Introduction

In this project, we investigate the public dataset, ncaa_basketball, by applying fundamental SQL commands in Big Query of the Google Cloud Platform (GCP) to explore the glory of Duke MBB.

There are many tables in this dataset, and we would like to use the “mbb_players_games_sr” to help us understand Duke players’ average point per game (PPG). There are four steps to achieve this target, which are listed in the next section.

Implementation

Step 1: calculate a Duke player’s total scored points in his college career with below command:

```
(
select mb.full_name, count(mb.game_id) as player_game_played
from `bigquery-public-data.ncaa_basketball.mbb_players_games_sr` mb
where mb.played = true and mb.team_market = "Duke"
group by mb.full_name
order by player_game_played desc
) mb1
```

Step 2: calculate a Duke player’s total attended games in his college career with below command:

```
(
select mb.full_name, sum(mb.points) as player_total_points
from `bigquery-public-data.ncaa_basketball.mbb_players_games_sr` mb
where mb.played = true and mb.team_market = "Duke"
group by mb.full_name
order by player_total_points desc
) mb1
```

Step 3: join the two tables above based on player’s name with below command:

```
(
select mb1.full_name, mb1.player_game_played, mb2.player_total_points
```

```
    from (
        ...
        ) mb1
    join (
        ...
        ) mb2
    on mb1.full_name = mb2.full_name
    ) mb3
```

Step 4: calculate the average point per game of each Duke player with below command:

```
select mb3.full_name,
round(mb3.player_total_points/mb3.player_game_played,1) as player_avg_point
from (
    ...
) mb3
order by player_avg_point desc;
```

Summary

The partial result, the top 15 PPG Duke players, of the SQL query is displayed as the below screenshots. We can find out that many of them turn out entering NBA, and some even become all-stars.

Row	full_name	player_avg_point
1	Marvin Bagley III	21.0
2	Jabari Parker	19.1
3	Brandon Ingram	17.3
4	Jahlil Okafor	17.3
5	Jayson Tatum	16.8
6	Rodney Hood	16.1
7	Luke Kennard	15.7

8	Gary Trent Jr.	14.5
9	Grayson Allen	14.1
10	Quinn Cook	13.6
11	Wendell Carter Jr.	13.5
12	Justise Winslow	12.6
13	Tyus Jones	11.8
14	Frank Jackson	10.9
15	Trevon Duval	10.3

Appendix

Complete SQL command:

```
select mb3.full_name,
round(mb3.player_total_points/mb3.player_game_played,1) as player_avg_point
from (
    select mb1.full_name, mb1.player_game_played, mb2.player_total_points
    from (
        select mb.full_name, count(mb.game_id) as player_game_played
        from `bigquery-public-data.ncaa_basketball.mbb_players_games_sr` mb
        where mb.played = true and mb.team_market = "Duke"
        group by mb.full_name
        order by player_game_played desc
        ) mb1
    join (
        select mb.full_name, sum(mb.points) as player_total_points
        from `bigquery-public-data.ncaa_basketball.mbb_players_games_sr` mb
        where mb.played = true and mb.team_market = "Duke"
        group by mb.full_name
        order by player_total_points desc
        ) mb2
```

```
    on mb1.full_name = mb2.full_name
) mb3
order by player_avg_point desc;
```