In [7]:
```
# Function to read CSV data into a DataFrame and
# returns the DataFrame object

filepath = 'DataFiles/income.csv'
import pandas as pd

def readCSVdata(filepath):
    return pd.read_csv(filepath)
readCSVdata(filepath)
```

Out[7]:

	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 [8]:
```
incomedf=readCSVdata(filepath)

# Function to print all column names in a single line
# GEOID State 2005 2006 2007 2008 2009 2010 2011 2012 2013

def printDataFrameColumns(df):
    columns=df.columns
    for column in columns:
        print(column,end=" ")
    return
printDataFrameColumns(incomedf)
```

```
GEOID State 2005 2006 2007 2008 2009 2010 2011 2012 2013
```

- Average Income of all states from 2005 to 2013

In [17]:
```
incomedf=readCSVdata(filepath)
filepath='DataFiles/income.csv'

def average(df):
    sum=0
    count=0
    for i in range(len(df.values)):
        for j in range(2,len(df.columns)):
            sum=sum+df.values[i][j]
            count+=1
    print(sum//count)
average(incomedf)

```

```
48525
```

- State with highest average income in the last three years

In [40]:
```
def stateHighestAverage(df):
    avg={}
    for i in range(len(df.values)):
        sum=0
        count=0
        for j in range(8,len(df.columns)):
            sum=sum+df.values[i][j]
            count=count+1
        avg[df.values[i][1]]=sum//count
    print(avg)
    a=avg.values()
    maxavg=max(a)
    for item in avg.items():
        if item[1]==maxavg:
            print(item[0],':',maxavg)
stateHighestAverage(incomedf)
```

```
{'Alabama': 42478, 'Alaska': 60738, 'Arizona': 48755, 'Arkansas': 40079, 'Calif
ornia': 55971}
Alaska : 60738
```

In [41]:
```
def stateHighestAverage(df):
    avg={}
    for i in range(len(df.values)):
        sum=0
        count=0
        for j in range(4,8):
            sum=sum+df.values[i][j]
            count=count+1
        avg[df.values[i][1]]=sum//count
    print(avg)
    a=avg.values()
    minavg=min(a)
    for item in avg.items():
        if item[1]==minavg:
            print(item[0],':',minavg)
stateHighestAverage(incomedf)
```

```
{'Alabama': 41900, 'Alaska': 61608, 'Arizona': 46691, 'Arkansas': 38876, 'Calif
ornia': 55791}
Arkansas : 38876
```

In [50]:
```
def stateHighestAverage(df):
    avg={}
    for i in range(len(df.values)):
        sum=0
        count=0
        for j in range(8,len(df.columns)):
            sum=sum+df.values[i][j]
            count=count+1
        avg[df.values[i][1]]=sum//count
    print(avg)
    calvalue=avg['California']
    for item in avg.items():
        if item[1]<calvalue:
            print(item[0],end=" ")
stateHighestAverage(incomedf)
```

```
{'Alabama': 42478, 'Alaska': 60738, 'Arizona': 48755, 'Arkansas': 40079, 'Calif
ornia': 55971}
Alabama Arizona Arkansas
```

In [72]:
```
def getColumnIndex(df,columnkey):
    for i in range(len(df.columns)):
        if df.columns[i]==columnkey:
            columnindex=i
    return columnindex
getColumnIndex(incomedf,'2009')

def descending(df,columnkey):
    income={}
    incomevalues=[]
    columnindex=getColumnIndex(df,columnkey)
    for i in range(len(df.values)):
        income[df.values[i][1]]=df.values[i][columnindex]
    incomevalues=sorted(income.values(),reverse=True)
    for i in incomevalues:
        for item in income.items():
            if item[1]==i:
                print(item[0],":",i)
descending(incomedf,'2009')


```

```
Alaska : 61604
California : 56134
Arizona : 45739
Alabama : 39980
Arkansas : 36538
```

In [75]:
```
def stateLowestIncome(df):
    totalincome={}
    for i in range(len(df.values)):
        sum=0
        count=0
        for j in range(2,len(df.columns)):
            sum=sum+df.values[i][j]
        totalincome[df.values[i][1]]=sum
    print(totalincome)
    a=totalincome.values()
    minincome=min(a)
    for item in totalincome.items():
        if item[1]==minincome:
            print(item[0],':',minincome)
stateLowestIncome(incomedf)
```

```
{'Alabama': 370138, 'Alaska': 540959, 'Arizona': 424933, 'Arkansas': 349460, 'C
alifornia': 498154}
Arkansas : 349460
```

In [ ]:
```
1
```