

GROUP I

ECE

**DASHBOARD CREATION FOR
IPL MATCHES**

Department of Management Studies
Indian Institute of technology
(Indian school of mines)
Dhanbad

Certificate

This project entitled “**DASHBOARD CREATION FOR IPL MATCHES**” submitted by Mr. Froz Husain Mr.Gaurav, Mr.Gupta Saket Devendranath, Mr. Hritik Aggarwal, Mr.Jayant Anand and Mr.Suryansh is approved for the course project of subject Business Analytics .

Prof .Shravan

Course Instructor

Business Analytics

Department of Management Studies

DECLARATION

I declare that this written submission of the project titled “**DASHBOARD CREATION FOR IPL MATCHES**” represents my ideas in my own word and where other ideas and/or words are used , have been adequately mentioned in the references.

Team Members

Froz Husain	Sophomore, ECE
Gaurav	Sophomore, ECE
Gupta Saket Devendranath	Sophomore, ECE
Hritik Aggarwal	Sophomore, ECE
Jayant Anand	Sophomore, ECE
Suryansh Singh	Sophomore, ECE

ACKNOWLEDGEMENT

We would like to offer our sincere gratitude to Prof Shrawan Trivedi for his able guidance and encouragement at various stages of this project. We would also like to express gratitude to Ajay Sir, who introduced us to the correct methodology of work and encouraged us to function as a professional during the course of completion of this project.

We would also like to thank all of them who have directly or indirectly guided us in achieving this goal, this includes several websites and YouTube channels, which we referred for learning various concepts involved in the project. We also thank our classmate for their comments and suggestions, which have immensely helped us to shape this project.

Group I

(Electronics & Communication Engineering)

Summary

A dashboard provides a centralized, interactive means of monitoring, analyzing, measuring and extracting a wealth of business understandings from relevant datasets in various key areas while displaying aggregated information intuitively and visually.

The dataset that was chosen for the construction of the dashboard was about the cricket matches played in the IPL series. The main reason behind the selection of this dataset was the popularity and varsity of the game which sources this dataset.

The dataset has 17 columns and 757 rows describing the match id, the IPL season in which match was played, the city in which the match was played, the date of the match, the teams who played the match, the team who won the toss, the decision of the toss winning team to either bat or field first, the result of the match, the no of runs by which the match was won, the number of wickets by which the match was won, the player of the match, the venue of the match and the names of first umpire and second umpire.

Some of the important conclusion that can be drawn from the dashboard are:

1-The home teams didn't have a huge bias to blabber about, but with about 33% of the games done after, home teams have won a surprisingly high number of them.

2-The team that wins the toss prefers chasing due to the dew factor. The 8 seam Kookaburra ball easily gets wets due to the dew and when the moisture soaks in there is no coming back, the ball becomes merely a soap bar in the hands of a spinner.

TABLE OF CONTENT

Cover page	i
Title page	ii
Certificate	iii
Declaration	iv
Acknowledgement	v
Abstract	vi

1. Introduction

2. Literature Review

2.1 - Why we choose this data

2.2 - Brief description of the dataset

3. Methodology

3.1 - FEATURES USED TO CREATE DASHBOAR

3.2 - Analyzing the Dataset

3.2.2 - HOME ADVANTAGE (Come on Lads!)

3.2.3 - Biasness Of Toss

4. Result Analysis and Conclusion

5. References

Introduction

The data generated by the various departments and sub-departments of an organisation can be considered as the heart of the organisation. The data not only help the company to draw the insights of the working and progression of the company but also enable the firm to eradicate inconsistencies and mistakes from their system. Nowadays the data can be considered as one of the corporate assets which play a very significant role in strategic as well as analytic models of the company.

But, in this digital world, each of the firms is generating tonnes and tonnes of data every week, which without proper management and organisations can be valued the same as a digital pile of random numbers and alphabets. It is the management and organisation techniques which simplify the intricate and complex layers of the data. The data is managed and processed for generating possible useful insights from it by storing them effectively in a kind of electronic notebook or an account called a spreadsheet.

A spreadsheet is an application which is commonly used for organization, analysis and storage of data. These are pretty much similar to the physical accounting worksheets, however, at the same time promises much faster and efficient handling of data. The program operates on data entered in cells of a table. These worksheets may come handy when the data have to be operated with certain mathematical functions and operations. These programs. These applications also assist in the effective representation of data with the help of different kind of charts and graphs, which makes the data more understandable. Besides this, some of the spreadsheet options also provide the tool for analysis and prediction of data, which massively saves the human effort of rigorous calculation and repeated investigations.

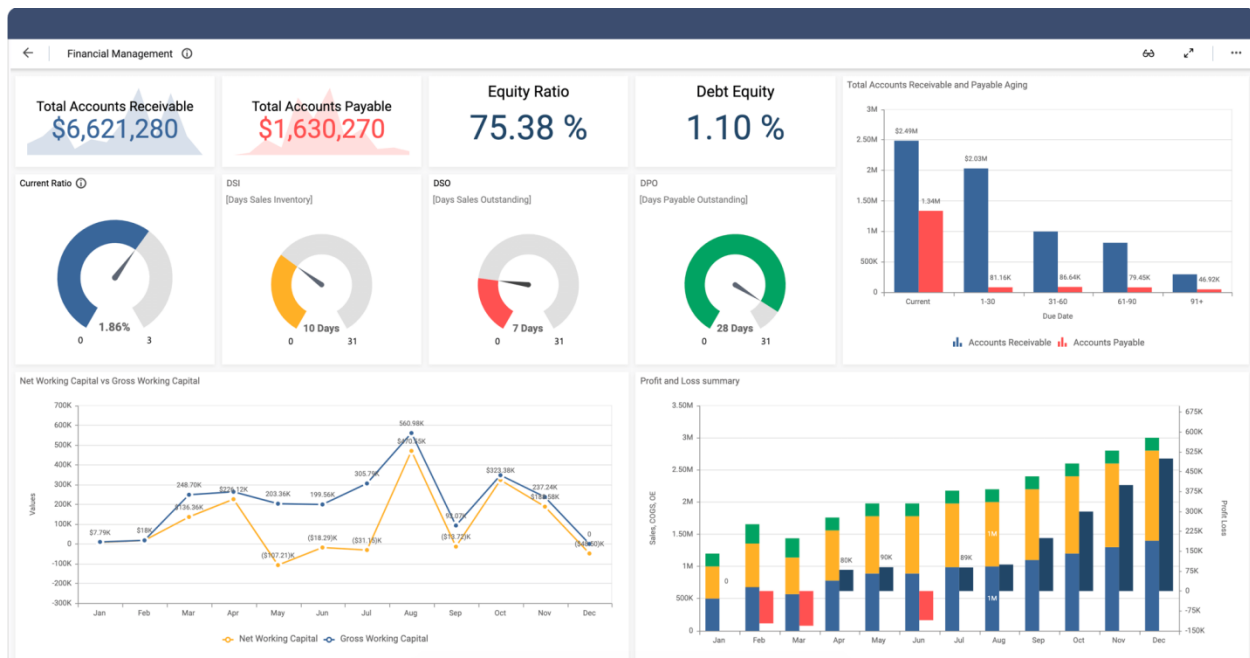
It can be confidently said that these spreadsheets have taken over the earlier paper-based systems, due to its simple yet radical ability to control, operate and analysis of the data. these spreadsheets have provided a modern solution to various financial and statistical problems with the help of built-in functions.

LANPAR was the first electronic spreadsheet introduced in 1969. and after that, there was no looking back, various software firm came up with their version of electronic spreadsheets including Autoplan by General Electric, Financial Planning and Control System by IBM. The list was extended by the introduction of VisiCalc and SuperCalc. VisiCalc rapidly got popularity due to its features of modern spreadsheet applications along with interactive user interface, automatic recalculation, status and formula lines, range copying with relative and absolute references, formula building by selecting referenced cells. however, very soon in 1982, VisiCalc was displaced by IBM's Lotus 1-2-3 and MS-DOS Spreadsheets.

In 1985, Microsoft came up with MS- Excel, which today holds the largest market share on the Windows and macOS platforms. Since 1885, Microsoft has released 16 versions of Excel, to cope up with growing statistical, engineering and financial needs in the industry. Excel also introduced various additional features over time including the introduction of the Pivot Tables and Scenario Manager, interactive interface with Decision Support System (DSS). Microsoft also the programming aspect to excel with the introduction of *Visual Basic for Applications*, which provides useful tools to perform various numerical methods over complex differential and integral equations. It also introduced auto-fill, the ability to drag and expand the selection box to automatically copy a cell or row contents to adjacent cells or rows, adjusting the copies intelligently by automatically incrementing cell references or contents. Along with this, the 16th version of MS-Excel i.e Excel-2016 also offers intelligent cell recomputation with extensive graphing capabilities.

One of the information management tools that can be created using Excel is a dashboard. A dashboard can be considered as a graphical summary or interface of facts and figures. These are generally used for tracking of metrics, data points and

performance indicators. Dashboards provide an easy to grasp representation of data and are often used to simplify complex data sets with the help of pictorial visualizations. It can not only help in the easy interpretation but also helps to gauge exactly how well a department is performing when compared to others. Dashboards can be used for information, operational, analytical or even strategic performance analysis and monitoring. it also helps the manager to gain total visibility of all systems in a glance with the identification of correlations and outliers, which are difficult to detect otherwise.



A typical Dashboard (Src: boldbi.com)

LITERATURE REVIEW

Why we chose this data

There are lots of data all around, many of them can be easily organized and studied and many of them can not be. The main aim of having data is to extract information from it and it solely depends on the type of data that we have, so the choice of data is also a major concern. A dashboard is generally made to monitor multiple data at a time and to extract useful information from them in a easier and faster way. It saves time and increases the productivity of the business. The choice of data for dashboard creation must be such that it can be easily reflected in the form of dashboard and useful information about the data can be obtained from the dashboard itself.

As we all know that one of the most-craziest game for today's generation is cricket. People of all age group from children to adult are equally interested in this game. Now a days cricket is not just a game but also people wants to make it their passion, not just for money but for the name and fame that the players get. When the question comes to IPL every thing else remain aside as it is the most loving form of cricket among the public. It consists of different teams with players from all over the world mixed up together which makes it far more interesting than regular cricket matches. It occurs once every year and as this year it is going to come soon so soon there will be discussions all around and people would be crazy among themselves choosing their favorite team and betting with each other for the matches.

Now the question arises that what are the criteria that people should follow to choose their favorite team. Generally, people look at the previous trends, previous matches, performances of their favorite players in previous matches etc. So, the previous data for IPL matches are such important for the general public. The main reason for selecting this data is the popularity and the diversity of the data. A dashboard of a data is much more useful if the data is vast and diverse as

we can generate a fairly large amount of information from the data and the data for IPL matches is of one such type. It consists of many entities like teams that played the matches, result of the matches, location of the matches etc. With the help of a dashboard we can represent these data in a much simpler way using graphs, pie-charts etc from which one can easily compare among different entities and derive information from the data in much lesser time and effort.

By choosing this data set we were able to generate many relations between the different entities present in the data set and by representing them in the form of dashboard we can derive various conclusions like which player is more consistent, which stadium is more preferable for a particular team, probabilities of the umpires giving correct decision etc. These conclusions are very important for the IPL officials for setting up of the IPL matches in the coming years. There are huge amount of data for IPL matches and to read such a large amount of data and to derive conclusions from it is a very tough job. Thus in this case Dashboard is a very important tool.

As stated earlier the main reason for choosing this data is its vastness and popularity. Also the data is much more diverse than other data sets which makes it more worthy to be represented in the form of dashboard. Also there must be relations between different entities which can be represented on the dashboard and conclusions can be derived regarding that. All these conditions are fulfilled by the data for IPL matches that we have chosen for our project. So choosing this data was fruitful for us as we able to derive many important conclusions from the data and these conclusions are very important for the IPL lovers and IPL officials for setting up of upcoming matches making them much more fair and interesting than before.

Brief description of the dataset

The dataset chosen by our team describes the results of various cricket matches played in the IPL series.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
1	id	season	city	date	team1	team2	toss	win/loss	decision	result	dl_apple	winner	win_by	r_win_by	w_player	of	venue	umpire1	umpire2									
2	1	2017	Hyderabad		Sunrisers	Royal Cha	Royal Cha	field	normal	0	Sunrisers	35	0	Yuvraj Sin	Rajiv Gan	AY Dandel	NI Ulong											
3	2	2017	Pune		Mumbai Ii	Rising Pur	Rising Pur	field	normal	0	Rising Pur	0	7	SPD Smith	Maharash	A Nand Ki	S Ravi											
4	3	2017	Rajkot		Gujarat Li	Kolkata Kr	Kolkata Kr	field	normal	0	Kolkata Kr	0	10	CA Lynn	Saurashtra	Nitin Men	CK Nandan											
5	4	2017	Indore		Rising Pur	Kings XI P	Kings XI P	field	normal	0	Kings XI P	0	6	GJ Maxwe	Holkar Cri	AK Chaud	C Shamshuddin											
6	5	2017	Bangalore		Royal Cha	Delhi Dar	Royal Cha	bat	normal	0	Royal Cha	15	0	KM Jadhav	M Chinnaswamy	Stadium												
7	6	2017	Hyderabad		Gujarat Li	Sunrisers	Sunrisers	field	normal	0	Sunrisers	0	9	Rashid Kh	Rajiv Gan	A Deshm	NI Ulong											
8	7	2017	Mumbai		Kolkata Kr	Mumbai Ii	Mumbai Ii	field	normal	0	Mumbai Ii	0	4	N Rana	Wankhedi	Nitin Men	CK Nandan											
9	8	2017	Indore		Royal Cha	Kings XI P	Royal Cha	bat	normal	0	Kings XI P	0	8	AB Patel	Holkar Cri	AK Chaud	C Shamshuddin											
10	9	2017	Pune		Delhi Dar	Rising Pur	Rising Pur	field	normal	0	Delhi Dar	57	0	SV Sampat	Maharash	AY Dandel	S Ravi											
11	10	2017	Mumbai		Sunrisers	Mumbai Ii	Mumbai Ii	field	normal	0	Mumbai Ii	0	4	JJ Bumrah	Wankhedi	Nitin Men	CK Nandan											
12	11	2017	Kolkata		Kings XI P	Kolkata Kr	Kolkata Kr	field	normal	0	Kolkata Kr	0	8	SP Narine	Eden Gar	C Deshm	NI Ulong											
13	12	2017	Bangalore		Royal Cha	Mumbai Ii	Mumbai Ii	field	normal	0	Mumbai Ii	0	4	KA Pollard	M Chinnaswamy	Stadium	AK Chaudhary											
14	13	2017	Rajkot		Rising Pur	Gujarat Li	Gujarat Li	field	normal	0	Gujarat Li	0	7	AJ Tye	Saurashtra	A Nand Ki	S Ravi											
15	14	2017	Kolkata		Kolkata Kr	Sunrisers	Sunrisers	field	normal	0	Kolkata Kr	17	0	RV Uthap	Eden Gar	AY Dandel	NI Ulong											
16	15	2017	Delhi		Delhi Dar	Kings XI P	Delhi Dar	bat	normal	0	Delhi Dar	51	0	CJ Anders	Feroz Sha	YC Barde	Nitin Menon											
17	16	2017	Mumbai		Gujarat Li	Mumbai Ii	Mumbai Ii	field	normal	0	Mumbai Ii	0	6	N Rana	Wankhedi	A Nand Ki	S Ravi											
18	17	2017	Bangalore		Rising Pur	Royal Cha	Royal Cha	field	normal	0	Rising Pur	27	0	BA Stokes	M Chinnaswamy	Stadium	C Shamshuddin											
19	18	2017	Delhi		Delhi Dar	Kolkata Kr	Delhi Dar	bat	normal	0	Kolkata Kr	0	4	NM Coult	Feroz Sha	Nitin Men	CK Nandan											
20	19	2017	Hyderabad		Sunrisers	Kings XI P	Kings XI P	field	normal	0	Sunrisers	5	0	B Kumar	Rajiv Gan	AY Dandel	A Deshmukh											
21	20	2017	Rajkot		Royal Cha	Gujarat Li	Gujarat Li	field	normal	0	Royal Cha	21	0	CH Gayle	Saurashtra	S Ravi	VR Sharma											
22	21	2017	Hyderabad		Sunrisers	Delhi Dar	Sunrisers	bat	normal	0	Sunrisers	15	0	KS William	Rajiv Gan	CB Gaffan	NI Ulong											
23	22	2017	Indore		Kings XI P	Mumbai Ii	Mumbai Ii	field	normal	0	Mumbai Ii	0	8	JC Buttler	Holkar Cri	M Erasmus	C Shamshuddin											
24	23	2017	Kolkata		Kolkata Kr	Gujarat Li	Gujarat Li	field	normal	0	Gujarat Li	0	4	SK Raina	Eden Gar	CB Gaffan	Nitin Menon											
25	24	2017	Mumbai		Mumbai Ii	Delhi Dar	Delhi Dar	field	normal	0	Mumbai Ii	14	0	MJ McCle	Wankhedi	A Nand Ki	S Ravi											
26	25	2017	Pune		Sunrisers	Rising Pur	Rising Pur	field	normal	0	Rising Pur	0	6	MS Dhoni	Maharash	AY Dandel	A Deshmukh											
27	26	2017	Rajkot		Kings XI P	Gujarat Li	Gujarat Li	field	normal	0	Kings XI P	26	0	HM Amla	Saurashtra	AK Chaud	M Erasmus											
28	27	2017	Kolkata		Kolkata Kr	Royal Cha	Royal Cha	field	normal	0	Kolkata Kr	82	0	NM Coult	Eden Gar	CB Gaffan	CK Nandan											
29	28	2017	Mumbai		Rising Pur	Mumbai Ii	Mumbai Ii	field	normal	0	Rising Pur	3	0	BA Stokes	Wankhedi	A Nand Ki	S Ravi											
30	29	2017	Pune		Rising Pur	Kolkata Kr	Kolkata Kr	field	normal	0	Kolkata Kr	0	7	RV Uthap	Maharash	AY Dandel	NI Ulong											
31	30	2017	Bangalore		Royal Cha	Gujarat Li	Gujarat Li	field	normal	0	Gujarat Li	0	7	AJ Tye	M Chinnaswamy	Stadium	C Shamshuddin											
32	31	2017	Kolkata		Delhi Dar	Kolkata Kr	Kolkata Kr	field	normal	0	Kolkata Kr	0	7	G Gambhir	Eden Gar	NI Ulong	S Ravi											
33	32	2017	Chandigar		Sunrisers	Kings XI P	Kings XI P	field	normal	0	Sunrisers	26	0	Rashid Kh	Punjab Cri	Nitin Men	CK Nandan											
34	33	2017	Pune		Rising Pur	Royal Cha	Royal Cha	field	normal	0	Rising Pur	61	0	LH Ferguson	Maharash	KN Anant	M Erasmus											
35	34	2017	Rajkot		Gujarat Li	Mumbai Ii	Gujarat Li	bat	tie	0	Mumbai Ii	0	0	KH Pandey	Saurashtra	AK Chaud	CB Gaffan											
36	35	2017	Chandigar		Delhi Dar	Kings XI P	Kings XI P	field	normal	0	Kings XI P	0	10	Sandeep	Punjab Cri	YC Barde	CK Nandan											
37	36	2017	Hyderabad		Sunrisers	Kolkata Kr	Kolkata Kr	field	normal	0	Sunrisers	48	0	DA Warner	Rajiv Gan	AY Dandel	S Ravi											
38	37	2017	Mumbai		Royal Cha	Mumbai Ii	Royal Cha	bat	normal	0	Mumbai Ii	0	5	RG Sharma	Wankhedi	AK Chaud	CB Gaffan											
39	38	2017	Pune		Gujarat Li	Rising Pur	Rising Pur	field	normal	0	Rising Pur	0	5	BA Stokes	Maharash	M Erasmus	C Shamshuddin											
40	39	2017	Delhi		Sunrisers	Delhi Dar	Delhi Dar	field	normal	0	Delhi Dar	0	6	Mohamm	Feroz Sha	YC Barde	Nitin Menon											
41	40	2017	Kolkata		Kolkata Kr	Rising Pur	Rising Pur	field	normal	0	Rising Pur	0	4	BA Tripath	Eden Gar	KN Anant	A Nand Kishore											
42	41	2017	Delhi		Gujarat Li	Delhi Dar	Delhi Dar	field	normal	0	Delhi Dar	0	7	RR Pant	Feroz Sha	M Erasmus	Nitin Menon											
43	42	2017	Bangalore		Kings XI P	Royal Cha	Royal Cha	field	normal	0	Kings XI P	19	0	Sandeep	M Chinnaswamy	Stadium	C Shamshuddin											

Matches.csv

The dataset chosen by our team describes the results of various cricket matches in the IPL series.

The dataset has 17 columns and 757 rows describing the match id, the IPL season in which match was played, the city in which the match was played, the date of

match, the teams who played the match, the team who won the toss, the decision of the toss winning team to either bat or field first, the result of the match, the no of runs by which the match was won, the number of wickets by which the match was won, the player of the match, the venue of the match and the names of first umpire and second umpire.

In the id column, every match is assigned a unique integer starting from one till 11415. It acts as a unique identifier (UID) for the dataset. The values in the "season" column are integer, ranging from 2017-2019. It's usually the year in which the match was played. The date column stores the date of the match. In the next two columns describes the name of the teams that are going to play that match.

The next column "toss winners" stores the name of the team who won the toss. For predicting the results of a match the results of the toss play an important role. The judgment of the team who won the toss, to bat or to field first is given in the next column "toss decision". The strategy of playing teams varies with the opponent and the field. Based on it the teams usually decide to bat or to field first.

The city column describes the place where the match was played. Match results can also be affected by the place where the match was played making it an important column for predicting the result of a match. Each team plays one home and one away game, success at their home venue is a key factor during the league stage. It is often that the opening and the final matches of the season are played at the home ground of the final winners of the previous season. The description of the ground in which the match was played is listed in the "venue" column.

The "results" column describes how the results were. For a strict win, its value is normal, for tie its value is a tie. The next column tells about how many runs the winning team secured more than the losing one, in case if the team who batted first wins the game, else the values are zero.

The next column tells about how many wickets the winning team secured more than the losing one, in case if the team who fielded first wins the game, else the values are zero.

The next column describes the player of the match. The name of the player who performed best is written in this column. The prestige of the player who is awarded by the title of the player of the match increases.

Umpires play an important role in maintaining fairness in the game. The next two columns store the name of the first and second umpire.

The dataset described above was only a small sample of a few cricket matches which may seem important only for the fans but the real fact is that a lot of this data is used for planning strategies for a match. The relation between the result and various columns of such a small dataset can conclude that by changing various factors like toss results, venue of the match, etc the results of the match can be altered. By studying a much larger dataset and applying the modern-day techniques of machine learning one can predict the outcomes of the match with great precision. With deep insight into this data, one will not only be able to predict but also prescribe the strategies for better results. If the same technique is applied in other sports the country can be helped in training the sportsman by predicting the best suitable conditions for the training.

METHODOLOGY

FEATURES USED TO CREATE DASHBOARD:

PIVOT TABLE:

Pivot tables are one of Excel's strongest features which make it unique in its own way. The axis table allows you to fetch values in large data points.

PivotTable is an interactive and user-friendly way to quickly and efficiently capture large amounts of data. We can use PivotTable to analyze numerical data in detail and answer unexpected questions about our data.

The axis table is a table of statistics, which captures more comprehensive data of the table. This summary may include totals, averages or other calculations that come together in a meaningful way in axis groups.

The fields to be created appear on the right side of the worksheet.

Each field from the list can be extracted from this property, which has four modes.

a-> filter

b-> column label

c-> lower label

d-> summative value

Here is example of pivot table that we created using our dataset,

Row Labels	Count of id
2008	58
2009	57
2010	60
2011	73
2012	74
2013	76
2014	60
2015	59
2016	60
2017	59
2018	60
2019	60
Grand Total	756

Pivot Chart

Pivot Charts represents PivotTables graphically by adding visualizations and graphics to the summary data in a PivotTable, and allow us to easily see comparisons, patterns, and trends. Both of PivotTables and Pivot Charts enable us to judge informed decisions about very critical data or information in our enterprise.

With the help of these features we can connect to other data sources such as SQL Server tables, SQL Server Analysis Services cubes, Azure Marketplace, Office Data Connection files, XML , Access databases, and simple text files to create PivotTables, or use current PivotTables to generate new tables.

Pivot Charts basically display data series, categories, data markers, and axes as standard charts. we can change chart types and title, legend placement, data label, chart position and other options.

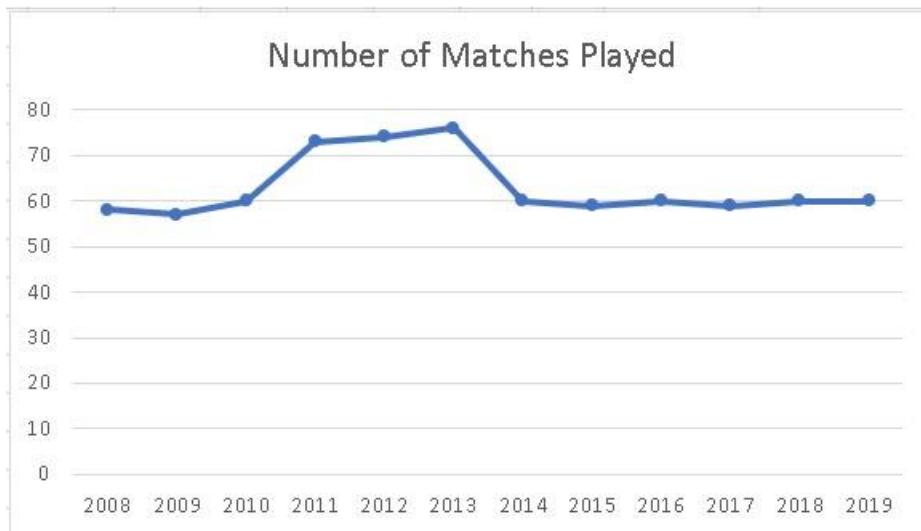
There are different types of pivot charts available in Excel such as

- 1.Column chart
- 2.line chart
- 3.pie chart
- 4.bar
- 5.Area
- 6.XY scatter
- 7.Stock
- 8.Surface
- 9.Radar
- 10.Treemap
- 11.Sunburst
- 12.Histogram
- 13.Box and Whisker
14. Waterfall
15. Combo

for each chart type, there are different subtypes are available. for example in line chart there are simple line chart, stacked line chart ,3d line chart ,line chart with markers and so on.

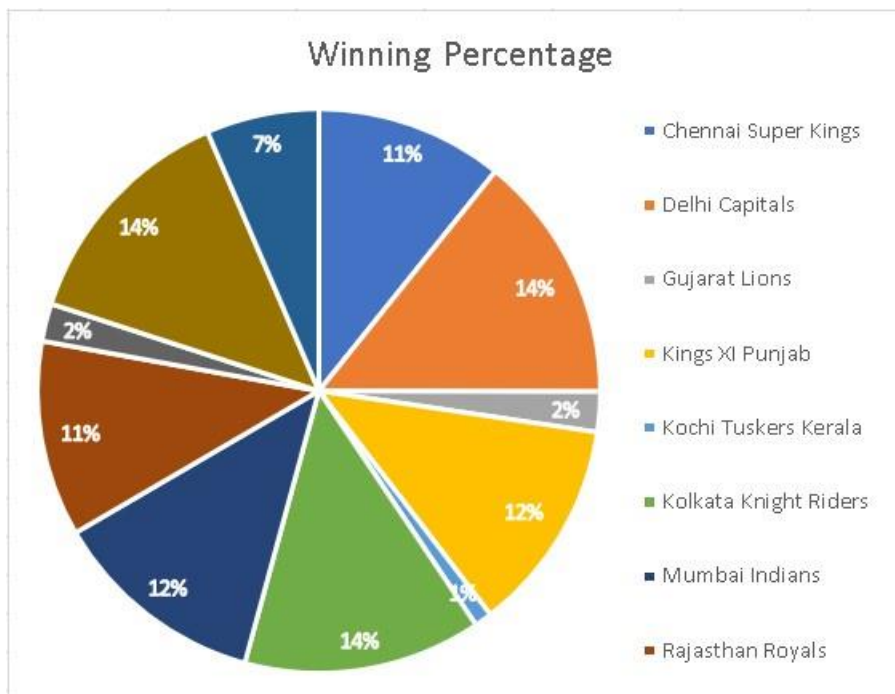
Here is example of our pivot chart between number of matches played between different teams in IPL and in which year. The year ranges from 2008 to 2019.

A=line chart



B=pie chart

Here we created the pie chart which shows the winning percentage of all different team in IPL.



DIFFERENCE BETWEEN PIVOT CHART AND STANDARD CHARTS

If well we know about the standard charts, we will find that most of the task or operations in Pivot Charts are very similar. However, there are some major differences in these two:

Row - Column Orientation -> As in case of the standard chart, we can change the row-column, but in the case of pivot chart orientation of the pivot chart can be change using the Data Source Select dialog box. Instead, we can pivot the rows and column labels of the corresponding pivot Table to achieve the same effect.

Chart Type -> we can convert PivotChart to any other chart type except few charts like XY, stack or bubble chart.

Source Data -> Standard maps are directly linked or maintain their connection to worksheet cells, but pivot charts only depend on the data source of their associated pivot Table. Unlike the standard chart, we cannot change the chart data range in the PivotChart's Data Source Select dialog box.

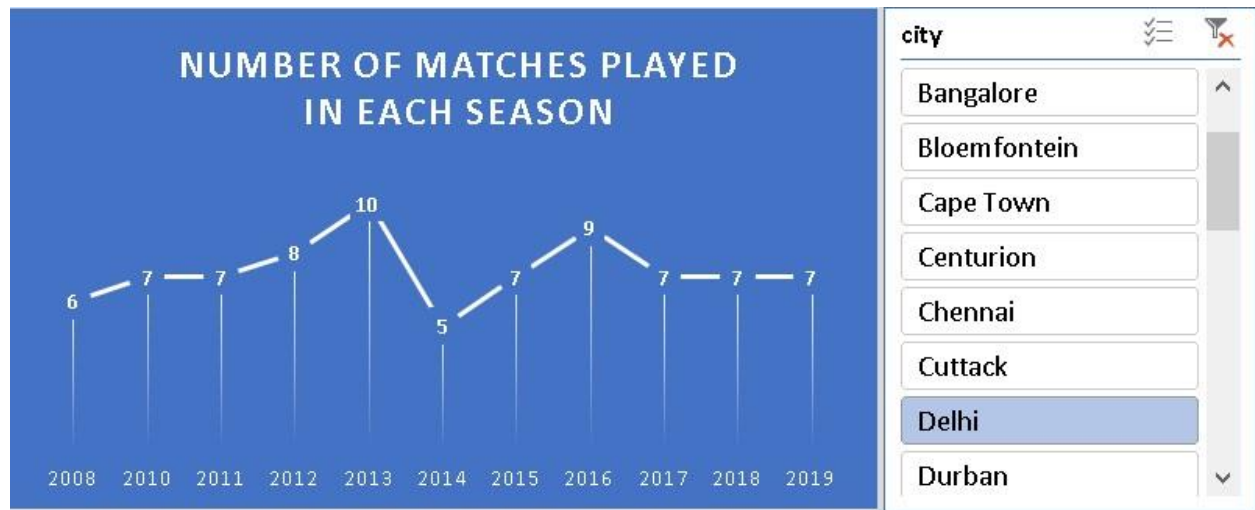
Formatting -> When we refresh a PivotChart, most of the charting formats, layout, and style that we add are saved. However, trendlines, data labels, error bars, and other changes to the data set are not saved. Standard maps do not lose this configuration once implemented.

Although we cannot re-size the labels directly on the pivot chart, we can increase the text font size to effectively resize the labels.

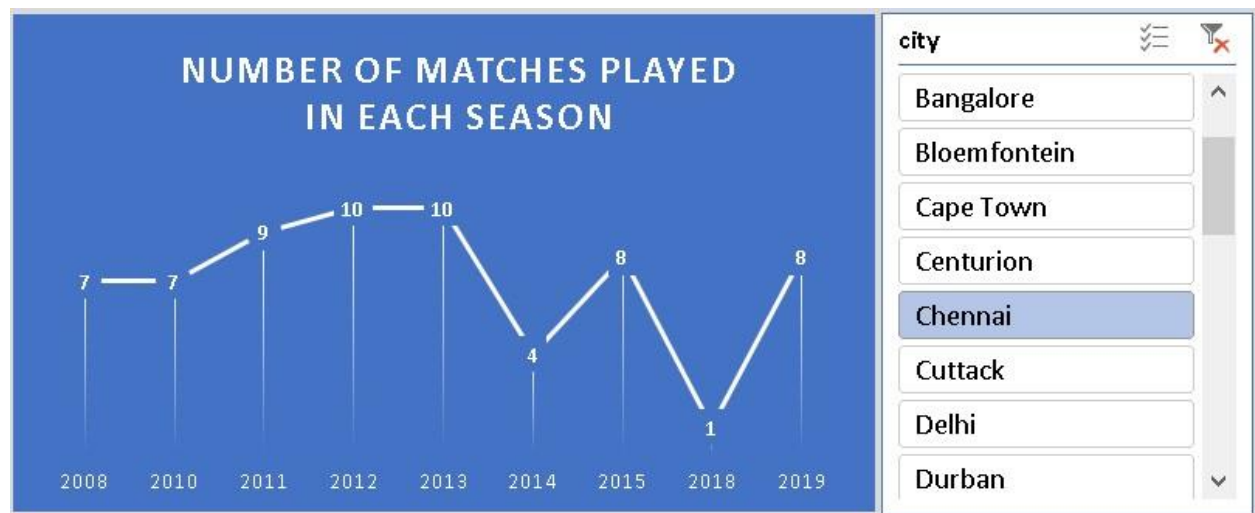
SLICERS:

Slicers provide us different options to filter the or pivot table .In addition to fast filtering It also help to show the current state which make it very easy to understand for user. In short, we can say that it is a visual filtering tool for pivot table.

Here we use slicers to show the number of matches played by different teams in IPL in Delhi for each year from 2008 to 2019.



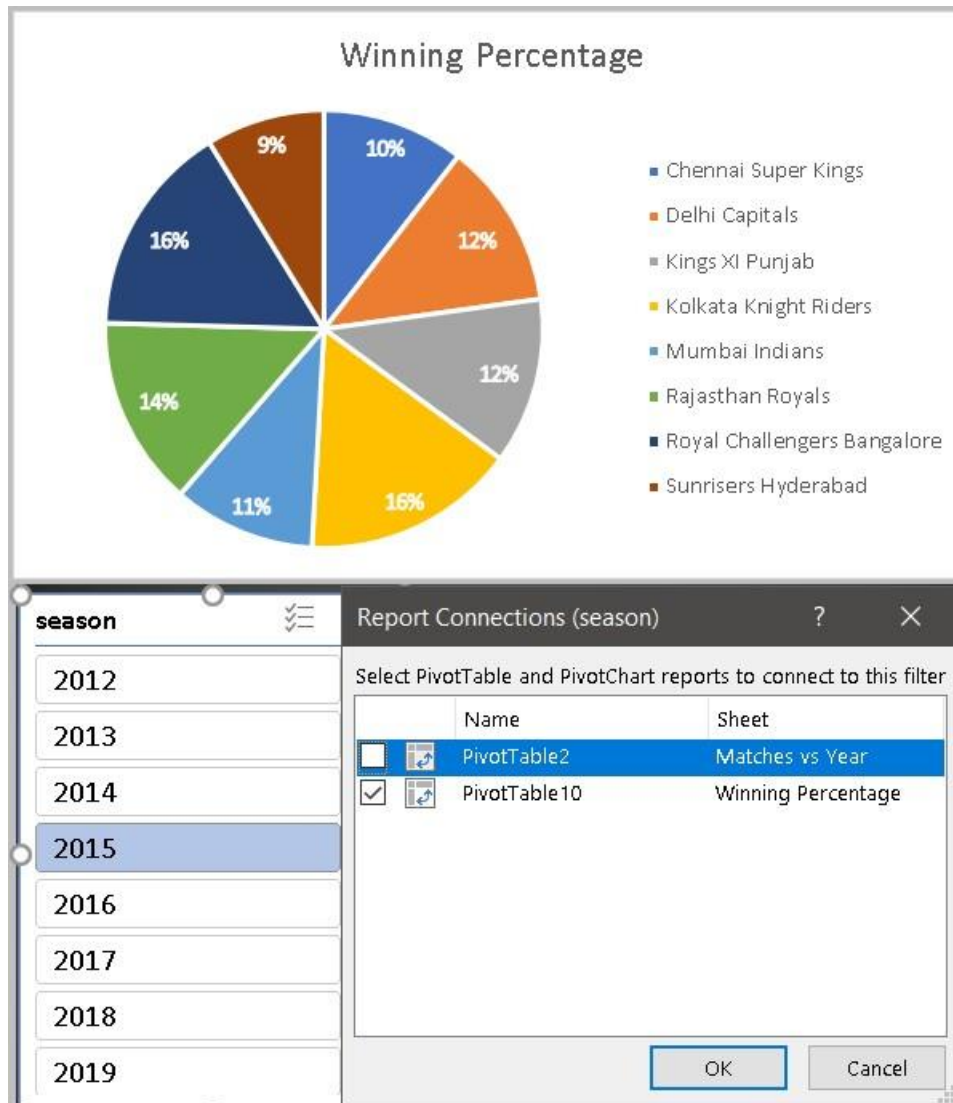
Similarly here we showed the number of matches played in city Chennai from year 2008 to 2019



REPORT CONNECTION:

Report connection helps us to maintain the connection between the pivot table and the pivot charts. With the help of this we can dynamically change the pivot chart corresponding to different values in pivot table.

Here we shows that the pie chat which is connected to corresponding pivot table (left bottom in the picture) with the help of report connection the pie chart changes its value and shows the winning percentage of different IPL team that played during that particular here we given the example of pie chart and the year is 2015, If we change the year the corresponding pie chart also changes.



Analyzing the Dataset

HOME ADVANTAGE (Come on Lads!)

Does home turf resonates with glory in IPL? You'd reckon “No Mate!” , given that the matches are all held in land of dust bowls or flat tracks with no bone in the steak(Barring for the god forsaken 2009 IPL in the Protea arena). Most of the prodigies participating are so called “Desi”, many of whom had their first-class / age-group adventure at their home grounds. Travelling plays its part, but it's the same for all so Gauss won't be disappointed as the home matches are not held in arrow.

In the previous three editions when the tournament was held in India, the home teams didn't have a huge bias to blabber about, but with about a 33% of the games done after, home teams have won an surprisingly high number of games.

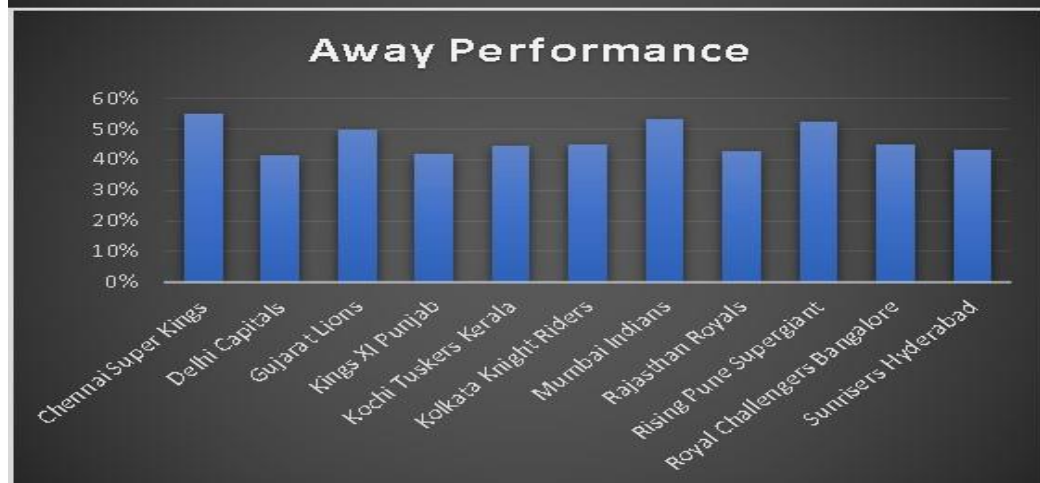
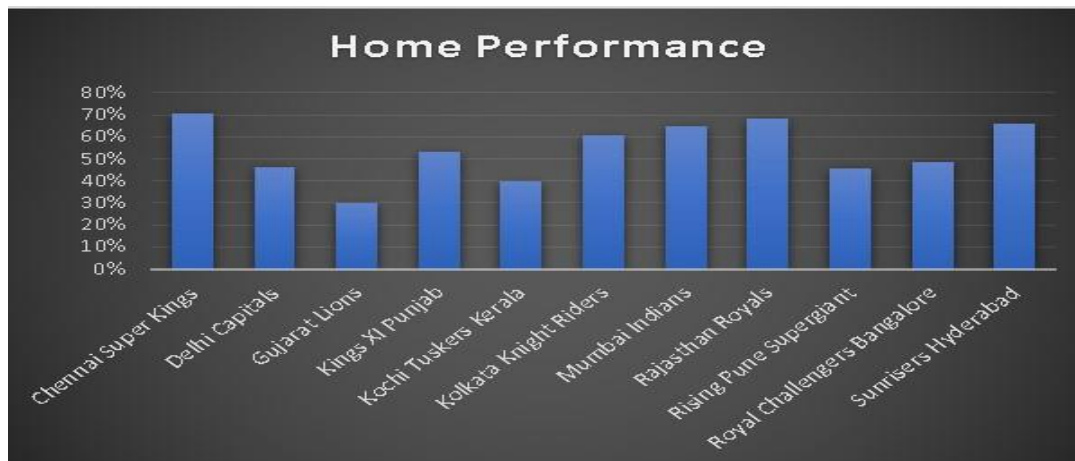
In fact, last season playing at home was like shooting in one's foot, with teams biting the dust more often than in front of their home support (32-40). That was a blip, though, as home teams have generally had a slight positive bias: in the first three occurrences, they collectively tasted glory “ $30/2=15$ ” times they had to bear drubbing at their own arena.

Now, As the “Greatest cricket analyst of our generation cum test cricketer cum game plan genius” Mr.Aakash Chopra says-

“Numbers do tell a story” except the fact “For MS Dhoni numbers lie, and his presence is enough although the strike rate is below 80 and average below 35 in the past 4 years ,he is the best finisher currenty,spare some thoughts for Stokes, Butler and Mr. 360”.

Let's Analyse:

Row Labels	Sum of HP	Row Labels	Sum of AP
Chennai Super Kings	0.704918033	Chennai Super Kings	0.553398058
Delhi Capitals	0.460526316	Delhi Capitals	0.415841584
Gujarat Lions	0.3	Gujarat Lions	0.5
Kings XI Punjab	0.534246575	Kings XI Punjab	0.417475728
Kochi Tuskers Kerala	0.4	Kochi Tuskers Kerala	0.444444444
Kolkata Knight Riders	0.608108108	Kolkata Knight Riders	0.451923077
Mumbai Indians	0.646341463	Mumbai Indians	0.533333333
Rajasthan Royals	0.680851064	Rajasthan Royals	0.43
Rising Pune Supergiant	0.454545455	Rising Pune Supergiant	0.526315789
Royal Challengers Bangalore	0.486842105	Royal Challengers Bangalore	0.451923077
Sunrisers Hyderabad	0.66	Sunrisers Hyderabad	0.431034483
Grand Total	5.936379119	Grand Total	5.155689574



Now it is clear from the data that the teams fare well then, they play in front of their home support. Well you know, actually you don't, what it feels like when you charge for your run up and the crowd is behind you, you bowl better than ever, for example Mitchell Johnson a bowler who bowls in the mid 145 k's range got the ball dancing on this tune in excess of 155. The English were petrified to their death when Johnson used to hit the short length.

Now good teams play well whether its Mars or Moon. Similar is the case for "THALA" boys or more apt "Dad's Army", but as you can see from the graph that they suck the blood of their opponents at their slow and sluggish, to be honest exasperating track by fielding world class spinners like the one and only "Turbunator" and the man who runs faster than Bolt after taking wickets "Parasakthi Express".

Similarly the flat track suits the hard hitters of Mumbai who have the likes of "Big lad"-Pollard, "Ro-hitman Sharma", Pandya brothers and Cutting backed by their superb death attack comprising of "Sling"-Malinga and Jasprit "Gumrah".

Kolkata Knight Riders also take their turning track to their strengths, by shackling the batsmen with the mystery of Sunil Narine and "the ever 16" Piyush Chawla with "Dre" Russel muscling the balls to all corners of the small Eden Gardens.

For the ever underrated Sunrisers Hyderabad, The good top order makes enough runs for their impeccable bowling attack comprising the likes of "PARAKRAM" promoter Rashid Khan, and Bhuvneshwar Kumar.

But, you might be wondering why the "Red Army" is at the disadvantage, even after having the like of "SUPERMAN", MR.360 Abraham Benjamin Devilliers and "RUN MACHINE" -Virat Kohli, past players like "UNIVERSE BOSS" Chris Gayle and "All-Rounder" K.L. Rahul. The blame here goes to their perennial poor death bowling having the likes of Super over disappointment "Tim Southee" and Mohammad Siraj, or even the ever learning Umesh Yadav, you are bound to get a beating.

In Contrast, Delhi Capitals suffer because of their slow and uneven track which doesn't suit their "Indian Superstars" like Rishabh Pant, Shreyas Iyer and Rishabh Pant.

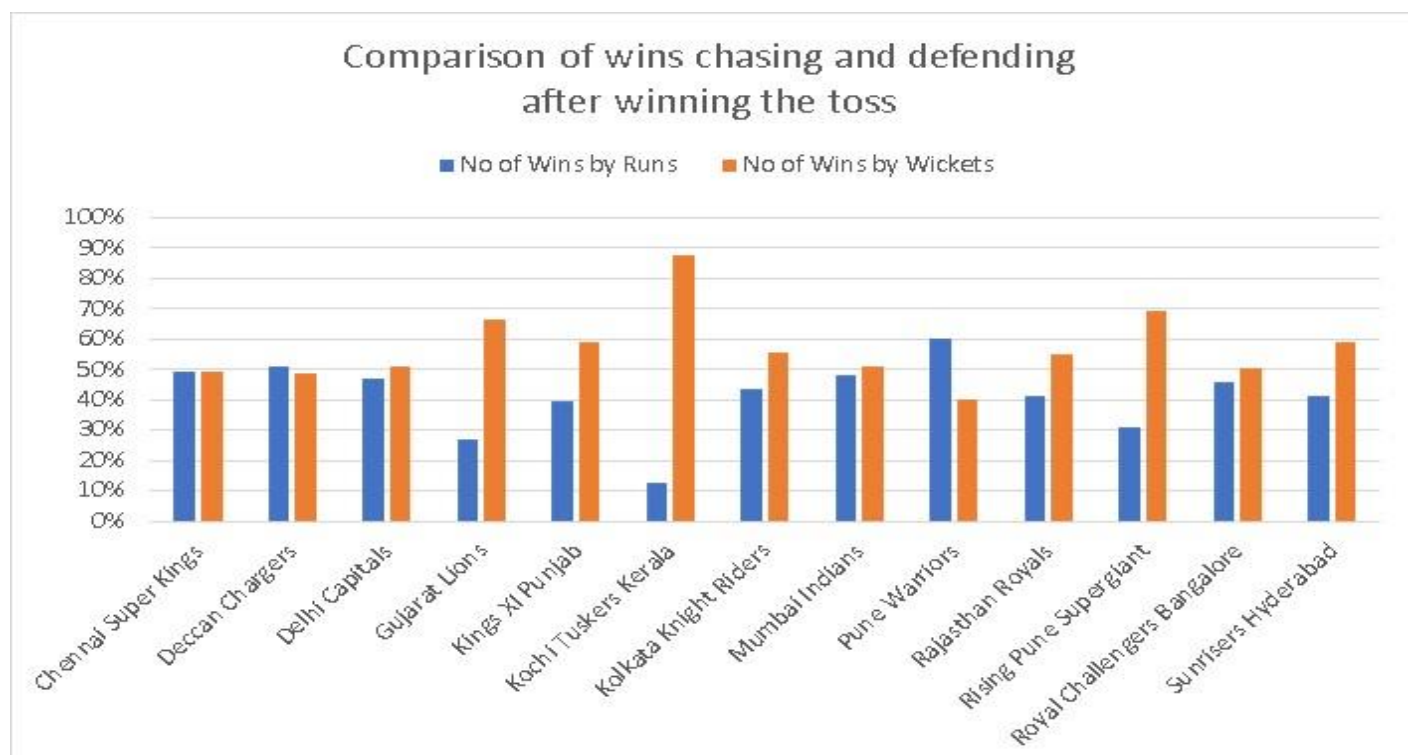
Biasness of Toss

Now let's, look at another biasness in IPL that has happened in the recent years. The team that wins the toss prefers chasing due to the dew factor. The 8 seam Kookaburra ball easily gets wets due to the dew and when the moisture soaks in there is no coming back, the ball becomes merely a soap bar in the hands of a spinner.

The dew effect is at max At the Wankhede Stadium where during the 2016 World T20,even targets in excess of 230 were not safe and were easily chased with 5 balls to spare .Isn't it stupendous?

The similar advice was given to Ms Dhoni to field an Extra seamer in The Semifinals, but to our utter dismay it went unheard."SIXER KING" Yuvraj was injured and out of the tournament and due to the heavy dew at Wankhede ,Ashwin and Jadeja were mocked by the humongous sixes of Lendl Simmons ,Virat kohli was forced to bowl and we lost ,no Surprise. Well nobody can read the mind of MS Dhoni.The target of 193 was mocked in style, in contrast to the Single and Doubles that MS Dhoni took, the Calypso guys gave the leather quite a beating.

Now, let's Analyze in reference to IPL:



As you can see the graph, with the sole exception of the now defunct Deccan Chargers, all the teams prefer chasing.

For, Gujarat Lions their strong batting attack comprising of Brendon McCullum,"MR IPL" Suresh Raina,Dwyane smith ,Aaron Finch,DK and others....chasing was far more better than relying on the pas this prime bowler Pravin Kumar to win you matches.

Similarly, for Rps their nice batting attack comprising of Ajinkya Rahane,Steven Smith "The sandpaper guy",Rahul Tripathi,Ben Stokes barring MS.

For Sahara Pune Warriors, they swim in the direction opposite to the current. They neither had good batting, neither good bowling. Chris Gayle had a feast at their attack once and all the bowlers were vandalized like hell.

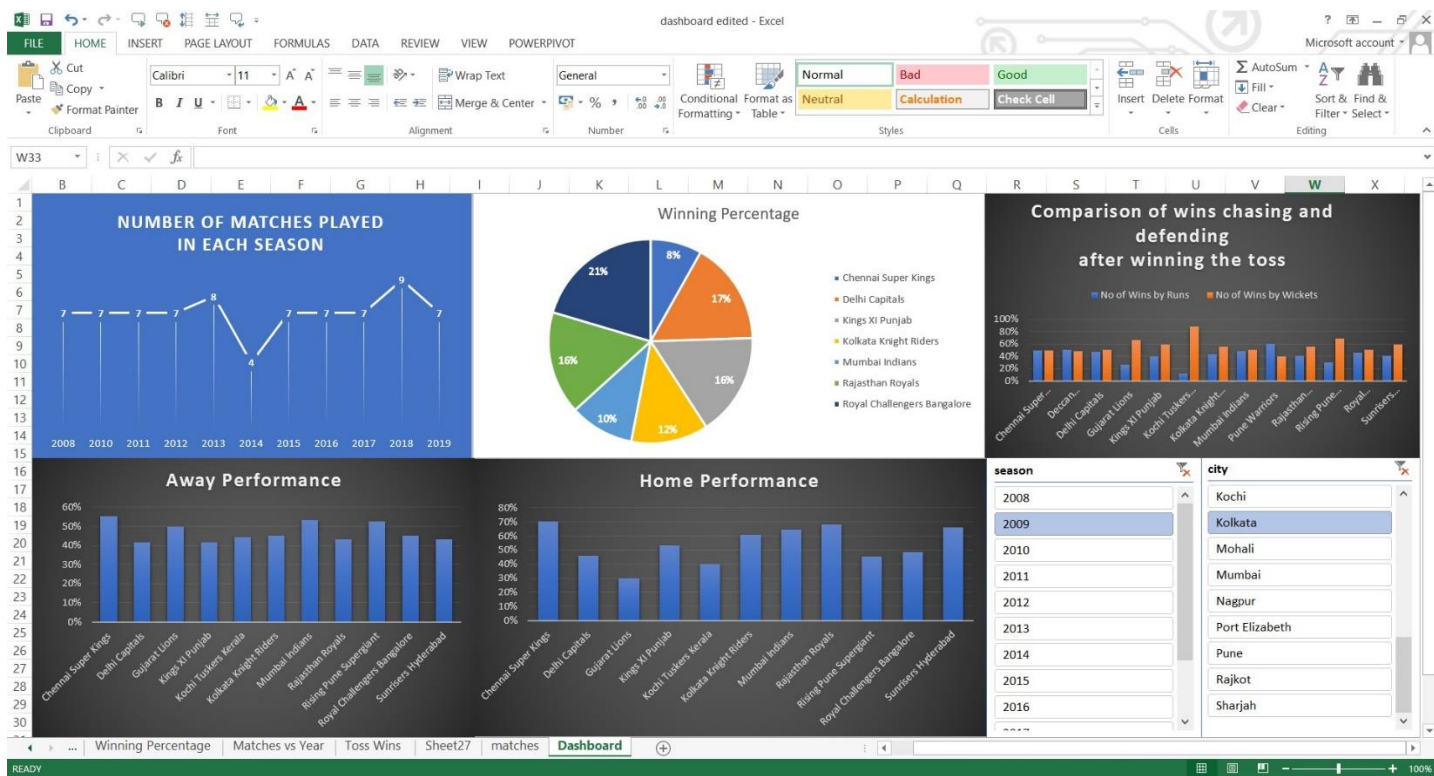
Although, biasness level is low but it's still enough to classify that winning the toss puts you in the driving seat due to dew factor.

Result and Analysis

Result

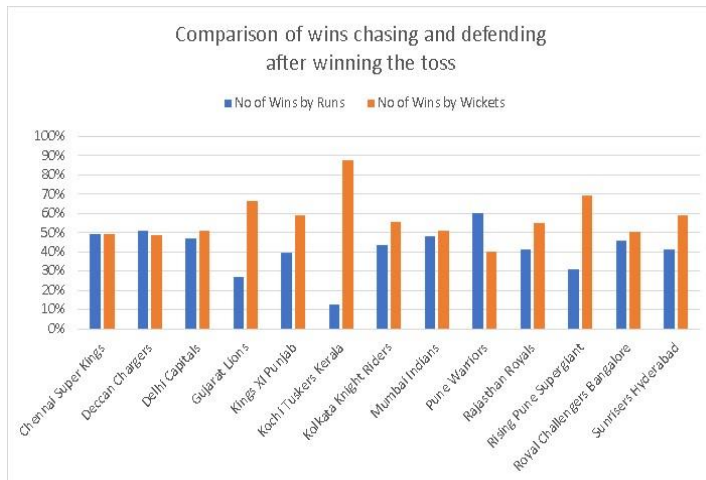
The result is the creation of a dashboard that shows how various factors affect the winning chances of team such as choosing chasing or defending after winning the toss, performance variation due to playing in home or away. These trends can greatly help in predicting the winning chances of team.

The following dashboard depicts the above mention trends with tools to choosing the session and place.



The dashboard contain 5 charts showing various trends and 2 slicers providing option to choose season and city.

Analysis

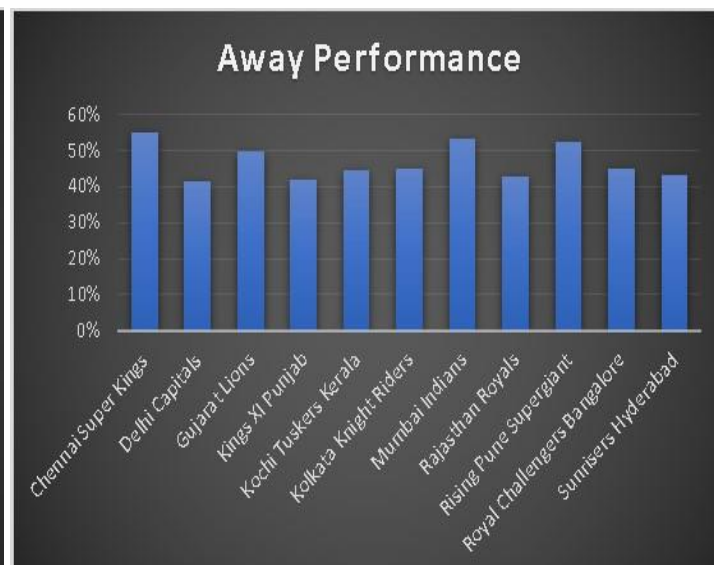


Analysis 1:

Analyzing the dashboard, helps in deciding various factors that may increase the chances of winning of a particular team. For instance, taking the help of “*Comparison of wins chasing and defending after winning the toss*” chart it is clearly visible that team “Rising

Pune Supergiant” has a high rate of winning when opt to choose chasing rather than defending as clearly shown in the chart. So using the help of this chart it can be predicted if the team is better in chasing or defending.

Analysis 2:



Also from the two charts “*Home Performance*” and “*Away Performance*” it can be seen that certain teams has high rate of winning and certain teams has the

opposite true. The two charts are shown above. It can be clearly seen that the teams like "Chennai Super Kings", "Rajasthan Royals", "Sunrise Hyderabad" has higher Home wins while teams like "Gujarat Lions", "Kolkata Knight Riders" have lower winning rate when plays the match at home.

CONCLUSION

From the two case studies being done with the aid of dashboard ,some interesting patterns were observed ,the use of dashboard gives us a visual graphical and enlightens us with knowledge that could be helpful.

From the first case study,it shows that the teams prepare the pitches suitable to their team's liking.So the opponent can expect what to expect when going into someone's den.they could prepare well in advance and analyse the possible X1 according to pitch and past records and can find a way to pierce through the team's armour.

From the second case study, we can infer that under some conditions, the team must plan both scenarios of the toss outcome as data shows the biasness in the results. Although toss has luck factor when can prepare well in advance and not be arrogant in team planning. Horses for courses. As in the past, the greatest example of not respecting the toss analysis was very devastating for Team India at the World T20 2016 semis where an arrogant decision by the Captain led the whole world cup campaign into damnation.

So, as we can see, stats never lie and with the help of dashboard analysis we can see hidden patterns with a more layman scope.

RECOMMENDATIONS

Many conclusions were drawn from the dashboard created from the dataset of IPL matches. We came to know about many factors that influenced the matches and the teams that had command over those factors performed surprisingly well in the matches. These factors are of great importance for the teams playing these matches they should take care of it well before the matches to increase their odds of winning.

While our dashboard model was completed and studied there are various recommendations that are equally important to be taken care of. As we know a dashboard represents its data in such a format that it delivers a large amount of information regarding the data in very less time, So data is a very important tool for dashboard creation. How much accurate the data is, is very important. Alterations with the data accuracies may lead to false correlations among the data and it may deliver false information to the reader. Thus, in this way we can see the significance of accuracy of the dataset and due to this it is highly recommended that the dataset should be accurate.

Many different tools could be used like Machine Learning, Data Mining etc. which increases the understanding and helps in better analysis of the data. These techniques can even be used for gathering accurate data. There are also various other software's and tools which could be used to make the dashboard which are even more accurate and useful. Due to unavailability of these softwares and tools and lack of knowledge at them we created this dashboard on Microsoft Excel. These high efficiency tools and softwares are highly recommended for business professionals working on dashboards.

References

websites: -

KAGGLE:-

<https://www.kaggle.com/nowke9/ipldata>

<https://storage.googleapis.com/kagglesdsdata/datasets/29843/429151/matches.csv?GoogleAccessId=web-data@kaggle-161607.iam.gserviceaccount.com>