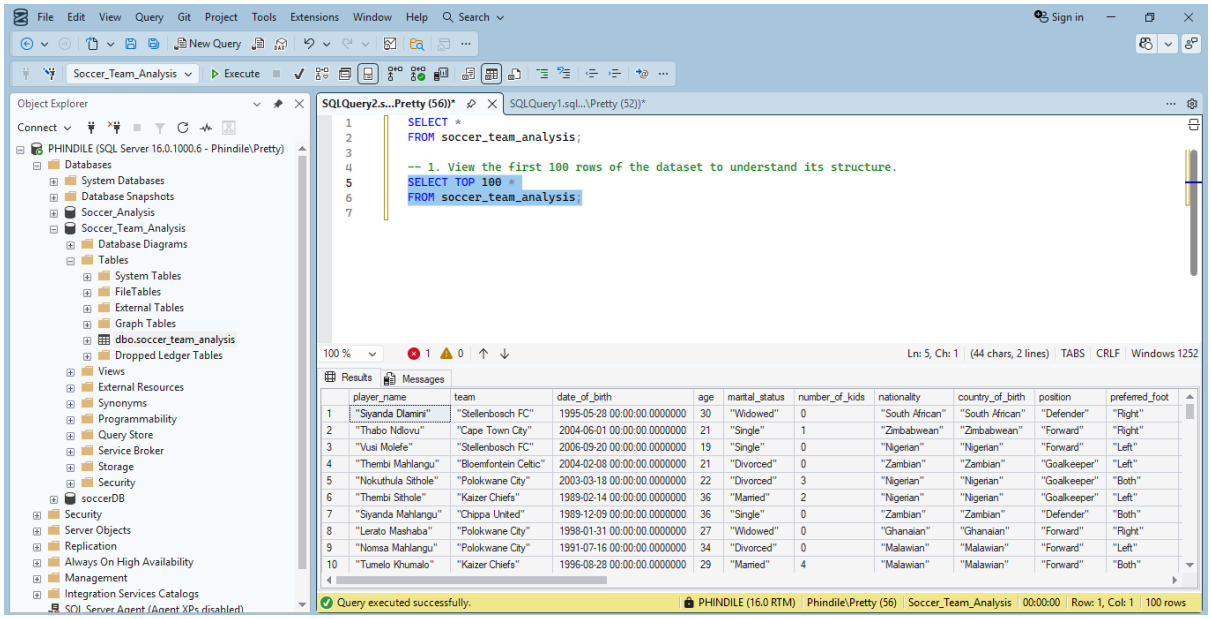
South African Soccer Dataset – SQL

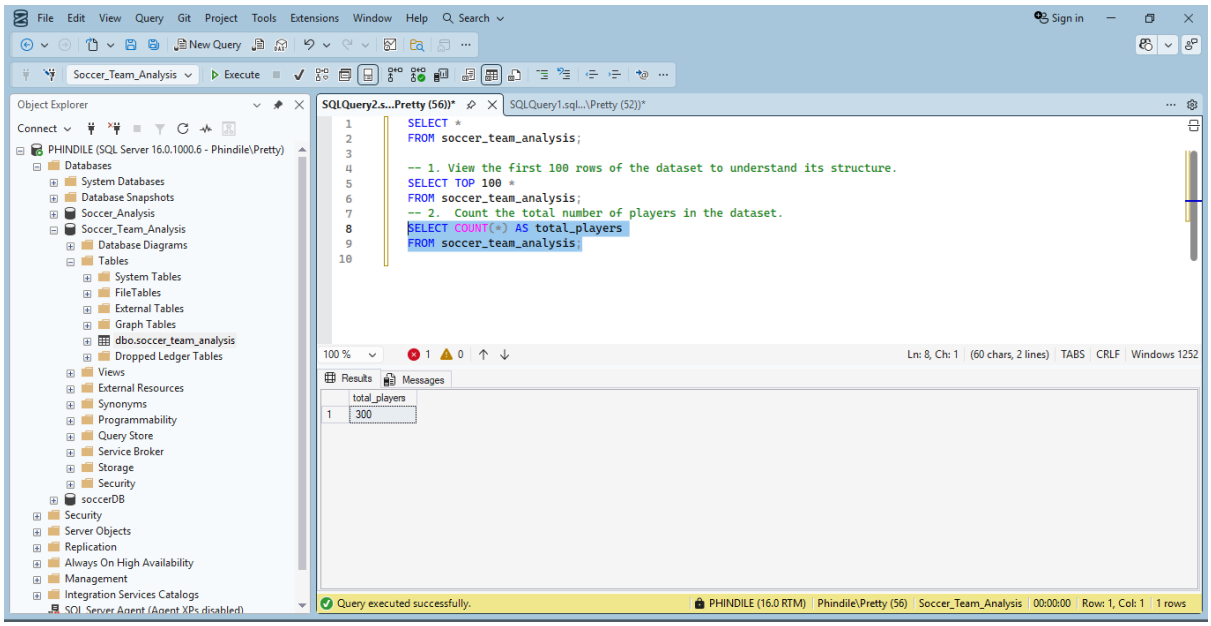
1. View the first 100 rows of the dataset to understand its structure.



```
SELECT *
FROM soccer_team_analysis;

-- 1. View the first 100 rows of the dataset to understand its structure.
SELECT TOP 100 *
FROM soccer_team_analysis;
```

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

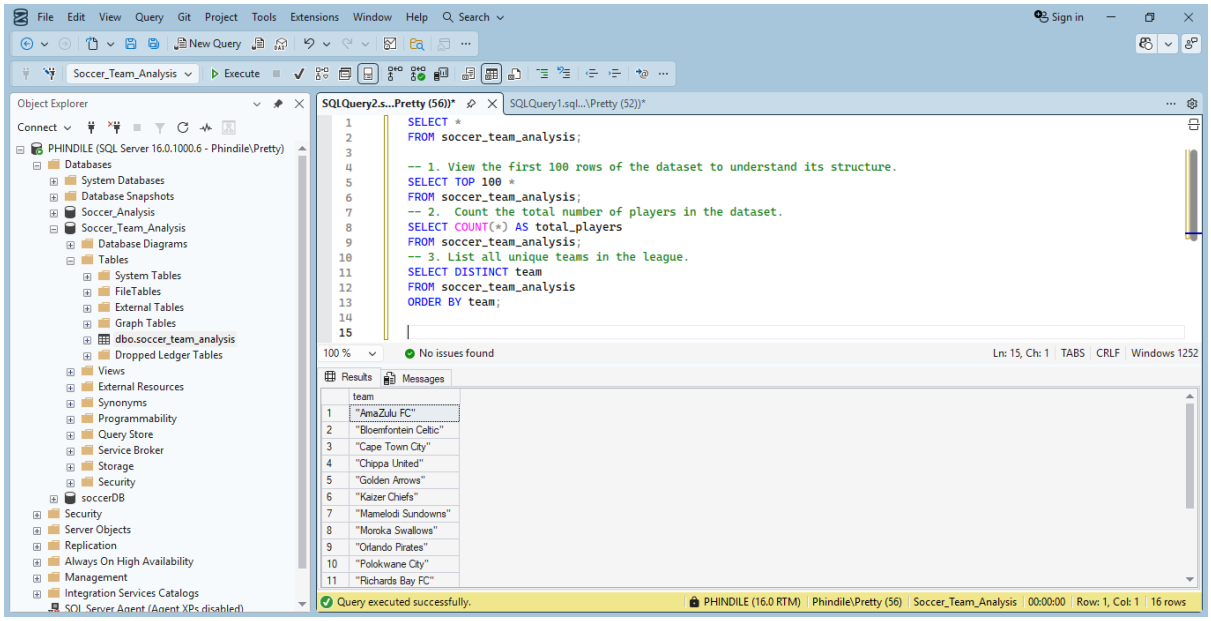
2. Count the total number of players in the dataset.



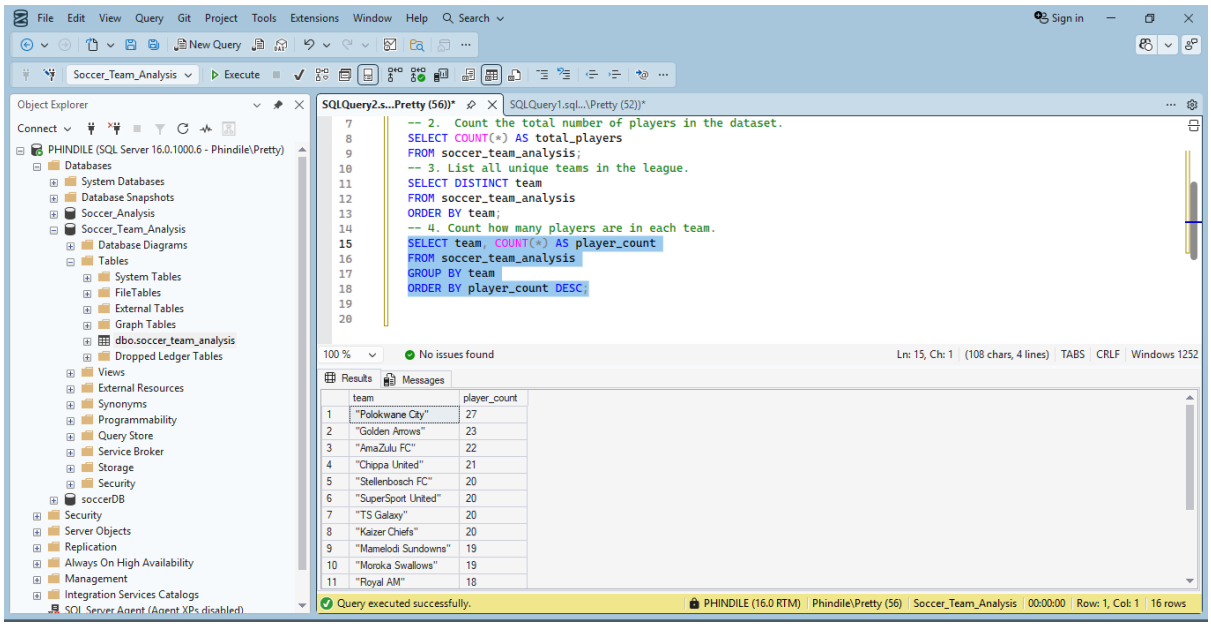
```
SELECT *
FROM soccer_team_analysis;

-- 1. View the first 100 rows of the dataset to understand its structure.
SELECT TOP 100 *
FROM soccer_team_analysis;
-- 2.  Count the total number of players in the dataset.
SELECT COUNT(*) AS total_players
FROM soccer_team_analysis;
```

| | total_players |
|---|---|
| 1 | 300 |

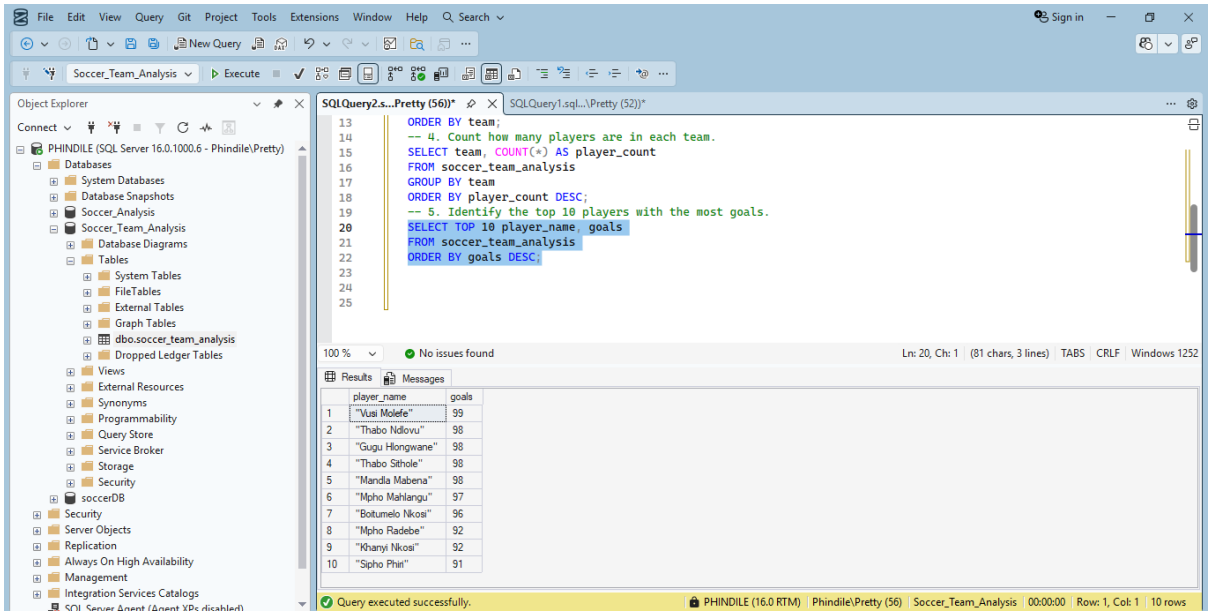
3. List all unique teams in the league.



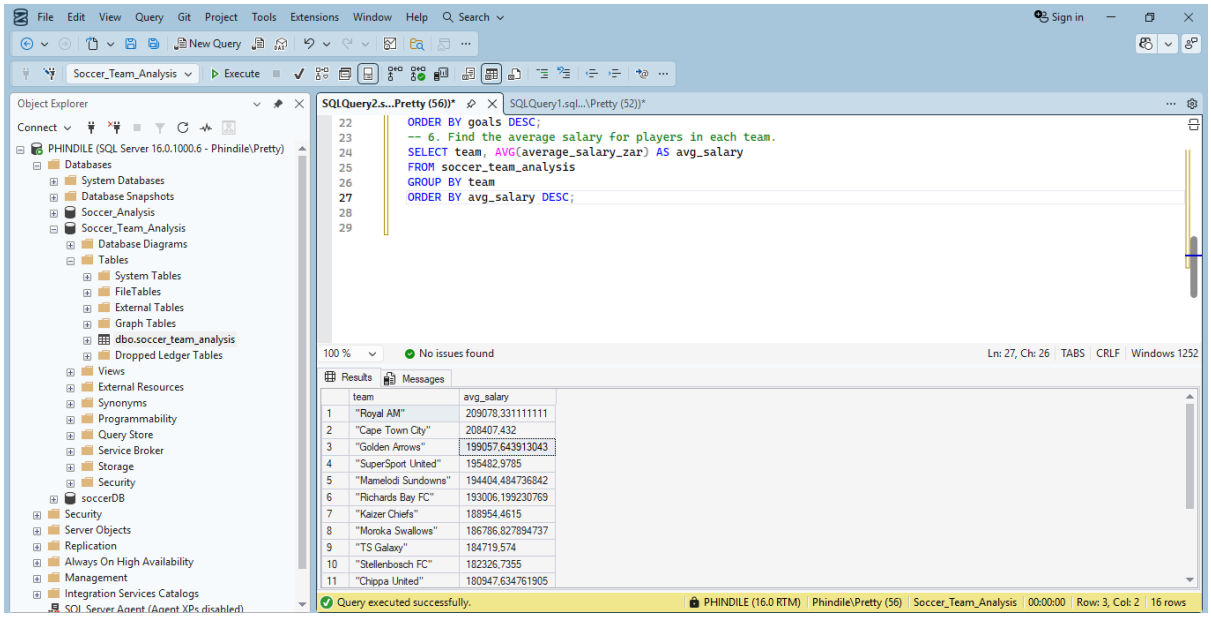

4. Count how many players are in each team.



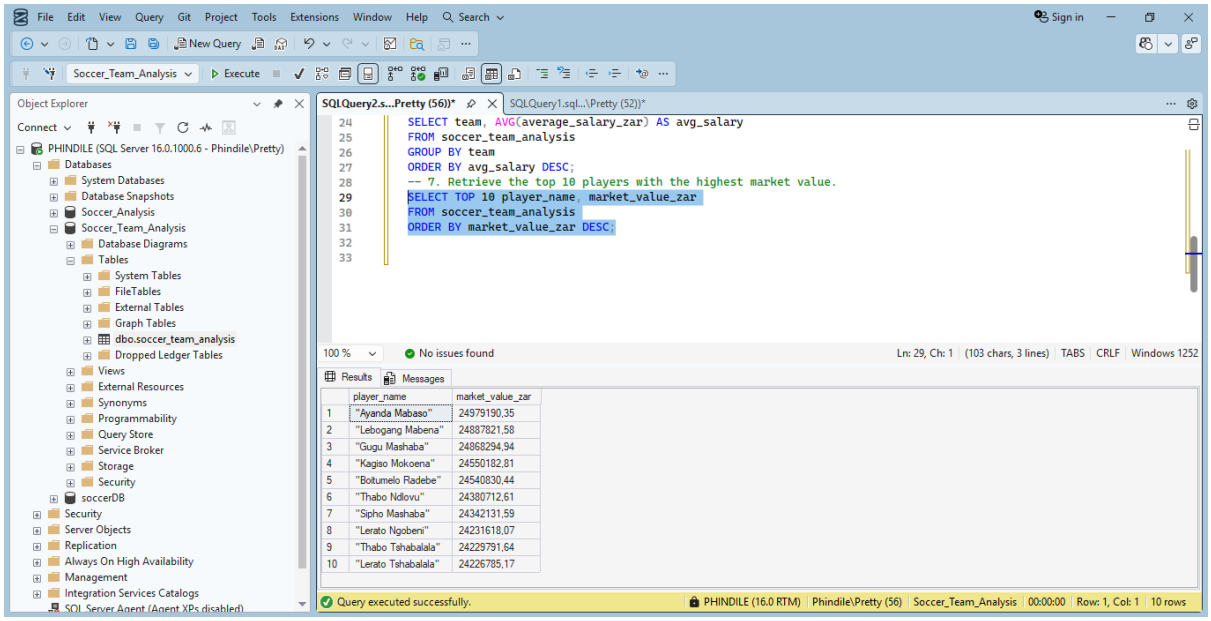

5. Identify the top 10 players with the most goals.



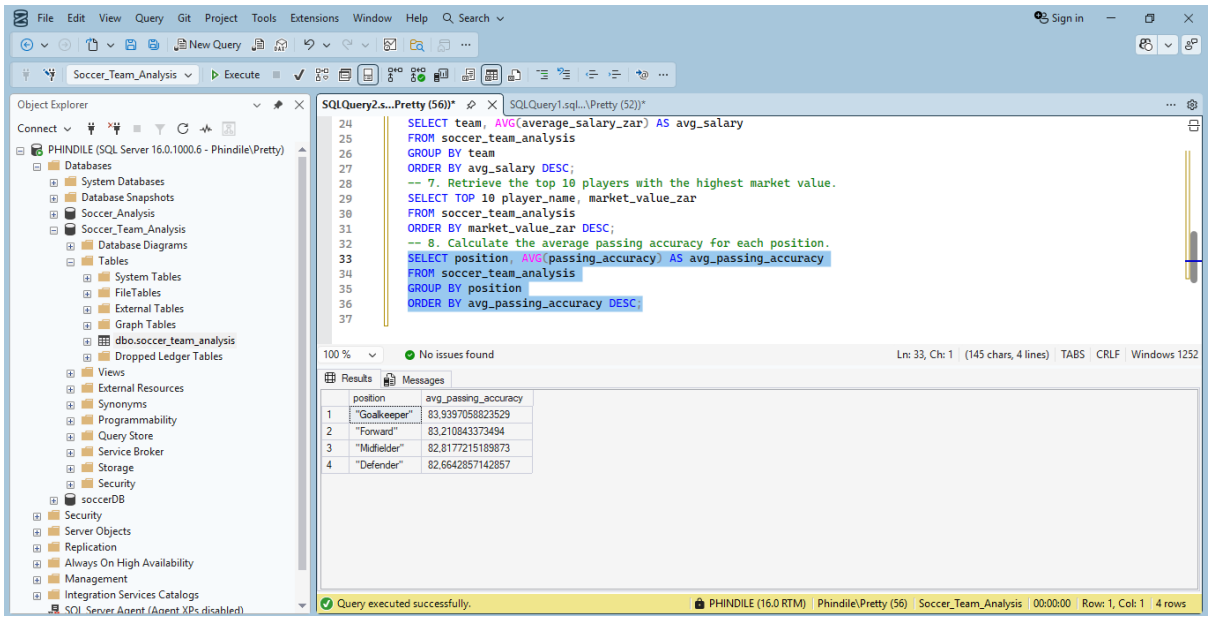

6. Find the average salary for players in each team.



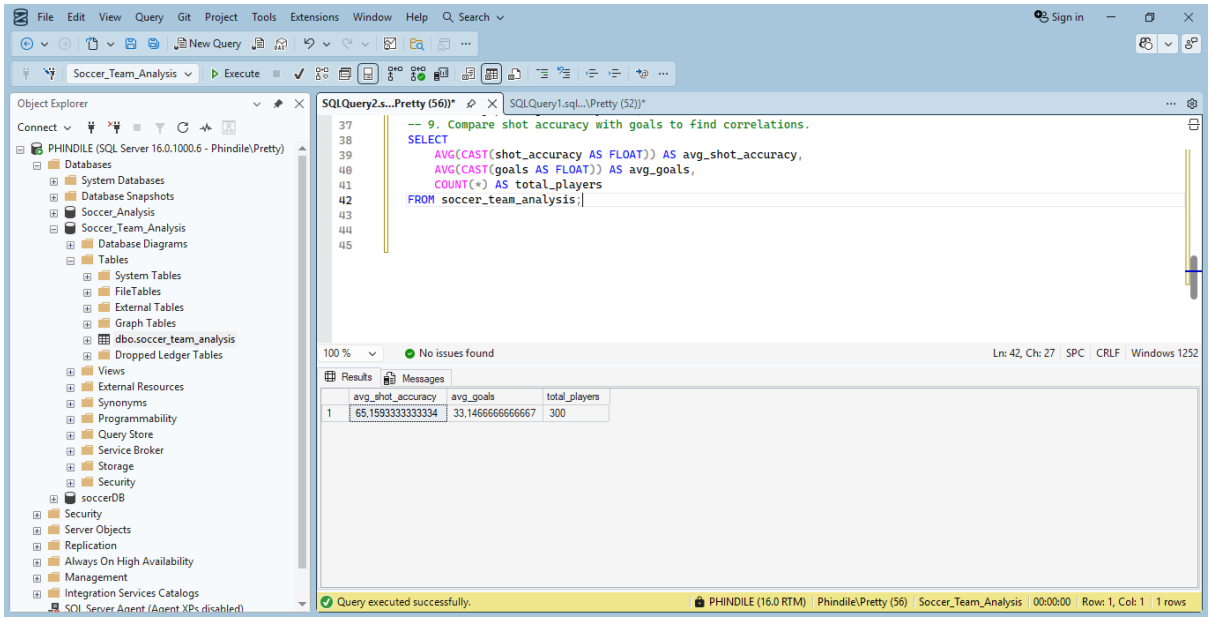

7. Retrieve the top 10 players with the highest market value.



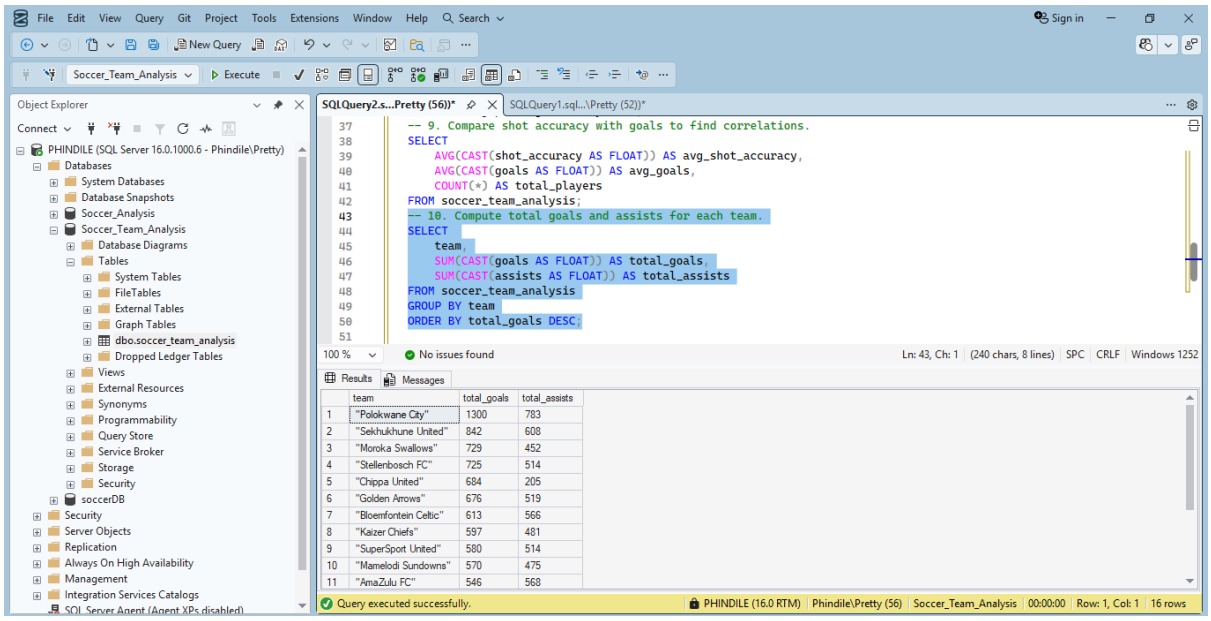

8. Calculate the average passing accuracy for each position.



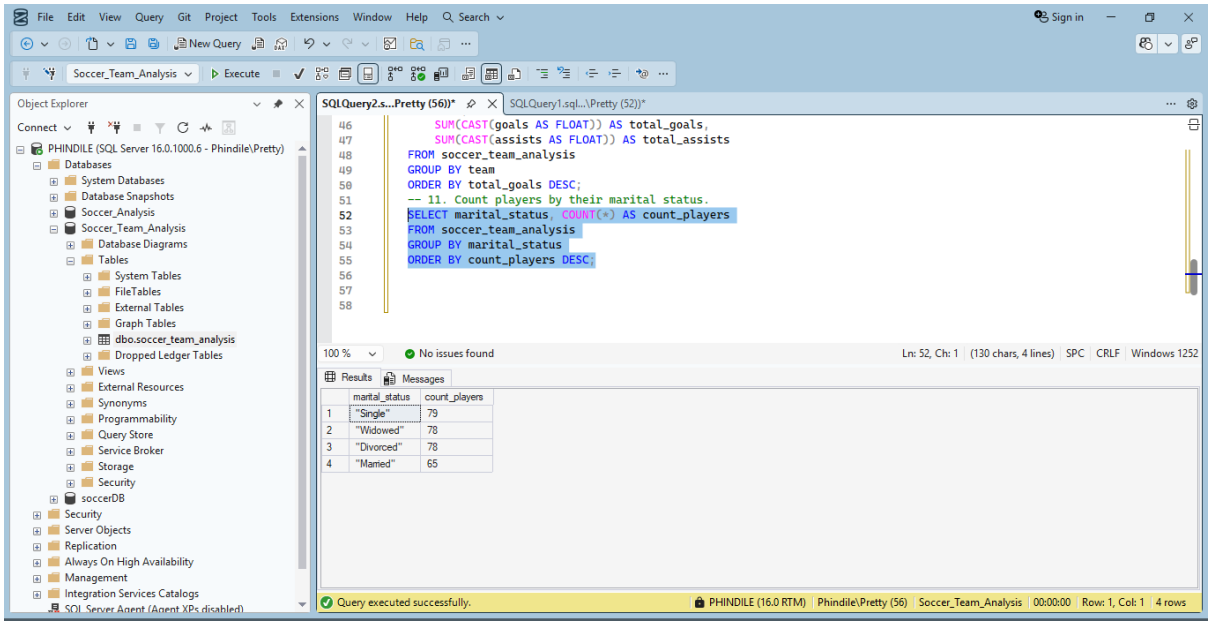

9. Compare shot accuracy with goals to find correlations.



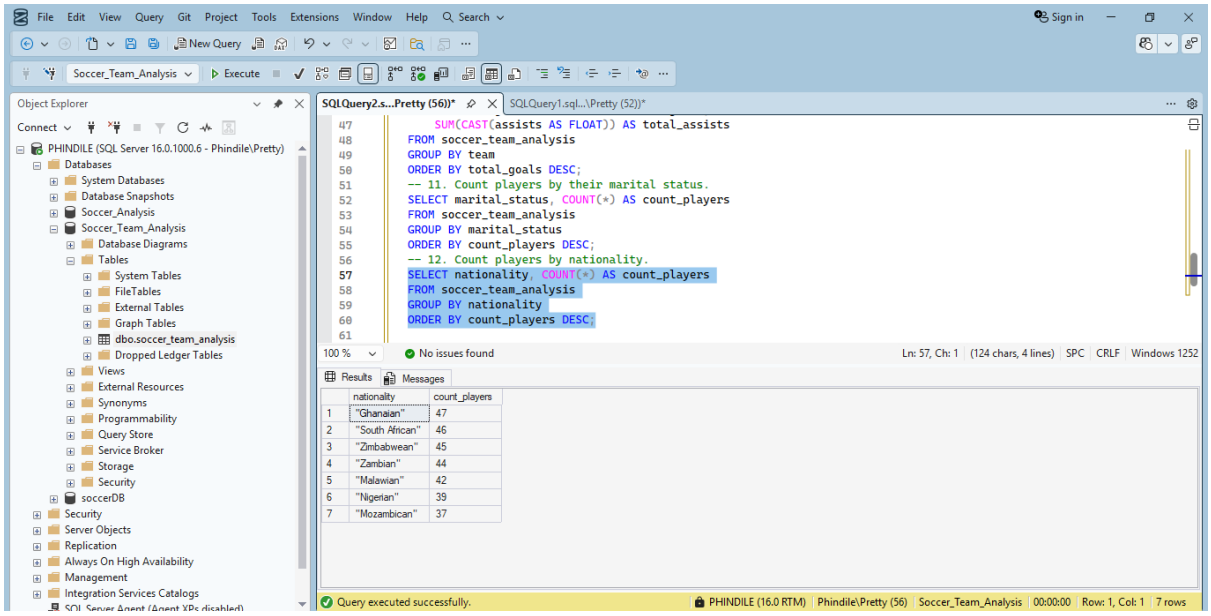

10. Compute total goals and assists for each team



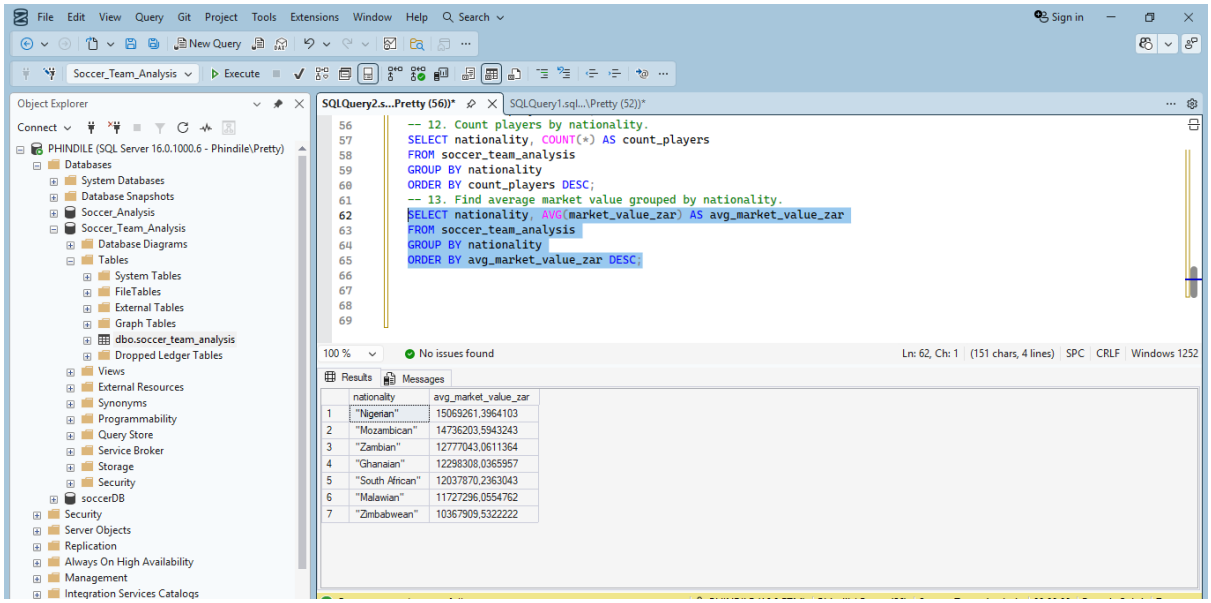

11. Count players by their marital status.



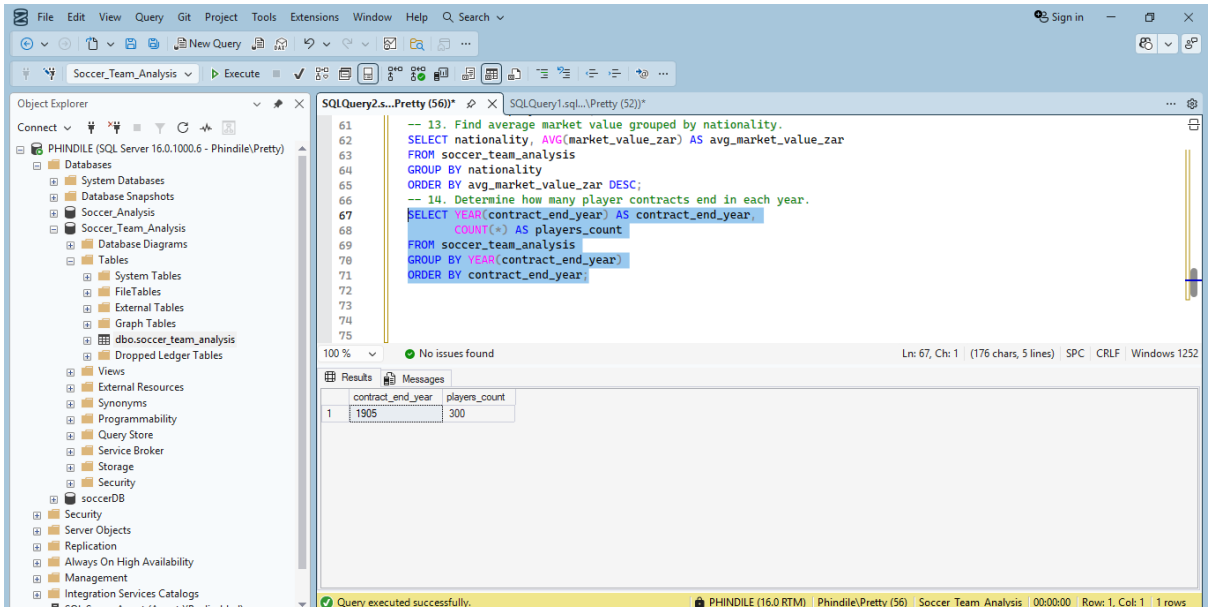

12. Count players by nationality.



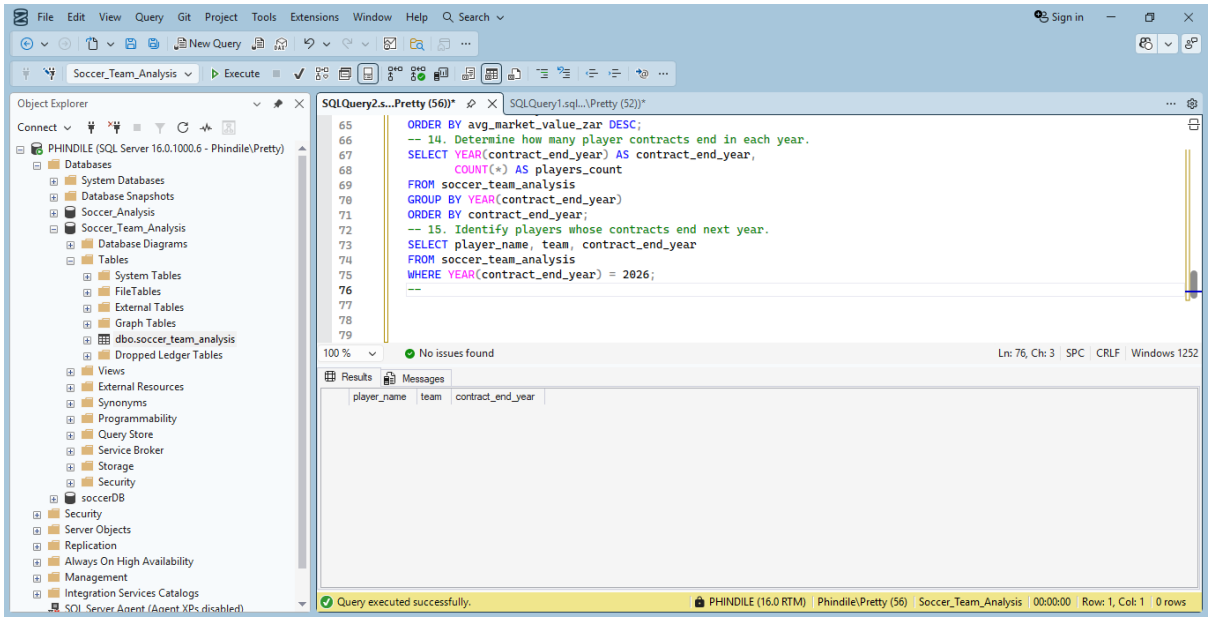

13. Find average market value grouped by nationality



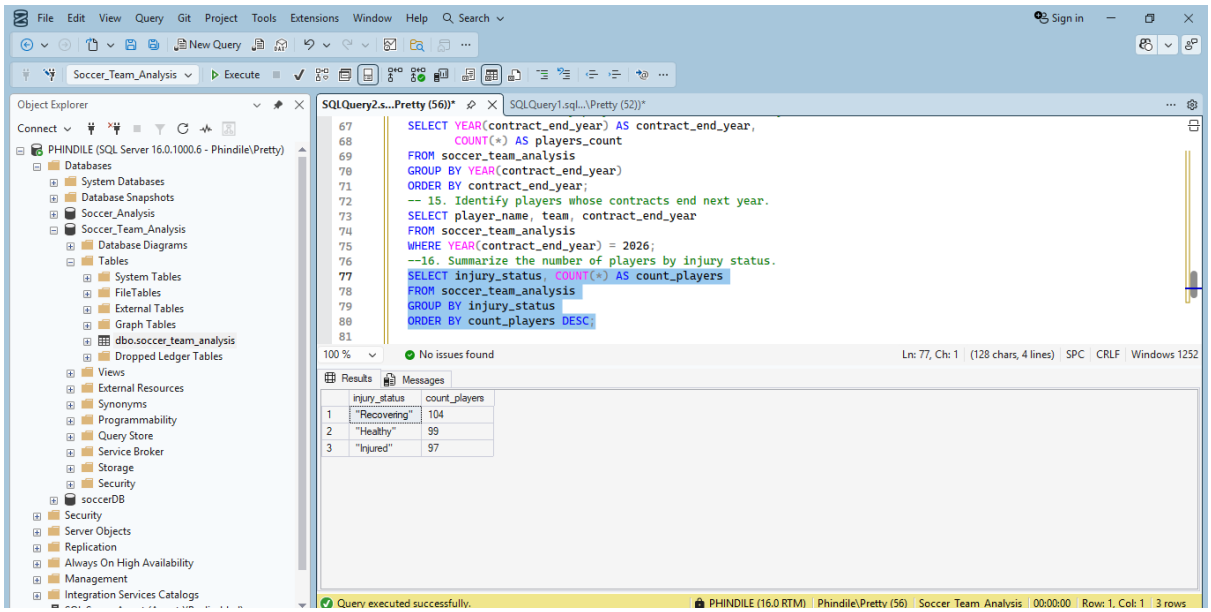

14. Determine how many player contracts end in each year.



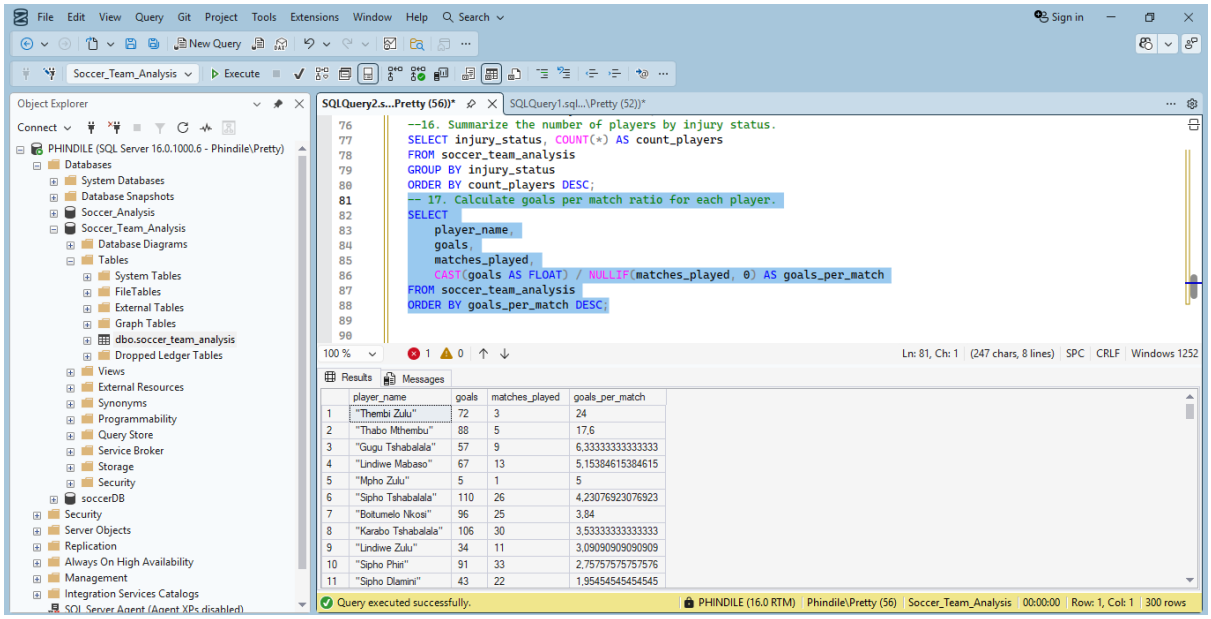

15. Identify players whose contracts end next year.




16. Summarize the number of players by injury status.



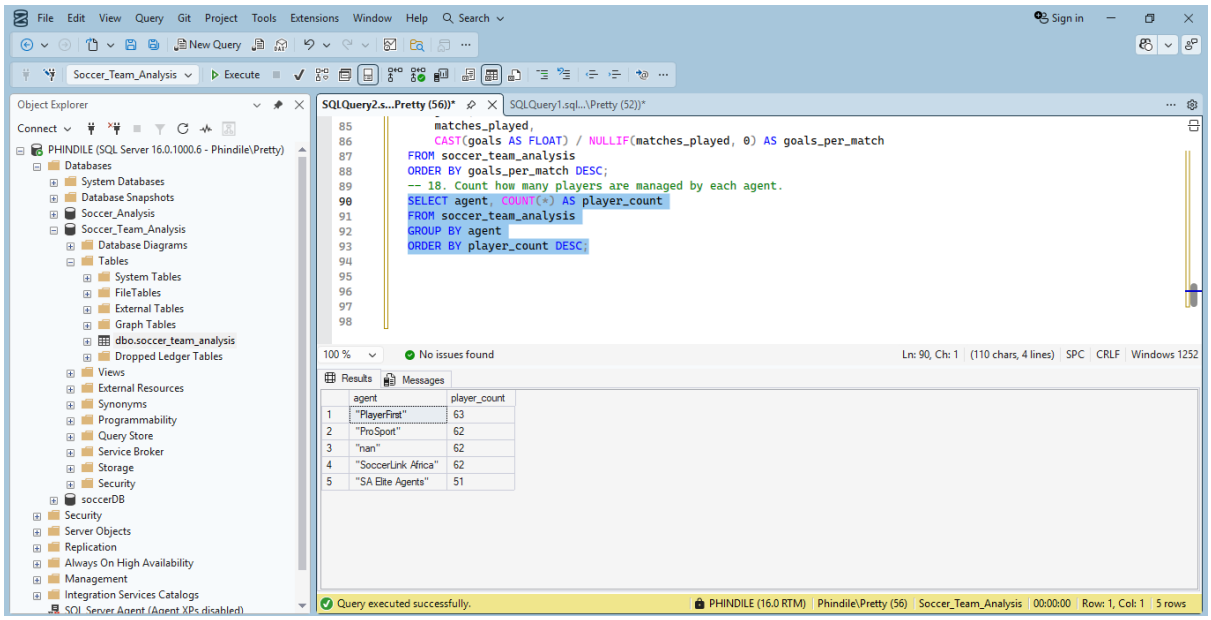

17. Calculate goals per match ratio for each player.



```
--16. Summarize the number of players by injury status.
SELECT injury_status, COUNT(*) AS count_players
FROM soccer_team_analysis
GROUP BY injury_status
ORDER BY count_players DESC;
-- 17. Calculate goals per match ratio for each player.
SELECT
    player_name,
    goals,
    matches_played,
    CAST(goals AS FLOAT) / NULLIF(matches_played, 0) AS goals_per_match
FROM soccer_team_analysis
ORDER BY goals_per_match DESC;
```

| | player_name | goals | matches_played | goals_per_match |
|---|---|---|---|---|
| 1 | "Thembi Zulu" | 72 | 3 | 24 |
| 2 | "Thabo Mthembu" | 88 | 5 | 17,6 |
| 3 | "Gugu Tshabalala" | 57 | 9 | 6,33333333333333 |
| 4 | "Lindiwe Mabaso" | 67 | 13 | 5,15384615384615 |
| 5 | "Mpho Zulu" | 5 | 1 | 5 |
| 6 | "Sipho Tshabalala" | 110 | 26 | 4,23076923076923 |
| 7 | "Boitumelo Nkosi" | 96 | 25 | 3,84 |
| 8 | "Karabo Tshabalala" | 106 | 30 | 3,53333333333333 |
| 9 | "Lindiwe Zulu" | 34 | 11 | 3,09090909090909 |
| 10 | "Sipho Phiri" | 91 | 33 | 2,75757575757576 |
| 11 | "Sipho Dlamini" | 43 | 22 | 1,95454545454545 |

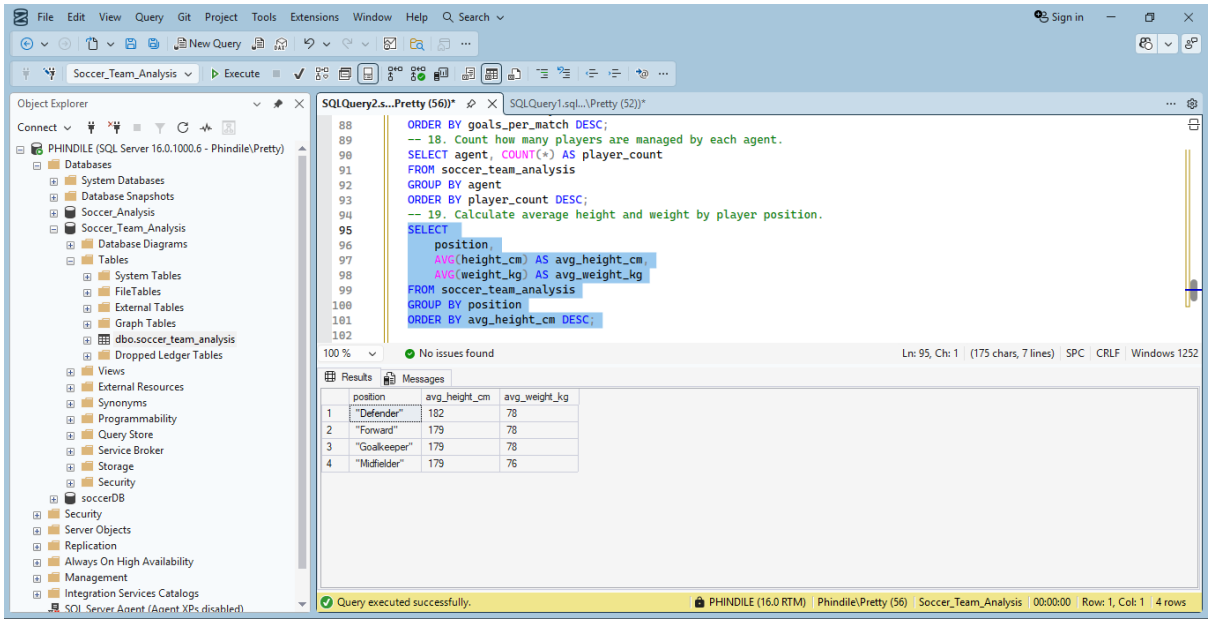
18. Count how many players are managed by each agent.



```
    matches_played,
    CAST(goals AS FLOAT) / NULLIF(matches_played, 0) AS goals_per_match
FROM soccer_team_analysis
ORDER BY goals_per_match DESC;
-- 18. Count how many players are managed by each agent.
SELECT agent, COUNT(*) AS player_count
FROM soccer_team_analysis
GROUP BY agent
ORDER BY player_count DESC;
```

| | agent | player_count |
|---|---|---|
| 1 | "PlayerFirst" | 63 |
| 2 | "ProSport" | 62 |
| 3 | "nan" | 62 |
| 4 | "SoccerLink Africa" | 62 |
| 5 | "SA Elite Agents" | 51 |

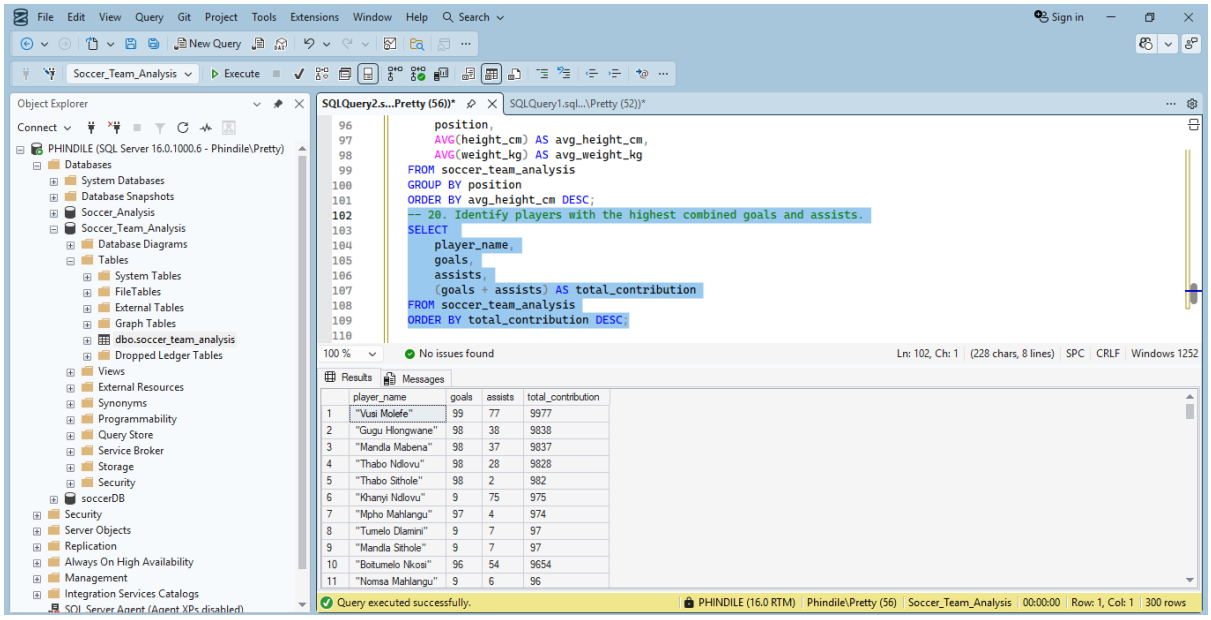
19. Calculate average height and weight by player position.



```
ORDER BY goals_per_match DESC;
-- 18. Count how many players are managed by each agent.
SELECT agent, COUNT(*) AS player_count
FROM soccer_team_analysis
GROUP BY agent
ORDER BY player_count DESC;
-- 19. Calculate average height and weight by player position.
SELECT
    position,
    AVG(height_cm) AS avg_height_cm,
    AVG(weight_kg) AS avg_weight_kg
FROM soccer_team_analysis
GROUP BY position
ORDER BY avg_height_cm DESC;
```

| | position | avg_height_cm | avg_weight_kg |
|---|---|---|---|
| 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.



```
    position,
    AVG(height_cm) AS avg_height_cm,
    AVG(weight_kg) AS avg_weight_kg
FROM soccer_team_analysis
GROUP BY position
ORDER BY avg_height_cm DESC;
-- 20. Identify players with the highest combined goals and assists.
SELECT
    player_name,
    goals,
    assists,
    (goals + assists) AS total_contribution
FROM soccer_team_analysis
ORDER BY total_contribution DESC;
```

| | player_name | goals | assists | total_contribution |
|---|---|---|---|---|
| 1 | "Vusi Molefe" | 99 | 77 | 9977 |
| 2 | "Gugu Hlongwane" | 98 | 38 | 9838 |
| 3 | "Mandla Mabena" | 98 | 37 | 9837 |
| 4 | "Thabo Ndlovu" | 98 | 28 | 9828 |
| 5 | "Thabo Sithole" | 98 | 2 | 982 |
| 6 | "Khanyi Ndlovu" | 9 | 75 | 975 |
| 7 | "Mpho Mahlangu" | 97 | 4 | 974 |
| 8 | "Tumelo Dlamini" | 9 | 7 | 97 |
| 9 | "Mandla Sithole" | 9 | 7 | 97 |
| 10 | "Boitumelo Nkosi" | 96 | 54 | 9654 |
| 11 | "Nomsa Mahlangu" | 9 | 6 | 96 |