

**Yog Chaudhary**

**11727095**

**ADTA 5240 Week (Harvesting, Storing, And Retrieving Data)**

**Professor: Dr. Zeynep Orhan**

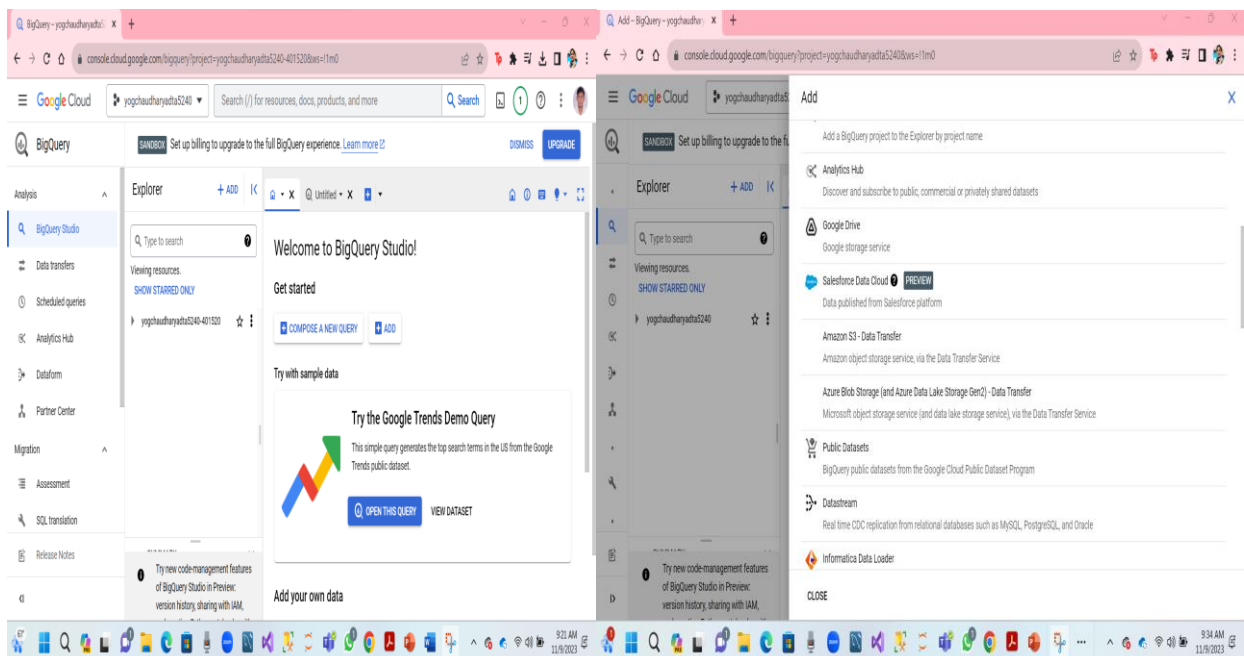
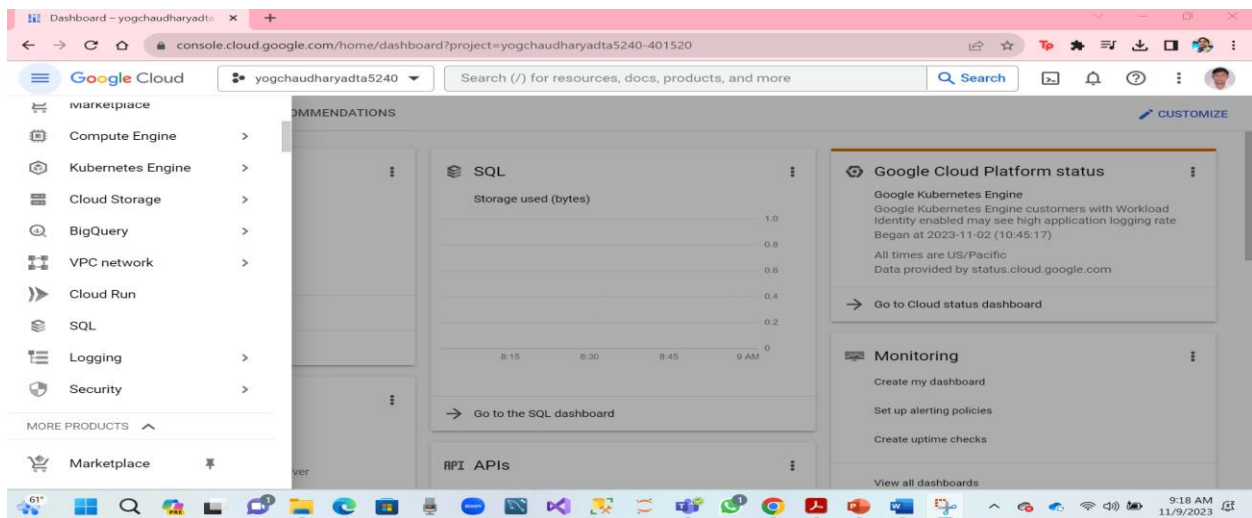
**University Of North Texas**

**Nov 09, 2023**

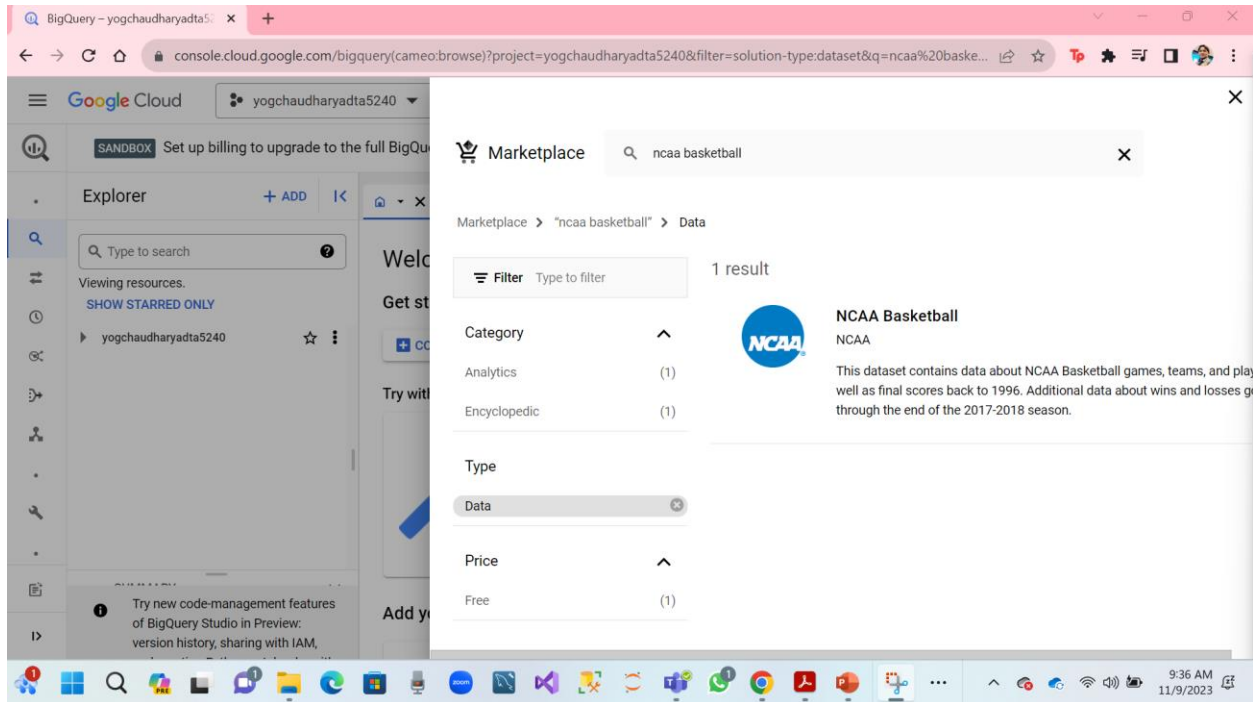
### ❖ Big query Queries:

Using the public Big Query dataset. The NCCA basketball data. This dataset contains data about the NCCA basketball games, items, and players.

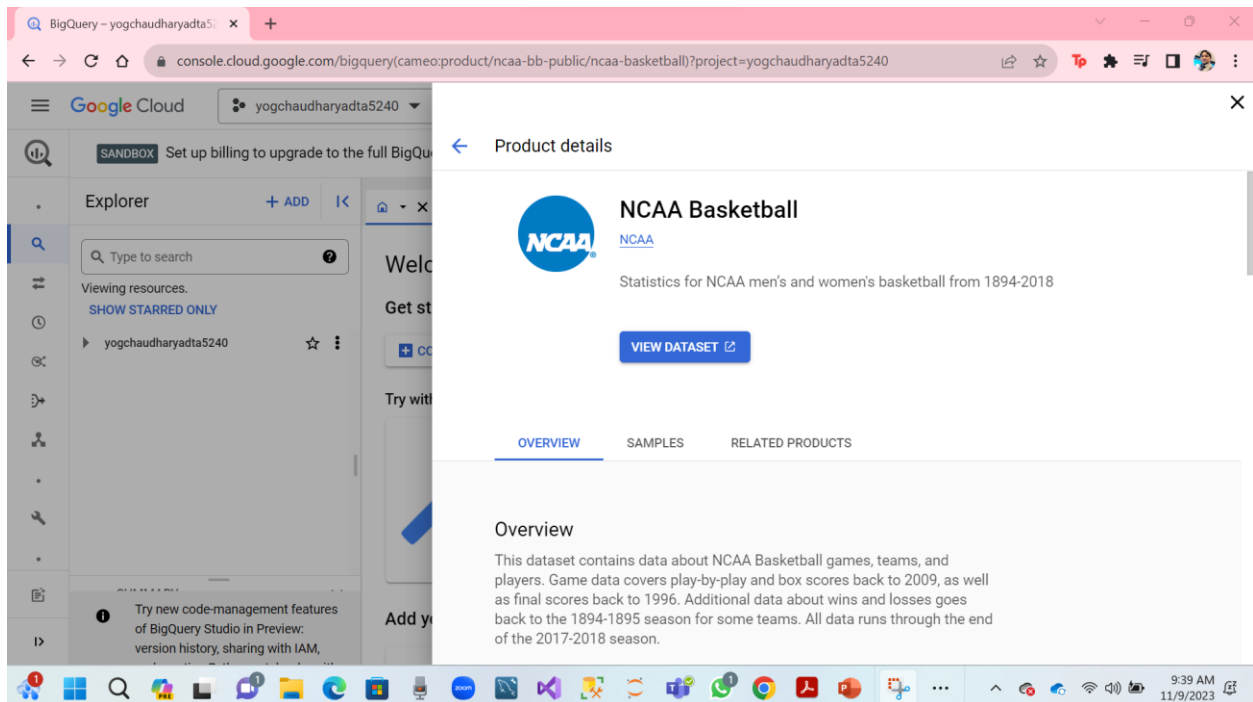
- First, I opened my GCP account and clicked on the navigation panel, then selected bid query.
- After clicking on that I was taking on this below page



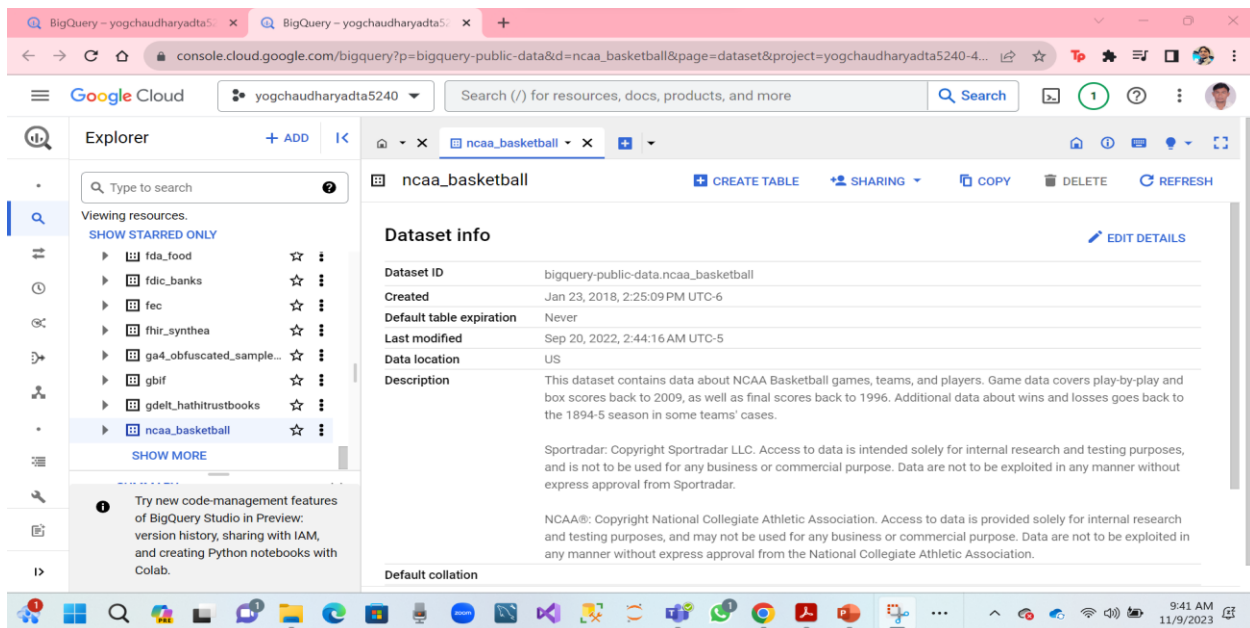
- I clicked on add data. Which has taken in this page. Then I selected a public dataset.
- Below is the screenshot of that.



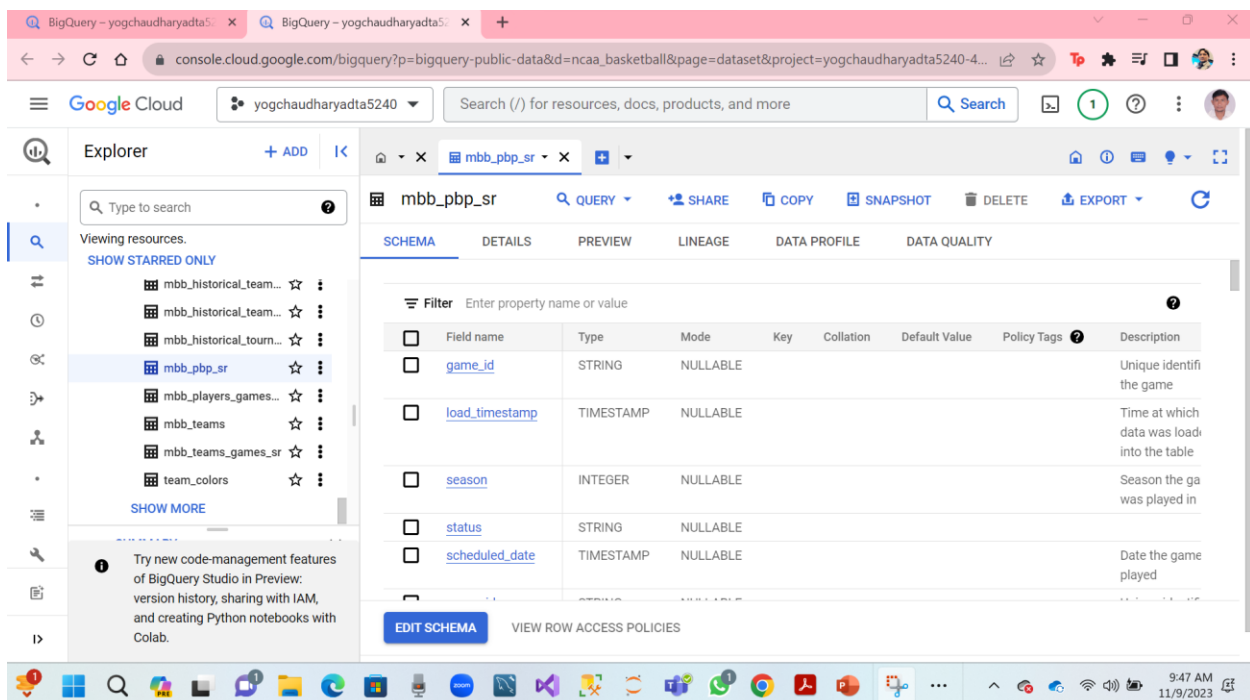
- Then I clicked on the NCAA basketball dataset.



- Here, I clicked on view dataset.



- Then I expanded the NCA basketball dataset.
- After, I click mbb\_pbp\_sr
- Click on open. It showed the schema of the predefined table.
- Here is a screenshot of that.



- After, I clicked preview to see more information about the dataset.

The screenshot shows the Google Cloud BigQuery interface. In the Explorer on the left, the dataset 'mbb\_pbp\_sr' is selected. A context menu is open over it, with the 'Query' option highlighted. The main panel shows the 'Preview' tab for this dataset, displaying a table with the following columns: Row, game\_id, load\_timestamp, season, status, and scheduled\_date. The table contains 9 rows of data for the 2017 season, all with a 'closed' status. The bottom status bar indicates 'Results per page: 50' and '1 - 50 of 4160393'.

- I clicked on 3 dots in mbb\_pbp\_sr and clicked on Query.
- I have used SQL to query the dataset.
- For selecting the columns, I have used this code.
- In this have used SELECT, WHERE, ORDER BY, LIMIT statements.

```
SELECT
game_clock,
points_scored,
team_name,
event_description,
FROM
`bigquery-public-data.ncaa_basketball.mbb_pbp_sr`
WHERE
season = 2014
AND home_name = 'Wildcats'
AND away_name = 'Fighting Irish'
AND points_scored IS NOT NULL
ORDER BY
timestamp DESC
LIMIT
10;
```

- Here is a screenshot of that.

The screenshot shows the Google Cloud BigQuery console. The Explorer panel on the left lists resources under the project 'yogchaudharyadta5240', including 'ncaa\_basketball'. The main panel displays a query titled 'Untitled' with the following SQL code:

```
1 # SELECT FROM `bigquery-public-data.ncaa_basketball.mbb_pbp_sr` LIMIT 1000
2 SELECT
3   game_clock,
4   points_scored,
5   team_name,
6   event_description,
7 FROM
8   `bigquery-public-data.ncaa_basketball.mbb_pbp_sr`
9 WHERE
10  season = 2014
11 AND home_name = 'Wildcats'
12 AND away_name = 'Fighting Irish'
13 AND points_scored IS NOT NULL
14 ORDER BY
```

The query results are displayed in a table with columns: Row, game\_clock, points\_scored, team\_name, and event\_description. The results show the first 10 rows of the query output.

Row	game_clock	points_scored	team_name	event_description
7	4:08	1.0	Wildcats	Karl-Anthony Towns makes fre...
8	4:08	2.0	Wildcats	Karl-Anthony Towns makes tw...
9	4:28	1.0	Fighting Irish	Jerian Grant makes free throw ...
10	5:05	2.0	Wildcats	Aaron Harrison makes two poi...

- Clicked on run to run the code. After clicking on that it shows the result.

The screenshot shows the Google Cloud BigQuery console with the same query as before. The query results are displayed in a table with columns: Row, game\_clock, points\_scored, team\_name, and event\_description. The results show the first 8 rows of the query output.

Row	game_clock	points_scored	team_name	event_description
1	00:06	1.0	Wildcats	Andrew Harrison makes free th...
2	00:06	1.0	Wildcats	Andrew Harrison makes free th...
3	1:12	2.0	Wildcats	Karl-Anthony Towns makes tw...
4	2:34	3.0	Fighting Irish	Jerian Grant makes three point jump shot (Pat Connaughton assists)
5	3:15	3.0	Wildcats	Aaron Harrison makes three point jump shot (Tyler Uilis assists)
6	3:45	1.0	Fighting Irish	Pat Connaughton makes free t...
7	4:08	1.0	Wildcats	Karl-Anthony Towns makes fre...
8	4:08	2.0	Wildcats	Karl-Anthony Towns makes tw...

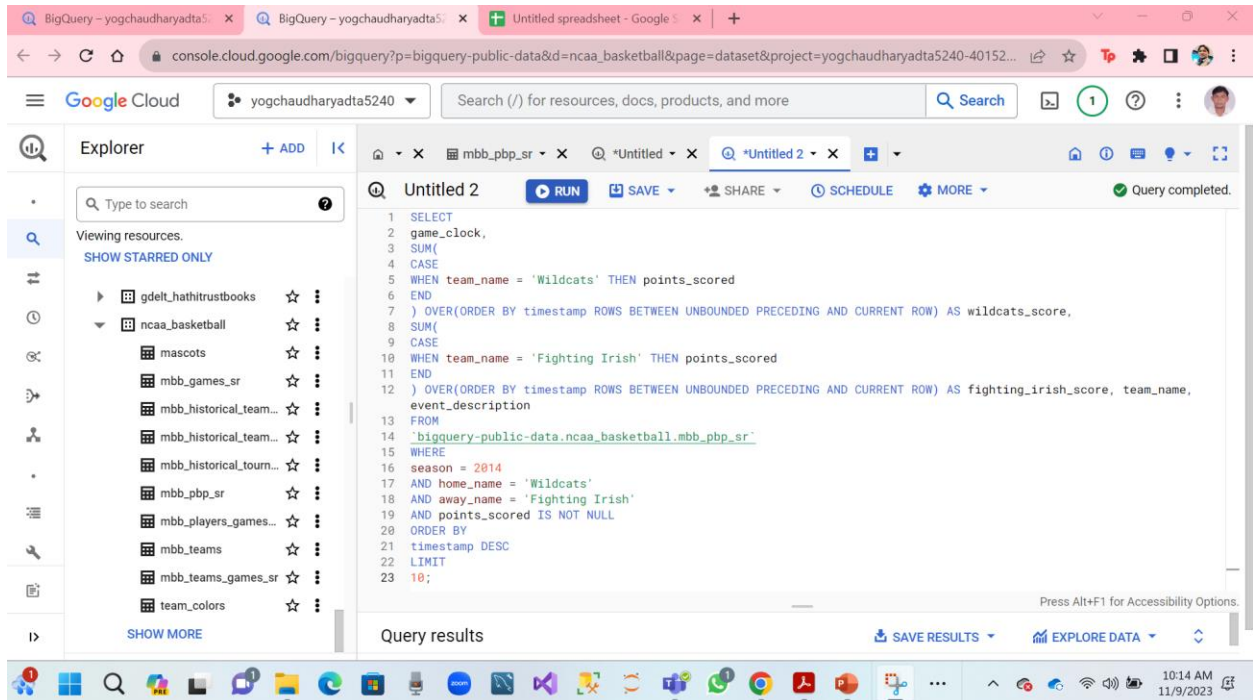
- Now I am changing the SQL code.
- By calculating the cumulative sum of scores for each team in the games.
- For clicked on composer new query sheet
- Run the query by adding two new columns: **wildcats\_score** and **fighting\_irish\_score**



```

SELECT
game_clock,
SUM(
CASE
WHEN team_name = 'Wildcats' THEN points_scored
END
) OVER(ORDER BY timestamp ROWS BETWEEN UNBOUNDED PRECEDING AND CURRENT ROW) AS
wildcats_score,
SUM(
CASE
WHEN team_name = 'Fighting Irish' THEN points_scored
END
) OVER(ORDER BY timestamp ROWS BETWEEN UNBOUNDED PRECEDING AND CURRENT ROW) AS
fighting_irish_score, team_name, event_description
FROM
`bigquery-public-data.ncaa_basketball.mbb_pbp_sr`
WHERE
season = 2014
AND home_name = 'Wildcats'
AND away_name = 'Fighting Irish'
AND points_scored IS NOT NULL
ORDER BY
timestamp DESC
LIMIT
10;

```



- Click on run to run the code. After clicking on that it shows results.

The screenshot displays the Google Cloud BigQuery console interface. The top navigation bar shows the Google Cloud logo and the project name 'yogchaudharyadta5240'. The left sidebar contains an 'Explorer' panel with a search bar and a list of resources under 'ncaa\_basketball', including 'mascots', 'mbb\_games\_sr', 'mbb\_historical\_team...', 'mbb\_historical\_tourn...', 'mbb\_pbp\_sr', 'mbb\_players\_games...', 'mbb\_teams', 'mbb\_teams\_games\_sr', and 'team\_colors'. The main area shows 'Query results' for a query titled '\*Untitled 2'. The results are displayed in a table with columns: 'Row', 'game\_clock', 'wildcats\_score', 'fighting\_irish\_score', 'team\_name', and 'event\_description'. The table contains 9 rows of data, showing game progress and scores for Wildcats and Fighting Irish. A 'Load more' button is visible at the bottom of the table. The bottom of the screen shows a Windows taskbar with various application icons and the system clock indicating 10:16 AM on 11/9/2023.

Row	game_clock	wildcats_score	fighting_irish_score	team_name	event_description
1	00:06	68.0	66.0	Wildcats	Andrew Harrison makes free th...
2	00:06	67.0	66.0	Wildcats	Andrew Harrison makes free th...
3	1:12	66.0	66.0	Wildcats	Karl-Anthony Towns makes tw...
4	2:34	64.0	66.0	Fighting Irish	Jerian Grant makes three point jump shot (Pat Connaughton assists)
5	3:15	64.0	63.0	Wildcats	Aaron Harrison makes three point jump shot (Tyler Ullis assists)
6	3:45	61.0	63.0	Fighting Irish	Pat Connaughton makes free t...
7	4:08	61.0	62.0	Wildcats	Karl-Anthony Towns makes fre...
8	4:08	60.0	62.0	Wildcats	Karl-Anthony Towns makes tw...
9	4:28	58.0	62.0	Fighting Irish	Jerian Grant makes free throw ...

- Hence, Finally, we can see the score throughout the game by using this code.