CSc 496, Homework #6: "Nomentum".

Due date: Thursday, November 7th, 2024. **No late assignments will be accepted.**

In this assignment you will analyze momentum in football.

Football fans have a belief in momentum. Hardly a game goes by without the announcers commenting on teams having momentum, teams taking back the momentum, etc. Does momentum exist? If it does, it ought to be identifiable in a dispassionate way. (It is not dispassionate to look at game play-by-play after the fact and state that the momentum did or did not change at a certain point in a game based on knowing what subsequently happened.)

Over 10 years ago, Bill Barnwell tackled this problem in an article for Grantland (which was a relatively short-lived ESPN spinoff site named for the early 20th century sportswriter Grantland Rice). The article was entitled "Nomentum" because Barnwell did not believe momentum existed. You can find the article here: `https://grantland.com/features/bill-barnwell-theory-momentum-football/`.

Your first job is to read the article. Note that the article studies (supposed) momentum due to change of possessions within NFL games, momentum due to being "hot" over over several weeks in the NFL, and momentum going into overtime games in the NHL based on one team scoring late to tie the game. You will only be studying the first of these scenarios.

You are to create limited play-by-play data (one CSV file per game) for seasons 2014–2023. This data should include the play id, id of first play of drive, time remaining in the half, yard line, down, distance, score differential, how the drive started, and what the drive result was. Note that these fields are, respectively, `play_id`, `drive_play_id_started`, `half_seconds_remaining`, `yardline_100` (which means distance from the goal line for the offense), `down`, `ydstogo`, `score_differential`, `drive_start_transition`, and `fixed_drive_result`. The CSV files should be placed into a subdirectory (from where the program is invoked) called `NFLCSV`; you need to create this directory if it does not yet exist.

The "Nomentum" program takes one command line parameter, which is a directory that all the play-by-play files (one per game) will reside.

With the play-by-play data in hand, you are to write your momentum analyzer. This program takes several command-line parameters, in this order:

- the directory that contains the play-by-play files you created in the previous step

- a minimum time remaining in either half for a drive to be considered. This is in seconds; for example, 300 means that five minutes are left in the half.

- a yard line that a team must be within for a drive to be considered. This is given from 0 to 100, where the number means how many yards from the end zone for the team with possession. For example, 95 means to consider drives in which a team starts with the ball on its own five yard line or worse.

- a maximum score differential for a drive to be considered

- a way in which a team took over possession: 0 for turnover (fumble or interception), 1 for punt, and 2 for downs (failed fourth down). A drive is only considered if it the event indicated is how the team took over possession.

The program iterates through all the play-by-play files and, whenever a drive starts such that the four constraints above are satisfied (the minimum time, yard line, maximum score differential, and how the team took over), this drive is considered. You are to keep track of all possible end results of each considered drive and then output the percentage of time that each occurred, in alphabetical order of the result. An example output is shown below.

```
End of half: 1 (0.86%)
Field goal: 19 (16.38%)
Missed field goal: 2 (1.72%)
Opp touchdown: 4 (3.45%)
Punt: 57 (49.14%)
Safety: 6 (5.17%)
Touchdown: 12 (10.34%)
Turnover: 12 (10.34%)
Turnover on downs: 3 (2.59%)
Total: 116
```

Following are some helpful tips:

- A drive start can be identified by lines where the play id and the drive id are equal. This does not work for drives started after a kickoff, but we are not considering such drives.

- The line that you identify as the drive start has the drive result; so you do not need to search for the last play of a drive.

Submit your python files on lectura using the `turnin` command; for this program, use the assignment name `csc496-f24-hw6`. Your programs should be named `createPBP.py` and `nomentum.py`.