Transaction Processing Assignment – Cursor Processing

Cursors are often used as a way to limit processing when you need to update a very large number of rows in a database. In large commercial database, it can often take hours to update hundreds of millions of rows. Data quality can often cause these processes to abend and need to be rerun. Cursors can also be useful in these instances since they can be used to identify and skip rows that have already been processed. The Baseball database does not lend itself to this type of processing, so while it may appear to be easier to simply update all rows instead of using a Cursor, a Cursor is being used as an example for this type of processing. For this assignment, you need to write a script that does the following:

1. Add 2 columns to the PEOPLE table. The columns should be UCID_Career_EqA and UCID_Date_Last_Update. As always, the UCID should be replaced with your actual UCID.
2. Creates an update cursor that contains the playerid and the sum of the player's Equivalent Average calculated using the BATTING . **Equivalent Average** (**EqA**) is a baseball metric invented by Clay Davenport and intended to express the production of hitters in a context independent of park and league effects.[1] It represents a hitter's productivity using the same scale as batting average. Thus, a hitter with an EqA over .300 is a very good hitter, while a hitter with an EqA of .220 or below is poor. An EqA of .260 is defined as league average.The formula for the Equivalent Average is

$$EqA = \frac{H + TB + 1.5 \cdot (BB + HBP) + SB + SH + SF}{AB + BB + HBP + SH + SF + CS + \frac{SB}{3}}$$

Another version of the formula is

```
Equivalent Average = (Hits + Total Bases + (1.5 x (Walks + Hit by Pitch)) + Stolen
Bases + Sacrifice Hits + Sacrifice Flies) ÷ (At Bats + Walks + Hit by Pitch +
Sacrifice Hits + Sacrifice Flies + Caught Stealing + (Stolen Bases ÷ 3))
```

```
Use the Total Bases Touched formula in question #4 of Assignment 4 for total
bases.
```

```
You can find all these columns in the Batting table. The database documentation in
Module 1 can help identify the columns to use.
```

3. Using the cursor created in #2, write a script that:
    a. Updates the new columns in the PEOPLE table when UCID_Date_Last_Update is not equal to the current date
    b. updates the UCID_Career_EqA with the column from the cursor and set UCID_Date_Last_Update to the current date.
    c. Selects the system variable @@Cursor_Rows after you open the CURSOR so you can see how many rows are in the cursor.

    d. Turns off the individual rows update counter by specifying SET NOCOUNT ON at the beginning of the Cursor processing.
    e. Writes the start and stop date and time as well as the # of records updated with the date and time for every 1,000 records updated
    f. Closes and deallocates the cursor as the last step in the script.
4. Include a query to be run after the Cursor processing is complete that selects the playerid, UCID_Career_EqA and UCID_Date_Last_Update from the PEOPLE table
5. Run the script created for steps #2 and #3 a second time and see what happens. Think of why you got these results when you run the script a second time.

For assistance please look in the Canvas site for this week (Week 10).