Day Objectives

- Practice on Income Dataset
  - Define functions for the following data points
  - Average Income of all states from 2005 to 2013
  - State with highest average income in the last three years
  - State with lowest average income from 2007 to 2010(inclusive)
  - Print the list of all states in the same line with average income less than California
  - Print the names of states based on descending order of income in the year 2009
  - State with the lowest recorded income from 2005 to 2013

In [ ]:

```
```

In [1]:

```
import pandas as pd
# Reading CSV files

filepath='DataFiles/Income.csv.txt'

incomedf=pd.read_csv(filepath)

incomedf
```

Out[1]:

	GEOID	State	2005	2006	2007	2008	2009	2010	2011	2012	2013
0	04000US01	Alabama	37150	37952	42212	44476	39980	40933	42590	43464	41381
1	04000US02	Alaska	55891	56418	62993	63989	61604	57848	57431	63648	61137
2	04000US04	Arizona	45245	46657	47215	46914	45739	46896	48621	47044	50602
3	04000US05	Arkansas	36658	37057	40795	39586	36538	38587	41302	39018	39919
4	04000US06	California	51755	55319	55734	57014	56134	54283	53367	57020	57528

In [ ]:

```
```

In [2]:

```
def getRowIndex(df,rowkey):
    for i in range(len(df.values)):
        if df.values[i][0]==rowkey or df.values[i][1]==rowkey:
            rowindex=i
    return rowindex

def getColumnIndex(df,columnkey):
    for i in range(len(df.columns)):
        if df.columns[i]==columnkey:
            columnindex=i
    return columnindex
```

In [ ]:

```

```

In [5]:

```
#import numpy as np
def AvgIncomeofDF(incomedf,sy,ey):
    Average=[]

    l=ey-sy+1
    for i in range(0,len(incomedf.values)):
        s=sum(incomedf.values[i][sy:ey+1])
        Average.append(s//l)
    return Average

sy=getColumnIndex(incomedf,'2005')
ey=getColumnIndex(incomedf,'2013')
AverageIncome=AvgIncomeofDF(incomedf,sy,ey)
print("Average income of all states from 2005 to 2013 is: ",sum(AverageIncome)//len(Ave
rageIncome))
```

```
Average income of all states from 2005 to 2013 is:  48524
```

In [ ]:

```

```

In [6]:

```
#State with highest average income in the last three years
def HighestAverageIncome(incomedf,sy,ey):
    AverageIncome=AvgIncomeofDF(incomedf,sy,ey)

    return AverageIncome


sy=getColumnIndex(incomedf,incomedf.columns[len(incomedf.columns)-3])
ey=getColumnIndex(incomedf,incomedf.columns[-1])
HighestIncome=HighestAverageIncome(incomedf,sy,ey)

print(max(HighestIncome))
incomedf.values[HighestIncome.index(max(HighestIncome))][1]
```

```
60738
```

Out[6]:

```
'Alaska'
```

In [7]:

```
# State with Lowest Average Income from 2007 to 2010
def LowestAverageIncome(incomedf,sy,ey):
    AverageIncome=AvgIncomeofDF(incomedf,sy,ey)

    return AverageIncome


sy=getColumnIndex(incomedf,'2007')
ey=getColumnIndex(incomedf,'2010')
LowestIncome=LowestAverageIncome(incomedf,sy,ey)

#HighestIncome.index(max(li))
incomedf.values[LowestIncome.index(min(LowestIncome))][1]
```

Out[7]:

```
'Arkansas'
```

In [8]:

```
# Print the list of all states in the same line with average income less than Californi
a

def AvgIncomeLessthanCalifornia(incomedf,sy,ey):
    AverageIncome=AvgIncomeofDF(incomedf,sy,ey)
    least=[]
    CaliforniaAvgIncome=AverageIncome[4]
    for income in AverageIncome:
        if income<CaliforniaAvgIncome:
            small=str(incomedf.values[AverageIncome.index(income)][1])+':'+str(income)
            least.append(small)
    return least

sy=getColumnIndex(incomedf,'2005')
ey=getColumnIndex(incomedf,'2013')


AvgIncomeLessthanCalifornia(incomedf,sy,ey)
```

Out[8]:

```
['Alabama:41126', 'Arizona:47214', 'Arkansas:38828']
```

In [ ]:

```

```

In [9]:

```
# Print the names of states based on descending order of income in the year 2009
def DescendingOrderIncome(incomedf,reqI):
    incomelist=[]
    for index in range(0,len(incomedf.values)):
        incomelist.append(incomedf.values[index][reqI])
    sortedincome=sorted(incomelist,reverse=True)
    for inc in sortedincome:
        print(incomedf.values[incomelist.index(inc)][1])

    return sortedincome


reqI=getColumnIndex(incomedf,'2009')
DescendingOrderIncome(incomedf,reqI)
```

```
Alaska
California
Arizona
Alabama
Arkansas
```

Out[9]:

```
[61604, 56134, 45739, 39980, 36538]
```

In [10]:

```
# State with the lowest recorded income from 2005 to 2013
def LowestRecordedIncome(incomedf,si,ei):
    minincome=[]
    for index in range(0,len(incomedf.values)):
        minincome.append(min(incomedf.values[index][si:]))
    overallmin=min(minincome)
    for i in range(0,len(incomedf.values)):
        for j in range(0,len(incomedf.columns)):
            if incomedf.loc[i][j]==overallmin:
                rowIndex=i
                break


    return incomedf.values[rowIndex][1],overallmin

si=getColumnIndex(incomedf,'2005')
ei=getColumnIndex(incomedf,'2013')
LowestRecordedIncome(incomedf,si,ei)
```

Out[10]:

```
('Arkansas', 36538)
```