Tableau Assignment on Indian Premier League

By

Syed Sha Khalid

The two data sets provide to us regarding IPL seasons from 2008 to 2017 are:

matches.csv - It contains match-level information for each and every match held in IPL from 2008 to 2017.

deliveries.csv - It contains ball-by-ball information for each of the matches.

These are the Dashboards asked to create:

1. **Match Statistics**
2. **Player Statistics**
3. **Team Statistics**

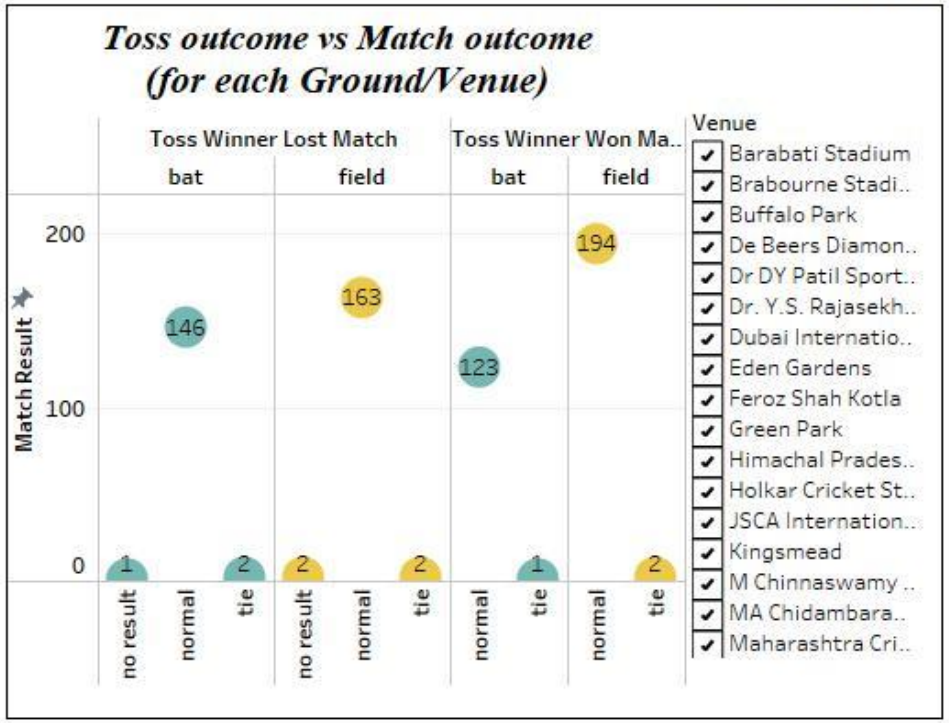
The same are explained in the upcoming slides

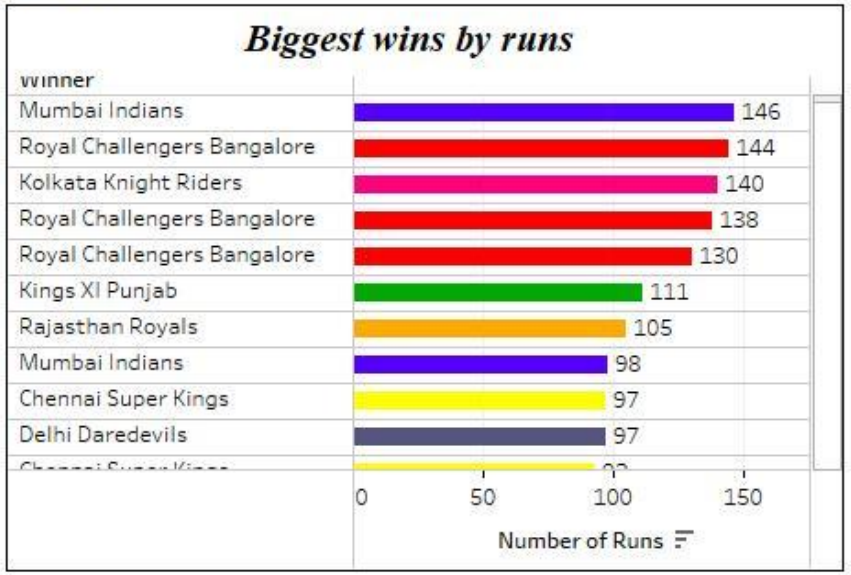
1. Match Statistics

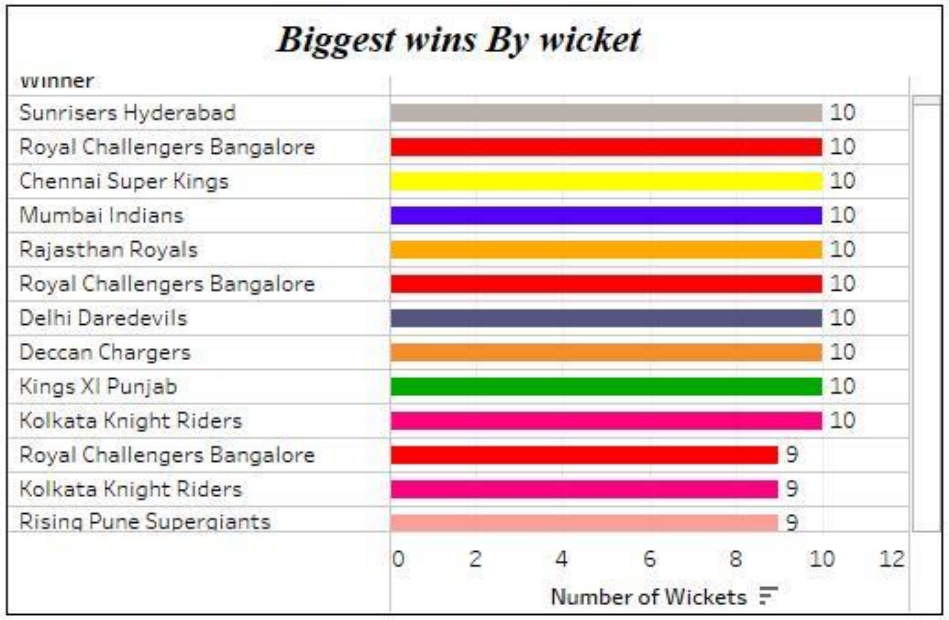
1. Match Statistics

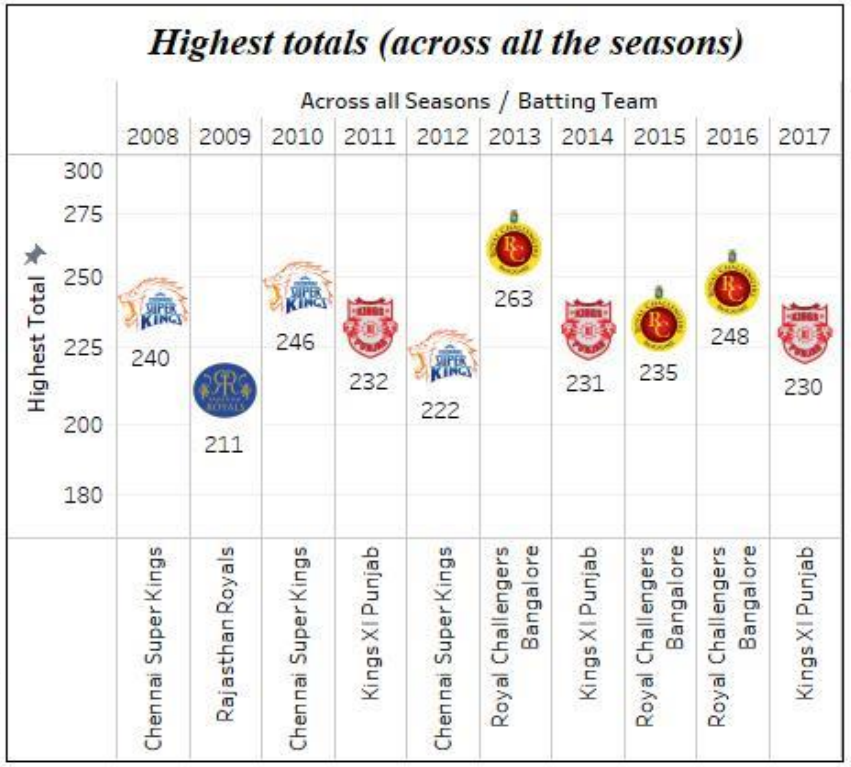
- Toss outcome vs Match outcome (for each Ground/Venue)
- Biggest wins (by runs and by wicket)
- Highest totals (across all the seasons)

The Details subtasks I explained in the upcoming slides









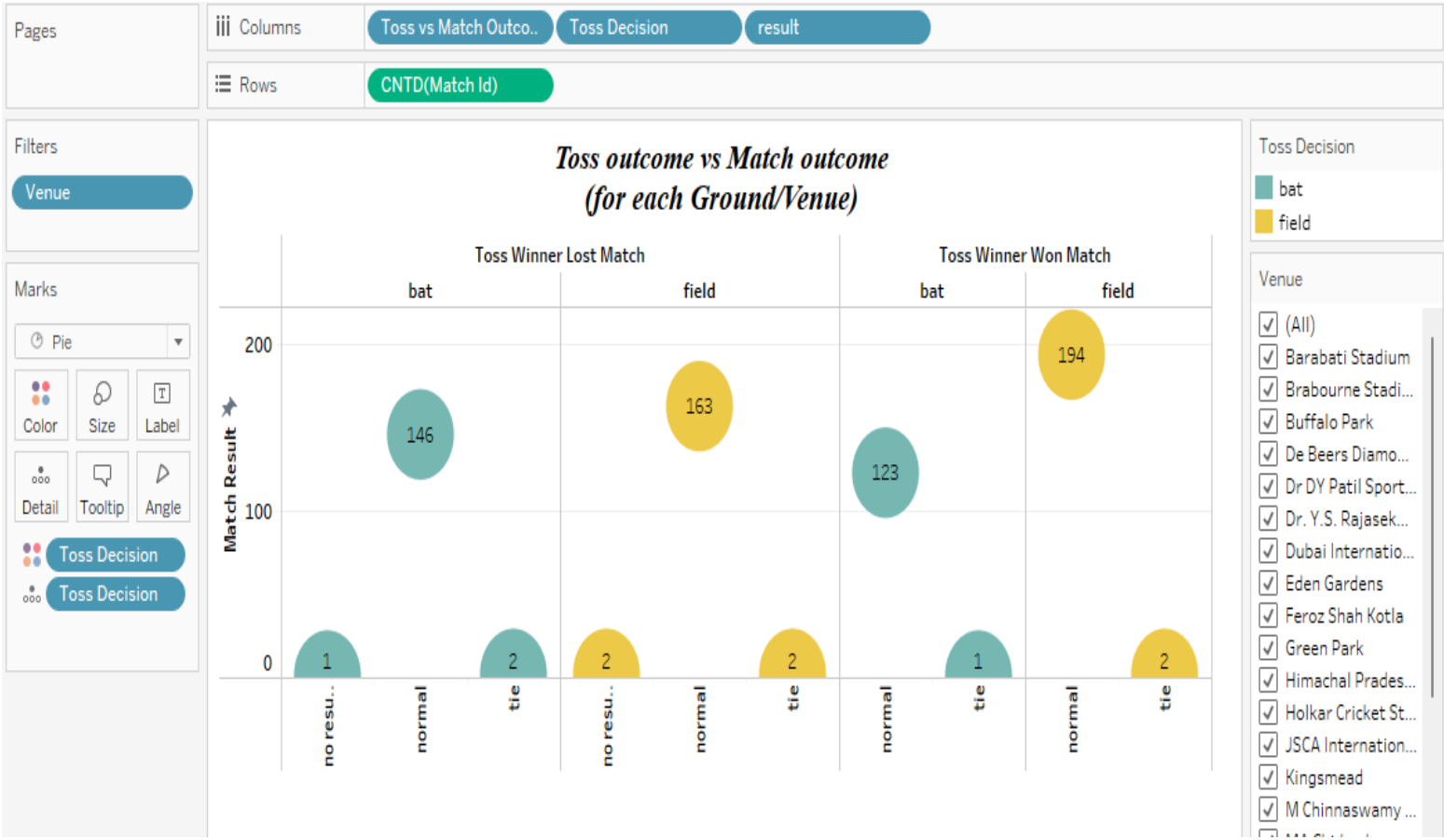
Sub Task-1: Toss outcome vs Match outcome (for each Ground/Venue)

Step 1: create “Toss vs Match Outcome”
in create calculated field Type this code

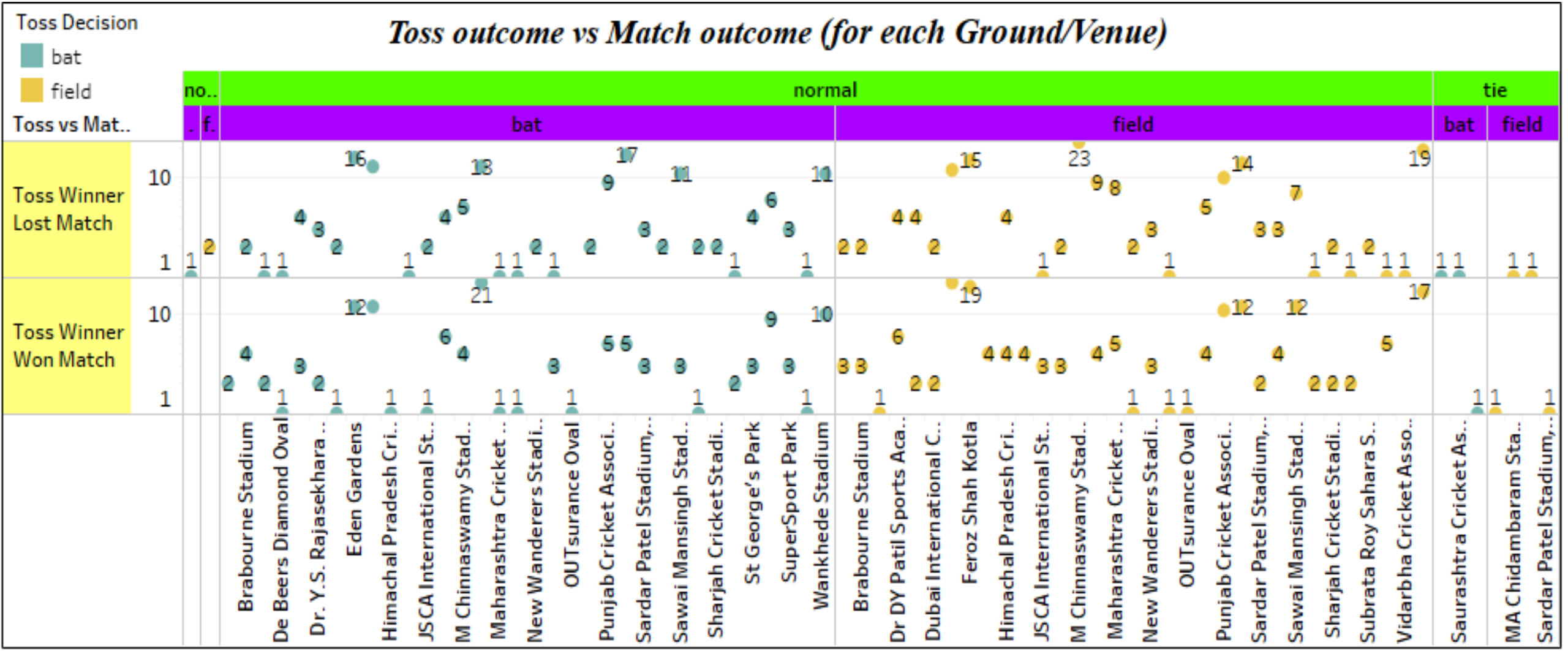
IF [toss_winner] = [Winner]
THEN "Toss Winner Won Match"
ELSE "Toss Winner Lost Match"
END

Step 2: In columns take “Toss vs Match Outcome” , “Toss Decision”, “result”
Step 3: In Rows take “COUNTD([Match Id])”
Step 4: In filters use “Venue”

In order to show different colors for Bat & Field I used Toss Decision in Color and in order to the number of times match decision happened I taken Toss Decision as details

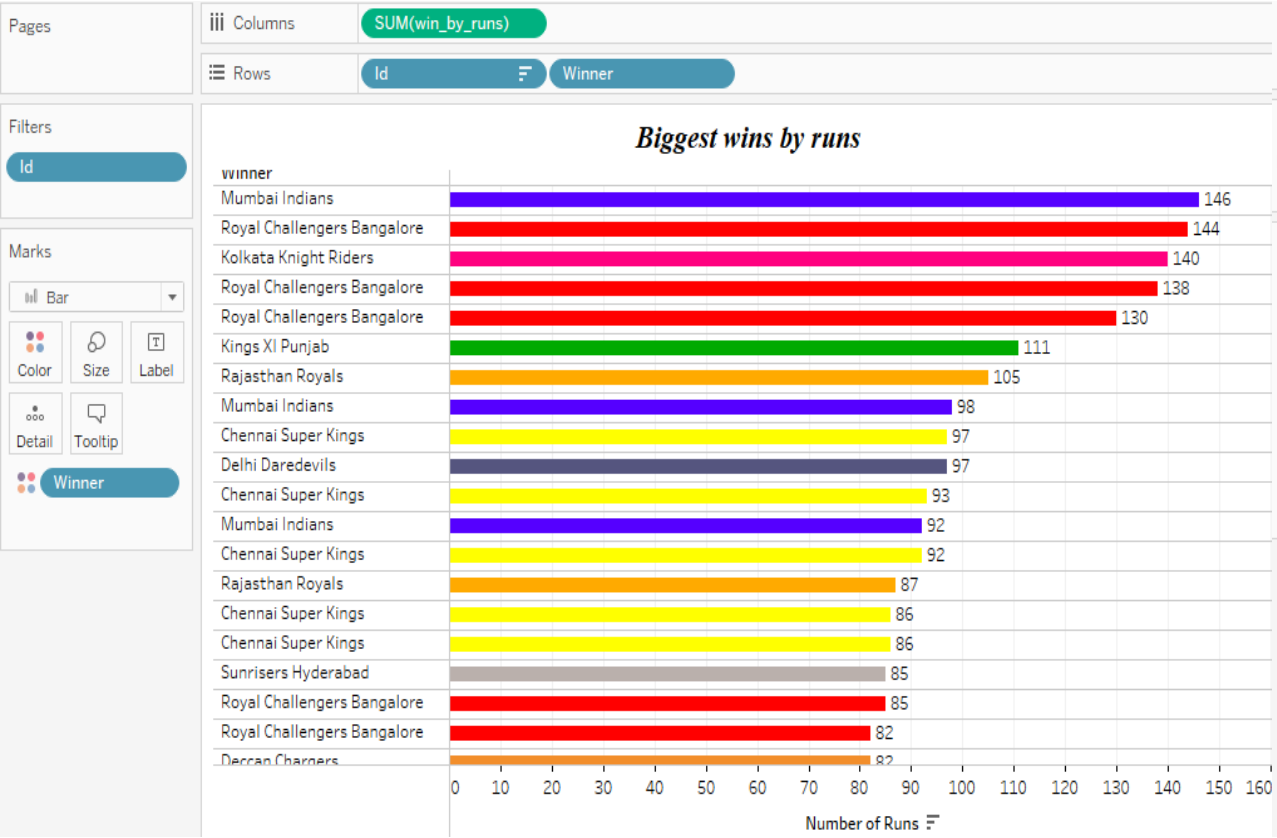


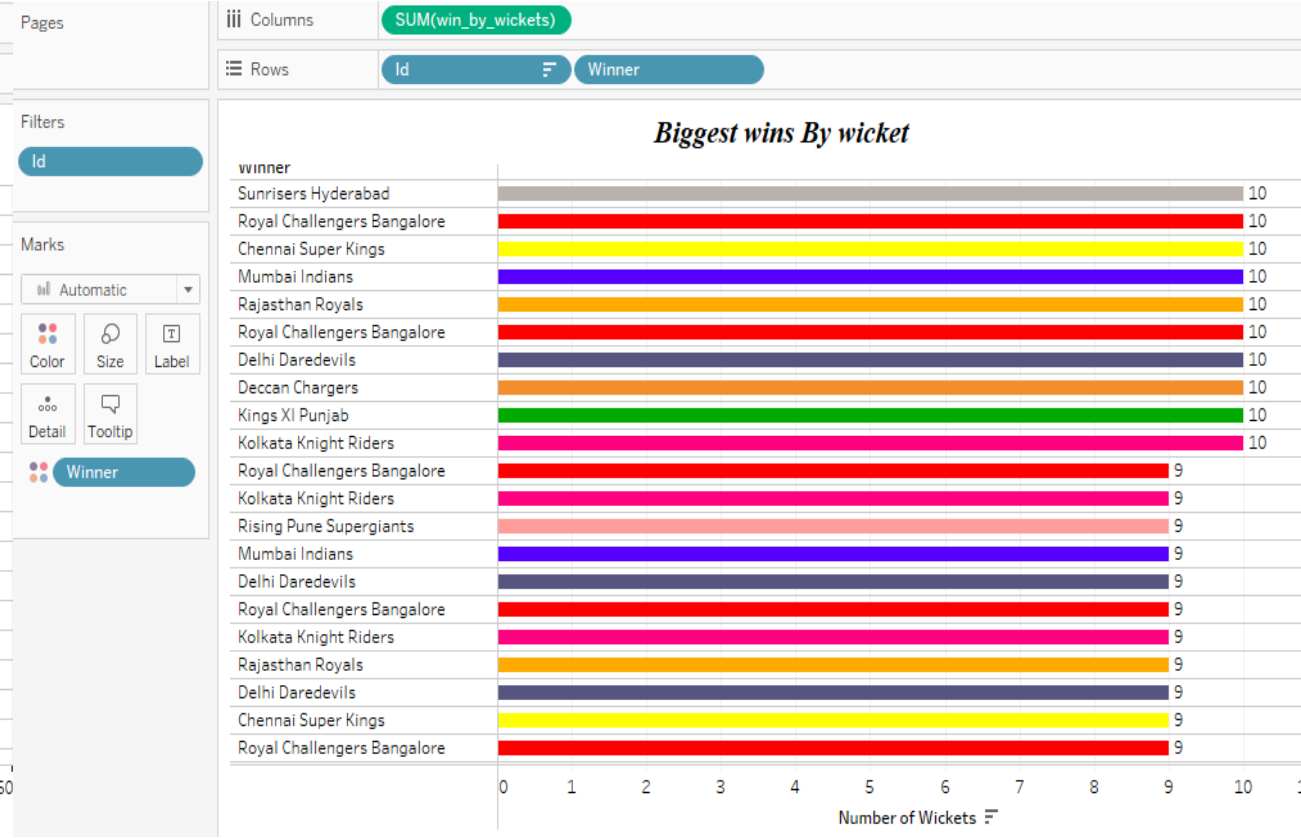
Here I found “Toss outcome vs Match outcome” for all the seasons in one graph if the one need to see performance of individual Ground/Venue then they can select the name of the Ground/Venue they want to see in filter.
In the next slide I made analysis for individual Ground/Venue also.



In Details you can get it in Syed Sha Khalid IPL Data Analysis .twbx (In Additional Dashboard) which is attach with this PPT.

Sub Task-2: Biggest wins (by runs and by wicket)

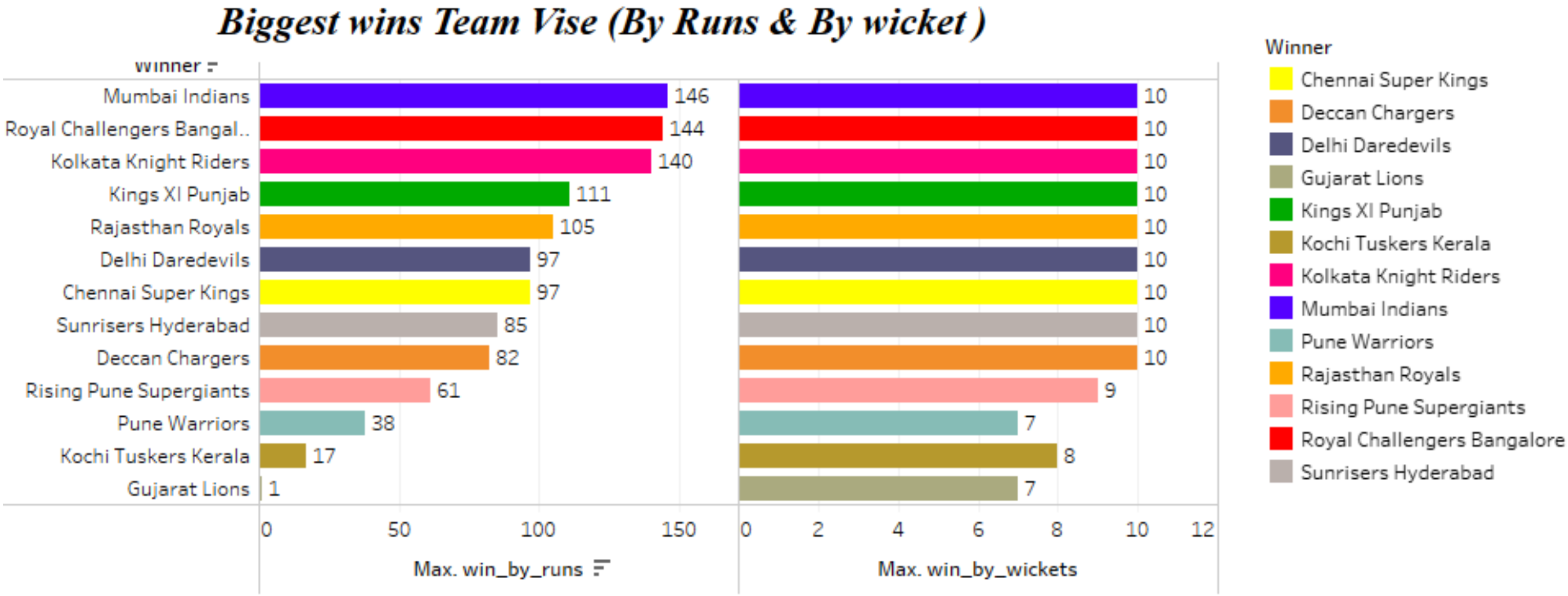




Step 1: In Row take Id, Winner

Step 2: In columns Take SUM([win_by_wickets]) for Biggest win by wickets or Take SUM([win_by_runs]) for Biggest win by runs.

- If you want to see top 10 then use “Id” in filter. Go to edit filter select Top and there in “By field” you can give to 10.
- Here you can see Biggest win by runs is Mumbai Indians won by 146 Runs and 9 teams are won by 10 wickets as shown in the fig.
- The Next slide shows team vise compression.



The Team Vies performance in all the seasons of IPL as per given data is show above its clear that the Biggest win by runs is Mumbai Indians won by 146 Runs and 9 teams are won by 10 wickets.

Sub Task-3: Highest totals (across all the seasons)

Step 1: Create a calculated field name Across all Seasons, since seasons column is already available but it is numeric form so I convert it into string for better view. One can use date also for this purpose.

Step 2: Create a calculated field name Match Innings it can be created by concatenating Match Id and Innings by using code : STR([Match Id])+"_"+STR([Inning])

Step 3: In column take Across all Seasons, Match Innings and Batting team.

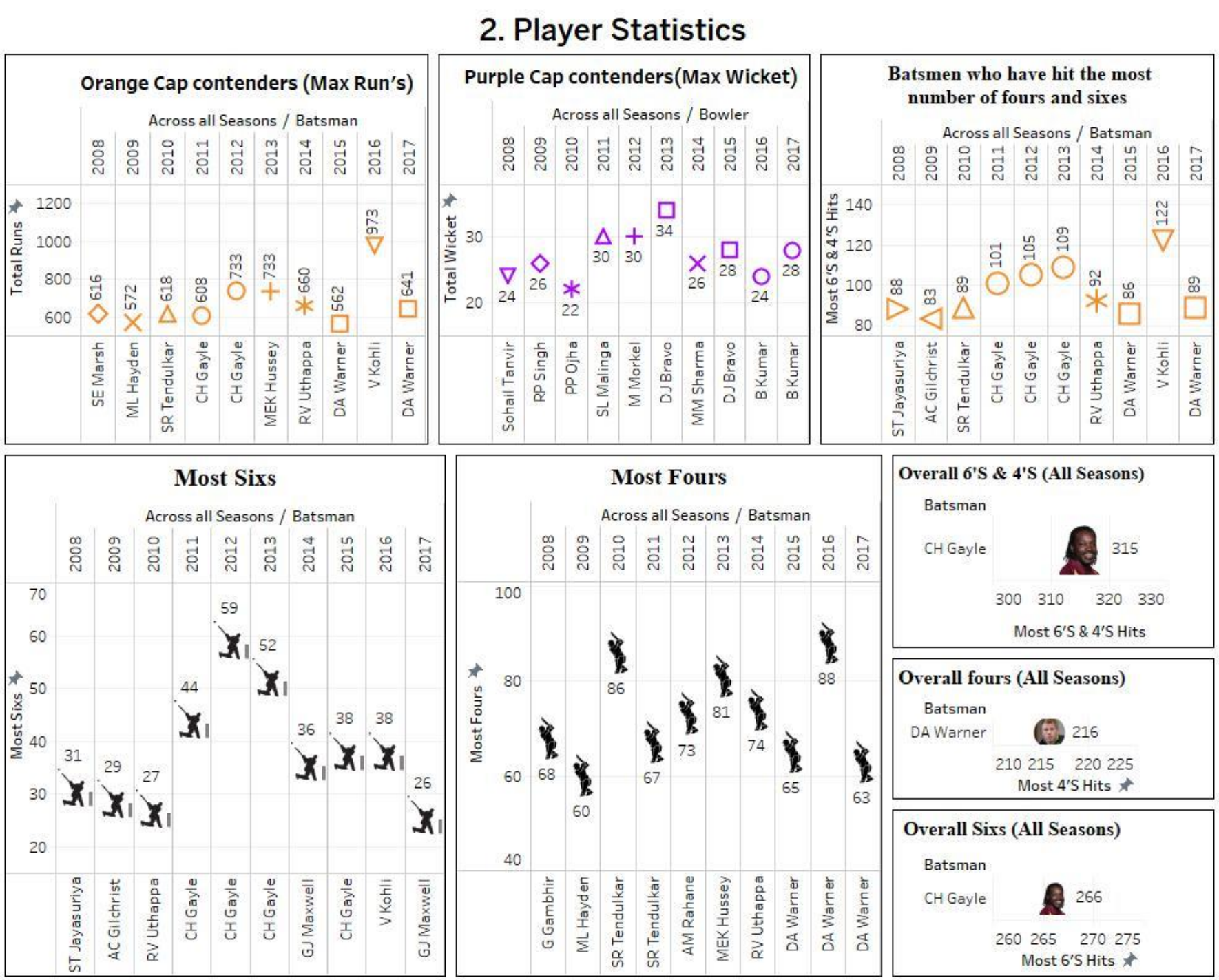
Step 4: In Rows take SUM([Total Runs])

Step 5: In filter you can use Across all Seasons and Match Innings their I filter and kept only those teams who scored highest total in all seasons of IPL.

2. Player Statistics

- Orange Cap contenders (The batsmen who have scored the maximum number of runs in a particular season)
- Purple Cap contenders (The bowlers who have taken the maximum number of wickets in a particular season)
- Batsmen who have hit the most number of fours and sixes (per season and overall)

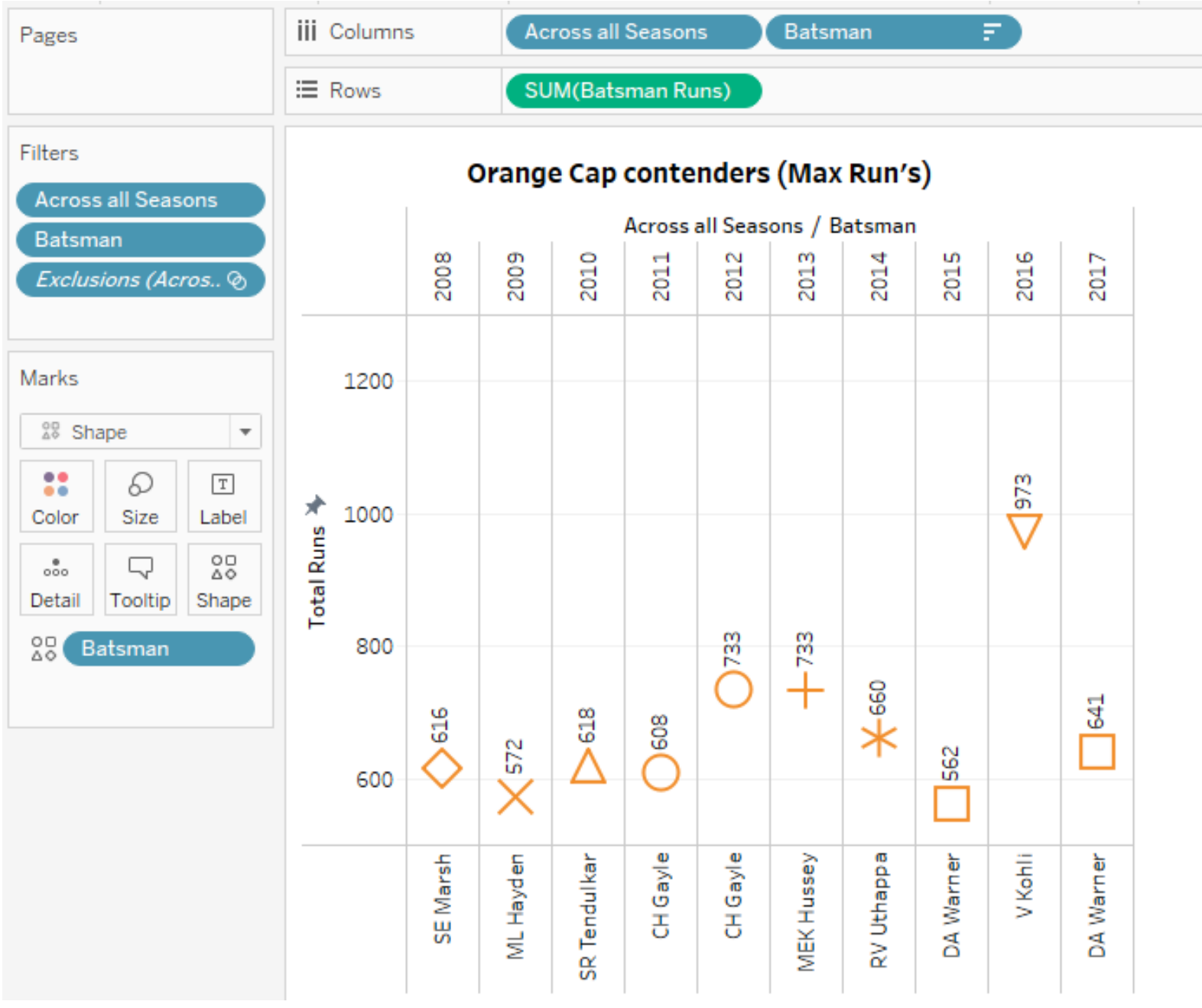
The Details subtasks I explained in the upcoming slides



Sub Task-1:Orange Cap contenders (The batsmen who have scored the maximum number of runs in a particular season)

Step 1: In Column take Across all Seasons, Batsman.
Step 2: In Rows take SUM([Batsman Runs])
Step 3: In filter I used “Across all Seasons” and “Batsman” in order to get the required seasons Orange cap contenders I used Exclusions.

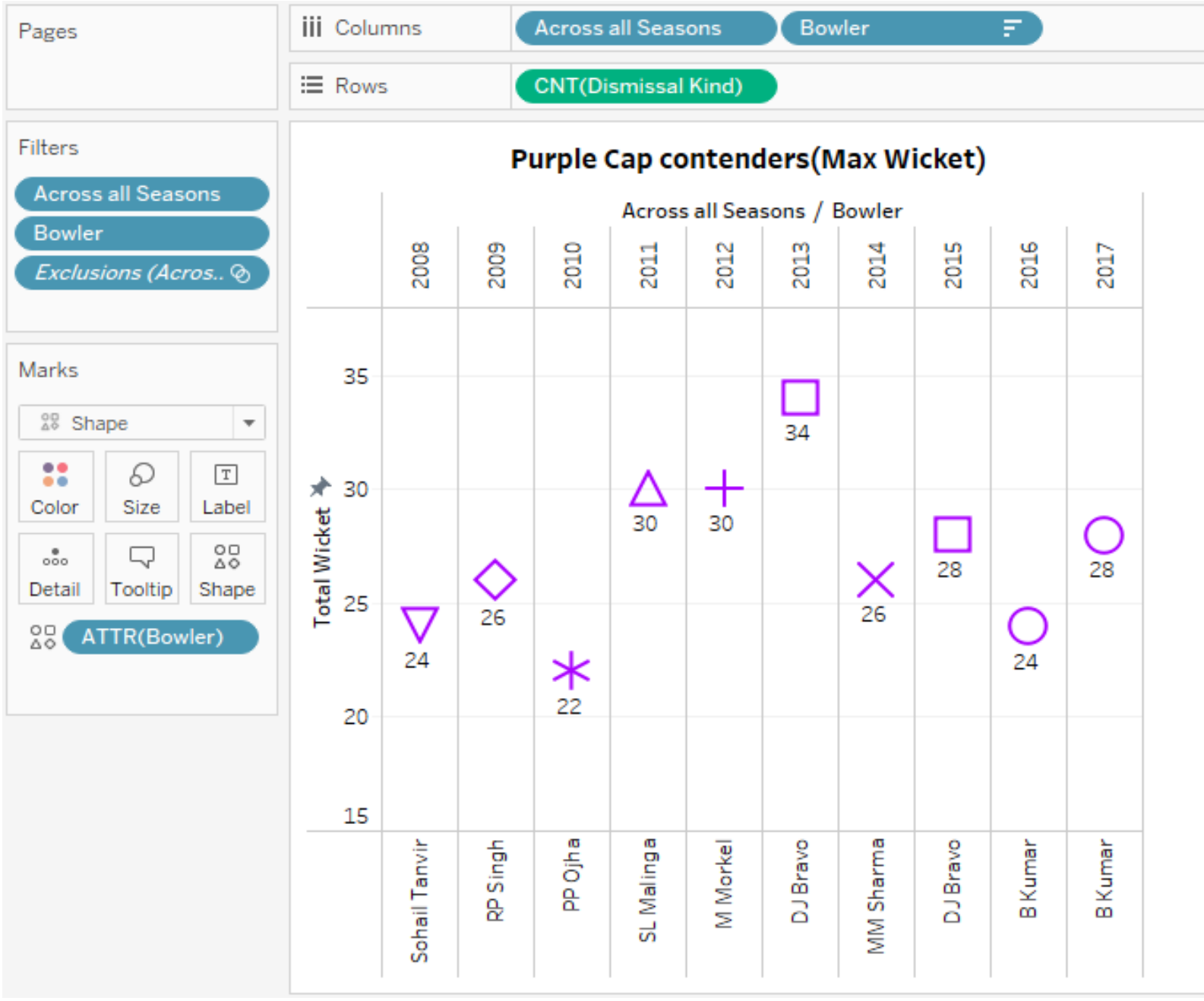
It is clear that Orange Cap contender for season 2008, 2009 …etc. is SE Marsh, ML Hayden,…etc. respectively as shown in fig.



Sub Task-2:Purple Cap contenders (The bowlers who have taken the maximum number of wickets in a particular season)

Step 1: In Column take Across all Seasons, Batsman.
Step 2: In Rows take COUNT([Dismissal Kind]) here we can use count of player_dismissed also.
Step 3: In filter I used “Across all Seasons” and “Bowler” in order to get the required seasons Purple cap contenders I used Exclusions.

It is clear that Purple Cap contender for season 2008, 2009 …etc. is Sohail Tanvir, RP Singh,…etc. respectively as shown in fig.



Sub Task-3:Batsmen who have hit the most number of fours and sixes (per season and overall)

Step 1: Create a calculated field name “Most 6 and 4 Hits” by using Batsman Runs column. Code used for creating :

IF [Batsman Runs] = 6
OR [Batsman Runs] = 4 THEN 1 ELSE 0 END

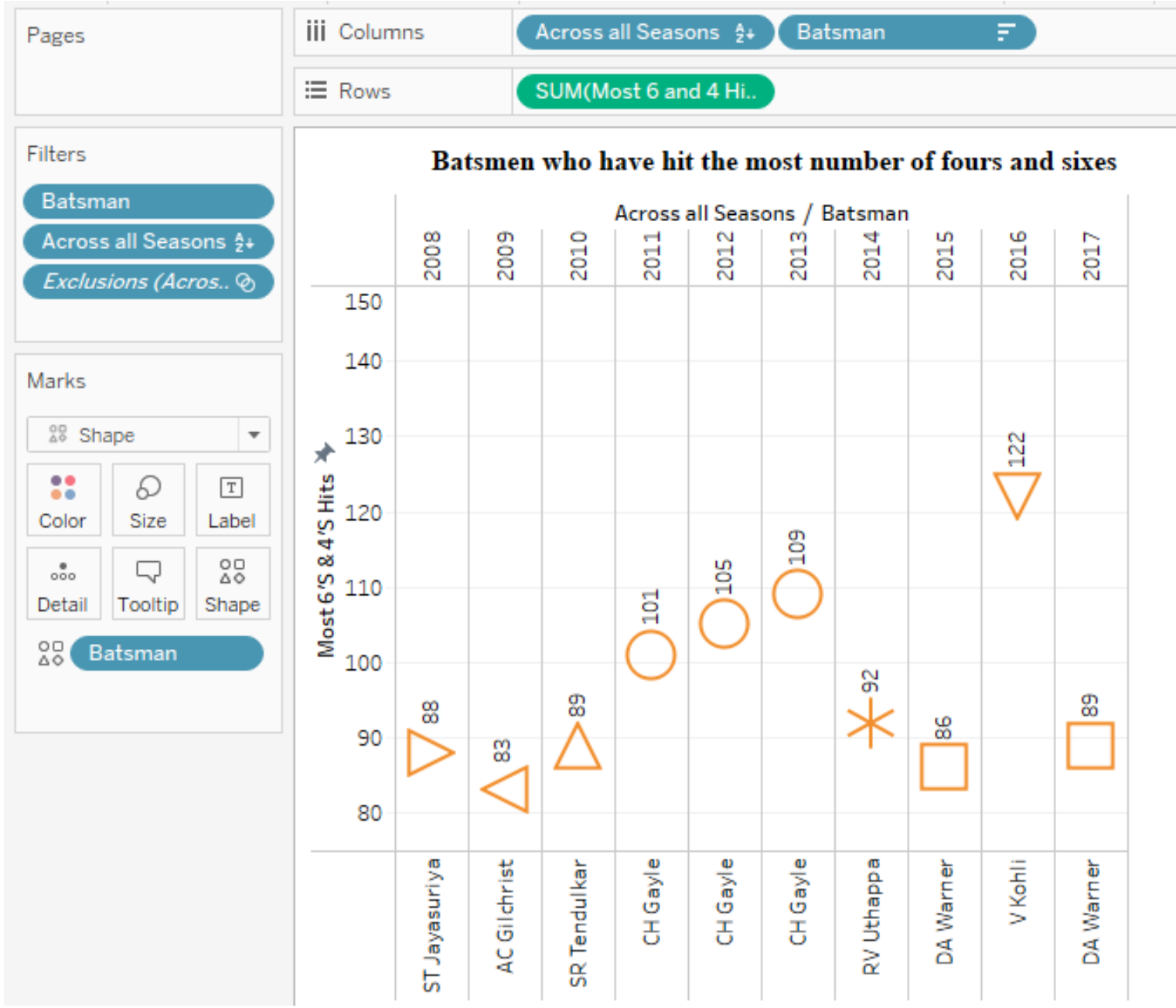
Step 2: In Column take Across all Seasons, Batsman.

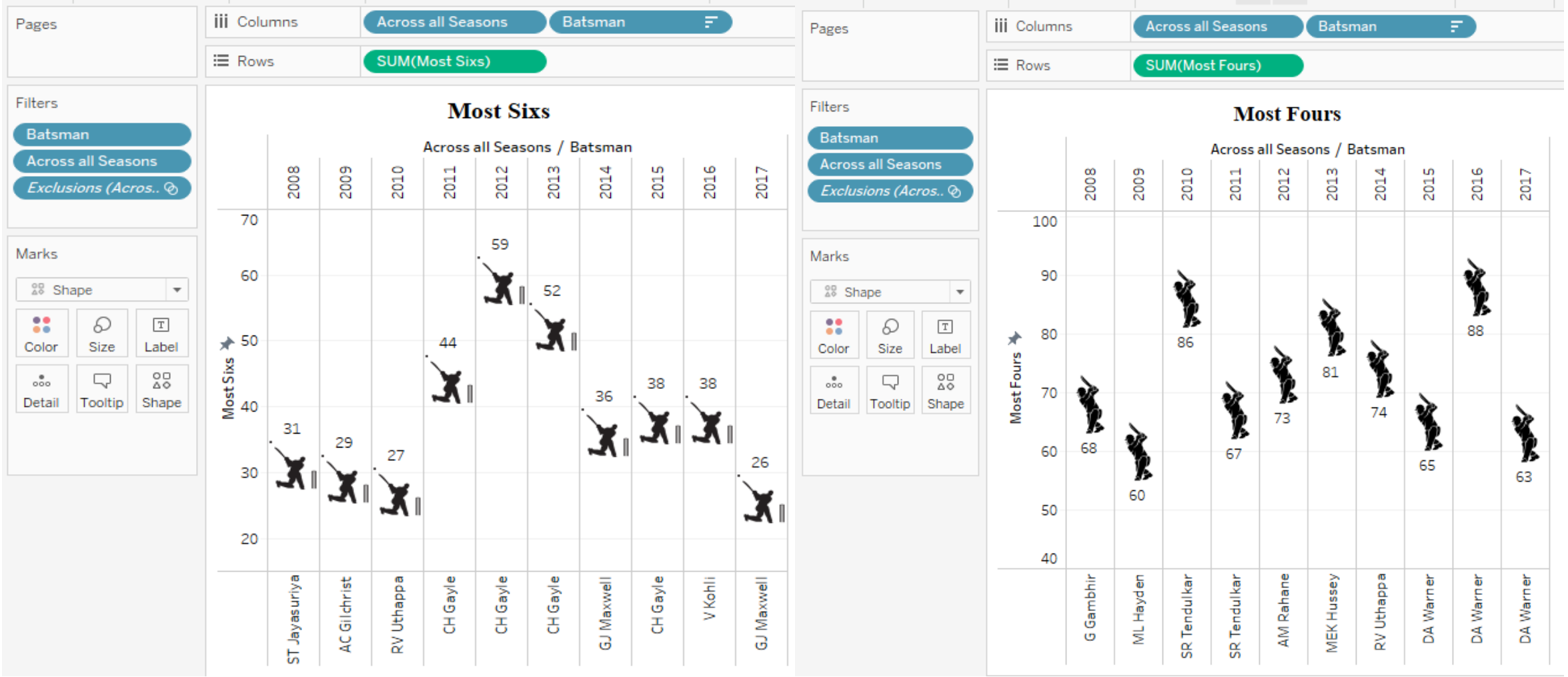
Step 3: In Rows take SUM([Most 6 and 4 Hits])

Step 4: In filter I used “Across all Seasons” and “Batsman” in order to get the required result for all seasons I used Exclusions.

It is clear that most number of fours and sixes hit for season 2008, 2009 …etc. is ST Jayasurya, AC Gilcrist,…etc. respectively as shown in fig.

This data I recorded for total boundaries (4’s & 6’s) hit by a player in different season, for particular 4’s hit by batsman & 6’s by a batsman for different seasons are done in up coming slides





Step 1: Create a calculated field name “Most 6” code: IF [Batsman Runs] = 6 THEN 1 ELSE 0 END & “4 Hits” IF [Batsman Runs] = 4 THEN 1 ELSE 0 END

Step 2: In Column take Across all Seasons, Batsman.

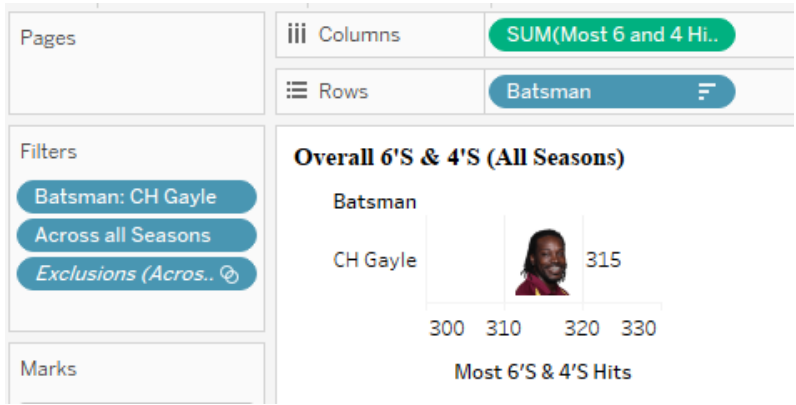
Step 3: In Rows take SUM([Most 6]) in order to find 4’S take SUM([Most 4])

Step 4: In filter I used “Across all Seasons” and “Batsman” in order to get the required result for all seasons I used Exclusions.

It is clear that most number of 6’S hit for season 2008, 2009 …etc. is ST Jayasurya, AC Gilcrist,…etc. respectively as shown in fig. also most number of 4’S hit for season 2008, 2009 …etc. is G Gambhir, ML Hyden,…etc. respectively as shown in fig.

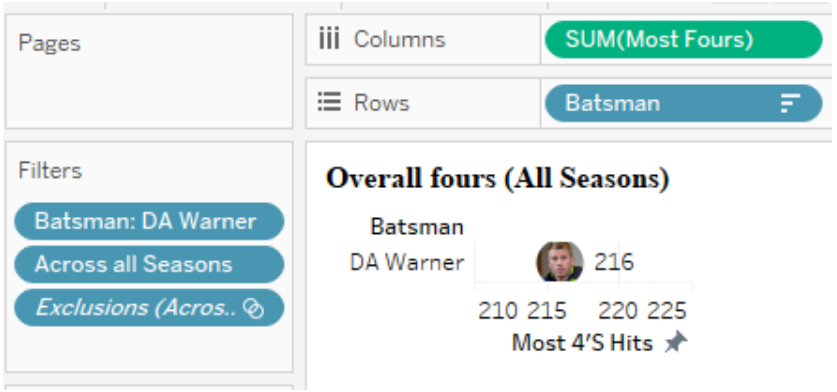
Batsmen who have hit the most number of Boundaries(4’S & 6’S) for all seasons

CH Gayle are the Batsman who hit Most boundaries In IPL all Seasons



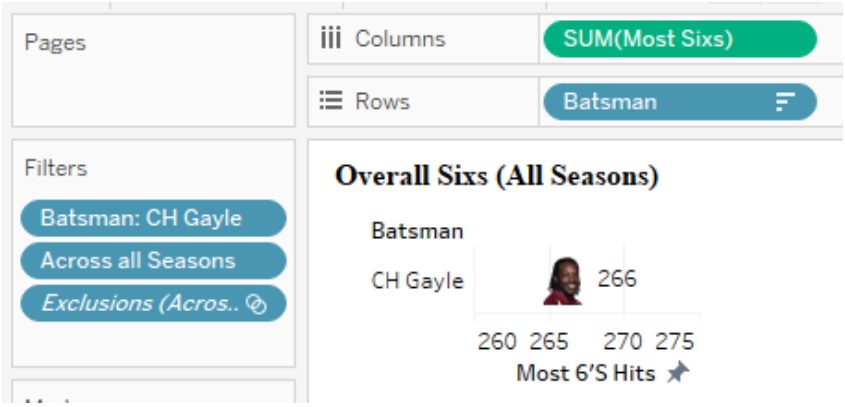
Step 1 : In column take SUM([Most 6 and 4 Hits])
Step 2: In Rows take Batsman
Step 3: In Filter use used “Across all Seasons” and “Batsman” then keep CH Gayle only by Exclusions all others

CH Gayle are the Batsman who hit Most 4’S In IPL all Seasons



Step 1 : In column take SUM([Most Fours])
Step 2: In Rows take Batsman
Step 3: In Filter use used “Across all Seasons” and “Batsman” then keep DA Warner only by Exclusions all others

CH Gayle are the Batsman who hit Most 6’S In IPL all Seasons



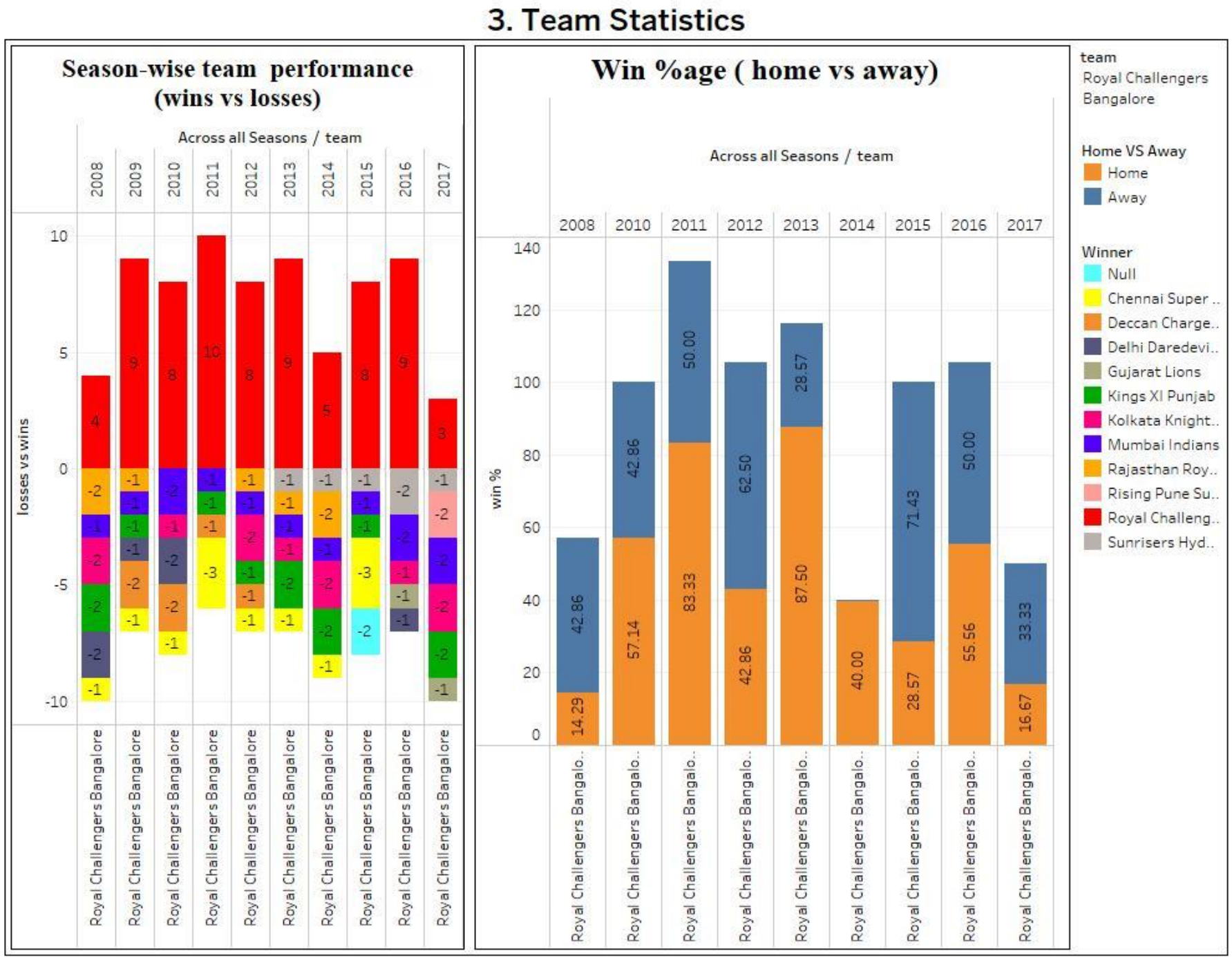
Step 1 : In column take SUM([Most Sixs])
Step 2: In Rows take Batsman
Step 3: In Filter use used “Across all Seasons” and “Batsman” then keep CH Gayle only by Exclusions all others

3. Team Statistics

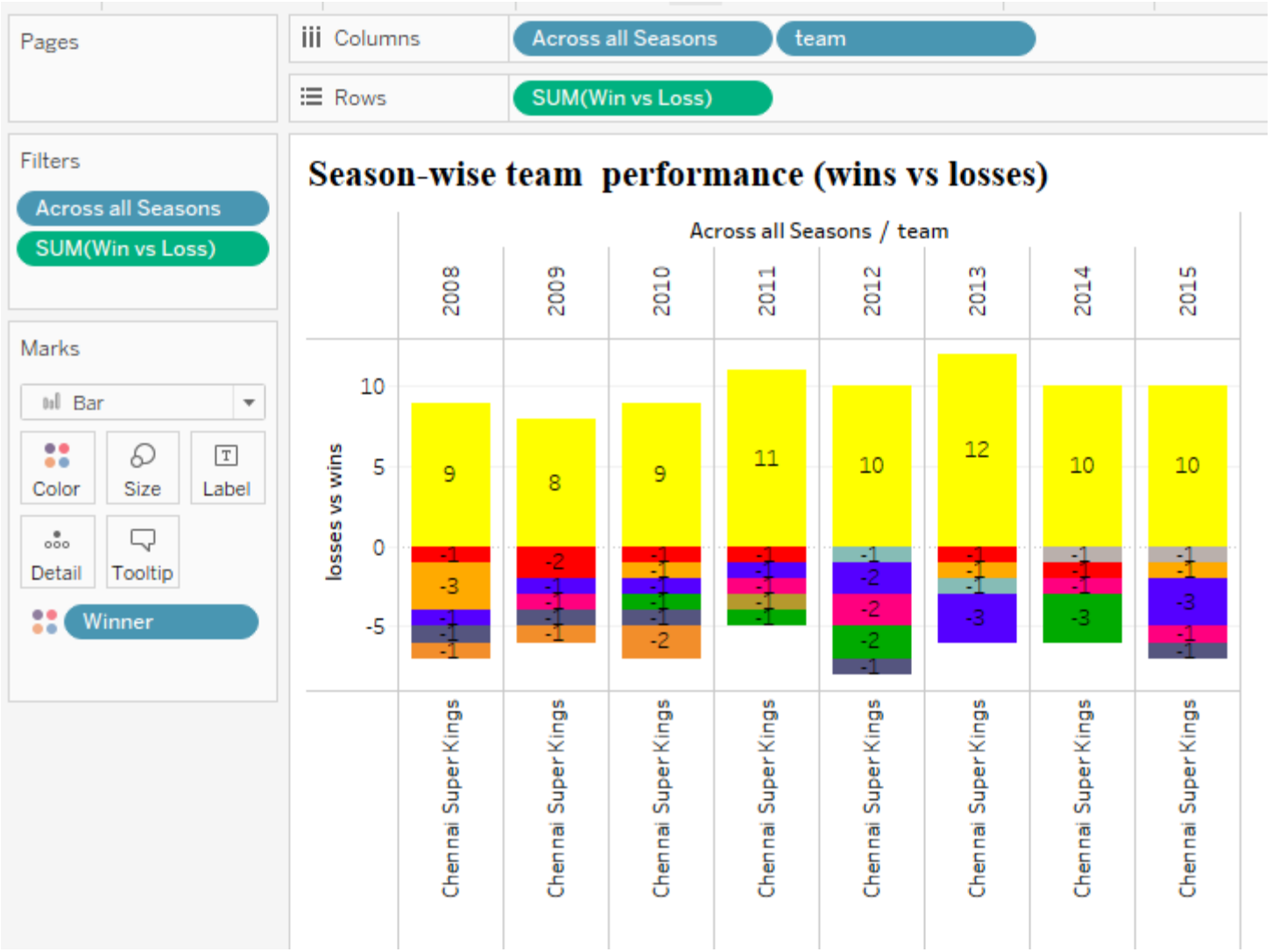
- Season-wise team performance (wins vs losses)
- Win %age (home vs away)

Here in both the graph data are showing for "Royal Challenger Bangalore" if we change parameter created as team we can get other team details also

- First graph showing for different seasons how many matches they won and how many matches they losses against which team also indicated by color change. As you can sea that in 2008 season "Royal Challenger Bangalore" won 4 matches and losses 10 matches against RR(2 Matches), MI(1 Match).. Etc. respectively
- Second graph shows win% home ground vs away ground. It is clear that RCB won 14.29% & 42.86% while playing in home & Away grounds respectively



Sub Task-1: Season-wise team performance (wins vs losses)



Step 1: Create a parameter name “team”, By selecting string then “list” in “when work Book open” select by Batting team.

Step 2: Create a calculated field name “win vs loss” By using code: IF [team] = [Winner] THEN 1 ELSE IF [team]=[team1] OR [team]=[team2] THEN -1 END END

Step 3: In column take Across all Seasons and team

Step 4: In Rows take SUM([Win vs Loss])

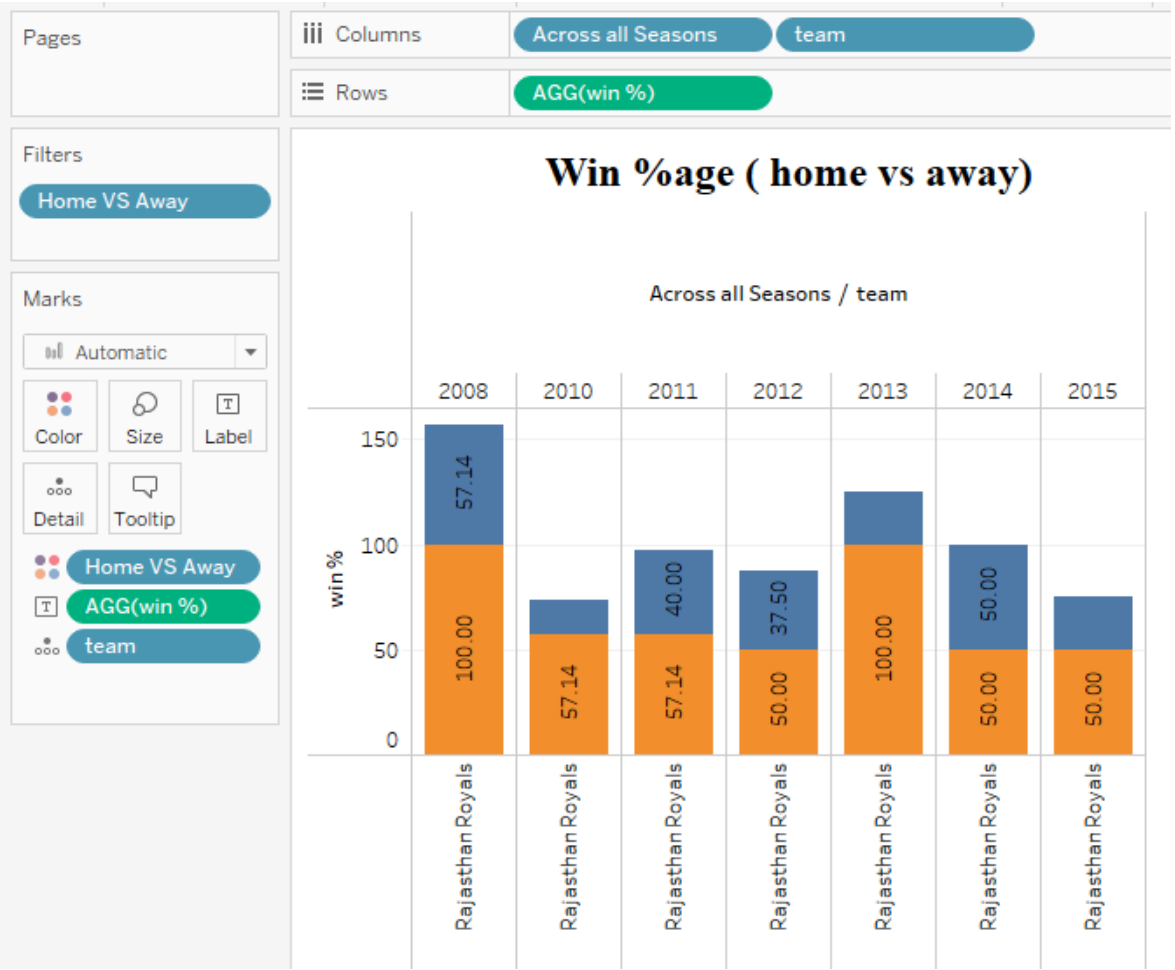
Step 4: In Filter use I used “Across all Seasons” and SUM([Win vs Loss])

Step5: In color we should indicate the wining column then only we can able to get the required results.

As you can sea in the graph data shown for CSK.

2016 & 2017 data are not available in the given data set hence it is not shown.

Sub Task-2: Win %age (home vs away)



Step 1: Create a parameter name “team”, By selecting string then “list” in “when work Book open” select by Batting team.

Step 2: Create a calculated field name “win %” By using code: COUNT(IF [Win vs Loss] = 1 THEN [Id] END) / COUNT(IF [Win vs Loss] = 1 OR [Win vs Loss] =-1 THEN [Id] END)*100

Step 3: Create a calculated field name “Home VS Away” By using code:

IF [home_team]="team1" AND [team]=[team1] THEN "Home“ ELSEIF [home_team] = "team1" AND [team] = [team2] THEN "Away“ ELSEIF [home_team] = "NA" THEN "NA" ELSEIF [home_team] = "team2" AND [team]=[team1] THEN "Away“ ELSEIF [home_team] = "team2" AND [team]=[team2] THEN "Home“ END

Step 4: In column take Across all Seasons & team

Step 5: In Rows take AGG(win %)

Step 4: In Filter use I used “Home VS Away”

As you can sea in the graph data shown for RR. 2009, 2016 & 2017 data are not available in the given data set hence it is not shown.

Thank You