



You are a Sports Analyst who specializes in Cricket. As the IPL mega auction 2018 is near and every team will start afresh with a new combination of players, you are being hired by the XYZ IPL team to help them with the best insights and give them an edge in the tournament. You have been provided with some data files which contains data(IPL 2013-2017) about the following:

- Ball by Ball details of every match
- Player Profile
- Match details
- Playing XI details
- Teams and their IDs

You are required to analyze the dataset and come up with key insights using any data visualization as your team owners will only have time to see the dashboard you create and make decisions on the basis of that. You are free to use your own metrics and come up with a viable strategy.

To help you with some tips:

- You have been provided with an open ground to come up with your creative ideas but only using the dataset provided.
- Some popular metrics for analysis can be player averages, strike rates etc.
- You can choose your metric which can help the owners reach a particular decision faster.
- For clarity, you can also come up with your best team as per your analysis which should satisfy the rules of IPL(min 18 players max 25 players with max 8 overseas players in a team)
- Playing team formation will require your justification to the panelists. However, coming up with your playing team is not mandatory.

Note: Coming up with unique metrics can attract brownie points provided you have the justification ready with you 😊

Data set
Ball_by_Ball

Variable name	Data type	Description
<i>Match_id</i>	<i>Numeric</i>	<i>Gives unique match id for every record based on the match to which it belongs</i>
<i>over_id</i>	<i>Numeric</i>	<i>Unique over id for each innings</i>
<i>Ball_id</i>	<i>Numeric</i>	<i>Unique ball id for each ball in an over</i>
<i>Innings_No</i>	<i>Numeric</i>	<i>Tells you if it is the first innings or second innings. 3 and 4 are present in the scenarios of super over</i>
<i>Team_Batting</i>	<i>String</i>	<i>Unique id of the team that is batting</i>
<i>Team_Bowling</i>	<i>String</i>	<i>Unique id of the team that is bowling</i>
<i>Striker_Batting_Position</i>	<i>Numeric</i>	<i>Gives the batting position of the batsman who is on strike</i>
<i>Extra_Type</i>	<i>String</i>	<i>The type of extra that is conceded if no extra is conceded in that then No Extras will be there</i>
<i>Runs_scored</i>	<i>Numeric</i>	<i>Number of runs scored in that particular ball</i>
<i>Extra_runs</i>	<i>Numeric</i>	<i>Number of runs conceded for that extra in that ball</i>
<i>Wides</i>	<i>Numeric</i>	<i>Runs of conceded because of wide in that ball</i>
<i>Legbyes</i>	<i>Numeric</i>	<i>Runs of conceded because of Legbyes in that ball</i>
<i>Noballs</i>	<i>Numeric</i>	<i>Runs of conceded because of Noball in that ball</i>
<i>Penalty</i>	<i>Numeric</i>	<i>Runs of conceded because of Penalty in that ball</i>

<i>Bolwer_extra</i>	<i>Numeric</i>	<i>Runs of conceded because of Extra in that ball that is accountable to bowler</i>
<i>Out_type</i>	<i>String</i>	<i>If batsman is out will give how batsman is out, will be Not Applicable if batsman is not out</i>
<i>Caught</i>	<i>Numeric</i>	<i>If the type of out is caught</i>
<i>Bowled</i>	<i>Numeric</i>	<i>If the type of out is Bowled</i>
<i>Run_out</i>	<i>Numeric</i>	<i>If the type of out is Run_out</i>
<i>LBW</i>	<i>Numeric</i>	<i>If the type of out is LBW</i>
<i>Retired_hurt</i>	<i>Numeric</i>	<i>If the type of out is Retired_hurt</i>
<i>stumped</i>	<i>Numeric</i>	<i>If the type of out is stumped</i>
<i>caught_and_bowled</i>	<i>Numeric</i>	<i>If the type of out is caught_bowled</i>
<i>hit_wicket</i>	<i>Numeric</i>	<i>If the type of out is hit_wicket</i>
<i>ObstructingField</i>	<i>Numeric</i>	<i>If the type of out is ObstructingField</i>
<i>Bowler_wicket</i>	<i>Numeric</i>	<i>If the type of out can be attributed to the bowler</i>
<i>Match_Date</i>	<i>String/Datetime</i>	<i>Date on which that particular match is held</i>
<i>Season</i>	<i>Numeric</i>	<i>Year of that IPL season</i>
<i>Striker</i>	<i>Numeric</i>	<i>Unique id of the player who is on strike</i>
<i>Non_Striker</i>	<i>Numeric</i>	<i>Unique id of the batsman who is on the non striker end</i>
<i>Bowler</i>	<i>Numeric</i>	<i>Unique id of the Bowler</i>
<i>Player_out</i>	<i>Numeric</i>	<i>Unique id of the batsman who is out for that ball</i>
<i>Fielders</i>	<i>Numeric</i>	<i>Unique id of the fielder who is involved to get the batsman out</i>

<i>Striker_Match_SK</i> <i>NonStriker_Match_SK</i> <i>Fielder_Match_SK</i> <i>Bowler_Match_SK</i> <i>PlayerOut_Match_SK</i>		Use NonStriker_Match_SK in "Ball_By_Ball.csv" to join with "Player_match.csv" using "Player_match_SK" to get the Player who is in the Non-striker's end when the ball is bowled in a particular match. Similarly for StrikerMatch_SK,BowlerMatch_SK,FielderMatch_SK,PlayerOutMatch_SK
<i>StrikerSK</i> <i>NonStriker_SK</i> <i>Fielder_SK</i> <i>Bowler_SK</i> <i>Player_out_SK</i>		With respect to this Dataset we can use NonStriker_SK in "Ball_By_Ball.csv" to join with "Player.csv" using "PLAYER_SK" get the Player who is in the Non-striker's end when ball is bowled. Similarly for Striker_SK,Bowler_SK,Fielder_SK,PlayerOut_SK.
<i>keeper_catch</i>	Numeric	Whether the catch taken is a keeper catch
<i>BattingTeam_SK</i> <i>BowlingTeam_SK</i>	Numeric	Use to link to TeamSK from Team.csv

Dateset: Player.csv

Variable name	Data type	Description
<i>Player_sk</i>	<i>numeric</i>	SK Stands for Surrogate Key(Machine generated Key) This is used to Join Facts and Dimensions. Use NonStriker_Match_SK in "Ball_By_Ball.csv" to join with "Player_match.csv" using "Player_match_SK" get the Player who is in the Non-striker's end when ball is bowled in a particular match. Similarly for StrikerMatch_SK,BowlerMatch_SK,FielderMatch_SK,Player OutMatch_SK.

<i>Player_Id</i>	<i>Nu meric</i>	<i>Gives unique identification of each player</i>
<i>Player_Name</i>	<i>Strin g</i>	<i>Name of the player</i>
<i>DOB</i>	<i>Date</i>	<i>Date of birth of the player</i>
<i>Batting_hand</i>	<i>Strin g</i>	<i>Left-hand batsman or Right-hand batsman</i>
<i>Bowling_skill</i>	<i>String</i>	<i>Type of skill of the bowler and bowling hand eg. Right arm medium, left arm fast medium</i>
<i>Country_Nam e</i>	<i>String</i>	<i>Name of country of the player</i>

Dataset: Player_match.csv

<i>Variable name</i>	<i>Data type</i>	<i>Description</i>
-----------------------------	-------------------------	---------------------------

<i>Player_match_SK</i>	<i>Numeric</i>	<i>SK Stands for Surrogate Key(Machine generated Key) This is used to Join Facts and Dimensions.</i>
<i>PlayerMatch_key</i>	<i>Numeric</i>	
<i>Match_Id</i>	<i>Numeric</i>	<i>Gives unique match if for every record based on the match to which it belongs</i>
<i>Player_Id</i>	<i>Numeric</i>	<i>Gives unique identification of each player</i>
<i>Player_Name</i>	<i>String</i>	<i>Name of the player</i>
<i>DOB</i>	<i>Date</i>	<i>Date of birth of player</i>
<i>Batting_hand</i>	<i>String</i>	<i>Left-hand batsman or Right-hand batsman</i>
<i>Bowling_skill</i>	<i>String</i>	<i>Type of skill of the bowler and bowling hand eg. Right arm medium, left arm fast medium</i>
<i>Country_Name</i>	<i>String</i>	<i>Name of country of the player</i>
<i>Role_Desc</i>	<i>String</i>	<i>Role of the player in the team</i>
<i>Player_team</i>	<i>String</i>	<i>Name of the team of the player</i>

<i>Opposit_Team</i>	<i>String</i>	<i>Name of the opposition team to the player's team for this match</i>
<i>Season_year</i>	<i>Date</i>	<i>Year in which this match took place</i>
<i>is_manofThematch</i>	<i>String</i>	<i>Whether the player won man of the match award for this match</i>
<i>Age_As_on_match</i>	<i>Numeric</i>	<i>Age of the player on the date of the match</i>
<i>IsPlayers_Team_won</i>	<i>String</i>	<i>Whether the player's team won this match</i>
<i>Batting_Status</i>	<i>NA</i>	<i>NA</i>
<i>Bowling_Status</i>	<i>NA</i>	<i>NA</i>
<i>Player_Captain</i>	<i>String</i>	<i>Name of the captain of player's team for this match</i>
<i>Opposit_captain</i>	<i>String</i>	<i>Name of the captain of opposition team for this match</i>
<i>Player_keeper</i>	<i>String</i>	<i>Name of the keeper of player's team for this match</i>
<i>Opposit_keeper</i>	<i>String</i>	<i>Name of the keeper of opposition team for this match</i>

Match.csv

Variable name	Data type	Description
Match_SK	numeric	To connect with the other data files
match_id	numeric	<i>Gives a unique match id for every record based on the match to which it belongs.</i>
Team1	string	<i>First team among the 2 teams who have played the match.</i>
Team2	string	<i>Second team among the 2 teams who have played the match.</i>
match_date	Date-time object	<i>Date of the match</i>
Season_Year	numeric	<i>IPL year</i>
Venue_Name	string	<i>Venue at which match is played. (This is the home ground for Team 1)</i>
City_Name	string	<i>Name of the City where the stadium is located.</i>
Country_Name	string	<i>Name of the Country</i>
Toss_Winner	string	<i>Which team had won the toss</i>
match_winner	string	<i>Which team won the match</i>
Toss_Name	string	<i>Decision taken after winning the toss</i>

Win_Type	string	<i>Whether the match is won by the defending or the chasing team.</i>
Outcome_Type	string	<i>What outcome did the match had</i>
ManOfMach	string	<i>Name of the player who got MOM</i>
Win_Margin	string	<i>Tells you about the margin by which the match is won.</i>
Country_id	numeric	<i>Ids of the different countries that hosted the match.</i>

Team.csv

<i>Variable name</i>	<i>Data type</i>	<i>Description</i>
<i>Team_SK</i>	<i>Numeric</i>	<i>Links to other sheets using team_SK column in those sheets</i>
<i>Team_Id</i>	<i>Numeric</i>	<i>Unique id for each team</i>
<i>Team_Name</i>	<i>String</i>	<i>Name of the team</i>

ALL THE BEST!!!