JakeDineen_5.4.5

August 8, 2018

```python
In [21]: import os
         import numpy as np
         import pandas as pd
         import csv
         cwd = os.getcwd()
         infile = os.listdir(cwd)[4]

         # create new empty list
         playersList = []

         with open(infile, 'rU') as csvfile:
             # the csv file reader returns a list of the csv items on each line
             ALReader = csv.reader(csvfile, dialect='excel', delimiter='\t')

             # from each line, a list of row items, put each element in a dictionary
             #   with a key representing the data
             for line in ALReader:
                 # skip lines without data
                 if line[0] == '' or line[0].startswith('American') or line[0].startswith('Team
                        or line[0].startswith('Source'):
                     continue
                 else:
                     try:
                         # create a dictionary for each player
                         player = {}
                         # add each piece of data under a key representing that data
                         player['team'] = line[0]
                         player['name'] = line[1]
                         player['sal'] = int(line[2].replace(',', ''))
                         player['position'] = line[3]

                         # add this player to the list
                         playersList.append(player)

                     except IndexError:
                         print('Error: ', line)
         csvfile.close()
```

1

```python
        print("Read", len(playersList), "player data")

        # Write a report text file with a title and the average of the salaries
        # First create an output file name
        outfile1 = infile.replace('tsv', '') + 'report.txt'
        # open the file for writing
        fout1 = open(outfile1, 'w')

        # write title at top of file
        fout1.write("American League Baseball players average salary in 2003\n\n")

        # comput the average salary over all players
        total_salary = 0.0
        for player in playersList:
            total_salary += player['sal']
        average_salary = total_salary / len(playersList)

        # write a report line as a string to the file
        fout1.write('Average salary = ${:,.2f}'.format(average_salary))
        fout1.close()

        # Write a file with all the players who made more than $10 million.
        # We write a comma separated file, using the csv writer to quote the player names wit
        # first create an output file name
        outfile2 = infile.replace('tsv', '') + 'million.csv'

        # open the file
        with open(outfile2, 'w', newline='') as csvfileout:
            # create a csv writer for a comman sep file, with quoting as needed
            ALwriter = csv.writer(csvfileout, delimiter=',', quoting=csv.QUOTE_MINIMAL)
            # write the header row as a list of column labels
            ALwriter.writerow(['Player', 'Team', 'Salary', 'Position'])
            for player in playersList:
                # select the players with salary under 310k
                if (player['position'] == 'Pitcher'):
                    # write the player as a list of data items
                    ALwriter.writerow([player['name'], player['team'], player['sal'], player[

        csvfileout.close()

Read 381 player data


C:\Users\jdine\Documents\Anaconda3\envs\tensorflowlatest\lib\site-packages\ipykernel_launcher.p
  # This is added back by InteractiveShellApp.init_path()


In [22]: pd.read_csv('albb.salaries.2003.million.csv')
```

	Player	Team	Salary	Position
0	Acevedo, Juan	New York Yankees	900000	Pitcher
1	Anderson, Jason	New York Yankees	300000	Pitcher
2	Clemens, Roger	New York Yankees	10100000	Pitcher
3	Contreras, Jose	New York Yankees	5500000	Pitcher
4	Hammond, Chris	New York Yankees	2200000	Pitcher
5	Hitchcock, Sterling	New York Yankees	6000000	Pitcher
6	Karsay, Steve	New York Yankees	5000000	Pitcher
7	Liever, Jon	New York Yankees	550000	Pitcher
8	Mussina, Mike	New York Yankees	12000000	Pitcher
9	Osuna, Antonio	New York Yankees	2400000	Pitcher
10	Pettitte, Andy	New York Yankees	11500000	Pitcher
11	Rivera, Mariano	New York Yankees	10500000	Pitcher
12	Weaver, Jeff	New York Yankees	4150000	Pitcher
13	Wells, David	New York Yankees	3250000	Pitcher
14	Appier, Kevin	Anaheim Angels	11500000	Pitcher
15	Callaway, Mickey	Anaheim Angels	302500	Pitcher
16	Donnelly, Brendan	Anaheim Angels	325000	Pitcher
17	Lackey, John	Anaheim Angels	315000	Pitcher
18	Ortiz, Ramon	Anaheim Angels	2266667	Pitcher
19	Percival, Troy	Anaheim Angels	7833333	Pitcher
20	Rodriquez, Francisco	Anaheim Angels	312500	Pitcher
21	Schoeneweis, Scott	Anaheim Angels	1425000	Pitcher
22	Sele, Aaron	Anaheim Angels	8166667	Pitcher
23	Shields, Scot	Anaheim Angels	305000	Pitcher
24	Washburn, Jarrod	Anaheim Angels	3875000	Pitcher
25	Weber, Ben	Anaheim Angels	375000	Pitcher
26	Wise, Matt	Anaheim Angels	302500	Pitcher
27	Burkett, John	Boston Red Sox	5500000	Pitcher
28	Embree, Alan	Boston Red Sox	3000000	Pitcher
29	Fossum, Casey	Boston Red Sox	324500	Pitcher
..	...	...	...	...
150	Moyer, Jamie	Seattle Mariners	6500000	Pitcher
151	Nelson, Jeff	Seattle Mariners	3983333	Pitcher
152	Pineiro, Joel	Seattle Mariners	440000	Pitcher
153	Rhodes, Arthur	Seattle Mariners	3500000	Pitcher
154	Sasaki, Kazuhiro	Seattle Mariners	8000000	Pitcher
155	Bowie, Micah	Oakland Athletics	304000	Pitcher
156	Bradford, Chad	Oakland Athletics	331000	Pitcher
157	Fikac, Jeremy	Oakland Athletics	316000	Pitcher
158	Foulke, Keith	Oakland Athletics	6000000	Pitcher
159	Halama, John	Oakland Athletics	750000	Pitcher
160	Hudson, Tim	Oakland Athletics	2700000	Pitcher
161	Lilly, Ted	Oakland Athletics	335000	Pitcher
162	Mecir, Jim	Oakland Athletics	3216667	Pitcher
163	Mulder, Mark	Oakland Athletics	2650000	Pitcher
164	Neu, Mike	Oakland Athletics	300000	Pitcher
165	Rincon, Ricardo	Oakland Athletics	1700000	Pitcher

```
166          Zito, Barry  Oakland Athletics   1000000  Pitcher
167   Cordero, Francisco      Texas Rangers    900000  Pitcher
168        Fultz, Aaron      Texas Rangers    600000  Pitcher
169    Garcia, Reynaldo      Texas Rangers    300000  Pitcher
170        Lewis, Colby      Texas Rangers    302500  Pitcher
171      Nitkowski, CJ      Texas Rangers    550000  Pitcher
172      Park, Chan Ho      Texas Rangers  13000000  Pitcher
173        Powell, Jay      Texas Rangers   3250000  Pitcher
174       Thomson, John      Texas Rangers   1300000  Pitcher
175      Urbina, Ugueth      Texas Rangers   4500000  Pitcher
176      Valdes, Ismael      Texas Rangers   2500000  Pitcher
177    Van Poppel, Todd      Texas Rangers   2500000  Pitcher
178        Yan, Esteban      Texas Rangers   1500000  Pitcher
179     Zimmerman, Jeff      Texas Rangers   3366667  Pitcher

[180 rows x 4 columns]
```