

1. View the first 100 rows of the dataset to understand its structure.

```
1  [ ]< SELECT TOP (100) [player_name]
2      ,[team]
3      ,[date_of_birth]
4      ,[age]
5      ,[marital_status]
6      ,[number_of_kids]
7      ,[nationality]
8      ,[country_of_birth]
9      ,[position]
10     ,[preferred_foot]
11     ,[height_cm]
12     ,[weight_kg]
13     ,[jersey_number]
14     ,[injury_status]
15     ,[agent]
16     ,[matches_played]
17     ,[minutes_played]
18     ,[goals]
19     ,[assists]
20     ,[tackles]
21     ,[interceptions]
22     ,[saves]
23     ,[clean_sheets]
24     ,[yellow_cards]
--
```

% No issues found Ln: 36, Ch: 3 | SPC | CRLF | Windows 1252

player_name	team	date_of_birth	age	marital_status	number_of_kids	nationality
Siyanda Dlamini	Stellenbosch FC	1995-05-28 00:00:00.0000000	30	Widowed	0	South African
Thabo Ndlovu	Cape Town City	2004-06-01 00:00:00.0000000	21	Single	1	Zimbabwean
Vusi Molefe	Stellenbosch FC	2006-09-20 00:00:00.0000000	19	Single	0	Nigerian
Thembi Mahlangu	Bloemfontein Celtic	2004-02-08 00:00:00.0000000	21	Divorced	0	Zambian
Nokuthula Sithole	Polokwane City	2003-03-18 00:00:00.0000000	22	Divorced	3	Nigerian
Thembi Sithole	Kaizer Chiefs	1989-02-14 00:00:00.0000000	36	Married	2	Nigerian
Siyanda Mahlangu	Chippa United	1989-12-09 00:00:00.0000000	36	Single	0	Zambian
Lerato Mashaba	Polokwane City	1998-01-31 00:00:00.0000000	27	Widowed	0	Ghanaian
Nomsa Mahlangu	Polokwane City	1991-07-16 00:00:00.0000000	34	Divorced	0	Malawian
Tendayi Kanganwa	Kaizer Chiefs	1996-09-20 00:00:00.0000000	29	Married	4	Malawian

2. Count the total number of players in the dataset

```
36
37     select COUNT(player_name)
38  from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced];
```

00 % No issues found Ln: 38, Ch: 18 | SPC | CRLF | Windows 1252

Results Messages

	player_name	team	date_of_birth	age	marital_status	number_of_kids	nationality
1	Siyanda Dlamini	Stellenbosch FC	1995-05-28 00:00:00.0000000	30	Widowed	0	South African
2	Thabo Ndlovu	Cape Town City	2004-06-01 00:00:00.0000000	21	Single	1	Zimbabwean
3	Vusi Molefe	Stellenbosch FC	2006-09-20 00:00:00.0000000	19	Single	0	Nigerian
4	Thembi Mahlangu	Bloemfontein Celtic	2004-02-08 00:00:00.0000000	21	Divorced	0	Zambian
5	Nokuthula Sithole	Polokwane City	2003-03-18 00:00:00.0000000	22	Divorced	3	Nigerian
6	Thembi Sithole	Kaizer Chiefs	1989-02-14 00:00:00.0000000	36	Married	2	Nigerian
7	Siyanda Mahlangu	Chippa United	1989-12-09 00:00:00.0000000	36	Single	0	Zambian
8	Lerato Mashaba	Polokwane City	1998-01-31 00:00:00.0000000	27	Widowed	0	Ghanaian
9	Nomsa Mahlangu	Polokwane City	1991-07-16 00:00:00.0000000	34	Divorced	0	Malawian
10	Tumelo Khumalo	Kaizer Chiefs	1996-08-28 00:00:00.0000000	29	Married	4	Malawian
11	Gugu Molefe	TS Galaxy	2007-01-29 00:00:00.0000000	18	Widowed	0	South African
12	Thembi Mokoena	Stellenbosch FC	1995-07-03 00:00:00.0000000	30	Widowed	0	Zimbabwean

(No column name)

1 300

3. List all unique teams in the league.

```
39
40     select distinct(team)
41      from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced];
```

10 % No issues found Ln: 41, Ch: 69 | SPC | CRLF | Windows 12

Results Messages

Nokuthula Sithole	Polokwane City	2003-03-18 00:00:00.0000000	22	Divorced	3	Nigerian
Thembi Sithole	Kaizer Chiefs	1989-02-14 00:00:00.0000000	36	Married	2	Nigerian
Siyanda Mahlangu	Chippa United	1989-12-09 00:00:00.0000000	36	Single	0	Zambian
Lerato Mashaba	Polokwane City	1998-01-31 00:00:00.0000000	27	Widowed	0	Ghanaian
Nomsa Mahlangu	Polokwane City	1991-07-16 00:00:00.0000000	34	Divorced	0	Malawian
0 Tumelo Khumalo	Kaizer Chiefs	1996-08-28 00:00:00.0000000	29	Married	4	Malawian
1 Gugu Molefe	TS Galaxy	2007-01-29 00:00:00.0000000	18	Widowed	0	South Africa

(No column name)

300

team
AmaZulu FC
Bloemfontein Celtic
Cape Town City
Chippa United
Golden Arrows
Kaizer Chiefs
Mamelodi Sundowns
Moroka Swallows

4. Count how many players are in each team.

```
42
43     | select team, count (player_name) as NumberOfPlayers
44     | from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
45     | group by team;
```

100 % No issues found Ln: 43, Ch: 1 (138 chars, 3 lines) SPC CRLF Windows

Results Messages

	team	NumberOfPlayers
1	AmaZulu FC	22
2	Bloemfontein Celtic	17
3	Cape Town City	15
4	Chippa United	21
5	Golden Arrows	23
6	Kaizer Chiefs	20
7	Mamelodi Sundowns	19
8	Moroka Swallows	19
9	Orlando Pirates	9
10	Polokwane City	27
11	Richards Bay FC	13
12	Royal AM	18
13	Sekhukhune United	17
14	Stellenbosch FC	20
15	SuperSport United	20
16	TS Galaxy	20

5. Identify the top 10 players with the most goals.

```
46
47     select top(10) player_name, goals
48     from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
49     order by goals desc;
50
```

100 % No issues found Ln: 47, Ch: 1 | (126 chars, 3 lines) | SPC | CRLF | W

Results Messages

	player_name	goals
1	Vusi Molefe	99
2	Thabo Ndlovu	98
3	Gugu Hlongwane	98
4	Thabo Sithole	98
5	Mandla Mabena	98
6	Mpho Mahlangu	97
7	Boitumelo Nkosi	96
8	Mpho Radebe	92
9	Khanyi Nkosi	92
10	Sipho Phiri	91

6. Find the average salary for players in each team.

```
select team, avg(average_salary_zar)
from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
group by team;
```

Range - Not Saved

No issues found Ln: 51, Ch: 2 | (122 chars, 3 lines) | SPC | CRLF | Wind

.17, Level 16, State 1, Line 51
 id data type nvarchar is invalid for avg operator.

ation time: 2025-11-25T23:26:07.8638902+02:00

7. Retrieve the top 10 players with the highest market value.

```
54
55     select top(10) player_name, market_value_zar
56     from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
57     order by market_value_zar desc;
```

100 % No issues found Ln: 55, Ch: 1 (147 chars, 3 lines) SPC CRLF

Results Messages

	player_name	market_value_zar
1	Nokuthula Baloyi	9906925.35
2	Gugu Mokoena	9868061.36
3	Khanyi Mthembu	9845842.09
4	Nomsa Hlongwane	975559.35
5	Kabelo Mahlangu	9606957.95
6	Kagiso Dlamini	9383118.97
7	Lindiwe Phiri	9352111.61
8	Mpho Radebe	9294272.32
9	Lindiwe Molefe	9277927.61
10	Thembi Tshabalala	9091048.14

8. Calculate the average passing accuracy for each position.

```
58
59     select position, avg(passing_accuracy)
60     from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
61     group by position;
```

100 % No issues found Ln: 61, Ch: 7 | SPC | CRLF | Windows

Messages

```
Msg 8117, Level 16, State 1, Line 59
Operand data type nvarchar is invalid for avg operator.

Completion time: 2025-11-25T23:30:41.7802173+02:00
```

9. Compare shot accuracy with goals to find correlations.

10. Compute total goals and assists for each team.

```
62
63     select team, sum(goals) as total_goals, sum(assists) as total_assists
64     from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
65     group by team;
```

100 % No issues found Ln: 65, Ch: 16 | (156 chars, 3 lines) | SPC | CRLF | Windows 1252

Messages

```
Msg 8117, Level 16, State 1, Line 63
Operand data type nvarchar is invalid for sum operator.

Completion time: 2025-11-25T23:34:09.5933375+02:00
```

11. Count players by their marital status.

```
66  
67 select marital_status, count (player_name) as NumberOfMarriedPlayers  
68 from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]  
69 group by marital_status;
```

100 % No issues found Ln: 67, Ch: 1 (165 chars, 3 lines) SPC CRLF

Results Messages

	marital_status	NumberOfMarriedPlayers
1	Divorced	78
2	Maried	65
3	Single	79
4	Widowed	78

12. Count players by nationality.

```
70  
71 select nationality, count (player_name) as NationalityOfPlayers  
72 from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]  
73 group by nationality;
```

100 % No issues found Ln: 71, Ch: 1 (157 chars, 3 lines) SPC CRLF

Results Messages

	nationality	NationalityOfPlayers
1	Ghanaian	47
2	Malawian	42
3	Mozambican	37
4	Nigerian	39
5	South African	46
6	Zambian	44
7	Zimbabwean	45

13. Find average market value grouped by nationality.

```
70
71 | select nationality, avg (market_value_zar) as AvgMarketValue
72 | from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
73 | group by nationality;
```

100 % No issues found Ln: 71, Ch: 1 (154 chars, 3 lines) SPC CRLF

Messages

Msg 8117, Level 16, State 1, Line 71
Operand data type nvarchar is invalid for avg operator.

Completion time: 2025-11-25T23:39:53.8926396+02:00

14. Determine how many player contracts end in each year.

```
74
75 | select contract_end_year, count (player_name) as OutOfContract
76 | from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
77 | group by contract_end_year;
```

100 % No issues found Ln: 75, Ch: 2 (161 chars, 3 lines) SPC CRLF

Results Messages

	contract_end_year	OutOfContract
1	2026	63
2	2027	70
3	2028	52
4	2029	50
5	2030	65

15. Identify players whose contracts end next year.

```

74
75   | select contract_end_year, player_name
76   | from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
77   | where contract_end_year = 2026;

```

100 % No issues found Ln: 75, Ch: 1 (140 chars, 3 lines) SPC CRLF

	contract_end_year	player_name
1	2026	Siyanda Dlamini
2	2026	Vusi Molefe
3	2026	Nokuthula Sithole
4	2026	Siyanda Mahlangu
5	2026	Nomsa Mahlangu
6	2026	Thembi Mokoena
7	2026	Thabo Sithole
8	2026	Kagiso Phiri
9	2026	Thembi Tshabalala
10	2026	Mandla Baloyi
11	2026	Tumelo Mokoena
12	2026	Thembi Mashaba
13	2026	Lerato Mabena
14	2026	Lindiwe Radebe
15	2026	Sipho Ndlovu
16	2026	Lebogang Phiri
17	2026	Sibusiso Radebe
18	2026	Sipho Khumalo
19	2026	Thabo Sithole

16. Summarize the number of players by injury status.

```

78
79   |< select injury_status, count (player_name) as injurystat
80   |< from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
81   |< group by injury_status;

```

Change - Not Saved

100 % No issues found Ln: 79, Ch: 1 (150 chars, 3 lines) SPC CRLF Win

	injury_status	injurystat
1	Healthy	99
2	Injured	97
3	Recovering	104

17. Calculate goals per match ratio for each player.

```
82
83     select player_name, (goals/matches_played) as ratio
84     from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced];
```

100 % No issues found Ln: 83, Ch: 29 | SPC | CRLF | V

Messages

Msg 8117, Level 16, State 1, Line 83
Operand data type nvarchar is invalid for divide operator.

Completion time: 2025-11-25T23:49:26.1744233+02:00

18. Count how many players are managed by each agent.

```
85
86     select agent, count (player_name) as agents
87     from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
88     group by agent;
```

100 % No issues found Ln: 86, Ch: 1 | (131 chars, 3 lines) | SPC | CRLF | W

Results Messages

	agent	agents
1	None	62
2	PlayerFirst	63
3	ProSport	62
4	SA Elite Agents	51
5	SoccerLink Africa	62

19.

19. Calculate average height and weight by player position.

```
89
90     select position, avg (height_cm) as Avgheight, avg (weight_kg) as AvgWeight
91     from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
92     group by position;
```

100 % No issues found Ln: 90, Ch: 1 | (166 chars, 3 lines) | SPC | CRLF | Windows

Results Messages

	position	Avgheight	AvgWeight
1	Defender	182	78
2	Forward	179	78
3	Goalkeeper	179	78
4	Midfielder	179	76

20. Identify players with the highest combined goals and assists.

```
92     group by position;
93
94     |> select player_name, (goals+assists)
95     |> from [Soccer_AnalysisDB1].[dbo].[ketro_sa_soccer_dataset_advanced]
96     |> group by plaayer_Name;
```

100 % 1 0 | ↑ ↓ Ln: 94, Ch: 2 | (129 chars, 3 lines) | SPC | CRLF

Messages

Msg 207, Level 16, State 1, Line 96
Invalid column name 'plaayer_Name'.

Completion time: 2025-11-25T23:54:47.5731308+02:00