Name :- MOHAN SAI KOTHAPALLI

Emp ID:- 2112951

United States - Crime Rates - 1960 - 2014

Introduction:

This time you will create a data

Special thanks to: https://github.com/justmarkham for sharing the dataset and materials.

Step 1. Import the necessary libraries

In [1]:
```
import pandas as pd
import numpy as np
```

Step 2. Import the dataset from this address.

Step 3. Assign it to a variable called crime.

In [2]:
```
Crime_Rate = pd.read_csv("https://raw.githubusercontent.com/guipsamora/pandas_exercises/master/04_Apply/US_Crime_Rates/US_Crime_Rates_1960_2014.csv")
Crime_Rate.head(10)
```

Out[2]:

	Year	Population	Total	Violent	Property	Murder	Forcible_Rape	Robbery	Aggravated_assault	Burglary	Larceny_Theft	Vehicle_T
0	1960	179323175	3384200	288460	3095700	9110	17190	107840	154320	912100	1855400	32
1	1961	182992000	3488000	289390	3198600	8740	17220	106670	156760	949600	1913000	33
2	1962	185771000	3752200	301510	3450700	8530	17550	110860	164570	994300	2089600	36
3	1963	188483000	4109500	316970	3792500	8640	17650	116470	174210	1086400	2297800	40
4	1964	191141000	4564600	364220	4200400	9360	21420	130390	203050	1213200	2514400	47
5	1965	193526000	4739400	387390	4352000	9960	23410	138690	215330	1282500	2572600	49
6	1966	195576000	5223500	430180	4793300	11040	25820	157990	235330	1410100	2822000	56
7	1967	197457000	5903400	499930	5403500	12240	27620	202910	257160	1632100	3111600	65
8	1968	199399000	6720200	595010	6125200	13800	31670	262840	286700	1858900	3482700	78
9	1969	201385000	7410900	661870	6749000	14760	37170	298850	311090	1981900	3888600	87

Step 4. What is the type of the columns?

In [3]:
```
Crime_Rate.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 55 entries, 0 to 54
Data columns (total 12 columns):
Year                 55 non-null int64
Population           55 non-null int64
Total                55 non-null int64
Violent              55 non-null int64
Property             55 non-null int64
Murder               55 non-null int64
Forcible_Rape        55 non-null int64
Robbery              55 non-null int64
Aggravated_assault   55 non-null int64
Burglary             55 non-null int64
Larceny_Theft        55 non-null int64
Vehicle_Theft        55 non-null int64
dtypes: int64(12)
memory usage: 5.3 KB
```

Have you noticed that the type of Year is int64. But pandas has a different type to work with Time Series. Let's see it now.

Step 5. Convert the type of the column Year to datetime64

In [4]:
```
Crime_Rate['Year'] = pd.to_datetime(Crime_Rate['Year'],format='%Y')
print(Crime_Rate.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 55 entries, 0 to 54
Data columns (total 12 columns):
Year                 55 non-null datetime64[ns]
Population           55 non-null int64
Total                55 non-null int64
Violent              55 non-null int64
Property             55 non-null int64
Murder               55 non-null int64
Forcible_Rape        55 non-null int64
Robbery              55 non-null int64
Aggravated_assault   55 non-null int64
Burglary             55 non-null int64
Larceny_Theft        55 non-null int64
Vehicle_Theft        55 non-null int64
dtypes: datetime64[ns](1), int64(11)
memory usage: 5.3 KB
None
```

In [5]:
```
Crime_Rate.head()
```

Out[5]:

	Year	Population	Total	Violent	Property	Murder	Forcible_Rape	Robbery	Aggravated_assault	Burglary	Larceny_Theft	Vehicle_
0	1960-01-01	179323175	3384200	288460	3095700	9110	17190	107840	154320	912100	1855400	32
1	1961-01-01	182992000	3488000	289390	3198600	8740	17220	106670	156760	949600	1913000	33
2	1962-01-01	185771000	3752200	301510	3450700	8530	17550	110860	164570	994300	2089600	36
3	1963-01-01	188483000	4109500	316970	3792500	8640	17650	116470	174210	1086400	2297800	40
4	1964-01-01	191141000	4564600	364220	4200400	9360	21420	130390	203050	1213200	2514400	47

Step 6. Set the Year column as the index of the dataframe

In [6]:
```
Crime_Rate = Crime_Rate.set_index('Year', drop = True)
Crime_Rate.head()
```

Out[6]:

Year	Population	Total	Violent	Property	Murder	Forcible_Rape	Robbery	Aggravated_assault	Burglary	Larceny_Theft	Vehicle_The
1960-01-01	179323175	3384200	288460	3095700	9110	17190	107840	154320	912100	1855400	3282
1961-01-01	182992000	3488000	289390	3198600	8740	17220	106670	156760	949600	1913000	3360
1962-01-01	185771000	3752200	301510	3450700	8530	17550	110860	164570	994300	2089600	3668
1963-01-01	188483000	4109500	316970	3792500	8640	17650	116470	174210	1086400	2297800	4083
1964-01-01	191141000	4564600	364220	4200400	9360	21420	130390	203050	1213200	2514400	4728

Step 7. Delete the Total column

In [7]:
```
del Crime_Rate['Total']
```

Step 8. Group the year by decades and sum the values

Pay attention to the Population column number, summing this column is a mistake

In [8]:
```
total = Crime_Rate.resample('10AS').sum()
population = Crime_Rate['Population'].resample('10AS').max()
total['Population'] = population
total
```

Out[8]:

Year	Population	Violent	Property	Murder	Forcible_Rape	Robbery	Aggravated_assault	Burglary	Larceny_Theft	Vehicle_Theft
1960-01-01	201385000	4134930	45160900	106180	236720	1633510	2158520	13321100	26547700	5292100
1970-01-01	220099000	9607930	91383800	192230	554570	4159020	4702120	28486000	53157800	9739900
1980-01-01	248239000	14074328	117048900	206439	865639	5383109	7619130	33073494	72040253	11935411
1990-01-01	272690813	17527048	119053499	211664	998827	5748930	10568963	26750015	77679366	14624418
2000-01-01	307006550	13968056	100944369	163068	922499	4230366	8652124	21565176	67970291	11412834
2010-01-01	318857056	6072017	44095950	72867	421059	1749809	3764142	10125170	30401698	3569080

Step 9. What is the most dangerous decade to live in the US?

In [9]:
```
total.idxmax(0)
```

Out[9]:
```
Population           2010-01-01
Violent              1990-01-01
Property             1990-01-01
Murder               1990-01-01
Forcible_Rape        1990-01-01
Robbery              1990-01-01
Aggravated_assault   1990-01-01
Burglary             1980-01-01
Larceny_Theft        1990-01-01
Vehicle_Theft        1990-01-01
dtype: datetime64[ns]
```

In [ ]:
```

```