

# Live Coding - Module 7

## Rachel Holman

```
In [1]: import numpy as np
import pandas as pd
import sqlite3
```

```
In [2]: nbadb = sqlite3.connect('nba.db')
```

```
In [3]: myquery = '''
SELECT *
FROM games
'''

pd.read_sql_query(myquery, nbadb)
```

```
Out[3]:
```

|      | game_id      | game_date  | OT  | season |
|------|--------------|------------|-----|--------|
| 0    | 202202170BRK | 2022-02-17 | 0   | 2022   |
| 1    | 202202170CHO | 2022-02-17 | 2   | 2022   |
| 2    | 202202170LAC | 2022-02-17 | 0   | 2022   |
| 3    | 202202170MIL | 2022-02-17 | 0   | 2022   |
| 4    | 202202170NOP | 2022-02-17 | 0   | 2022   |
| ...  | ...          | ...        | ... | ...    |
| 3192 | 202001080GSW | 2020-01-08 | 0   | 2020   |
| 3193 | 202008020HOU | 2020-08-02 | 0   | 2020   |
| 3194 | 201911060HOU | 2019-11-06 | 0   | 2020   |
| 3195 | 201912250GSW | 2019-12-25 | 0   | 2020   |
| 3196 | 202002200GSW | 2020-02-20 | 0   | 2020   |

3197 rows x 4 columns

Win/loss records in the 2021-2022 season

```
In [4]: myquery= '''
SELECT *
FROM team_game tg
INNER JOIN games g
    ON tg.game_id=g.game_id
'''

pd.read_sql_query(myquery, nbadb)
```

Out [4]:

|             | H_A | game_id      | Team_Abbrev | Team_Score | Team_pace | Team_efg_pct | Team_tov_ |
|-------------|-----|--------------|-------------|------------|-----------|--------------|-----------|
| <b>0</b>    | A   | 202202170BRK | WAS         | 117        | 94.5      | 0.627        |           |
| <b>1</b>    | H   | 202202170BRK | BRK         | 103        | 94.5      | 0.483        |           |
| <b>2</b>    | A   | 202202170CHO | MIA         | 111        | 88.8      | 0.471        |           |
| <b>3</b>    | H   | 202202170CHO | CHO         | 107        | 88.8      | 0.453        |           |
| <b>4</b>    | A   | 202202170LAC | HOU         | 111        | 103.7     | 0.533        |           |
| ...         | ... | ...          | ...         | ...        | ...       | ...          |           |
| <b>6389</b> | H   | 202002270GSW | GSW         | 86         | 104.8     | 0.481        |           |
| <b>6390</b> | A   | 202002290PHO | GSW         | 115        | 98.6      | 0.523        |           |
| <b>6391</b> | H   | 202003010GSW | GSW         | 110        | 100.2     | 0.522        |           |
| <b>6392</b> | A   | 202003030DEN | GSW         | 116        | 94.4      | 0.622        |           |
| <b>6393</b> | H   | 202003070GSW | GSW         | 118        | 90.9      | 0.606        |           |

6394 rows x 15 columns

```

In [6]: myquery= '''
SELECT *
FROM team_game tg
INNER JOIN games g
      ON tg.game_id=g.game_id
WHERE g.season=2022
'''
pd.read_sql_query(myquery, nbadb)

```

Out [6]:

|      | H_A | game_id      | Team_Abbrev | Team_Score | Team_pace | Team_efg_pct | Team_tov_p |
|------|-----|--------------|-------------|------------|-----------|--------------|------------|
| 0    | A   | 202202170BRK | WAS         | 117        | 94.5      | 0.627        | 13         |
| 1    | H   | 202202170BRK | BRK         | 103        | 94.5      | 0.483        | 13         |
| 2    | A   | 202202170CHO | MIA         | 111        | 88.8      | 0.471        | 13         |
| 3    | H   | 202202170CHO | CHO         | 107        | 88.8      | 0.453        | 13         |
| 4    | A   | 202202170LAC | HOU         | 111        | 103.7     | 0.533        | 13         |
| ...  | ... | ...          | ...         | ...        | ...       | ...          | ...        |
| 1761 | H   | 202201130NOP | NOP         | 113        | 97.0      | 0.515        | 13         |
| 1762 | A   | 202201250PHI | NOP         | 107        | 89.7      | 0.536        | 13         |
| 1763 | H   | 202201250PHI | PHI         | 117        | 89.7      | 0.524        | 13         |
| 1764 | A   | 202201210PHI | LAC         | 102        | 91.5      | 0.523        | 13         |
| 1765 | H   | 202201210PHI | PHI         | 101        | 91.5      | 0.473        | 13         |

1766 rows x 15 columns

In [8]:

```

myquery= '''
SELECT tg.game_id, tg.Team_Abbrev, tg.Opponent_Abbrev, tg.Team_Score
FROM team_game tg
INNER JOIN games g
      ON tg.game_id=g.game_id
WHERE g.season=2022
'''
pd.read_sql_query(myquery, nbadb)

```

Out[8]:

|      | game_id      | Team_Abbrev | Opponent_Abbrev | Team_Score |
|------|--------------|-------------|-----------------|------------|
| 0    | 202202170BRK | BRK         | WAS             | 103        |
| 1    | 202202170BRK | WAS         | BRK             | 117        |
| 2    | 202202170CHO | CHO         | MIA             | 107        |
| 3    | 202202170CHO | MIA         | CHO             | 111        |
| 4    | 202202170LAC | HOU         | LAC             | 111        |
| ...  | ...          | ...         | ...             | ...        |
| 1761 | 202201130NOP | NOP         | LAC             | 113        |
| 1762 | 202201250PHI | NOP         | PHI             | 107        |
| 1763 | 202201250PHI | PHI         | NOP             | 117        |
| 1764 | 202201210PHI | LAC         | PHI             | 102        |
| 1765 | 202201210PHI | PHI         | LAC             | 101        |

1766 rows x 4 columns

In [12]:

```

myquery= '''
SELECT tg.game_id, tg.Team_Abbrev, tg.Opponent_Abbrev, tg.Team_Score,
       sg.Team_Score AS Opponent_Score
FROM team_game tg
INNER JOIN games g
      ON tg.game_id=g.game_id
INNER JOIN team_game sg
      ON tg.game_id = sg.game_id AND tg.Opponent_Abbrev = sg.Team_Abbrev
WHERE g.season=2022
'''
pd.read_sql_query(myquery, nbadb)

```

Out[12]:

|      | game_id      | Team_Abbrev | Opponent_Abbrev | Team_Score | Opponent_Score |
|------|--------------|-------------|-----------------|------------|----------------|
| 0    | 202202170BRK | BRK         | WAS             | 103        | 117            |
| 1    | 202202170BRK | WAS         | BRK             | 117        | 103            |
| 2    | 202202170CHO | CHO         | MIA             | 107        | 111            |
| 3    | 202202170CHO | MIA         | CHO             | 111        | 107            |
| 4    | 202202170LAC | HOU         | LAC             | 111        | 142            |
| ...  | ...          | ...         | ...             | ...        | ...            |
| 1761 | 202201130NOP | NOP         | LAC             | 113        | 89             |
| 1762 | 202201250PHI | NOP         | PHI             | 107        | 117            |
| 1763 | 202201250PHI | PHI         | NOP             | 117        | 107            |
| 1764 | 202201210PHI | LAC         | PHI             | 102        | 101            |
| 1765 | 202201210PHI | PHI         | LAC             | 101        | 102            |

1766 rows x 5 columns

```
In [13]: myquery= '''
SELECT tg.game_id, tg.Team_Abbrev, tg.Opponent_Abbrev, tg.Team_Score,
       sg.Team_Score AS Opponent_Score, (tg.Team_Score > sg.Team_Score) AS win
FROM team_game tg
INNER JOIN games g
     ON tg.game_id=g.game_id
INNER JOIN team_game sg
     ON tg.game_id = sg.game_id AND tg.Opponent_Abbrev = sg.Team_Abbrev
WHERE g.season=2022
'''
pd.read_sql_query(myquery, nbadb)
```

```
Out[13]:
```

|      | game_id      | Team_Abbrev | Opponent_Abbrev | Team_Score | Opponent_Score | win |
|------|--------------|-------------|-----------------|------------|----------------|-----|
| 0    | 202202170BRK | BRK         | WAS             | 103        | 117            | 0   |
| 1    | 202202170BRK | WAS         | BRK             | 117        | 103            | 1   |
| 2    | 202202170CHO | CHO         | MIA             | 107        | 111            | 0   |
| 3    | 202202170CHO | MIA         | CHO             | 111        | 107            | 1   |
| 4    | 202202170LAC | HOU         | LAC             | 111        | 142            | 0   |
| ...  | ...          | ...         | ...             | ...        | ...            | ... |
| 1761 | 202201130NOP | NOP         | LAC             | 113        | 89             | 1   |
| 1762 | 202201250PHI | NOP         | PHI             | 107        | 117            | 0   |
| 1763 | 202201250PHI | PHI         | NOP             | 117        | 107            | 1   |
| 1764 | 202201210PHI | LAC         | PHI             | 102        | 101            | 1   |
| 1765 | 202201210PHI | PHI         | LAC             | 101        | 102            | 0   |

1766 rows x 6 columns

```
In [17]: myquery= '''
SELECT tg.Team_Abbrev,
       SUM((tg.Team_Score > sg.Team_Score)) AS win,
       SUM((tg.Team_Score < sg.Team_Score)) AS loss
FROM team_game tg
INNER JOIN games g
     ON tg.game_id=g.game_id
INNER JOIN team_game sg
     ON tg.game_id = sg.game_id AND tg.Opponent_Abbrev = sg.Team_Abbrev
WHERE g.season=2022
GROUP BY tg.Team_Abbrev
ORDER BY win DESC
'''
pd.read_sql_query(myquery, nbadb)
```

Out[17]:

|    | Team_Abbrev | win | loss |
|----|-------------|-----|------|
| 0  | PHO         | 48  | 10   |
| 1  | GSW         | 42  | 17   |
| 2  | MEM         | 41  | 19   |
| 3  | MIA         | 38  | 21   |
| 4  | CHI         | 38  | 21   |
| 5  | UTA         | 36  | 22   |
| 6  | MIL         | 36  | 24   |
| 7  | PHI         | 35  | 23   |
| 8  | DAL         | 35  | 24   |
| 9  | CLE         | 35  | 23   |
| 10 | BOS         | 34  | 26   |
| 11 | DEN         | 33  | 25   |
| 12 | TOR         | 32  | 25   |
| 13 | MIN         | 31  | 28   |
| 14 | BRK         | 31  | 28   |
| 15 | LAC         | 30  | 31   |
| 16 | CHO         | 29  | 31   |
| 17 | ATL         | 28  | 30   |
| 18 | WAS         | 27  | 31   |
| 19 | LAL         | 27  | 31   |
| 20 | POR         | 25  | 34   |
| 21 | NYK         | 25  | 34   |
| 22 | SAS         | 23  | 36   |
| 23 | NOP         | 23  | 36   |
| 24 | SAC         | 22  | 38   |
| 25 | IND         | 20  | 40   |
| 26 | OKC         | 18  | 40   |
| 27 | HOU         | 15  | 43   |
| 28 | ORL         | 13  | 47   |
| 29 | DET         | 13  | 45   |

Most average player in the NBA in 2021/2022

In [18]:

```
myquery= '''
SELECT *
FROM player_game pg
INNER JOIN games g
```

```

    ON pg.game_id=g.game_id
WHERE g.season=2022
'''
pd.read_sql_query(myquery, nbadb)

```

Out[18]:

|       | game_id      | Team_Abbrev | player_id | starter | mp    | fg  | fga | fg_pct | fg3 | fg3a | ... |
|-------|--------------|-------------|-----------|---------|-------|-----|-----|--------|-----|------|-----|
| 0     | 202202170BRK | BRK         | aldrila01 | 0       | 28:19 | 5   | 8   | 0.625  | 0   | 0    | ... |
| 1     | 202202170BRK | BRK         | brownbr01 | 1       | 34:57 | 3   | 10  | 0.300  | 0   | 1    | ... |
| 2     | 202202170BRK | BRK         | carteje01 | 0       | 19:28 | 5   | 11  | 0.455  | 3   | 9    | ... |
| 3     | 202202170BRK | BRK         | claxtni01 | 0       | 0:00  | 0   | 0   | 0.000  | 0   | 0    | ... |
| 4     | 202202170BRK | BRK         | curryse01 | 1       | 31:16 | 4   | 11  | 0.364  | 3   | 6    | ... |
| ...   | ...          | ...         | ...       | ...     | ...   | ... | ... | ...    | ... | ...  | ... |
| 22619 | 202201210PHI | PHI         | korkmfu01 | 1       | 38:53 | 2   | 7   | 0.286  | 1   | 6    | ... |
| 22620 | 202201210PHI | PHI         | maxeyty01 | 1       | 45:34 | 7   | 18  | 0.389  | 3   | 6    | ... |
| 22621 | 202201210PHI | PHI         | niangge01 | 0       | 30:13 | 3   | 11  | 0.273  | 1   | 7    | ... |
| 22622 | 202201210PHI | PHI         | powelmy01 | 0       | 0:00  | 0   | 0   | 0.000  | 0   | 0    | ... |
| 22623 | 202201210PHI | PHI         | sprinja01 | 0       | 0:00  | 0   | 0   | 0.000  | 0   | 0    | ... |

22624 rows x 64 columns

In [23]:

```

myquery= '''
SELECT pg.player_id,
       SUM(pg.fg) AS fg,
       SUM(pg.fga) AS fga,
       SUM(pg.fg3) AS fg3,
       SUM(pg.fg3a) AS fg3a,
       SUM(pg.ft) AS ft,
       SUM(pg.fta) AS fta,
       SUM(pg.trb) AS trb,
       SUM(pg.ast) AS ast,
       SUM(pg.stl) AS stl,
       SUM(pg.blk) AS blk,
       SUM(pg.tov) AS tov,
       SUM(pg.pf) AS pf,
       SUM(pg.pts) AS pts,
       SUM(pg.minutes) AS minutes
FROM player_game pg
INNER JOIN games g
    ON pg.game_id=g.game_id
WHERE g.season=2022
GROUP BY pg.player_id
       HAVING SUM(pg.minutes) >300
'''
pd.read_sql_query(myquery, nbadb)

```

Out [23]:

|     | player_id | fg  | fga  | fg3 | fg3a | ft  | fta | trb | ast | stl | blk | tov | pf  | pts  | minu      |
|-----|-----------|-----|------|-----|------|-----|-----|-----|-----|-----|-----|-----|-----|------|-----------|
| 0   | achiupr01 | 148 | 350  | 18  | 59   | 44  | 76  | 331 | 55  | 24  | 28  | 47  | 96  | 358  | 1107.0000 |
| 1   | adamsst01 | 154 | 282  | 0   | 0    | 85  | 153 | 537 | 180 | 49  | 38  | 89  | 106 | 393  | 1464.0160 |
| 2   | adebaba01 | 241 | 460  | 0   | 4    | 151 | 201 | 348 | 119 | 51  | 24  | 99  | 106 | 633  | 1134.1660 |
| 3   | aldrila01 | 227 | 406  | 12  | 39   | 71  | 83  | 224 | 33  | 13  | 42  | 38  | 70  | 537  | 894.1160  |
| 4   | alexani01 | 237 | 633  | 95  | 306  | 70  | 97  | 165 | 139 | 41  | 19  | 85  | 88  | 639  | 1322.1160 |
| ... | ...       | ... | ...  | ... | ...  | ... | ... | ... | ... | ... | ... | ... | ... | ...  | ...       |
| 383 | youngth01 | 81  | 142  | 0   | 7    | 10  | 23  | 102 | 61  | 25  | 8   | 33  | 41  | 172  | 403.3500  |
| 384 | youngtr01 | 495 | 1085 | 158 | 412  | 327 | 364 | 206 | 493 | 53  | 5   | 217 | 85  | 1475 | 1823.6000 |
| 385 | yurtsom01 | 112 | 208  | 1   | 7    | 34  | 54  | 250 | 44  | 15  | 18  | 36  | 71  | 259  | 593.0330  |
| 386 | zelleco01 | 51  | 90   | 0   | 4    | 38  | 49  | 125 | 22  | 8   | 6   | 19  | 56  | 140  | 354.5330  |
| 387 | zubaciv01 | 206 | 317  | 0   | 0    | 122 | 171 | 442 | 77  | 25  | 53  | 79  | 136 | 534  | 1332.5330 |

388 rows × 15 columns

In [27]:

```

myquery= '''
SELECT AVG(fg/minutes) AS meanfg,
       AVG(fga/minutes) AS meanfga,
       AVG(fg3/minutes) AS meanfg3,
       AVG(fg3a/minutes) AS meanfg3a,
       AVG(ft/minutes) AS meanft,
       AVG(fta/minutes) AS meanfta,
       AVG(trb/minutes) AS meantrb,
       AVG(ast/minutes) AS meanast,
       AVG(stl/minutes) AS meanstl,
       AVG(blk/minutes) AS meanblk,
       AVG(tov/minutes) AS meantov,
       AVG(pf/minutes) AS meanpf,
       AVG(pts/minutes) AS meanpts
FROM (SELECT pg.player_id,
            SUM(pg.fg) AS fg,
            SUM(pg.fga) AS fga,
            SUM(pg.fg3) AS fg3,
            SUM(pg.fg3a) AS fg3a,
            SUM(pg.ft) AS ft,
            SUM(pg.fta) AS fta,
            SUM(pg.trb) AS trb,
            SUM(pg.ast) AS ast,
            SUM(pg.stl) AS stl,
            SUM(pg.blk) AS blk,
            SUM(pg.tov) AS tov,
            SUM(pg.pf) AS pf,
            SUM(pg.pts) AS pts,
            SUM(pg.minutes) AS minutes
FROM player_game pg
INNER JOIN games g
    ON pg.game_id=g.game_id
WHERE g.season=2022
GROUP BY pg.player_id
    HAVING SUM(pg.minutes) >300)

```



```
'''
pd.read_sql_query(myquery, nbadb)
```

Out[27]:

|   | meanfg   | meanfga  | meanfg3  | meanfg3a | meanft   | meanfta  | meantrb  | meanast  | meanstl  |
|---|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 0 | 0.159503 | 0.350723 | 0.048402 | 0.140306 | 0.063702 | 0.083288 | 0.187059 | 0.094813 | 0.031873 |

In [28]:

```
myquery= '''
SELECT *
FROM (SELECT pg.player_id,
      SUM(pg.fg) AS fg,
      SUM(pg.fga) AS fga,
      SUM(pg.fg3) AS fg3,
      SUM(pg.fg3a) AS fg3a,
      SUM(pg.ft) AS ft,
      SUM(pg.fta) AS fta,
      SUM(pg.trb) AS trb,
      SUM(pg.ast) AS ast,
      SUM(pg.stl) AS stl,
      SUM(pg.blk) AS blk,
      SUM(pg.tov) AS tov,
      SUM(pg.pf) AS pf,
      SUM(pg.pts) AS pts,
      SUM(pg.minutes) AS minutes
FROM player_game pg
INNER JOIN games g
      ON pg.game_id=g.game_id
WHERE g.season=2022
GROUP BY pg.player_id
      HAVING SUM(pg.minutes) >300)
INNER JOIN (SELECT AVG(fg/minutes) AS meanfg,
      AVG(fga/minutes) AS meanfga,
      AVG(fg3/minutes) AS meanfg3,
      AVG(fg3a/minutes) AS meanfg3a,
      AVG(ft/minutes) AS meanft,
      AVG(fta/minutes) AS meanfta,
      AVG(trb/minutes) AS meantrb,
      AVG(ast/minutes) AS meanast,
      AVG(stl/minutes) AS meanstl,
      AVG(blk/minutes) AS meanblk,
      AVG(tov/minutes) AS meantov,
      AVG(pf/minutes) AS meanpf,
      AVG(pts/minutes) AS meanpts
FROM (SELECT pg.player_id,
      SUM(pg.fg) AS fg,
      SUM(pg.fga) AS fga,
      SUM(pg.fg3) AS fg3,
      SUM(pg.fg3a) AS fg3a,
      SUM(pg.ft) AS ft,
      SUM(pg.fta) AS fta,
      SUM(pg.trb) AS trb,
      SUM(pg.ast) AS ast,
      SUM(pg.stl) AS stl,
      SUM(pg.blk) AS blk,
      SUM(pg.tov) AS tov,
      SUM(pg.pf) AS pf,
      SUM(pg.pts) AS pts,
      SUM(pg.minutes) AS minutes
FROM player_game pg
```

```
INNER JOIN games g
  ON pg.game_id=g.game_id
WHERE g.season=2022
GROUP BY pg.player_id
  HAVING SUM(pg.minutes) >300))
...
pd.read_sql_query(myquery, nbadb)
```

Out [28]:

|     | player_id | fg  | fga  | fg3 | fg3a | ft  | fta | trb | ast | stl | ... | meanfg3a | meanft   | mea   |
|-----|-----------|-----|------|-----|------|-----|-----|-----|-----|-----|-----|----------|----------|-------|
| 0   | achiupr01 | 148 | 350  | 18  | 59   | 44  | 76  | 331 | 55  | 24  | ... | 0.140306 | 0.063702 | 0.08: |
| 1   | adamsst01 | 154 | 282  | 0   | 0    | 85  | 153 | 537 | 180 | 49  | ... | 0.140306 | 0.063702 | 0.08: |
| 2   | adebaba01 | 241 | 460  | 0   | 4    | 151 | 201 | 348 | 119 | 51  | ... | 0.140306 | 0.063702 | 0.08: |
| 3   | aldrila01 | 227 | 406  | 12  | 39   | 71  | 83  | 224 | 33  | 13  | ... | 0.140306 | 0.063702 | 0.08: |
| 4   | alexani01 | 237 | 633  | 95  | 306  | 70  | 97  | 165 | 139 | 41  | ... | 0.140306 | 0.063702 | 0.08: |
| ... | ...       | ... | ...  | ... | ...  | ... | ... | ... | ... | ... | ... | ...      | ...      | ...   |
| 383 | youngth01 | 81  | 142  | 0   | 7    | 10  | 23  | 102 | 61  | 25  | ... | 0.140306 | 0.063702 | 0.08: |
| 384 | youngtr01 | 495 | 1085 | 158 | 412  | 327 | 364 | 206 | 493 | 53  | ... | 0.140306 | 0.063702 | 0.08: |
| 385 | yurtsom01 | 112 | 208  | 1   | 7    | 34  | 54  | 250 | 44  | 15  | ... | 0.140306 | 0.063702 | 0.08: |
| 386 | zelleco01 | 51  | 90   | 0   | 4    | 38  | 49  | 125 | 22  | 8   | ... | 0.140306 | 0.063702 | 0.08: |
| 387 | zubaciv01 | 206 | 317  | 0   | 0    | 122 | 171 | 442 | 77  | 25  | ... | 0.140306 | 0.063702 | 0.08: |

388 rows x 28 columns

In [ ]: