SQL Case Study on IPL Dataset

Deepak Chawde

Information :

This dataset of IPL is of year three years

The analysis of this questions are give proper insight of the data

There are two sheets in the excel file
- Deliveries
- Matches

Agenda is to find the business insight form the dataset

Q1 WHAT ARE THE TOP 5 PLAYERS WITH THE MOST PLAYER OF THE MATCH AWARDS?

```sql
-- Q1 WHAT ARE THE TOP 5 PLAYERS WITH THE MOST PLAYER OF THE MATCH AWARDS?
select player_of_match, count(*) as no_of_match
from matches
group by  player_of_match
order by no_of_match desc
limit 5
```

Result Grid | Filter Rows:

player_of_match	no_of_match
YK Pathan	5
SE Marsh	5
SR Watson	4
BA Stokes	3
NM Coulter-Nile	3

Q2 HOW MANY MATCHES WERE WON BY EACH TEAM IN EACH SEASON?

```sql
-- Q2 HOW MANY MATCHES WERE WON BY EACH TEAM IN EACH SEASON?
select team1, count(winner) as no_of_winner, season
from matches
group by team1, season
order by season
```

team1	no_of_winner	season
Kolkata Knight Riders	9	2008
Chennai Super Kings	9	2008
Rajasthan Royals	6	2008
Mumbai Indians	6	2008
Deccan Chargers	8	2008
Kings XI Punjab	7	2008
Royal Challengers Bangalore	7	2008
Delhi Daredevils	6	2008
Mumbai Indians	1	2009
Royal Challengers Bangalore	2	2009
Kings XI Punjab	2	2009
Kolkata Knight Riders	1	2009
Chennai Super Kings	1	2009
Deccan Chargers	2	2009
Delhi Daredevils	1	2009
Rajasthan Royals	1	2009
Sunrisers Hyderabad	9	2017
Mumbai Indians	6	2017
Gujarat Lions	8	2017
Rising Pune Supergiant	8	2017

Q3 WHAT IS THE AVERAGE STRIKE RATE OF BATSMEN IN THE IPL DATASET?

```sql
-- Q3 WHAT IS THE AVERAGE STRIKE RATE OF BATSMEN IN THE IPL DATASET?
select batsman, count(*)as No_of_balls, sum(total_runs) as total_runs,(sum(total_runs)/count(*))*100 as strik_rate
from deliveries
group by batsman
order by total_runs desc
```

batsman	No_of_balls	total_runs	strik_rate
N Rana	154	205	133.1169
G Gambhir	133	198	148.8722
DA Warner	121	175	144.6281
KA Pollard	118	171	144.9153
SPD Smith	115	159	138.2609
BB McCullum	111	158	142.3423
MK Pandey	101	157	155.4455
CA Lynn	74	140	189.1892
SK Raina	112	139	124.1071
SV Samson	94	138	146.8085
KD Karthik	91	137	150.5495
HM Amla	106	134	126.4151
S Dhawan	106	124	116.9811
AM Rahane	88	121	137.5000
GJ Maxwell	61	120	196.7213
MC Henriques	87	118	135.6322
KM Jadhav	68	116	170.5882
AB de Villiers	71	116	163.3803
JC Buttler	85	115	135.2941
SW Billings	79	109	137.9747

```sql
-- Q4 WHAT IS THE NUMBER OF MATCHES WON BY EACH TEAM BATTING FIRST VERSUS BATTINGSECOND?
select Toss_decision, toss_winner
from matches
where toss_decision ="bat"


select batting_first , count(*) as match_wons
from(
        select case
                when win_by_runs >0 then team1
                else team2
                end
                as batting_first
        from matches
        where winner != "tie"
) as batting_first_team
group by batting_first
```

batting_first	match_wons
Sunrisers Hyderabad	8
Rising Pune Supergiant	10
Kolkata Knight Riders	17
Kings XI Punjab	18
Royal Challengers Bangalore	8
Mumbai Indians	20
Delhi Daredevils	16
Gujarat Lions	4
Chennai Super Kings	9
Rajasthan Royals	13
Deccan Chargers	5

Q5 WHICH BATSMAN HAS THE HIGHEST STRIKE RATE (MINIMUM 200 RUNS SCORED)?

```sql
-- Q5 WHICH BATSMAN HAS THE HIGHEST STRIKE RATE (MINIMUM 200 RUNS SCORED)?
select batsman , (sum(total_runs)/count(*))*100 as strickr_rate
from deliveries
group by batsman
having (sum(total_runs)/count(*))*100  >200
```

Result Grid | Filter Rows:

batsman	strickr_rate
A Choudhary	300.0000
HH Pandya	220.0000
TG Southee	233.3333
DT Christian	212.5000
CH Morris	256.5217
DL Chahar	233.3333

Q6 HOW MANY TIMES HAS EACH BATSMAN BEEN DISMISSED BY THE BOWLER 'MALINGA'?

```sql
-- Q6 HOW MANY TIMES HAS EACH BATSMAN BEEN DISMISSED BY THE BOWLER 'MALINGA'?
select batsman , count(*) as total_dissmiss
from deliveries
where player_dismissed is not null and  bowler = "SL Malinga"
group by batsman
order by total_dissmiss desc
-- having sum(batsman_runs)= 0
```

batsman	total_dissmiss
BB McCullum	11
CA Lynn	9
DA Warner	7
MK Pandey	6
KD Karthik	6
G Gambhir	5
S Dhawan	5
BCJ Cutting	5
NV Ojha	5
SK Raina	5
CR Woakes	4
JJ Roy	3
Rashid Khan	2
Ishan Kishan	2
SA Yadav	1
SP Narine	1
V Shankar	1

Q7 WHAT IS THE AVERAGE PERCENTAGE OF BOUNDARIES (FOURS AND SIXESCOMBINED) HIT BY EACH BATSMAN?

```sql
-- Q7 WHAT IS THE AVERAGE PERCENTAGE OF BOUNDARIES (FOURS AND SIXESCOMBINED) HIT BY EACH BATSMAN?
select batsman, avg(case
            when batsman_runs = 4 or batsman_runs = 6   then 1 else 0
            end
        )*100 as average
from deliveries
group by batsman
order by average desc
```

Result Grid | Filter Rows:

batsman	average
A Choudhary	50.0000
CH Morris	39.1304
DT Christian	37.5000
DL Chahar	33.3333
TG Southee	33.3333
HH Pandya	30.0000
BCJ Cutting	29.6296
GJ Maxwell	29.5082
Yuvraj Singh	26.7857
SP Narine	26.6667
WP Saha	25.8065
PJ Cummins	25.0000
JJ Roy	24.3902
CA Lynn	24.3243
RA Tripathi	23.9130
RV Uthappa	23.7288
DR Smith	22.9508
KM Jadhav	22.0588
DA Warner	21.4876
CJ Anderson	21.4286
Bipul Sharma	21.4286
AJ Finch	21.2121
G Gambhir	21.0526

Q8 WHAT IS THE AVERAGE NUMBER OF BOUNDARIES HIT BY EACH TEAM IN EACH SEASON?

```sql
-- Q8 WHAT IS THE AVERAGE NUMBER OF BOUNDARIES HIT BY EACH TEAM IN EACH SEASON?
select batting_team , avg(case
        when batsman_runs = 4 or batsman_runs = 6
        then 1 else 0
        end) as average, m.season
from deliveries
group by batting_team
```

Q8 WHAT IS THE AVERAGE NUMBER OF BOUNDARIES HIT BY EACH TEAM IN EACH SEASON?

```sql
select season,batting_team,avg(fours+sixes) as average_boundaries
from(select season,match_id,batting_team,
sum(case when batsman_runs=4 then 1 else 0 end)as fours,
sum(case when batsman_runs=6 then 1 else 0 end) as sixes
from deliveries,matches
where deliveries.match_id=matches.id
group by season,match_id,batting_team) as team_bounsaries
group by season,batting_team;
```

season	batting_team	average_boundaries
2017	Sunrisers Hyderabad	21.2500
2017	Royal Challengers Bangalore	18.0000
2017	Mumbai Indians	20.4000
2017	Rising Pune Supergiant	16.6000
2017	Gujarat Lions	21.0000
2017	Kolkata Knight Riders	22.7500
2017	Kings XI Punjab	19.5000
2017	Delhi Daredevils	23.3333

Q9 WHAT IS THE HIGHEST PARTNERSHIP (RUNS) FOR EACH TEAM IN EACH SEASON?

```sql
select bowler , bowling_team, count(*) as wickets, match_id
from deliveries
where player_dismissed is not null
group by bowling_team, bowler, match_id
order by wickets desc
```

bowler	bowling_team	wickets	match_id
VR Aaron	Kings XI Punjab	29	15
A Choudhary	Royal Challengers Bangalore	28	1
RD Chahar	Rising Pune Supergiant	28	4
MJ McClenaghan	Mumbai Indians	28	7
CR Woakes	Kolkata Knight Riders	28	7
AB Dinda	Rising Pune Supergiant	27	2
UT Yadav	Kolkata Knight Riders	27	11
SP Narine	Kolkata Knight Riders	27	11
P Kumar	Gujarat Lions	27	13
A Nehra	Sunrisers Hyderabad	27	14
CR Woakes	Kolkata Knight Riders	27	14
SL Malinga	Mumbai Indians	27	16

```sql
-- Q11 WHICH BOWLER HAS THE BEST BOWLING FIGURES (MOST WICKETS TAKEN) IN A SINGLEMATCH?
select  match_id, bowler , bowling_team, count(*) as wickets
from (
        select bowler , bowling_team, count(*) as wickets, match_id
        from deliveries
        where player_dismissed is not null
        group by bowling_team, bowler, match_id
        order by wickets desc
)a
group by bowling_team, bowler, match_id
order by match_id desc
```

Result Grid | Filter Rows: | Export: | Wrap Cell

match_id	bowler	bowling_team	wickets
17	S Aravind	Royal Challengers Bangalore	1
17	AF Milne	Royal Challengers Bangalore	1
17	S Badree	Royal Challengers Bangalore	1
17	P Negi	Royal Challengers Bangalore	1
17	SR Watson	Royal Challengers Bangalore	1
16	SL Malinga	Mumbai Indians	1
16	MJ McClenaghan	Mumbai Indians	1
16	JJ Bumrah	Mumbai Indians	1
16	Basil Thampi	Gujarat Lions	1
16	AJ Tye	Gujarat Lions	1
16	Harbhajan Singh	Mumbai Indians	1
16	MM Patel	Gujarat Lions	1
16	RA Jadeja	Gujarat Lions	1
16	P Kumar	Gujarat Lions	1
16	KH Pandya	Mumbai Indians	1
16	DR Smith	Gujarat Lions	1
16	HH Pandya	Mumbai Indians	1
15	VR Aaron	Kings XI Punjab	1
15	MM Sharma	Kings XI Punjab	1
15	CH Morris	Delhi Daredevils	1
15	Sandeep Sharma	Kings XI Punjab	1
15	AR Patel	Kings XI Punjab	1
15	Z Khan	Delhi Daredevils	1

Q12 HOW MANY MATCHES RESULTED IN A WIN FOR EACH TEAM IN EACH CITY?

```sql
-- Q12 HOW MANY MATCHES RESULTED IN A WIN FOR EACH TEAM IN EACH CITY?
select winner, city , count(*) as no_of_match_win
from matches
group by winner, city
order by no_of_match_win desc
```

winner	city	no_of_match_win
Mumbai Indians	Mumbai	9
Kings XI Punjab	Chandigarh	8
Kolkata Knight Riders	Kolkata	8
Delhi Daredevils	Delhi	8
Rajasthan Royals	Jaipur	7
Sunrisers Hyderabad	Hyderabad	6
Rising Pune Supergiant	Pune	5
Mumbai Indians	Bangalore	3
Kolkata Knight Riders	Bangalore	3
Chennai Super Kings	Chennai	3
Kings XI Punjab	Indore	2
Royal Challengers Ban...	Bangalore	2
Kings XI Punjab	Mumbai	2
Rising Pune Supergiant	Mumbai	2
Kings XI Punjab	Bangalore	2
Mumbai Indians	Kolkata	2
Mumbai Indians	Hyderabad	2
Rajasthan Royals	Mumbai	2
Deccan Chargers	Cape Town	2
Kolkata Knight Riders	Delhi	1
Royal Challengers Ban...	Rajkot	1
Mumbai Indians	Indore	1
Gujarat Lions	Kolkata	1

Q13 HOW MANY TIMES DID EACH TEAM WIN THE TOSS IN EACH SEASON?

```sql
-- Q13 HOW MANY TIMES DID EACH TEAM WIN THE TOSS IN EACH SEASON?
select toss_winner, season , count(*) as no_of_win
from matches
group by toss_winner, season
```

toss_winner	season	no_of_win
Royal Challengers Bangalore	2017	9
Rising Pune Supergiant	2017	6
Kolkata Knight Riders	2017	9
Kings XI Punjab	2017	4
Sunrisers Hyderabad	2017	5
Mumbai Indians	2017	11
Gujarat Lions	2017	7
Delhi Daredevils	2017	8
Royal Challengers Bangalore	2008	5
Chennai Super Kings	2008	5
Rajasthan Royals	2008	11
Mumbai Indians	2008	8
Deccan Chargers	2008	9
Kings XI Punjab	2008	8
Kolkata Knight Riders	2008	6
Delhi Daredevils	2008	6
Chennai Super Kings	2009	2
Royal Challengers Bangalore	2009	2
Delhi Daredevils	2009	2
Kolkata Knight Riders	2009	3
Deccan Chargers	2009	2

Q14 HOW MANY MATCHES DID EACH PLAYER WIN THE "PLAYER OF THE MATCH" AWARD?

```sql
-- Q14 HOW MANY MATCHES DID EACH PLAYER WIN THE "PLAYER OF THE MATCH" AWARD?
select player_of_match, count(*) as no_of_winner
from matches
group by player_of_match
order by no_of_winner desc
```

player_of_match	no_of_winner
YK Pathan	5
SE Marsh	5
SR Watson	4
BA Stokes	3
NM Coulter-Nile	3
MS Dhoni	3
AC Gilchrist	3
RV Uthappa	2
JD Unadkat	2
Rashid Khan	2
N Rana	2
CH Gayle	2
KH Pandya	2
Sandeep Sharma	2
SP Narine	2
SK Raina	2
AJ Tye	2
V Sehwag	2
ST Jayasuriya	2
SM Pollock	2

Q16 WHICH TEAM HAS THE HIGHEST TOTAL SCORE IN A SINGLE MATCH?

```sql
-- Q16 WHICH TEAM HAS THE HIGHEST TOTAL SCORE IN A SINGLE MATCH?

select sum(total_runs) as highest_score, batting_team, match_id
from deliveries
group by batting_team, match_id
order by highest_score desc
```

Result Grid | Filter Rows: | Expor

highest_score	batting_team	match_id
207	Sunrisers Hyderabad	1
205	Delhi Daredevils	9
188	Delhi Daredevils	15
187	Rising Pune Supergiant	2
184	Mumbai Indians	2
184	Kolkata Knight Riders	3
183	Gujarat Lions	3
180	Mumbai Indians	7
178	Kolkata Knight Riders	7
177	Mumbai Indians	16
176	Gujarat Lions	16
172	Royal Challengers Ban...	1
172	Gujarat Lions	13
172	Kolkata Knight Riders	14
171	Kolkata Knight Riders	11
171	Rising Pune Supergiant	13
170	Kings XI Punjab	11
164	Kings XI Punjab	4
163	Rising Pune Supergiant	4
159	Mumbai Indians	10
158	Sunrisers Hyderabad	10
157	Royal Challengers Ban...	5

Q17 WHICH BATSMAN HAS SCORED THE MOST RUNS IN A SINGLE MATCH?

```sql
-- Q17 WHICH BATSMAN HAS SCORED THE MOST RUNS IN A SINGLE MATCH?
select batsman , sum(total_runs) as highest_score, match_id
from deliveries
group by batsman, match_id
order by highest_score desc
```

batsman	highest_score	match_id
SV Samson	105	9
CA Lynn	102	3
AB de Villiers	92	8
SPD Smith	85	2
MK Pandey	84	7
G Gambhir	82	3
G Gambhir	80	11
DA Warner	78	6
KA Pollard	73	12
RV Uthappa	73	14
KM Jadhav	71	5
SK Raina	69	3
Yuvraj Singh	65	1
BB McCullum	65	16
V Kohli	63	12
N Rana	61	16
AM Rahane	60	2
RR Pant	60	5
HM Amla	60	8
SW Billings	58	15
KD Karthik	56	16
BA Stokes	53	4

THANK YOU