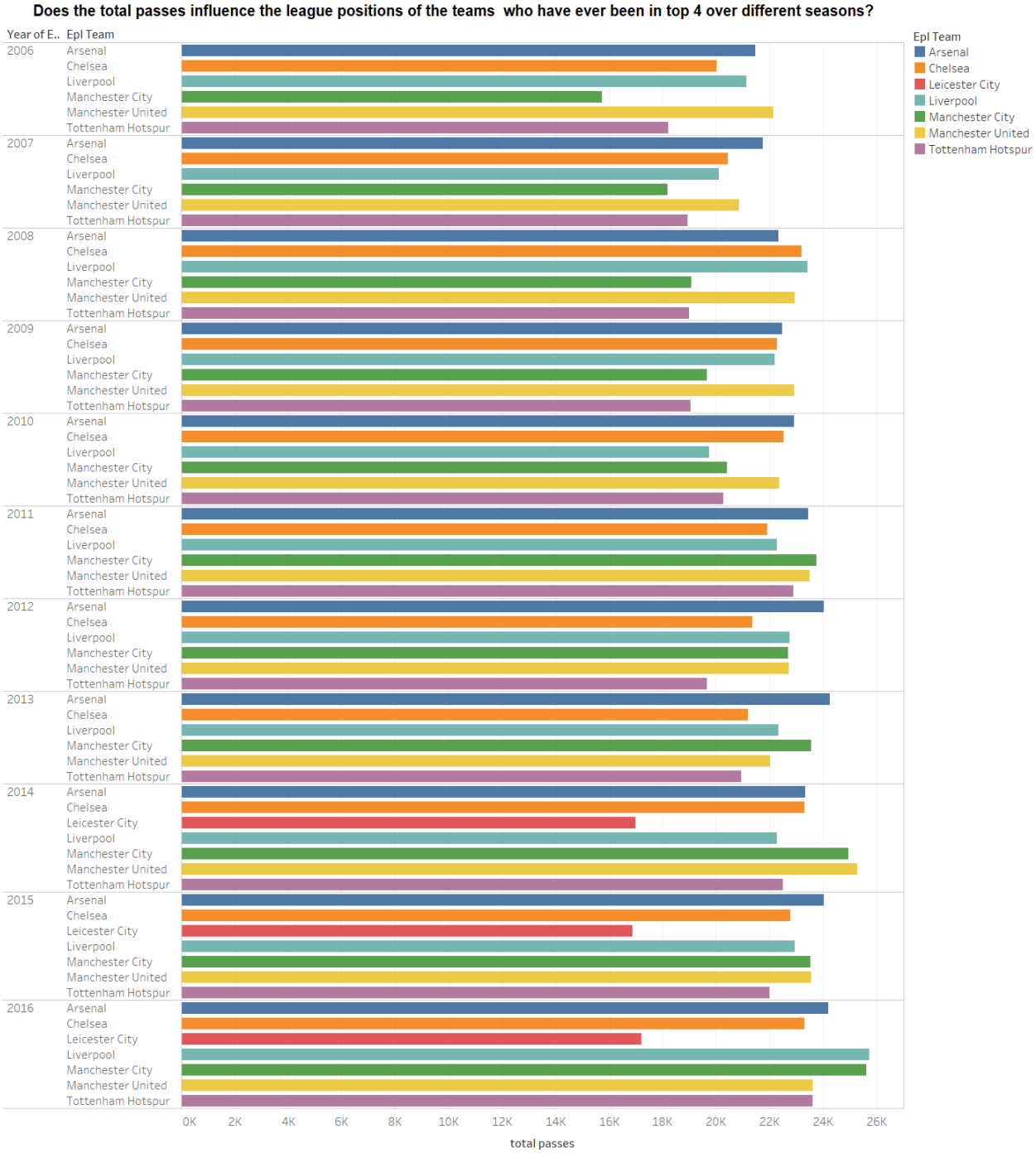
Visualization Description: The above stacked bar charts show how proportion of total pass, total through balls, total long balls and total cross of teams change over different seasons starting from 2006/2007 – 2016/2017

Important Insight:

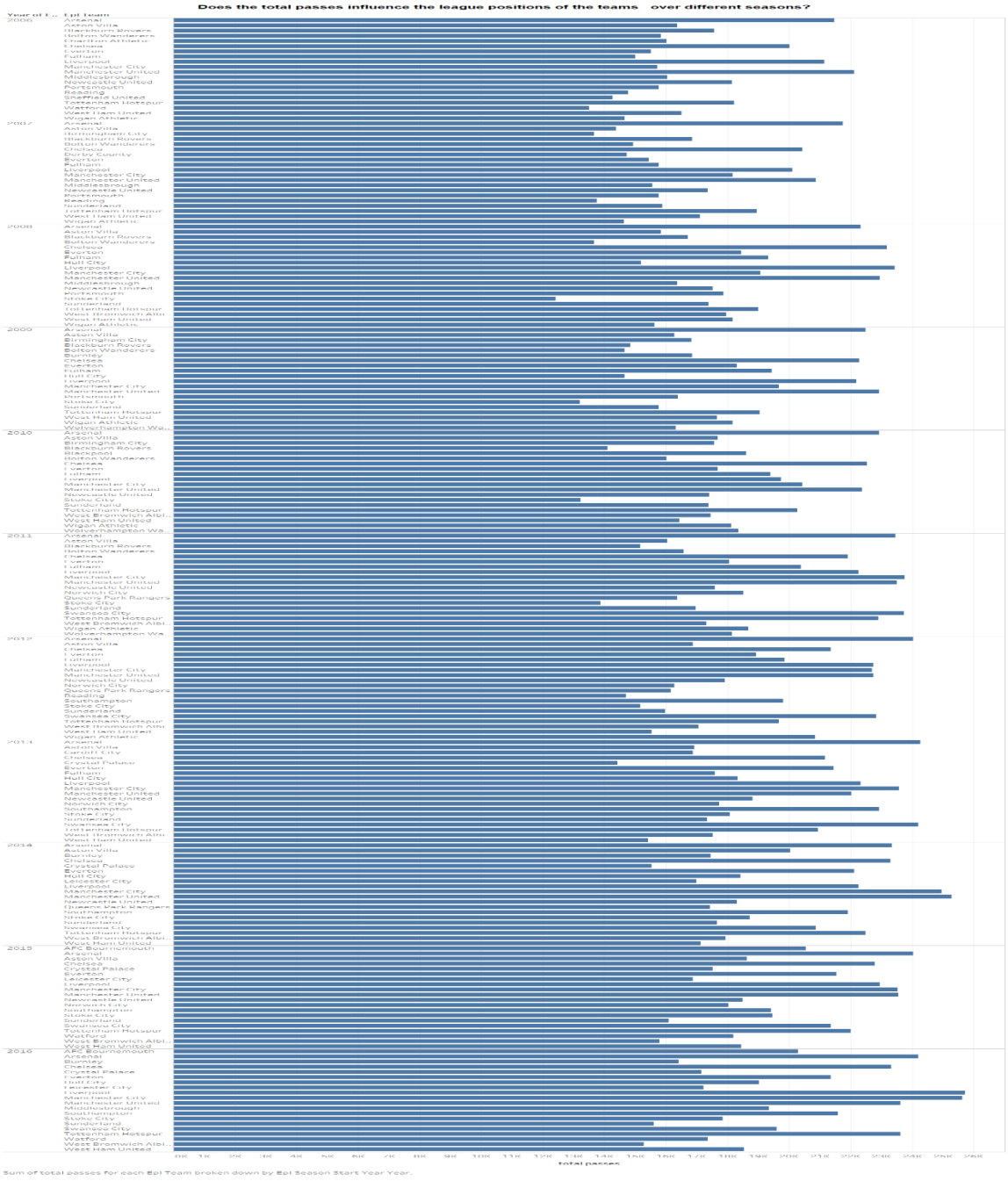
1. The number of short passes (total pass) has greater proportion as compared to number of through balls, long balls and crosses.
2. The proportion of total pass, total through balls, total long balls and total cross influences the league positions of the teams as generally teams placed at higher league positions have a higher proportion values as compared to the lower league position teams

Proxy Task 11: Does the total passes (sum of total pass, total through balls, total long balls and total cross) influence the league positions of the teams who have ever been in top 4 over different seasons?

Visualization:



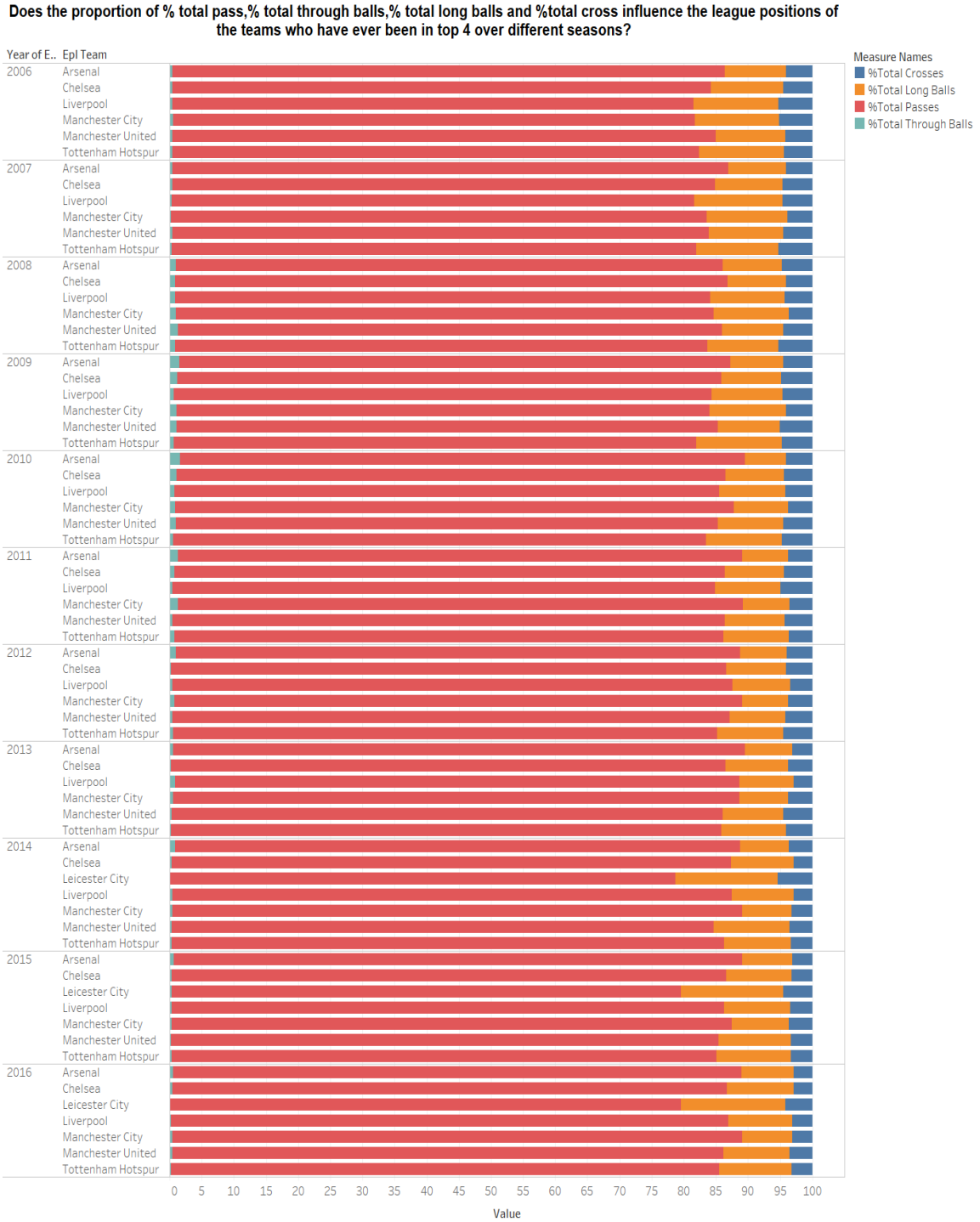
Proxy Task 12: Does the total passes (sum of total pass, total through balls, total long balls and total cross) influence the league positions of the teams over different seasons?
Visualization:



Visualization Description: The above bar charts show how total passes of teams change over different seasons starting from 2006/2007 – 2016/2017
Important Insight:
The total passes influence the league positions of the teams as generally teams placed at higher league positions have a higher value of total passes as compared to the lower league position teams

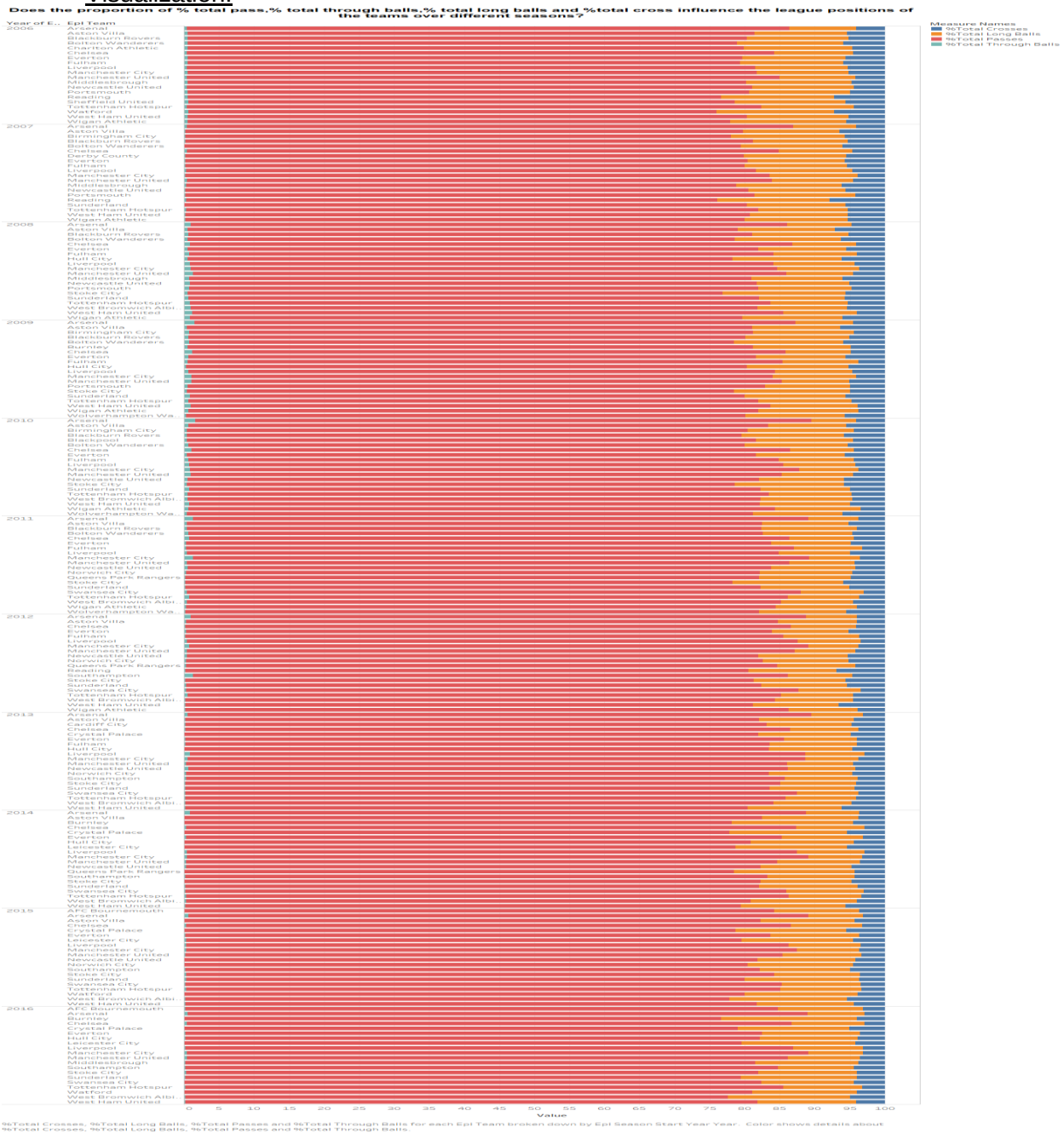
Proxy Task 13: Does the proportion of % total pass, % total through balls, % total long balls and % total cross influence the league positions of the teams who have ever been in top 4 over different seasons?

Visualization:



Proxy Task 14: Does the proportion of % total pass, % total through balls, % total long balls and % total cross influence the league positions of the teams over different seasons?

Visualization:

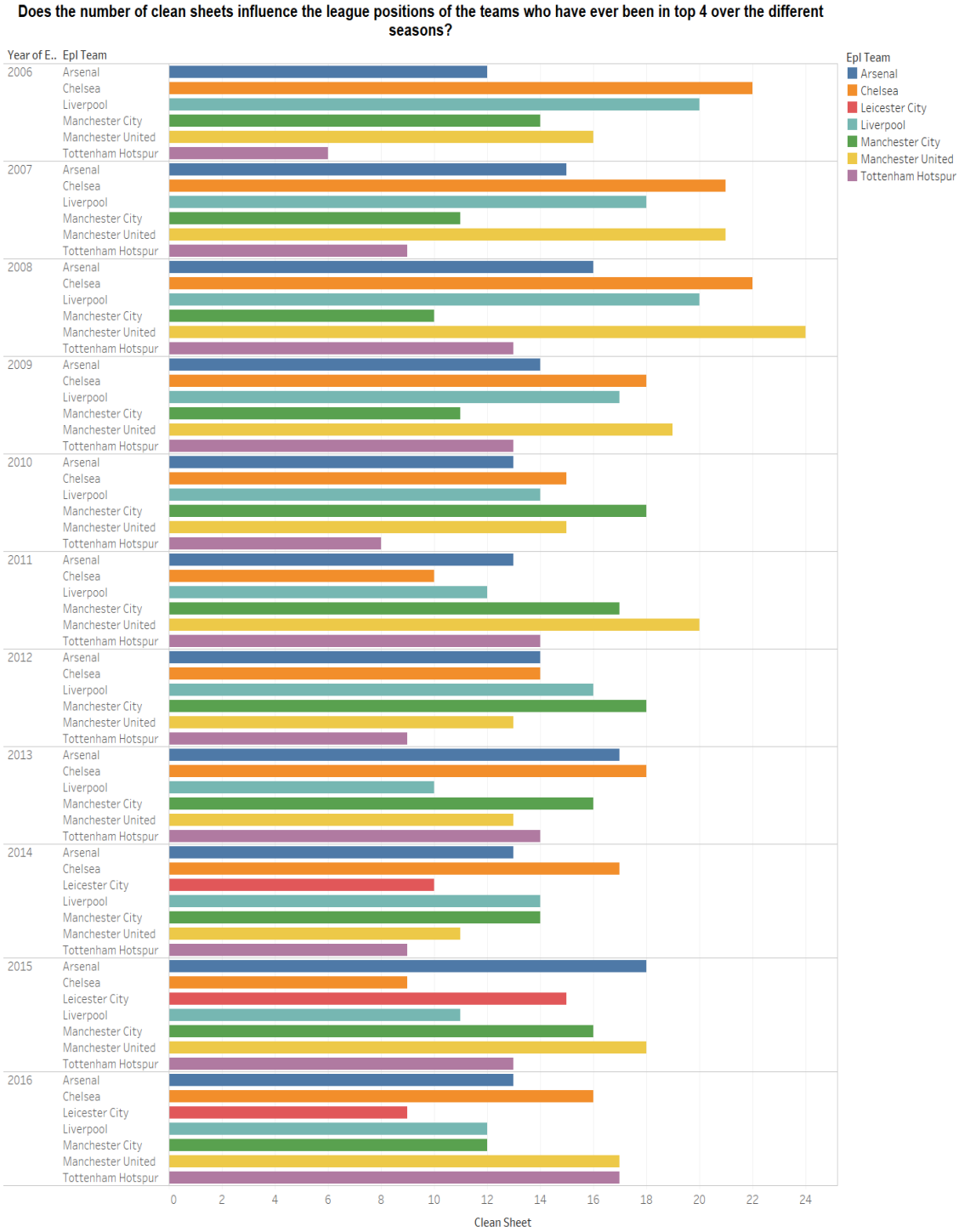


Visualization Description: The above stacked bar chart shows how proportion of % total pass, % total through balls, % total long balls and % total cross of teams changes over different seasons starting from 2006/2007 – 2016/2017

Important Insights:

The number of short passes (total pass) has greater proportion as compared to number of through balls, long balls and crosses.

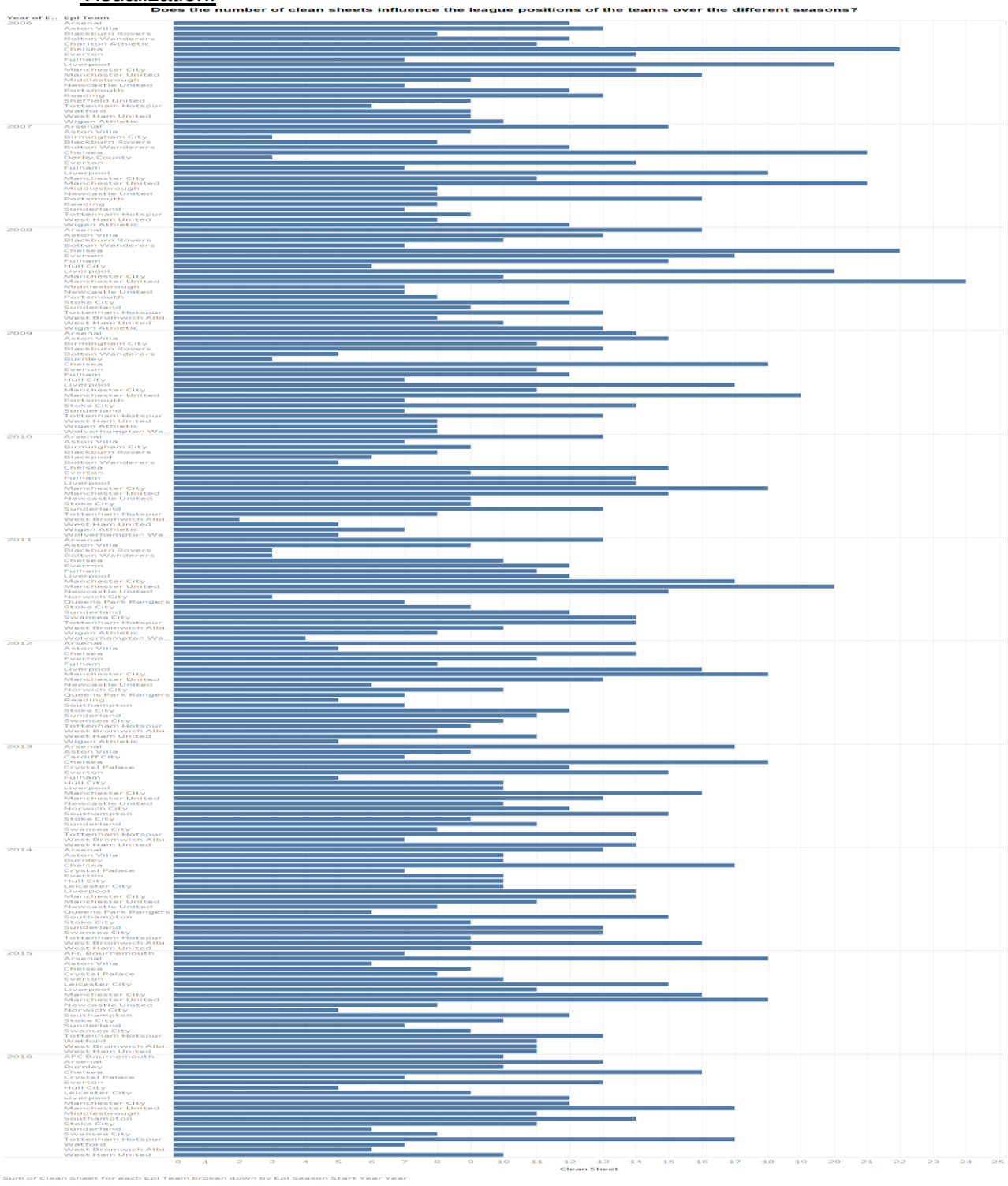
Proxy Task 15: Does the number of clean sheets influence the league positions of the teams who have ever been in top 4 over the different seasons?
Visualization:



Sum of Clean Sheet for each Epl Team broken down by Epl Season Start Year Year. Color shows details about Epl Team. The view is filtered on Epl Team, which keeps 7 of 37 members.

Proxy Task 16: Does the number of clean sheets influence the league positions of the teams over the different seasons?

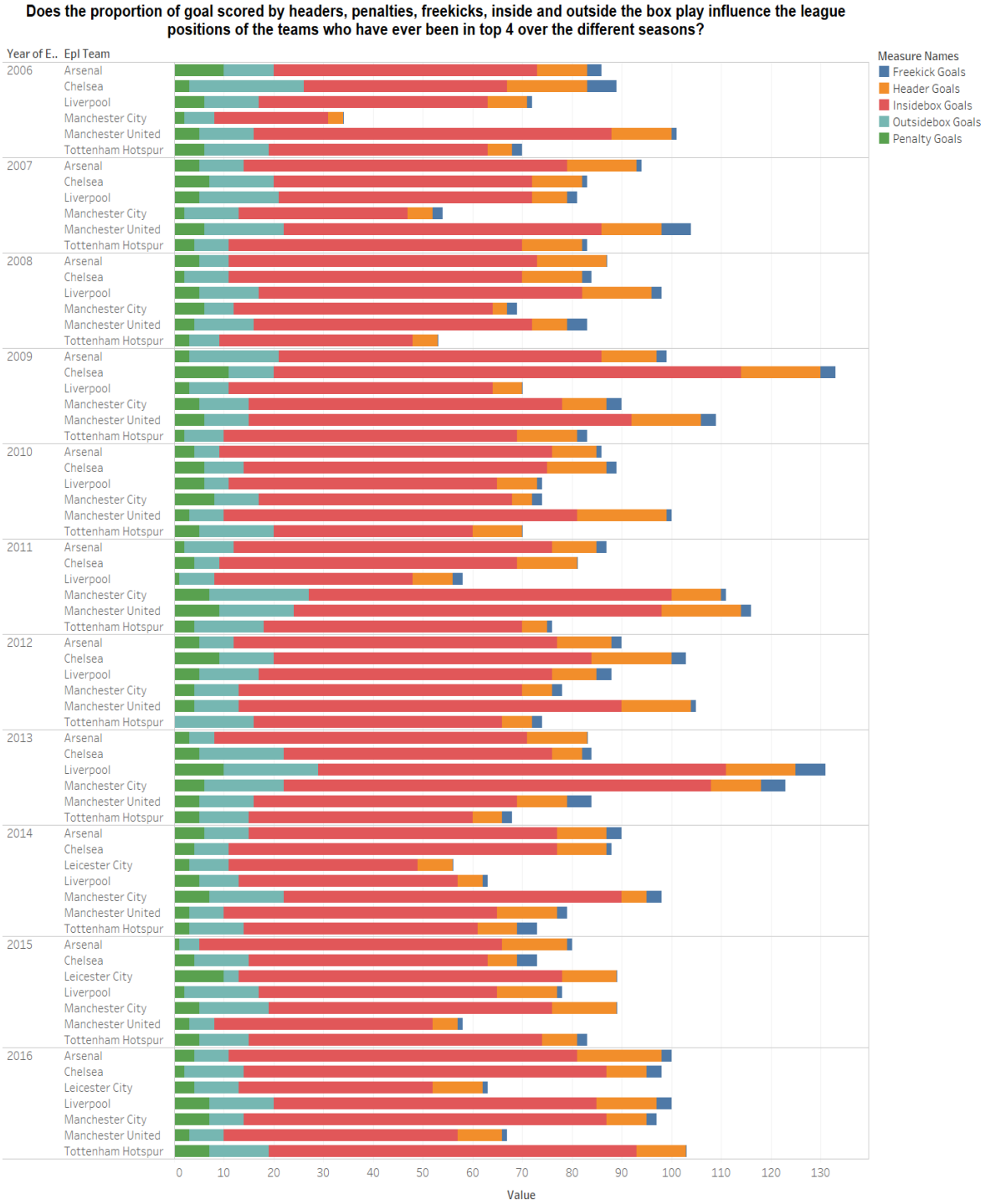
Visualization:



Visualization Description: The above bar charts show how the number of clean sheets of teams change over different seasons starting from 2006/2007 – 2016/2017

Proxy Task 17: Does the proportion of goal scored by headers, penalties, freekicks, inside & outside the box play influence the league positions of the teams who have ever been in top 4 over the different seasons?

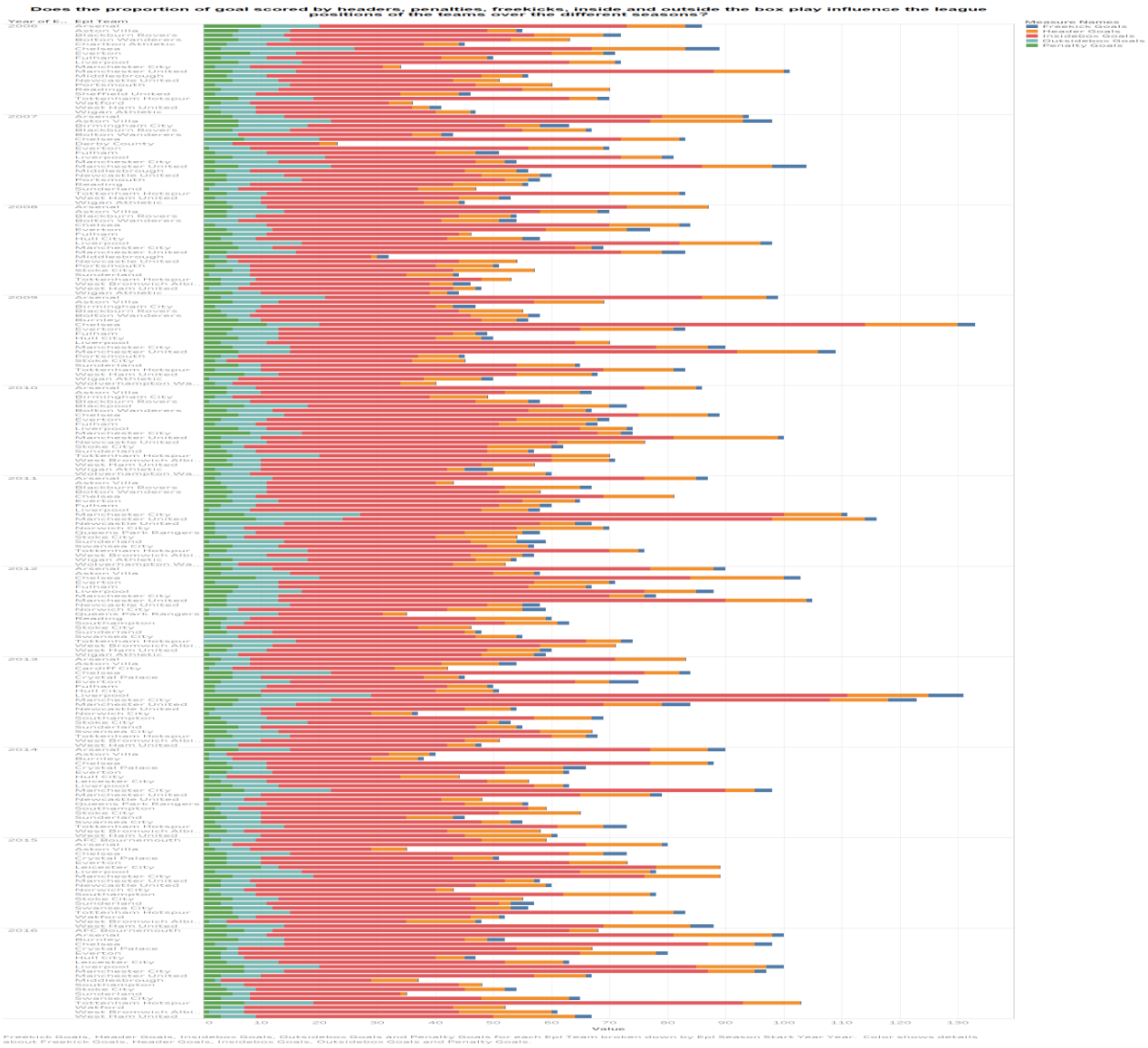
Visualization:



Freekick Goals, Header Goals, Insidebox Goals, Outsidebox Goals and Penalty Goals for each Epl Team broken down by Epl Season Start Year Year. Color shows details about Freekick Goals, Header Goals, Insidebox Goals, Outsidebox Goals and Penalty Goals. The view is filtered on Epl Team, which keeps 7 of 37 members.

Proxy Task 18: Does the proportion of goal scored by headers, penalties, freekicks, inside & outside the box play influence the league positions of the teams over the different seasons?

Visualization:



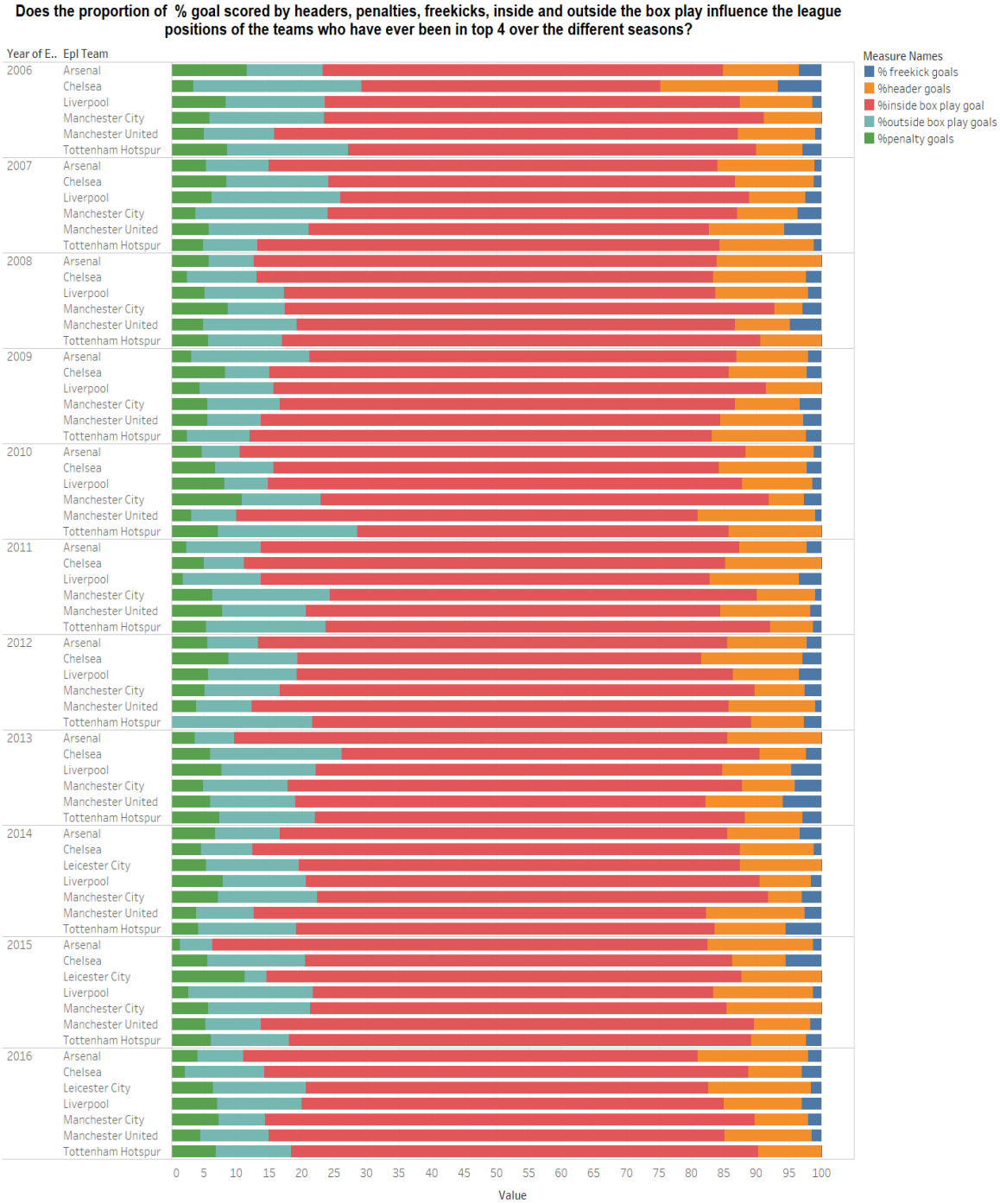
Visualization Description: The above stacked bar charts show how the proportion of goal scored by headers, penalties, freekicks, inside & outside the box play of teams changes over different seasons starting from 2006/2007 – 2016/2017

Important Insights:

1. The number of goals scored by inside box play has greater proportion as compared to number of goals scored by headers, penalties, freekicks and outside the box play
2. The proportion of goals scored by headers, penalties, freekicks, inside & outside the box influences the league positions of the teams as generally teams placed at higher league positions have a higher proportion values as compared to the lower league position teams

Proxy Task 19: Does the proportion of % goal scored by headers, penalties, freekicks, inside and outside the box play influence the league positions of the teams who have ever been in top 4 over the different seasons?

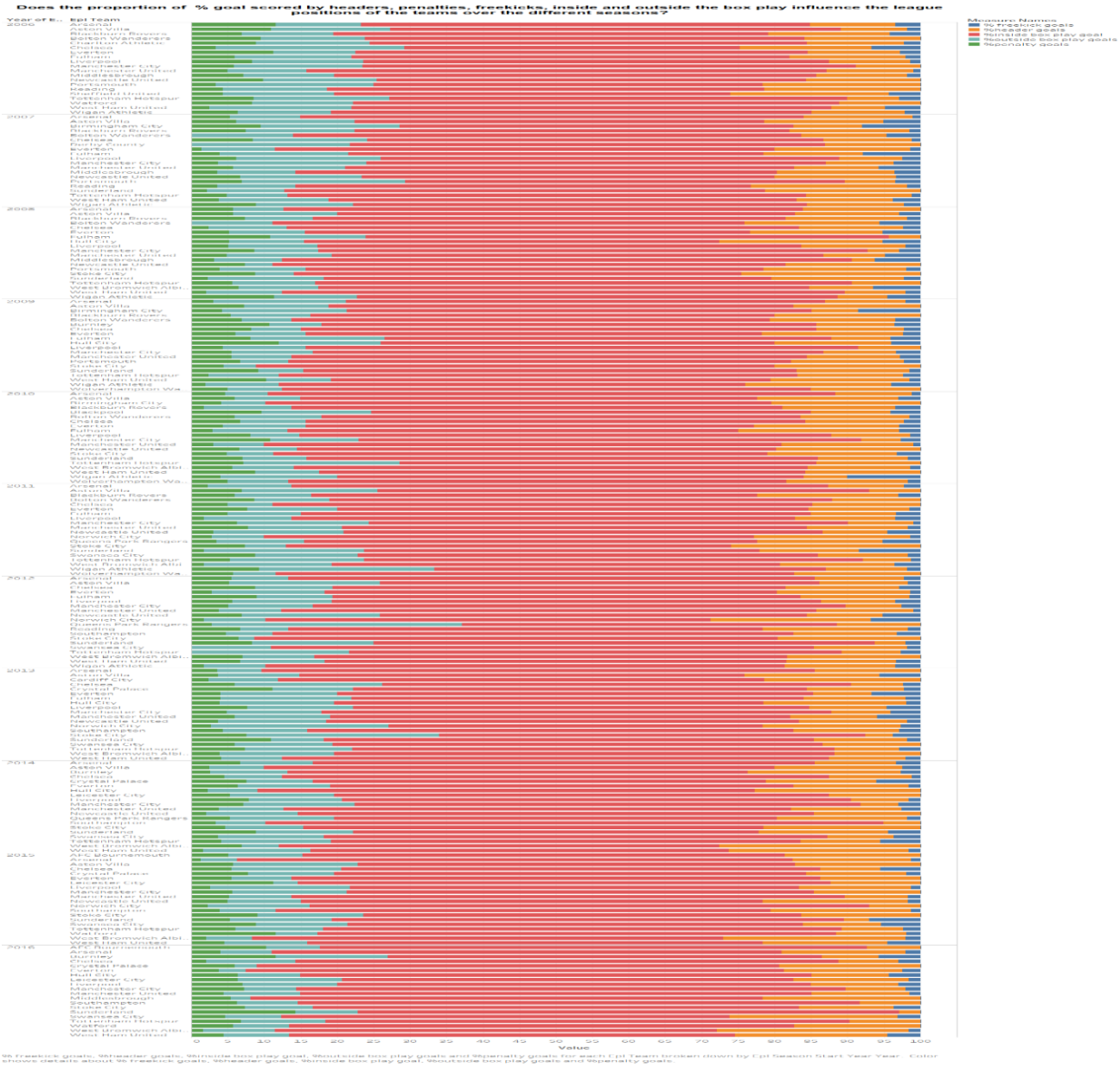
Visualization:



% freekick goals, %header goals, %inside box play goal, %outside box play goals and %penalty goals for each Epl Team broken down by Epl Season Start Year Year. Color shows details about % freekick goals, %header goals, %inside box play goal, %outside box play goals and %penalty goals. The view is filtered on Epl Team, which keeps 7 of 37 members.

Proxy Task 20: Does the proportion of % goal scored by headers, penalties, freekicks, inside and outside the box play influence the league positions of the teams over the different seasons?

Visualization:

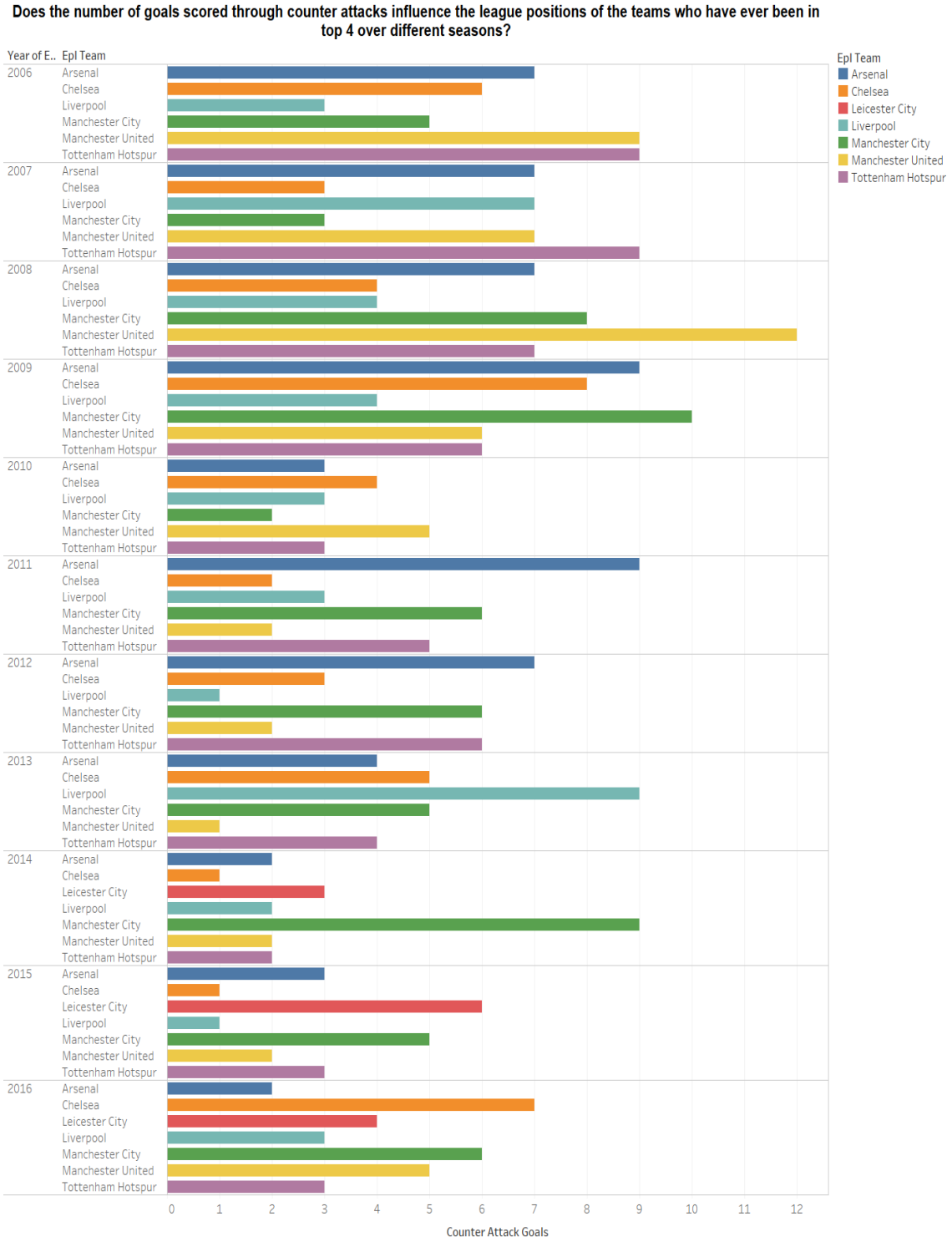


Visualization Description: The above stacked bar chart shows how the proportion of % goal scored by headers, penalties, freekicks, inside & outside the box play of teams changes over different seasons starting from 2006/2007 – 2016/2017

Important Insights:

The number of goals scored by inside box play has greater proportion as compared to number of goals scored by headers, penalties, freekicks and outside the box play

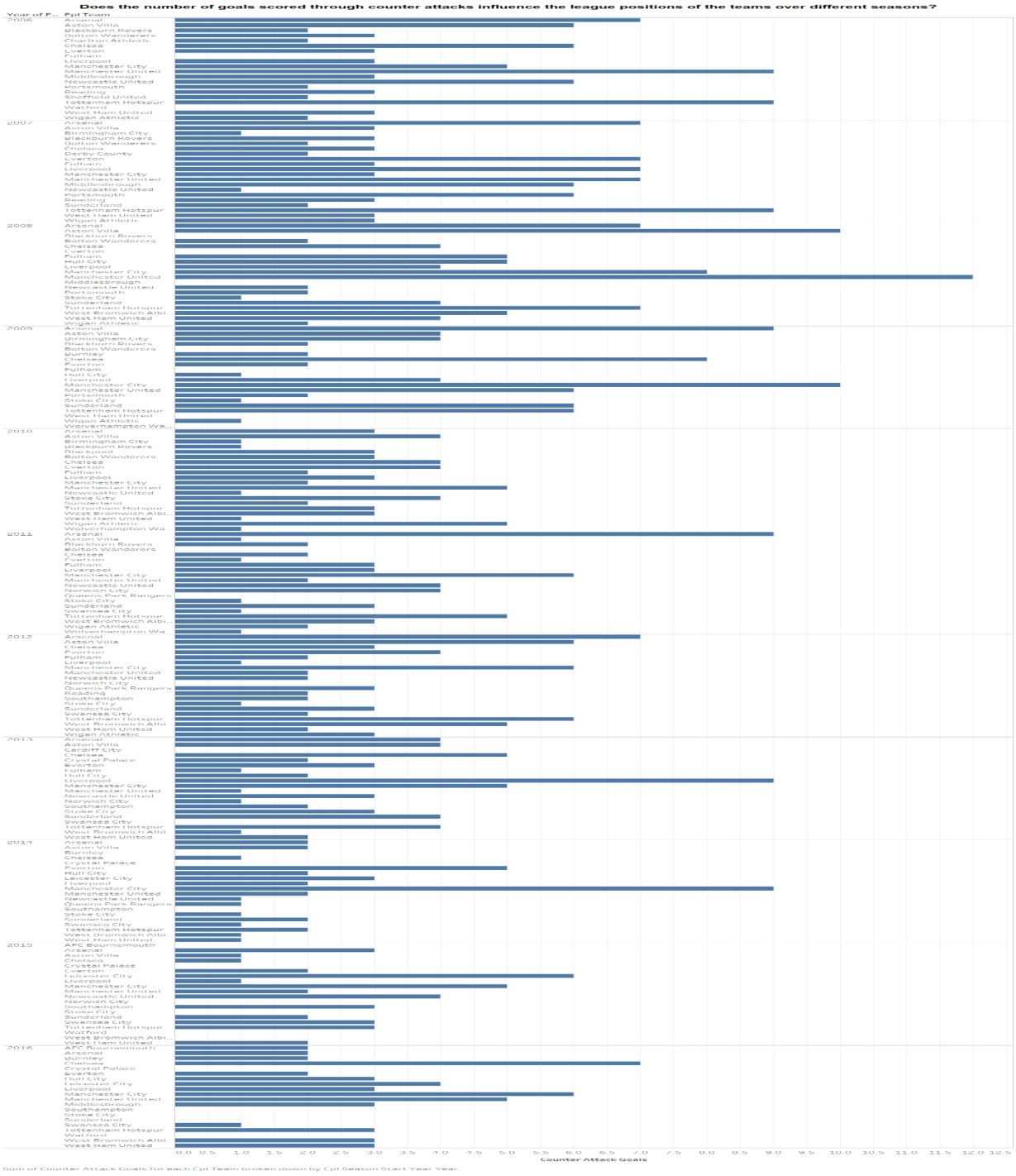
Proxy Task 21: Does the number of goals scored through counter attacks influence the league positions of the teams who have ever been in top 4 over different seasons?
Visualization:



Sum of Counter Attack Goals for each Epl Team broken down by Epl Season Start Year Year. Color shows details about Epl Team. The view is filtered on Epl Team, which keeps 7 of 37 members.

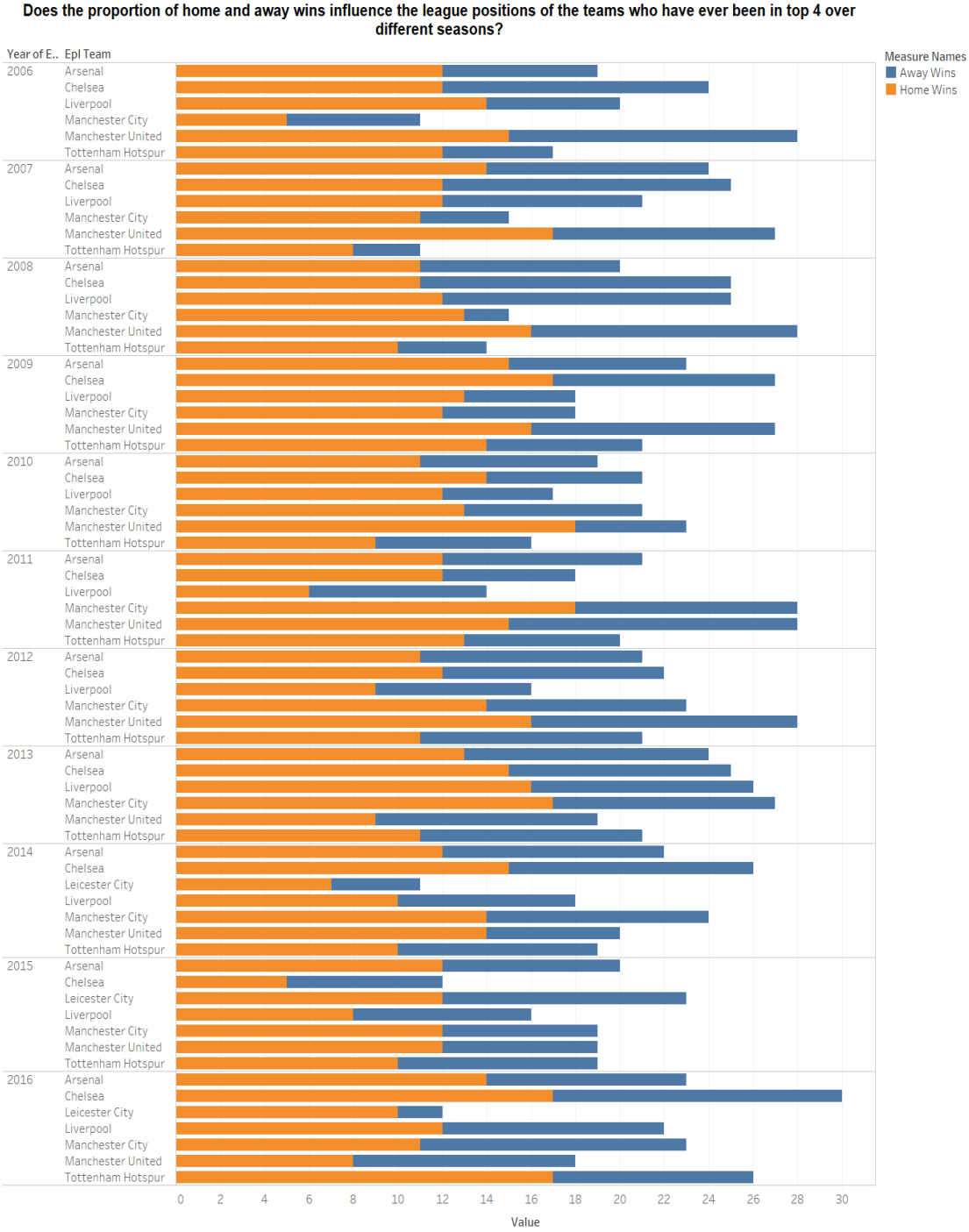
Proxy Task 22: Does the number of goals scored through counter attacks influence the league positions of the teams over different seasons?

Visualization:



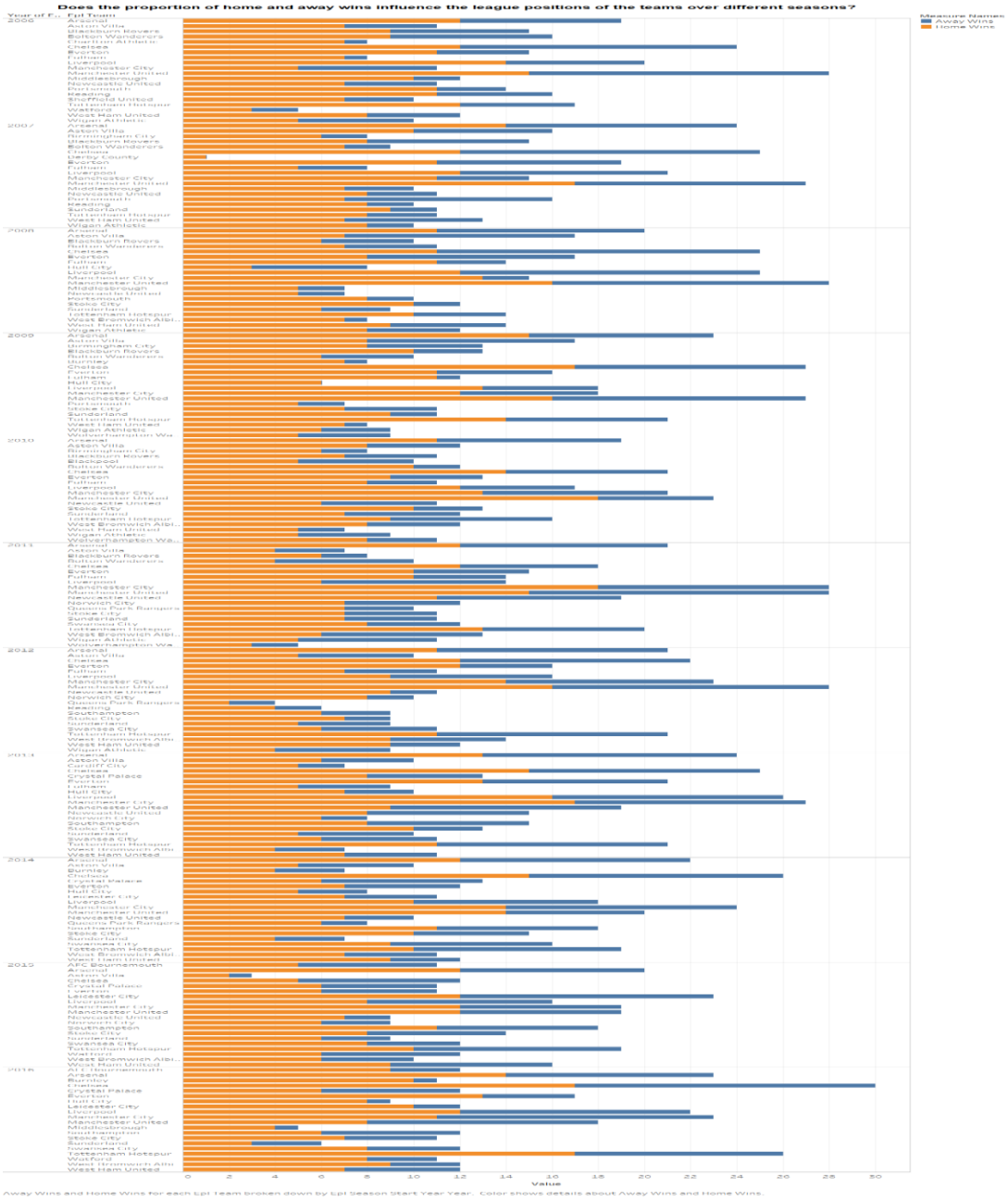
Visualization Description: The above bar chart shows how the goals scored through counter attack plays of teams changes over different seasons starting from 2006/2007 – 2016/2017

4. <u>Question:</u> Does the distribution of overall team form affect the league positions?
<u>Proxy Task 1</u>: Does the proportion of home and away wins influence the league positions of the teams who have ever been in top 4 over different seasons?
<u>Visualization</u>:



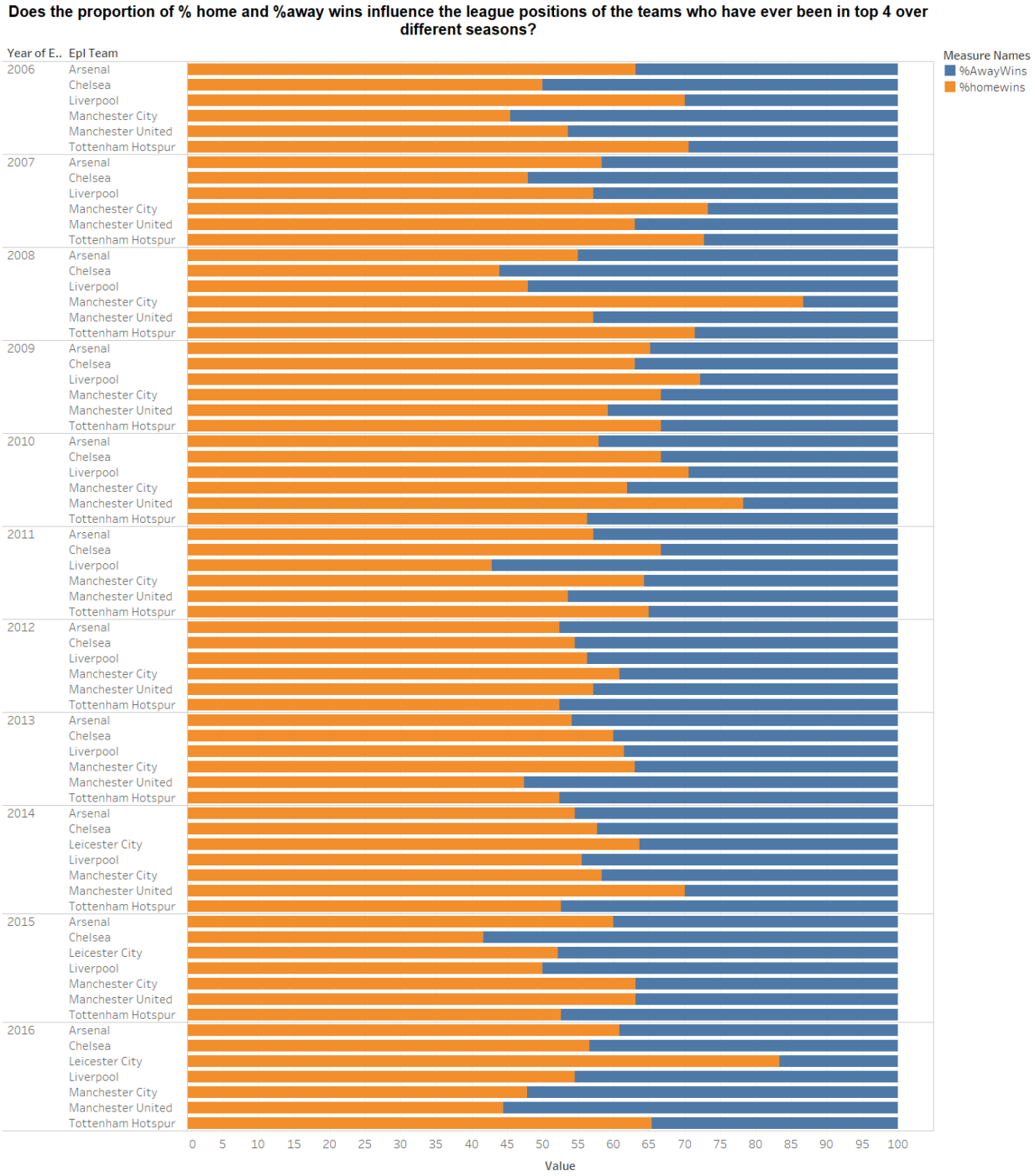
Away Wins and Home Wins for each Epl Team broken down by Epl Season Start Year Year. Color shows details about Away Wins and Home Wins. The view is filtered on Epl Team, which keeps 7 of 37 members.

Proxy Task 2: Does the proportion of home and away wins influence the league positions of the teams over different seasons?
Visualization:



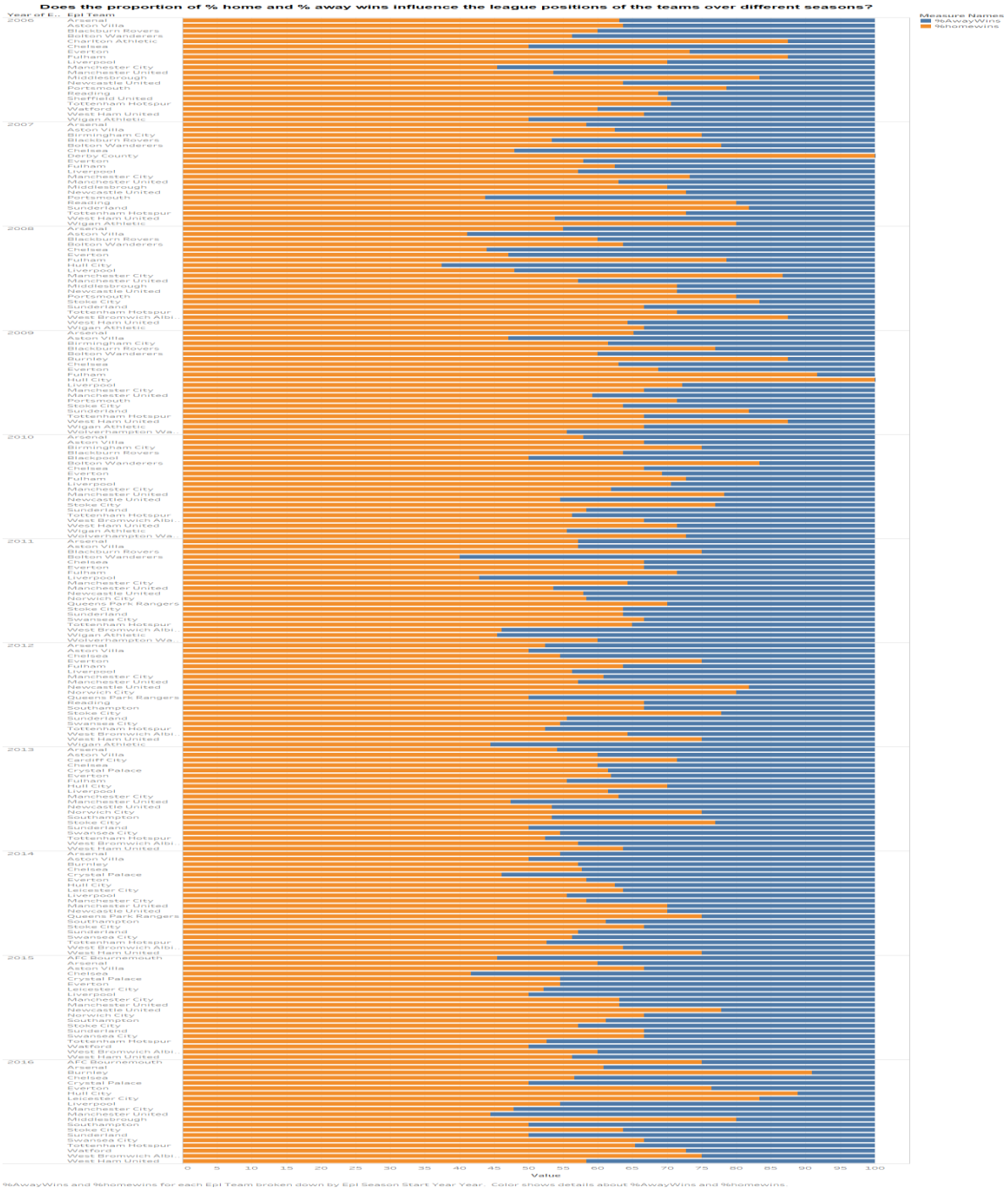
Visualization Description: The above stacked bar chart shows how the proportion of home and away wins of teams changes over different seasons starting from 2006/2007 – 2016/2017
Important Insights:
The proportion of home and away wins influences the league positions of the teams as generally teams placed at higher league positions have a higher proportion values as compared to the lower league position teams

Proxy Task 3: Does the proportion of % home and % away wins influence the league positions of the teams who have ever been in top 4 over different seasons?
Visualization:



%AwayWins and %homewins for each Epl Team broken down by Epl Season Start Year Year. Color shows details about %AwayWins and %homewins. The view is filtered on Epl Team, which keeps 7 of 37 members.

Proxy Task 4: Does the proportion of % home and % away wins influence the league positions of the teams over different seasons?
Visualization:



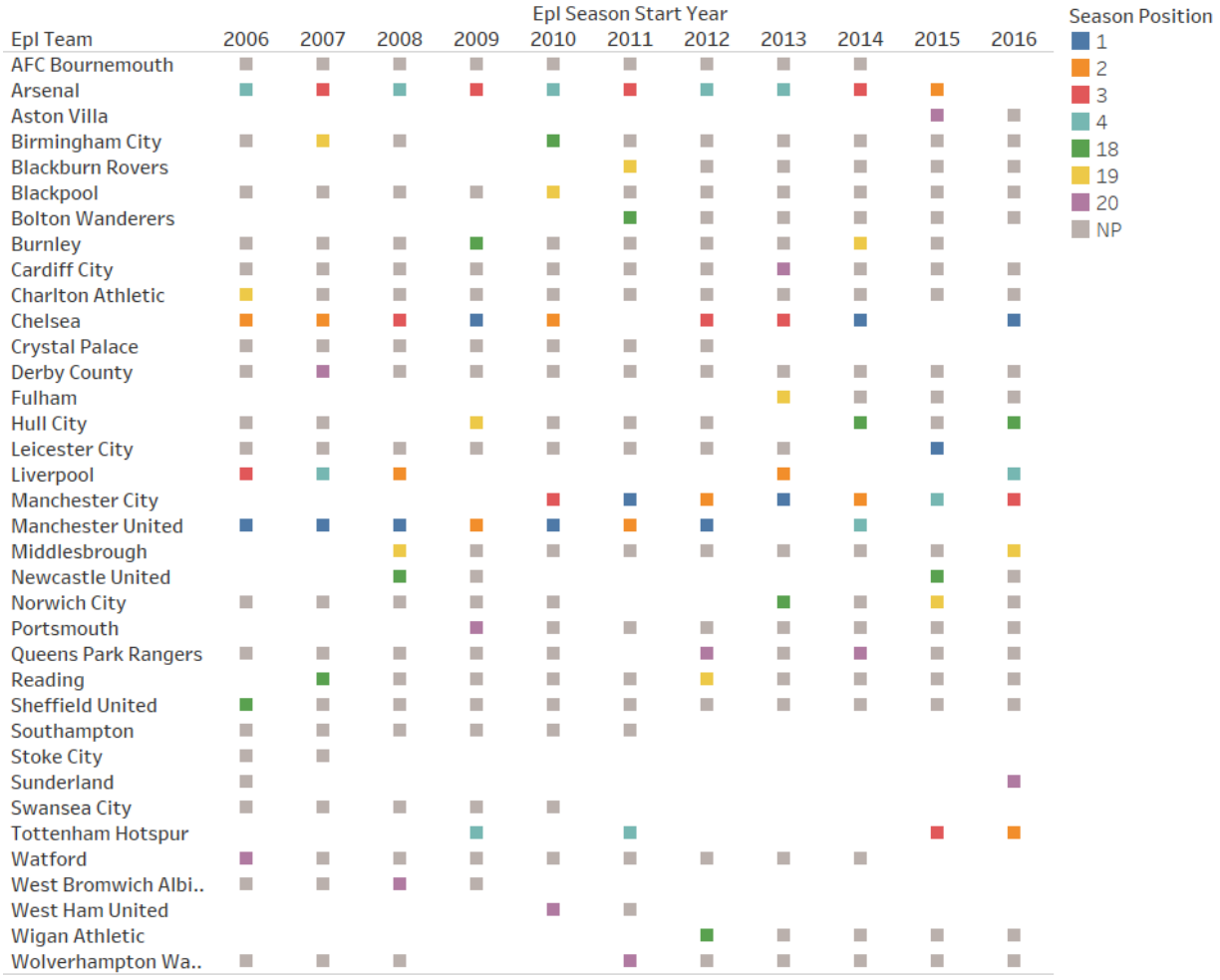
Visualization Description: The above stacked bar charts show how the proportion of % home And % away wins of teams changes over different seasons starting from 2006/2007 – 2016/2017

Storyboard

The story I am trying to visualize has the following raw structure as shown below:

1. Showing the distribution of the teams based on league positions over different seasons.
2. Showing the number of wins and draws and relating it with league position of the teams over different seasons.
3. Visualizing the Goal difference (Total Goals Scored – Total Goals Conceded) which affects league positions of teams over different seasons.
4. Visualizing the total passes (sum of short passes, crosses, through balls, long balls and backward passes) of the teams which might affect the league positions of teams over different seasons

What is the distribution of teams over different seasons based on league positions?

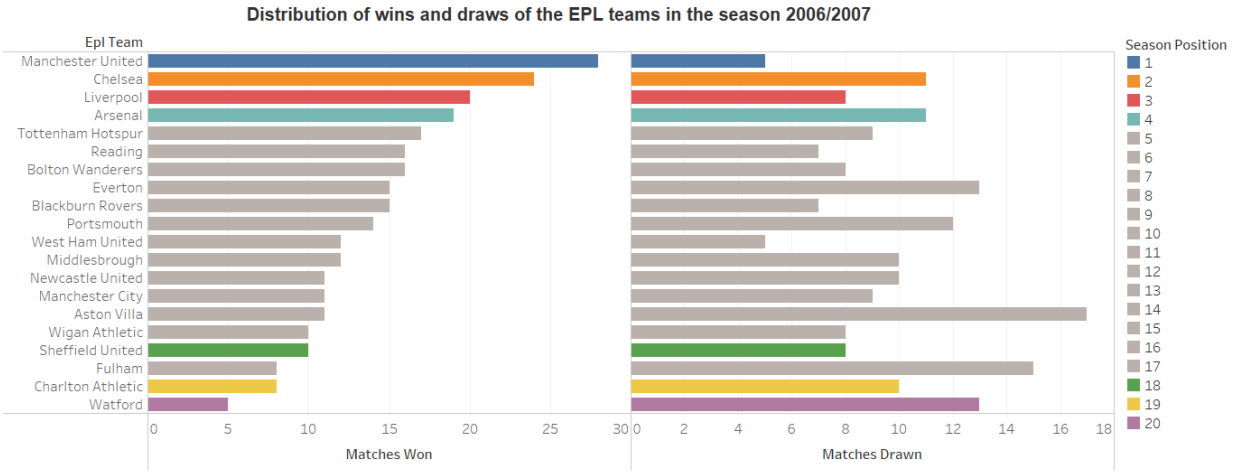


Season Position (color) broken down by Epl Season Start Year Year vs. Epl Team. The view is filtered on Season Position, which keeps 8 members.

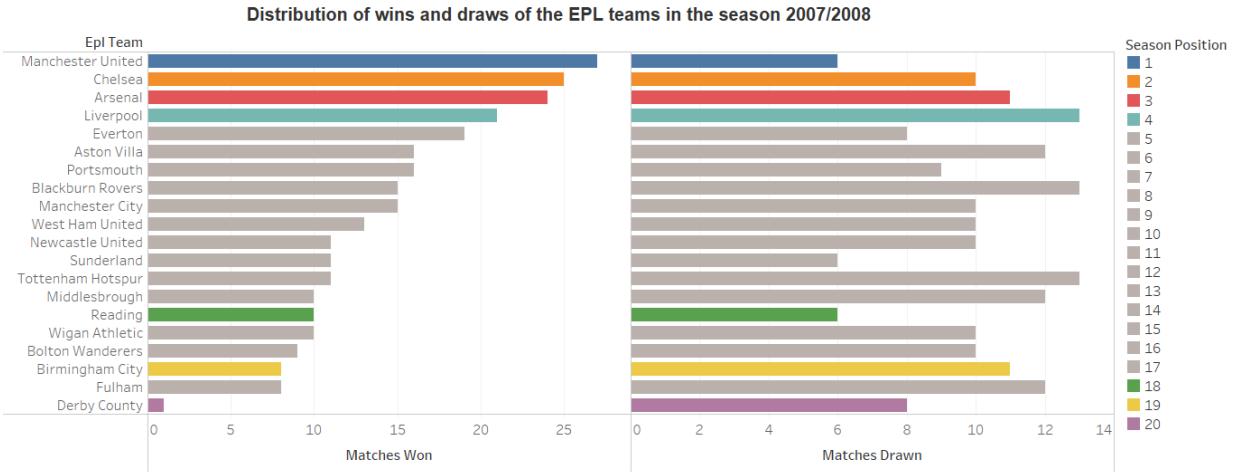
<u>Visualization Description</u>: The above matrix visualization shows how the top 4 and the bottom 3 teams change over different seasons starting from 2006/2007 – 2016/2017.Color Hue is used to encode the league positions highlighting the top 4 and bottom 3 teams every season. Grey color is used to indicate that the team didn't play in EPL that season

Important Insights:

1. Arsenal have consistently been in the top four over different seasons
2. Leicester City got promoted to EPL in 2014/2015 season and won the league in 2015/2016 season
3. Manchester City have shown consistent form after 2009/2010 season and have been in top four since then
4. Manchester United have also shown a drop-in form after the 2012/2013 season

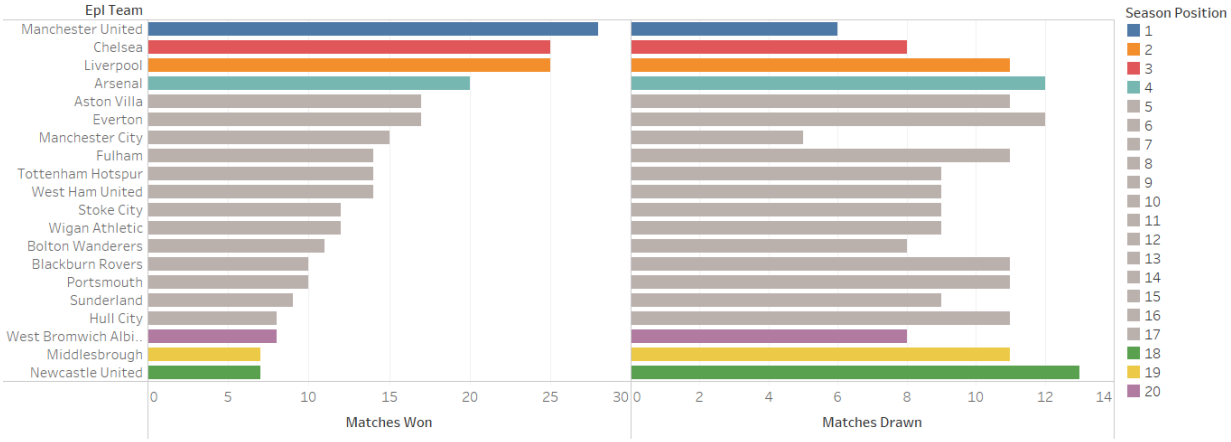


Sum of Matches Won and sum of Matches Drawn for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year and sum of Goal Difference. The Epl Season Start Year Year filter keeps 2006. The sum of Goal Difference filter keeps non-Null values only.



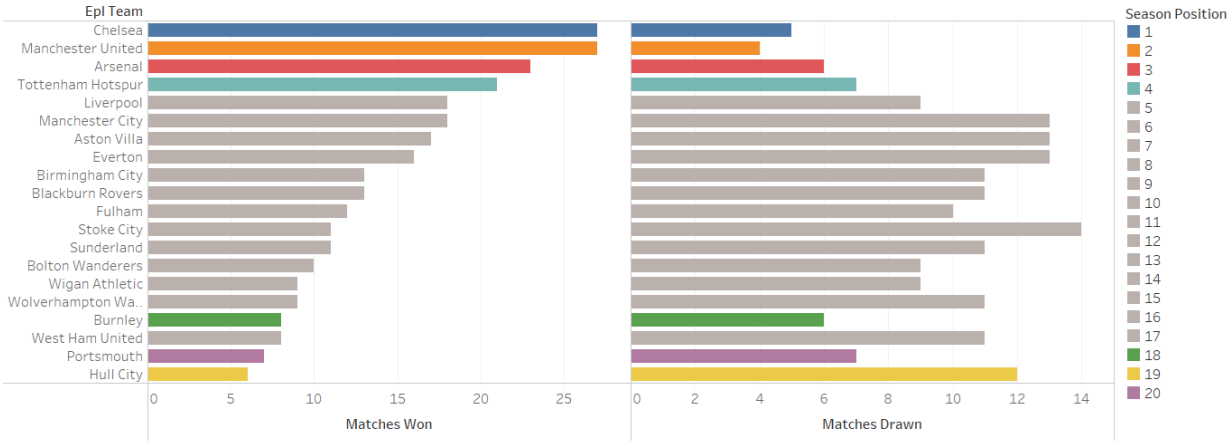
Sum of Matches Won and sum of Matches Drawn for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year and sum of Goal Difference. The Epl Season Start Year Year filter keeps 2007. The sum of Goal Difference filter keeps non-Null values only.

Distribution of wins and draws of the EPL teams in the season 2008/2009



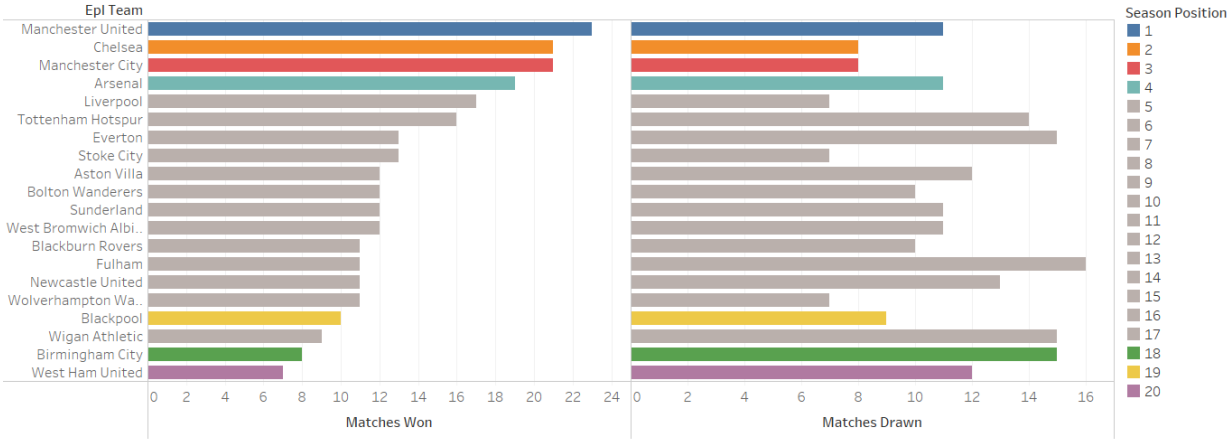
Sum of Matches Won and sum of Matches Drawn for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year and sum of Goal Difference. The Epl Season Start Year Year filter keeps 2008. The sum of Goal Difference filter keeps non-Null values only.

Distribution of wins and draws of the EPL teams in the season 2009/2010



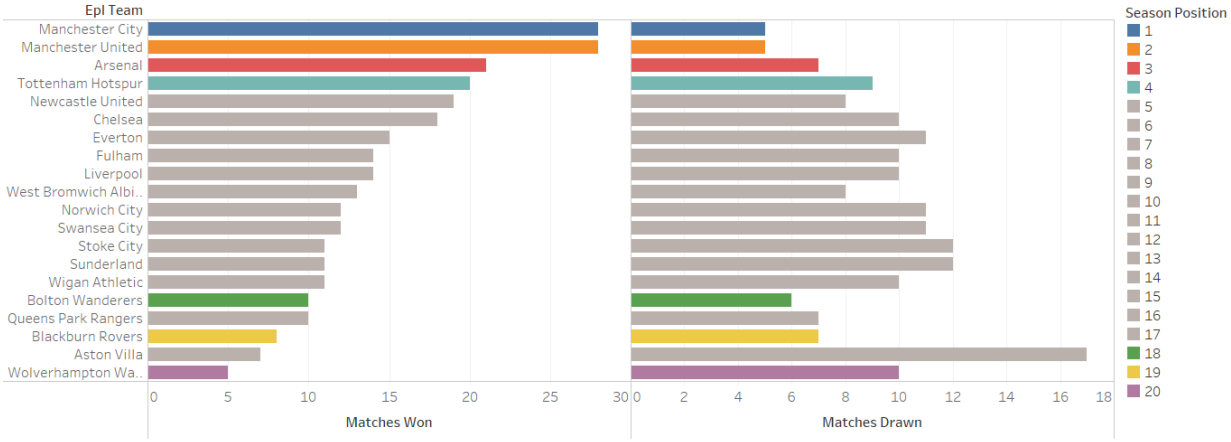
Sum of Matches Won and sum of Matches Drawn for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year and sum of Goal Difference. The Epl Season Start Year Year filter keeps 2009. The sum of Goal Difference filter keeps non-Null values only.

Distribution of wins and draws of the EPL teams in the season 2010/2011



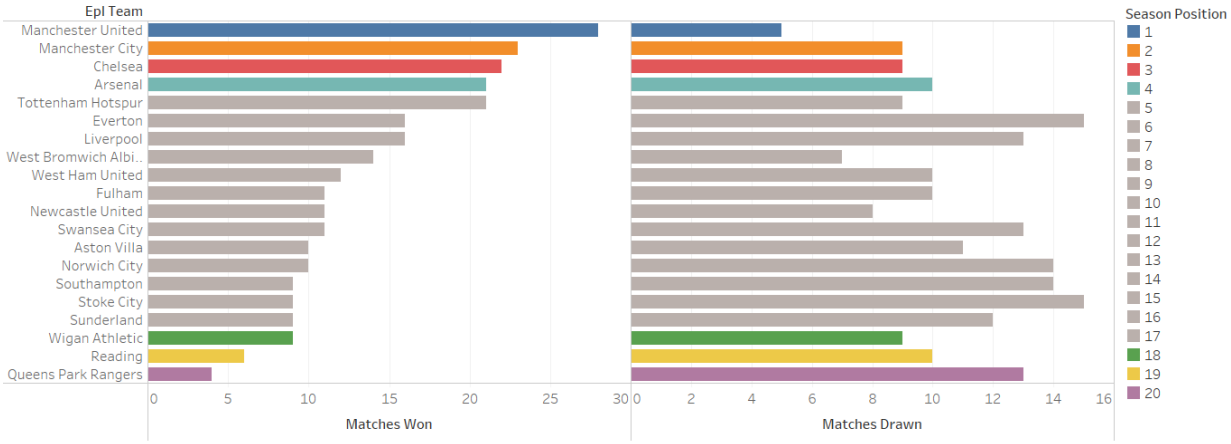
Sum of Matches Won and sum of Matches Drawn for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year and sum of Goal Difference. The Epl Season Start Year Year filter keeps 2010. The sum of Goal Difference filter keeps non-Null values only.

Distribution of wins and draws of the EPL teams in the season 2011/2012



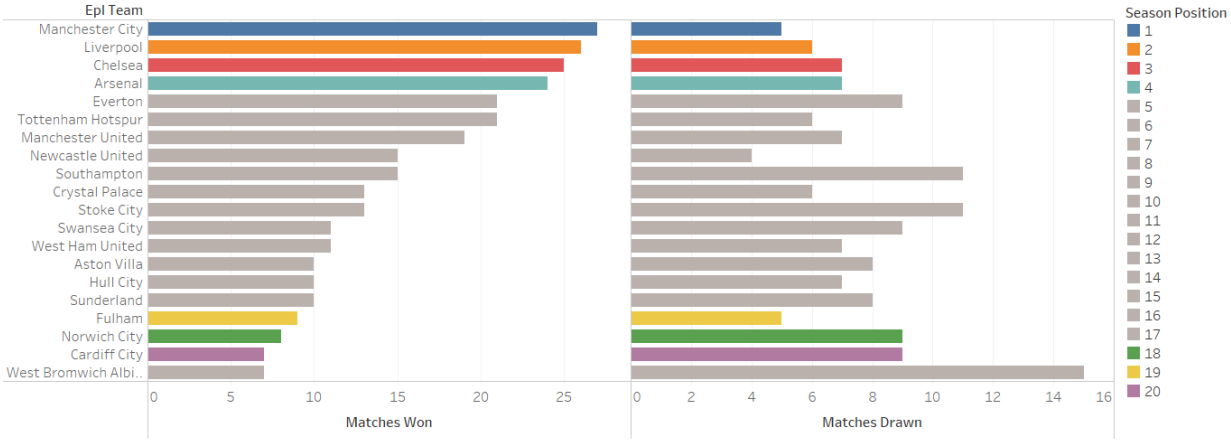
Sum of Matches Won and sum of Matches Drawn for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year and sum of Goal Difference. The Epl Season Start Year Year filter keeps 2011. The sum of Goal Difference filter keeps non-Null values only.

Distribution of wins and draws of the EPL teams in the season 2012/2013



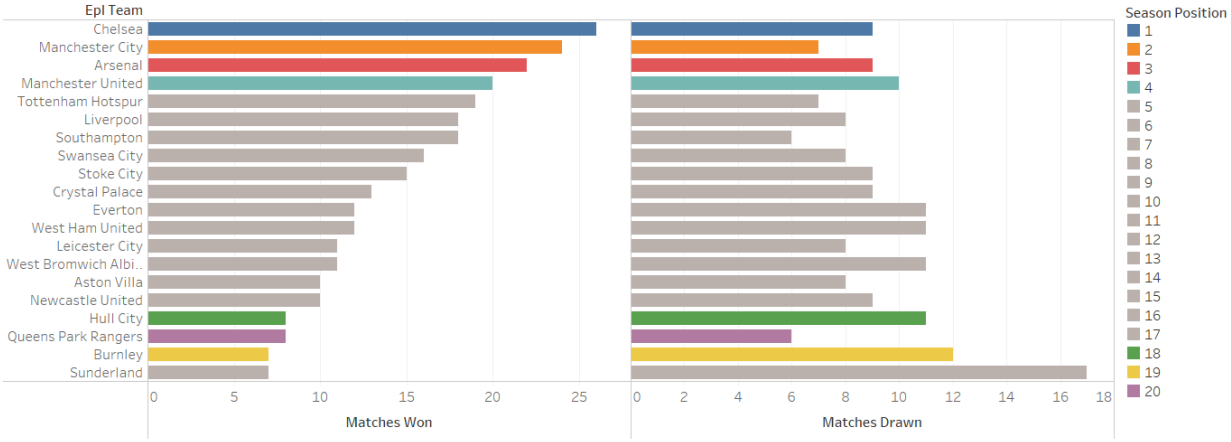
Sum of Matches Won and sum of Matches Drawn for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year and sum of Goal Difference. The Epl Season Start Year Year filter keeps 2012. The sum of Goal Difference filter keeps non-Null values only.

Distribution of wins and draws of the EPL teams in the season 2013/2014



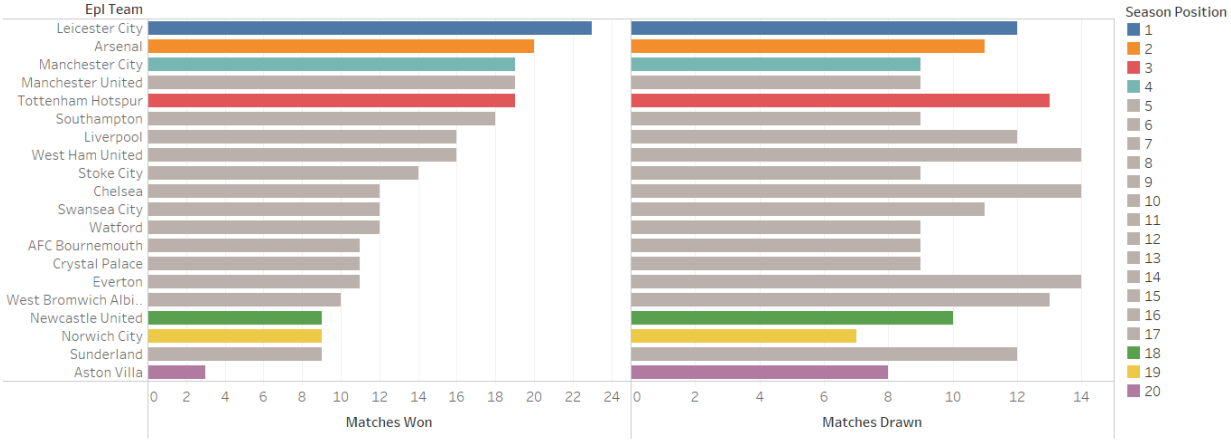
Sum of Matches Won and sum of Matches Drawn for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year and sum of Goal Difference. The Epl Season Start Year Year filter keeps 2013. The sum of Goal Difference filter keeps non-Null values only.

Distribution of wins and draws of the EPL teams in the season 2014/2015



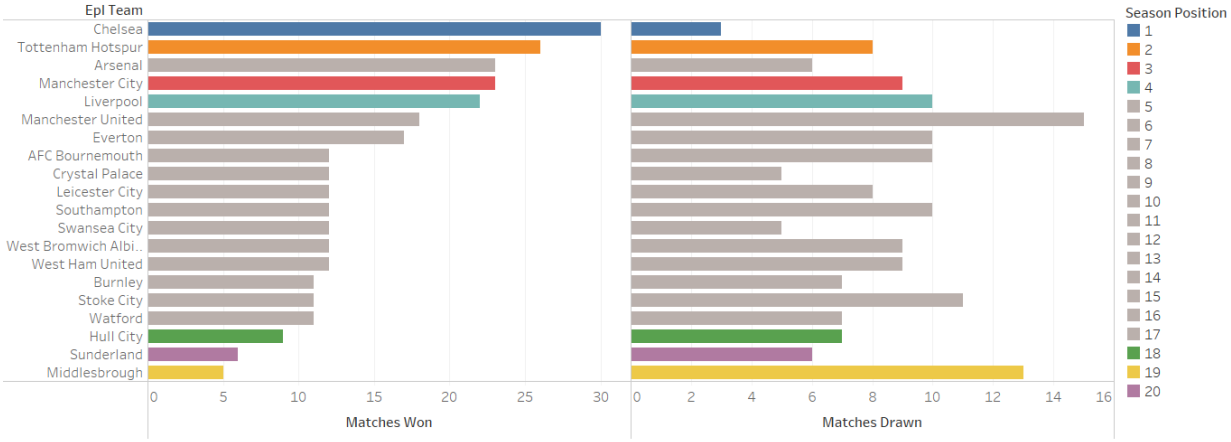
Sum of Matches Won and sum of Matches Drawn for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year and sum of Goal Difference. The Epl Season Start Year Year filter keeps 2014. The sum of Goal Difference filter keeps non-Null values only.

Distribution of wins and draws of the EPL teams in the season 2015/2016



Sum of Matches Won and sum of Matches Drawn for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year and sum of Goal Difference. The Epl Season Start Year Year filter keeps 2015. The sum of Goal Difference filter keeps non-Null values only.

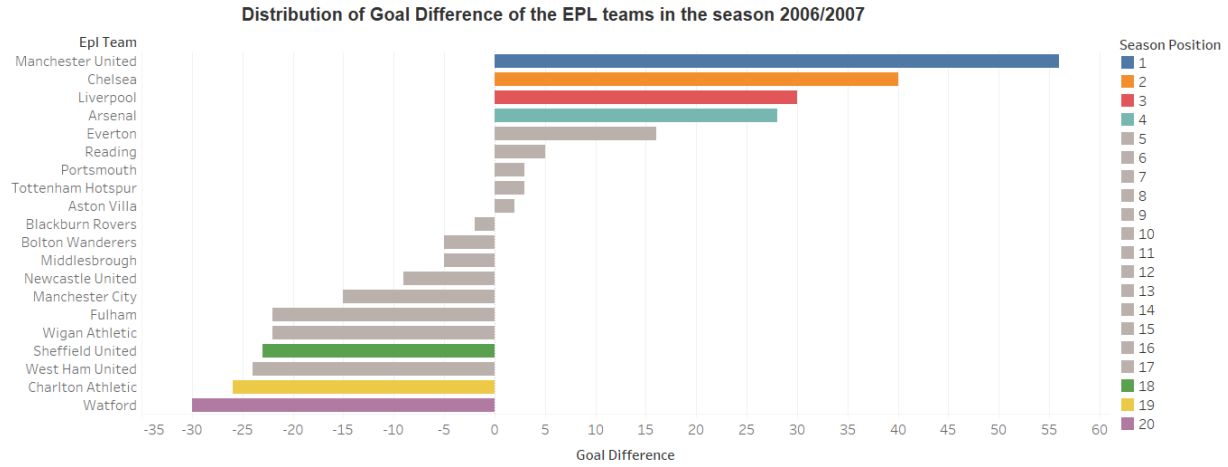
Distribution of wins and draws of the EPL teams in the season 2016/2017

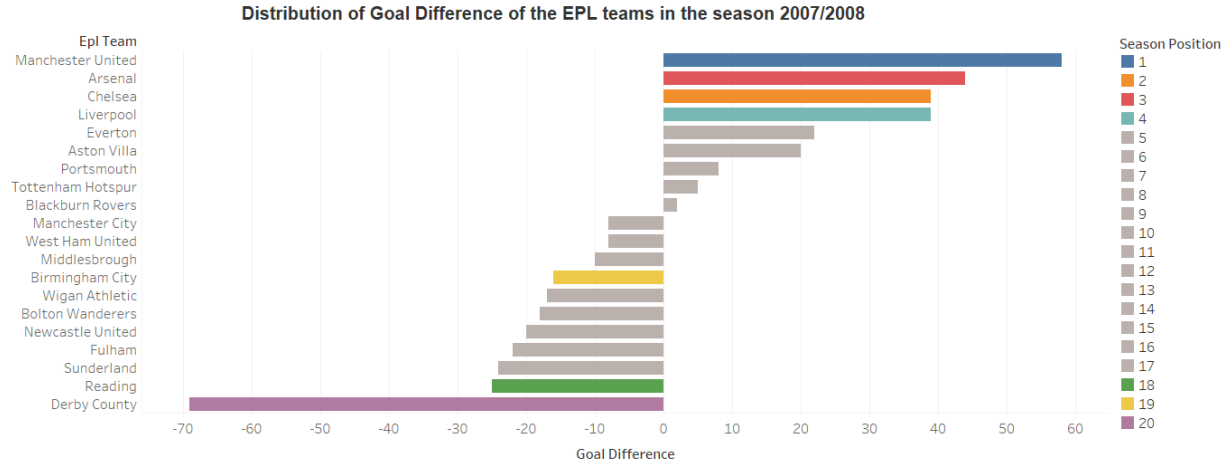


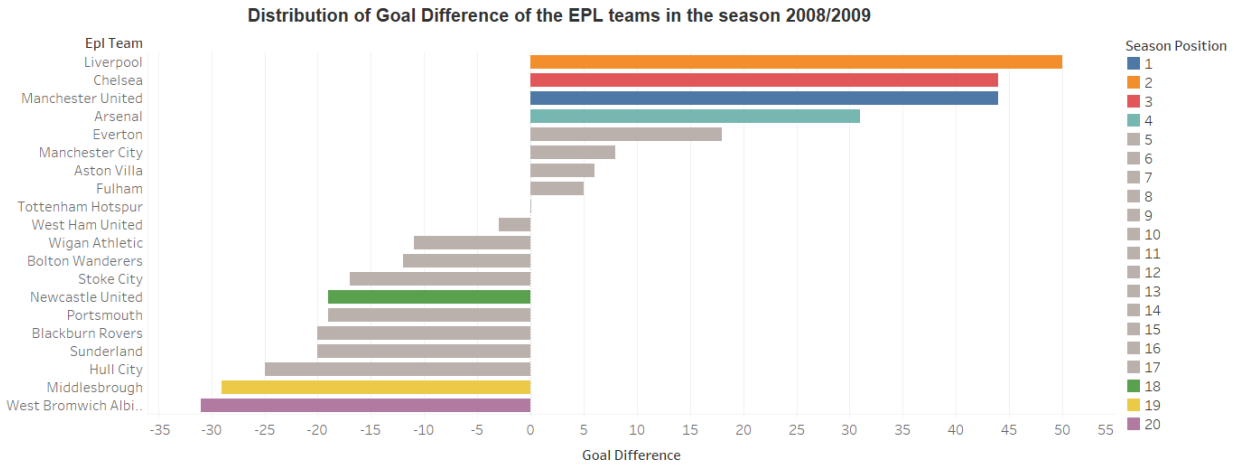
Sum of Matches Won and sum of Matches Drawn for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year and sum of Goal Difference. The Epl Season Start Year Year filter keeps 2016. The sum of Goal Difference filter keeps non-Null values only.

Visualization Description: The above bar charts shows how the number of wins and draws of the teams change over different seasons starting from 2006/2007 – 2016/2017 (One graph for every season). Color Hue is used to encode the league positions highlighting the top 4 and bottom 3 teams every season.

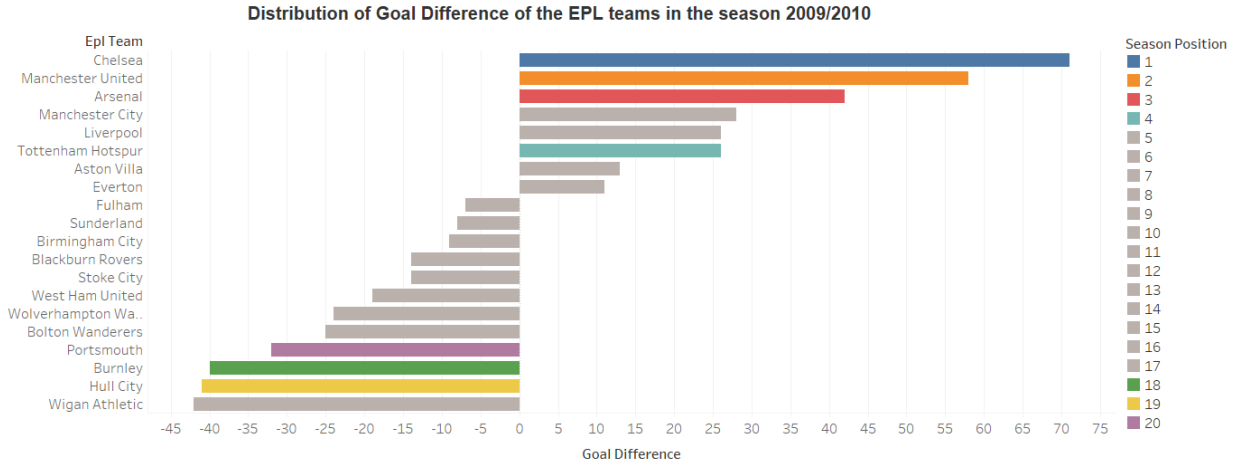
Important Insight:
The number of wins generally influence the league position of the teams over different seasons more rather than both number of wins and number of draws. This is more evident when we consider only the top 4 teams and bottom 3 teams.



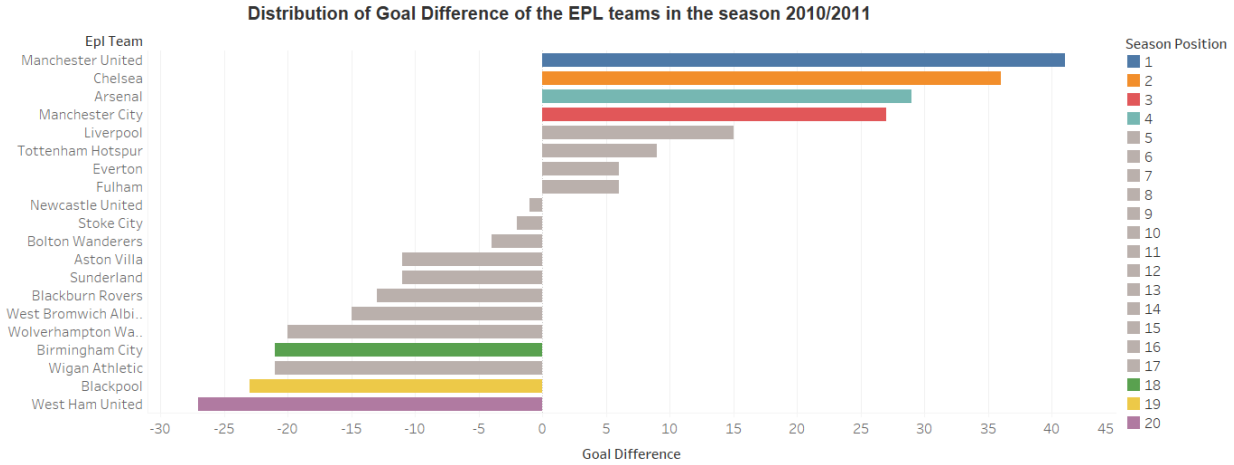




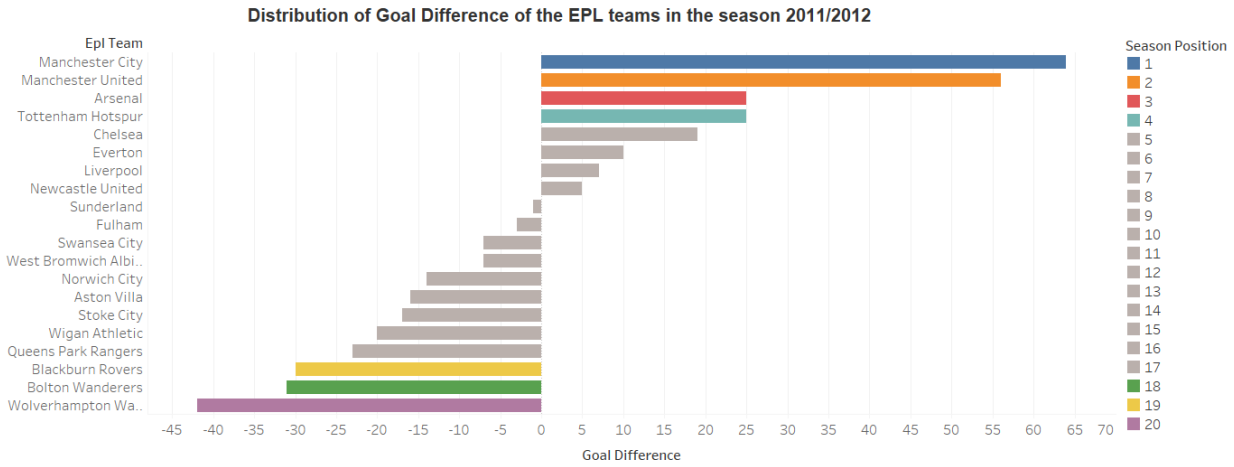
Sum of Goal Difference for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2008. The view is filtered on sum of Goal Difference, which keeps non-Null values only.



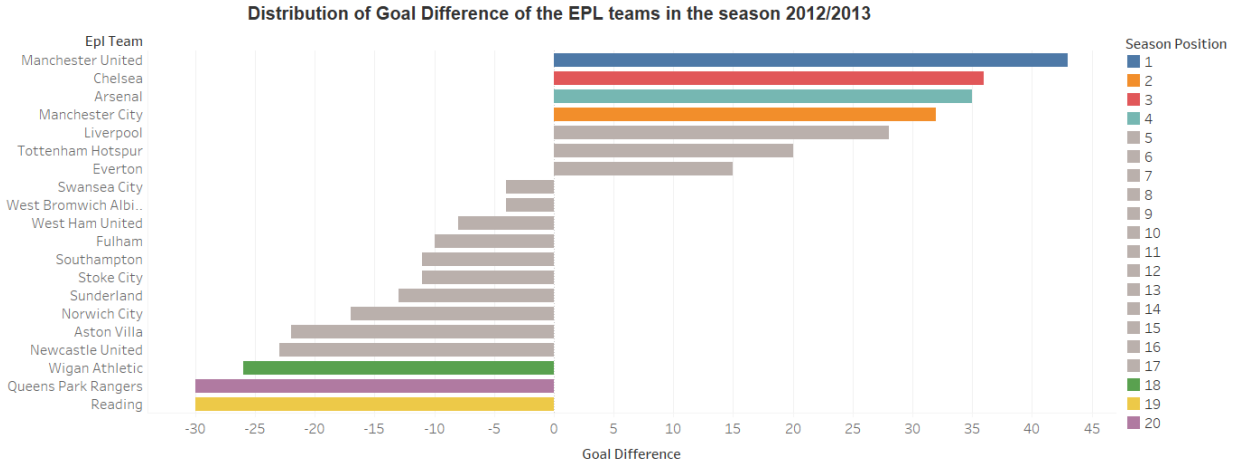
Sum of Goal Difference for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2009. The view is filtered on sum of Goal Difference, which keeps non-Null values only.



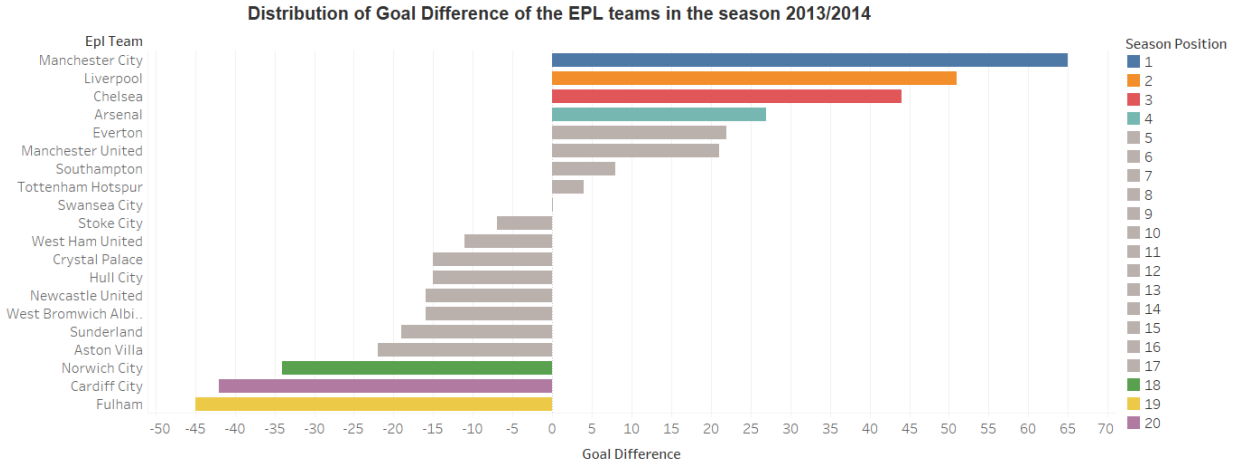
Sum of Goal Difference for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2010. The view is filtered on sum of Goal Difference, which keeps non-Null values only.



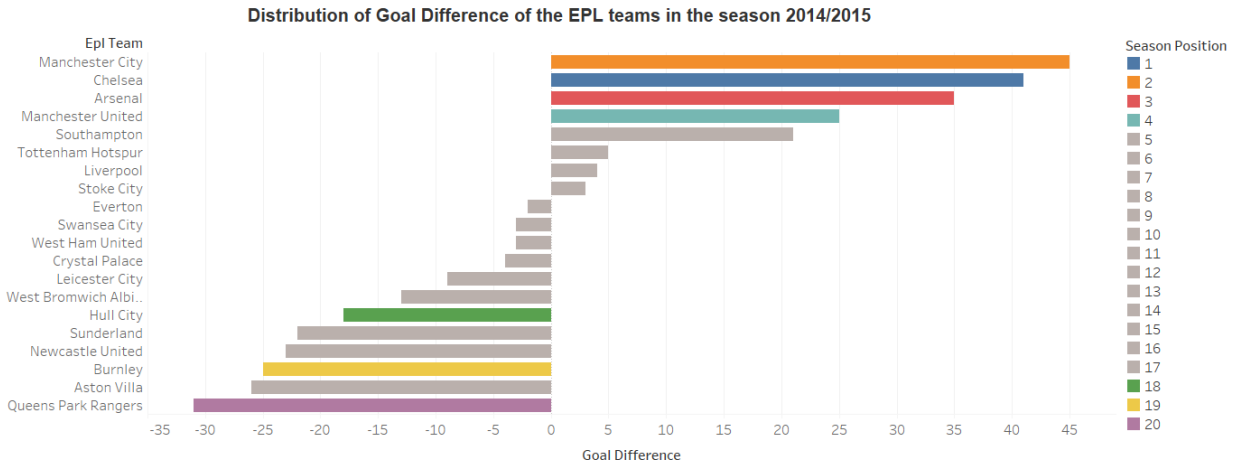
Sum of Goal Difference for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2011. The view is filtered on sum of Goal Difference, which keeps non-Null values only.



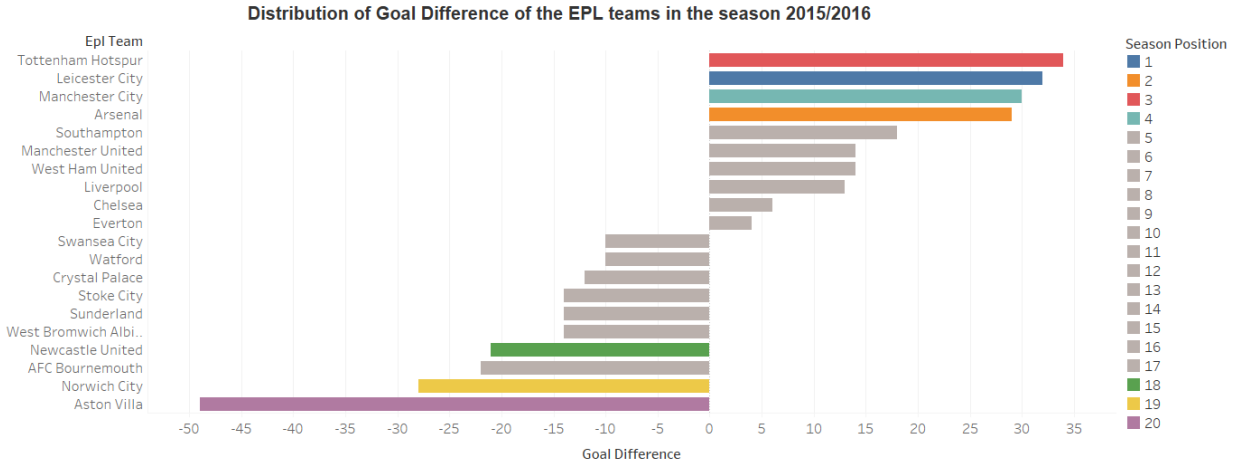
Sum of Goal Difference for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2012. The view is filtered on sum of Goal Difference, which keeps non-Null values only.



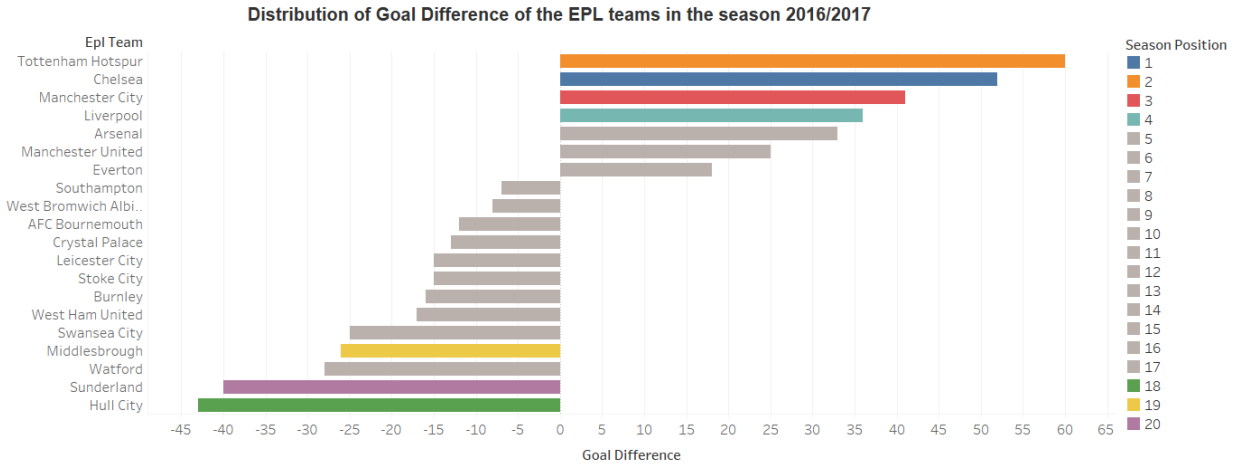
Sum of Goal Difference for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2013. The view is filtered on sum of Goal Difference, which keeps non-Null values only.



Sum of Goal Difference for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2014. The view is filtered on sum of Goal Difference, which keeps non-Null values only.



Sum of Goal Difference for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2015. The view is filtered on sum of Goal Difference, which keeps non-Null values only.

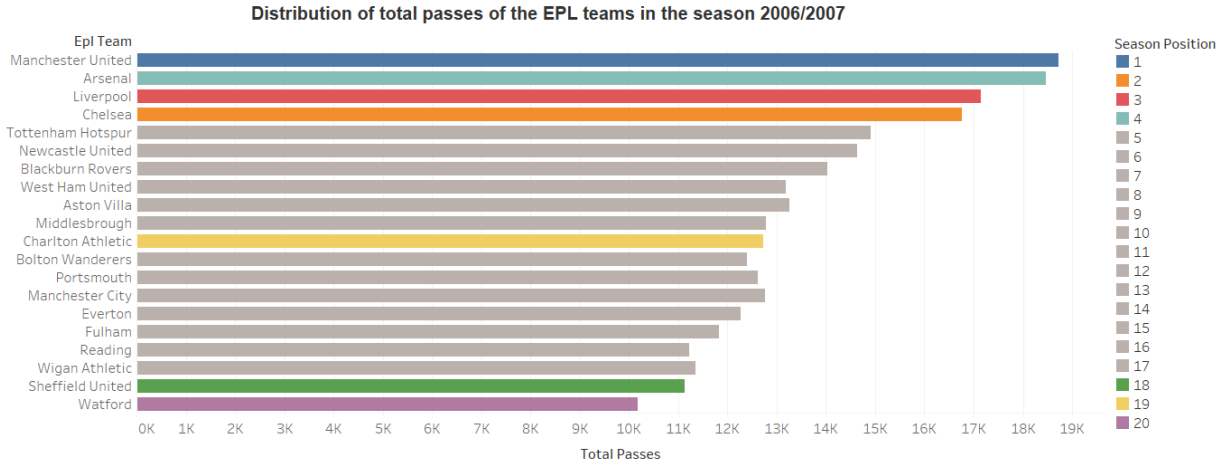


Sum of Goal Difference for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2016. The view is filtered on sum of Goal Difference, which keeps non-Null values only.

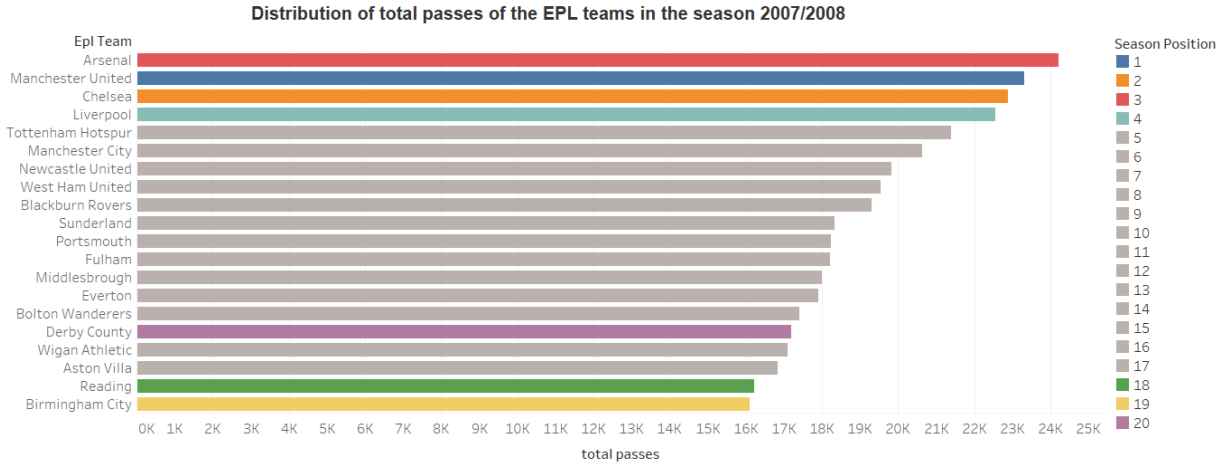
Visualization Description: The above bar charts shows how the goal difference of the teams change over different seasons starting from 2006/2007 – 2016/2017 (One graph for every season). Color Hue is used to encode the league positions highlighting the top 4 and bottom 3 teams every season.

Important Insight:

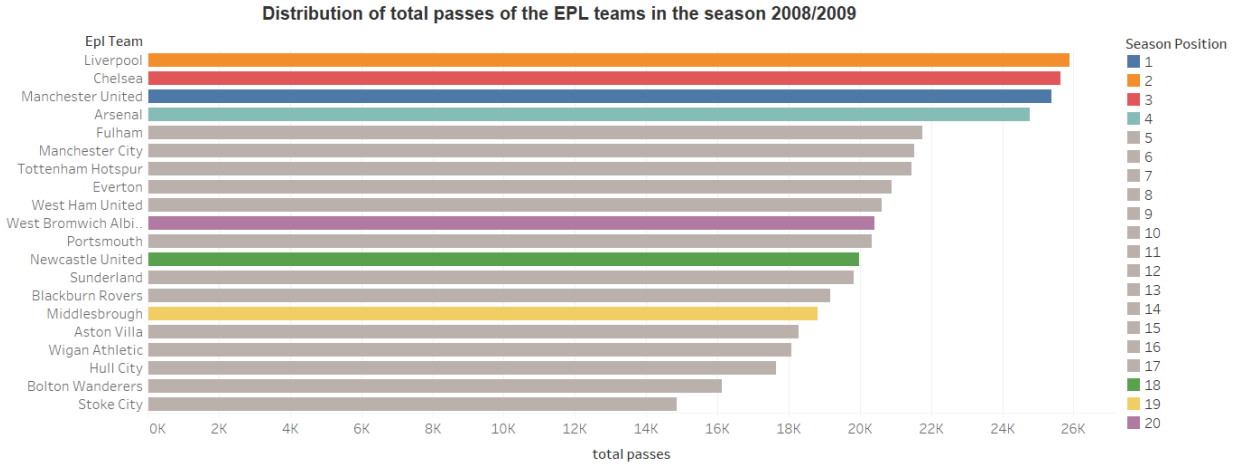
1. A good goal difference influences the league position as you can see a range of goal difference values (from negative to positive values)
2. Top four teams usually have a high positive goal difference and the bottom three teams have a high negative goal difference
3. Manchester City had a negative goal difference in season 2006/2007 and 2007/2008 and improved on their goal difference since then which helped them improve their league position.
4. Leicester city had a negative goal difference since getting promoted in EPL except for the season 2015/2016 when the won the league (i.e. were present in the top four)



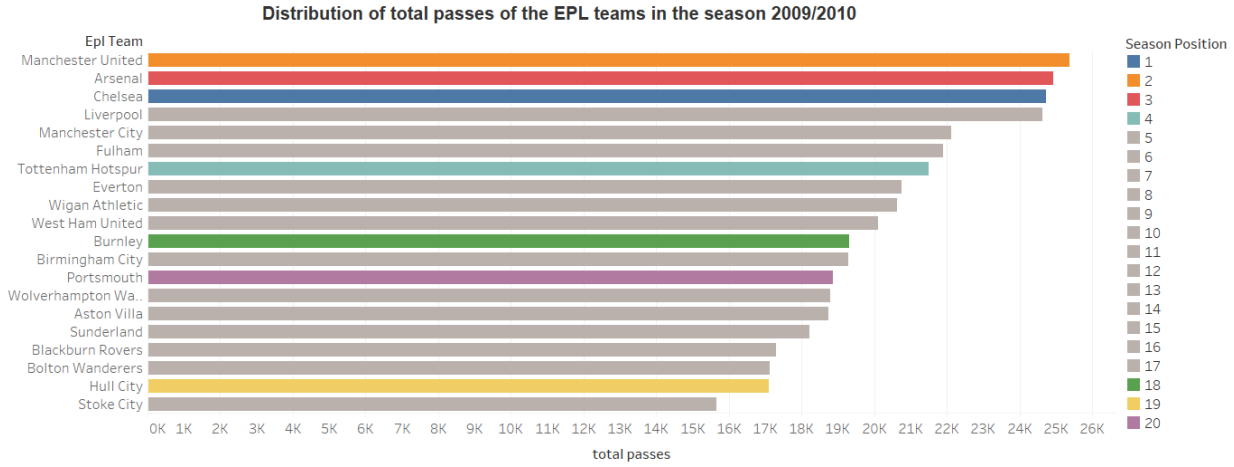
Sum of Total Passes for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year and sum of total passes. The Epl Season Start Year Year filter keeps 2006. The sum of total passes filter keeps non-Null values only.



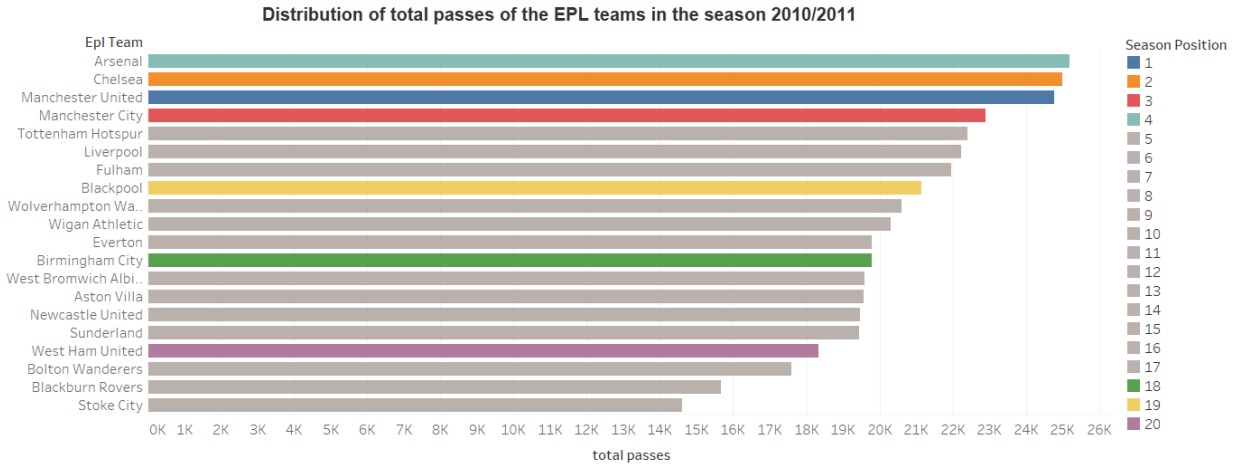
Sum of total passes for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2007. The view is filtered on sum of total passes, which keeps non-Null values only.



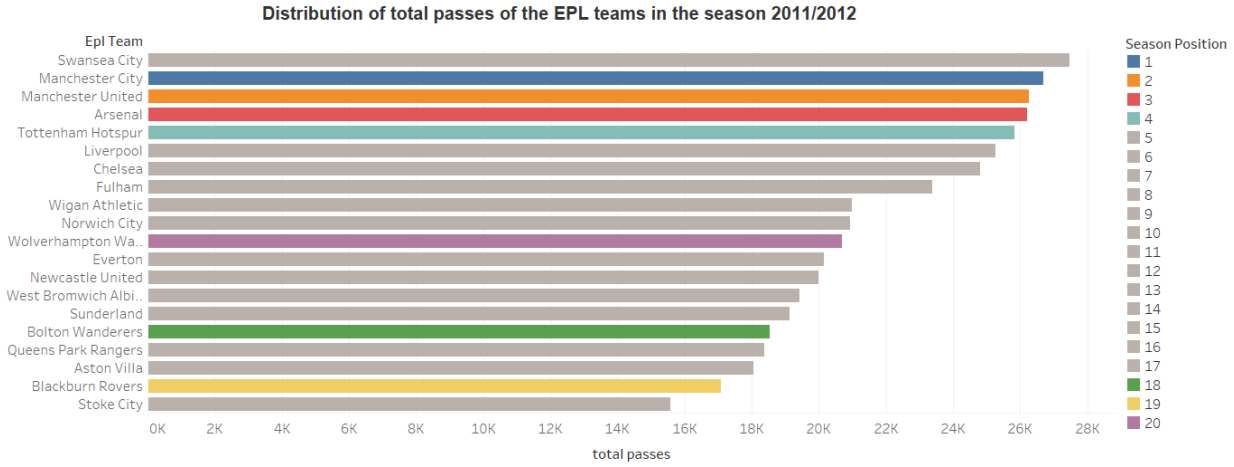
Sum of total passes for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2008. The view is filtered on sum of total passes, which keeps non-Null values only.



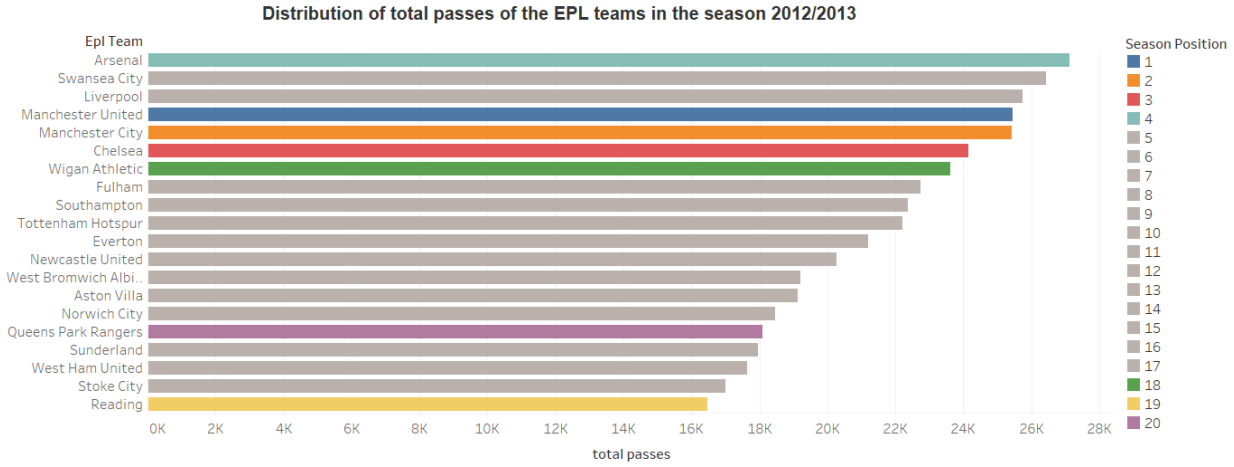
Sum of total passes for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2009. The view is filtered on sum of total passes, which keeps non-Null values only.



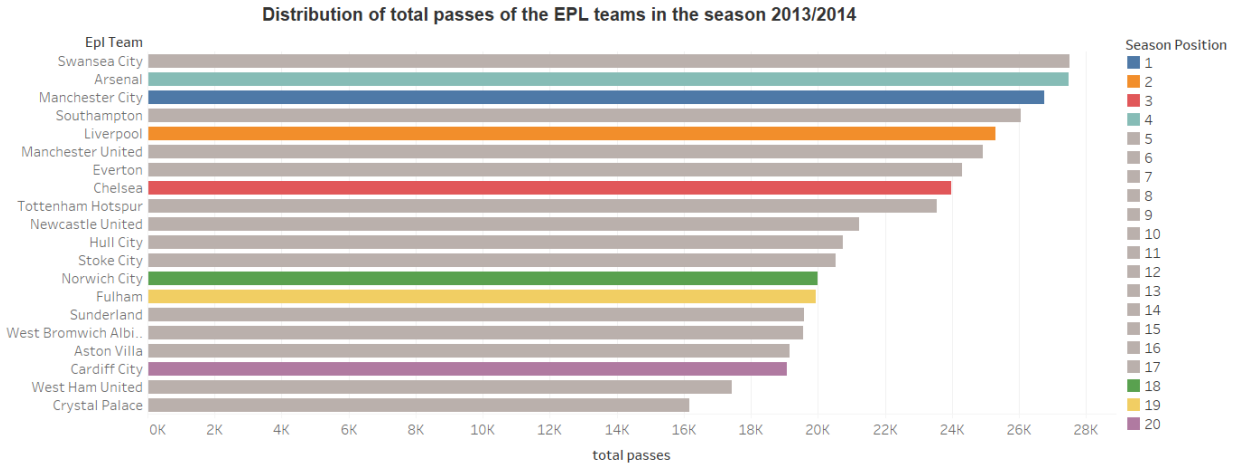
Sum of total passes for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2010. The view is filtered on sum of total passes, which keeps non-Null values only.



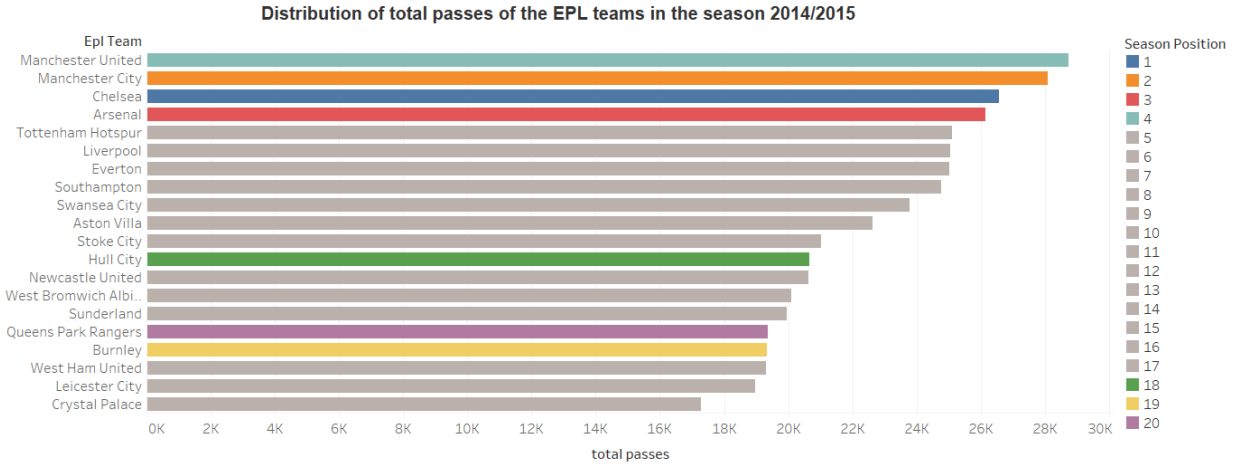
Sum of total passes for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2011. The view is filtered on sum of total passes, which keeps non-Null values only.



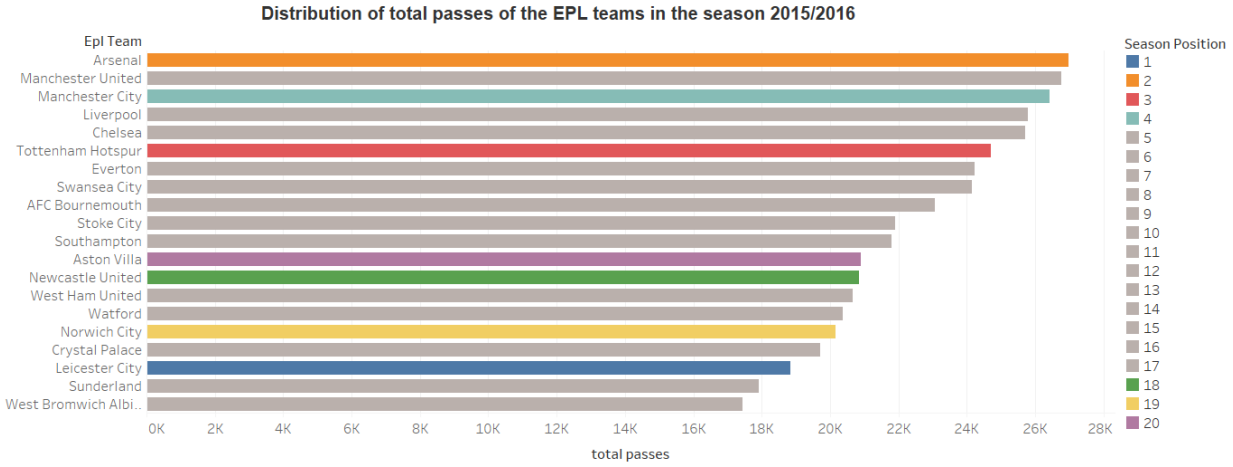
Sum of total passes for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2012. The view is filtered on sum of total passes, which keeps non-Null values only.



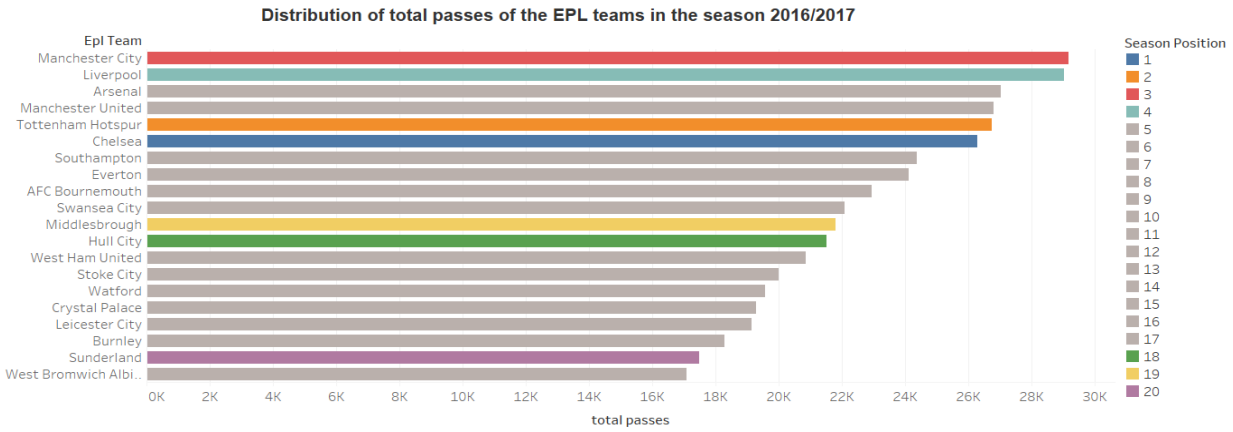
Sum of total passes for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2013. The view is filtered on sum of total passes, which keeps non-Null values only.



Sum of total passes for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2014. The view is filtered on sum of total passes, which keeps non-Null values only.



Sum of total passes for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2015. The view is filtered on sum of total passes, which keeps non-Null values only.



Sum of total passes for each Epl Team. Color shows details about Season Position. The data is filtered on Epl Season Start Year Year, which keeps 2016. The view is filtered on sum of total passes, which keeps non-Null values only.

Visualization Description: The above bar charts shows how the total passes (sum of short passes, crosses, through balls, long balls and backward passes) of the teams change over different seasons starting from 2006/2007 – 2016/2017 (One graph for every season). Color Hue is used to encode the league positions highlighting the top 4 and bottom 3 teams every season.

Important Insight:
The total passes influence the league positions of the teams as generally teams placed at higher league positions have a higher value of total passes as compared to the lower league position teams

Changelog

Additional Dataset Used:
We are using additional dataset as the league standing information for the season 2015-2016 is missing.
Link: https://www.kaggle.com/ronman11/english-premier-league-tables
File Name: epl20152016.csv
Description: Dataset contains EPL league stats for season 2015/2016
Important Attributes:

Attribute Name	Attribute Description	Attribute Type	Attribute Semantic
Team	Name of the Team	Categorical	
W	Number of matches won	Quantitative	Sequential
D	Name of the matches drawn	Quantitative	Sequential
L	Name of the matches lost	Quantitative	Sequential
A	Goals conceded by the team	Quantitative	Sequential
F	Goals scored by the team	Quantitative	Sequential
D	Goal Difference of Each team	Quantitative	Divergent
P	Total points of the team	Quantitative	Sequential

Implementation

Updates done based on feedback received:

1. In the first Matrix visualization graph, we are using the quantitative scale for visualizing the league standings of teams over different seasons. Two different color hue with varying intensity are used for top 4 and bottom 4 teams. All other teams are grayed out. To show some interesting patterns we are sorting the teams displayed in ascending order, based on mean of the league standings over different seasons.
2. For the second set of graphs, we are planning to visualize the proportion of wins, draws and losses of teams for every season. The ranking of the team will be given by the position of the team on the graph. Color hue will be used to distinguish between wins, draws and

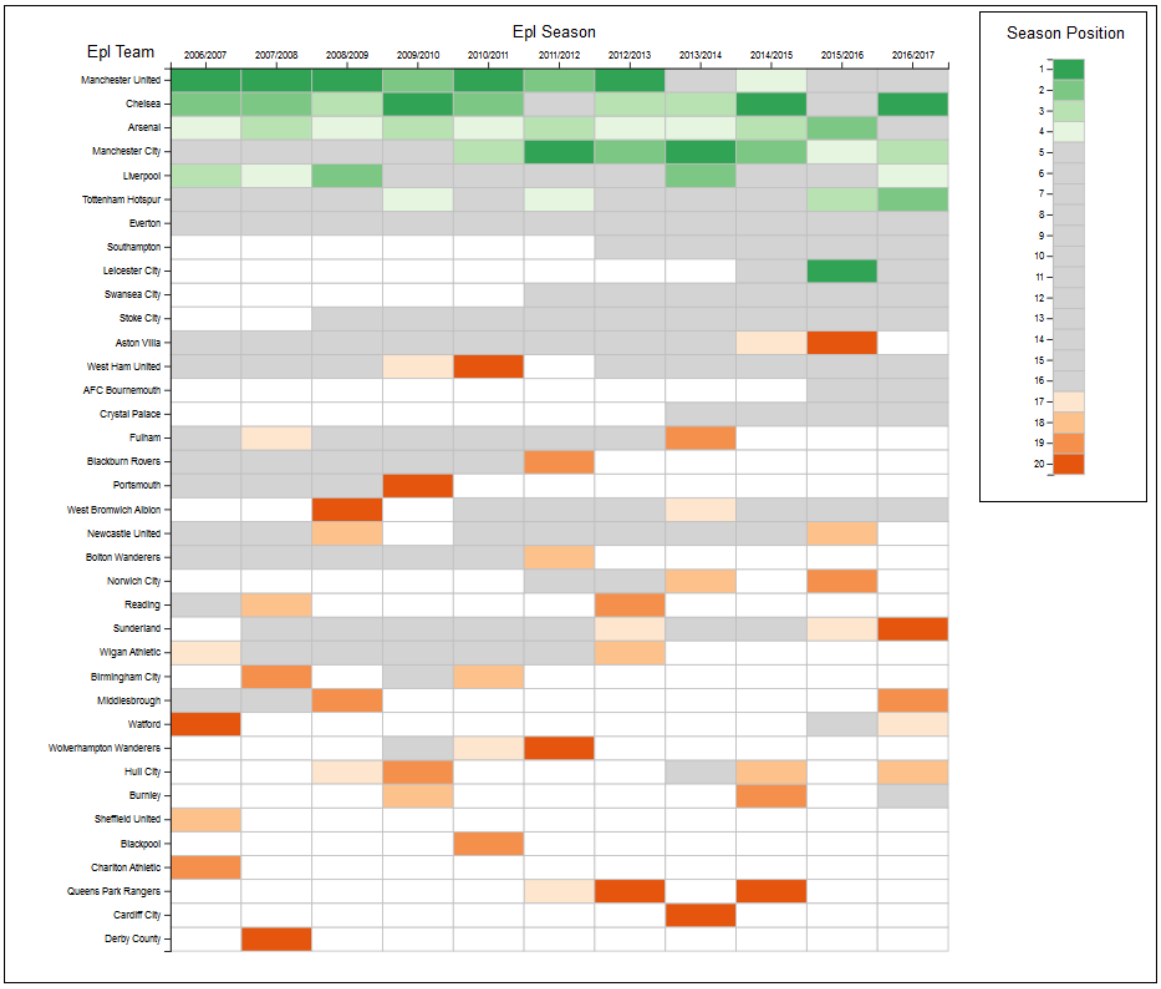
losses. We will be using interaction to give the user an option to select which year data to be displayed.

3. For third and fourth set of graphs, the ranking of the teams will be given by the position of the team on the graph. We will be using interaction to give the user an option to select which year data to be displayed.
4. We are making the second, third and fourth set of graphs smaller so that they can be displayed side by side.

<u>List of Items Implemented</u>:

1. The first graph i.e. Matrix Visualization.

What is the distribution of teams over different seasons based on league positions ?



2. Second, Third and Fourth Set of Graphs along with time interaction:

GitHub Link:

https://github.com/NYU-VIS-FALL2018/storytelling-group-3-online

Web Page Link:

https://nyu-vis-fall2018.github.io/storytelling-group-3-online/