Pandas

- pandas is an high performence in Data Analysis Tool
- More Flexible
- · Pandas is store the Data in Different ways
- · Pandas is Mainly uses in Missing Data
- indexing,slicing
- Pandas Data Structures in three types
 - series
 - One column Multiple rows Ex:List
 - DataFrame
 - Multiple columns and multiple rows Ex: list,Dict,sets
 - panel

```
In [1]:
                 # importing Libraries
         H
                 import pandas as pd
              3
                 import numpy as np
              4
                 d= pd.Series([1,2,3,4])
In [2]:
              1
                 d
   Out[2]: 0
                 1
                  2
            2
                 3
            3
                 4
            dtype: int64
In [3]:
                 d= pd.Series([1,2,3,4],index = ['a','b','c','d'])
                 d
In [4]:
   Out[4]: a
                 1
                 2
                 3
                 4
            dtype: int64
In [5]:
                 d.index
   Out[5]: Index(['a', 'b', 'c', 'd'], dtype='object')
In [6]:
         H
              1
                 d.values
   Out[6]: array([1, 2, 3, 4], dtype=int64)
```

```
1 d.sum()
 In [7]: ▶
    Out[7]: 10
 In [8]: ▶
                  d.mean()
    Out[8]: 2.5
 In [9]: ▶
                  a=pd.Series(np.arange(1,10))
               2
                  а
    Out[9]: 0
                  1
                  2
             2
                  3
             3
                  4
                  5
             5
                  6
             6
                  7
             7
                  8
             dtype: int32
In [10]: ▶
                  a.max()
   Out[10]: 9
In [11]: ▶
                  a.min()
   Out[11]: 1
In [12]:
                  a.cumsum()
   Out[12]: 0
                   1
             1
                   3
             2
                   6
             3
                  10
             4
                  15
             5
                  21
             6
                  28
             7
                  36
                  45
             dtype: int32
                 d=pd.date_range('01-01-2021',periods = 30)
In [15]:
```

1 d

In [16]:

```
Out[16]: DatetimeIndex(['2021-01-01', '2021-01-02', '2021-01-03', '2021-01-04',
                              '2021-01-05', '2021-01-06', '2021-01-07', '2021-01-08',
                              '2021-01-09', '2021-01-10', '2021-01-11', '2021-01-12', '2021-01-13', '2021-01-14', '2021-01-15', '2021-01-16',
                              '2021-01-17', '2021-01-18', '2021-01-19', '2021-01-20'
                              '2021-01-21', '2021-01-22', '2021-01-23', '2021-01-24',
                              '2021-01-25', '2021-01-26', '2021-01-27', '2021-01-28',
                              '2021-01-29', '2021-01-30'],
                             dtype='datetime64[ns]', freq='D')
In [23]:
                  a= pd.date range(start='01-01-2021',end='02-01-2021',freq = '60S')
In [24]:
                1
                  а
           H
    Out[24]: DatetimeIndex(['2021-01-01 00:00:00', '2021-01-01 00:01:00',
                              '2021-01-01 00:02:00', '2021-01-01 00:03:00',
                              '2021-01-01 00:04:00', '2021-01-01 00:05:00',
                              '2021-01-01 00:06:00', '2021-01-01 00:07:00',
                              '2021-01-01 00:08:00', '2021-01-01 00:09:00',
                              '2021-01-31 23:51:00', '2021-01-31 23:52:00',
                              '2021-01-31 23:53:00', '2021-01-31 23:54:00',
                              '2021-01-31 23:55:00', '2021-01-31 23:56:00',
                              '2021-01-31 23:57:00', '2021-01-31 23:58:00',
                              '2021-01-31 23:59:00', '2021-02-01 00:00:00'],
                             dtype='datetime64[ns]', length=44641, freq='60S')
          2.DataFrames
                  df = pd.DataFrame([[1,2],[3,4]])
In [25]:
           H
               1
               2
                  df
    Out[25]:
                 0 1
               0 1 2
               1 3 4
                  df = pd.DataFrame({'ravi':[1,2,3,4],'anil':[5,6,7,8]},index=[1,2,3,4])
In [30]:
```

```
In [31]:
                1
                  df
    Out[31]:
                  ravi anil
               1
                    1
                         5
               2
                    2
                         6
               3
                         7
                    3
                    4
                         8
                   df = pd.DataFrame({'ravi':[70,80,80,90],'anil':[60,50,70,80],'alekya':[9]
In [33]:
                   df
In [34]:
    Out[34]:
                  ravi anil alekya
               0
                   70
                        60
                               90
               1
                   80
                        50
                               80
               2
                   80
                        70
                               70
               3
                   90
                        80
                               90
In [35]:
                   df.index = ['python', 'machine learning', 'data analysis', 'Django']
In [36]:
                   df
    Out[36]:
                               ravi
                                    anil alekya
                                70
                                     60
                                            90
                        python
               machine learning
                                80
                                     50
                                            80
                   data analysis
                                80
                                     70
                                            70
                       Django
                                90
                                     80
                                            90
                   df.loc['machine learning']
In [37]:
    Out[37]: ravi
                         80
              anil
                         50
              alekya
                         80
              Name: machine learning, dtype: int64
                   df.loc['Django']
In [42]:
    Out[42]: ravi
                         90
              anil
                         80
                         90
              alekya
              Name: Django, dtype: int64
```

```
df.iloc[3]
In [45]:
    Out[45]: ravi
                          90
               anil
                          80
               alekya
                          90
               Name: Django, dtype: int64
                    df
In [46]:
    Out[46]:
                                ravi anil alekya
                                 70
                                      60
                                             90
                        python
                machine learning
                                 80
                                      50
                                             80
                   data analysis
                                 80
                                      70
                                             70
                        Django
                                 90
                                      80
                                             90
In [51]:
                    df.iloc[2:3,1:2]
    Out[51]:
                             anil
                data analysis
                              70
                    a= pd.DataFrame({'ravi':[70,80,80,90,100,40,60],'anil':[60,50,70,80,100,
In [52]:
            H
In [53]:
                 1
                    а
    Out[53]:
                   ravi
                       anil
                            alekya girsha vijay
                0
                    70
                        60
                                90
                                       70
                                             90
                1
                    80
                        50
                                80
                                       80
                                             80
                2
                    80
                        70
                                70
                                       90
                                             70
                3
                    90
                        80
                                90
                                      110
                                             60
                4
                   100
                        100
                                70
                                       40
                                             50
                5
                    40
                        50
                                       30
                                             40
                                60
                    60
                        60
                                40
                                       60
                                             30
                    a.index=[1,2,3,4,5,6,7]
In [54]:
```

n	[55]:	M	1	а				
	(Out[55	5]:		ravi	anil	alekya	girsha	vijay
				1	70	60	90	70	90
				2	80	50	80	80	80
				3	80	70	70	90	70
				4	90	80	90	110	60
				5	100	100	70	40	50
				6	40	50	60	30	40
				7	60	60	40	60	30
In	ſ	59]:	M	1	a.:	ilocſ	0:7:2,	0:5:21	
	(Out[59)]:				ya vijay	_	
				1	70	ę	90 90		
				3	80	7	70 70		
				5	100	7	70 50		
				7	60	4	40 30		
Tn	г	61]:	M	1	а				
	(Out[61	L]:				alekya	girsha	vijay
				1	70	60	90	70	90
				2	80	50	80	80	80
				3	80	70	70	90	70
				4	90	80	90	110	60
				5	100	100	70	40	50
				6	40	50	60	30	40
				7	60	60	40	60	30
_	_	50 3				., -		- 63	
In	[62]:	M	1	a.:	iloc[4:5,2:	5:2]	
	(Out[62	2]:		aleky	/a vij	jay		
				5	7	70	50		
Tn	Г	75]:	H	1	da	ta =	{'roll	number	':['17

```
In [76]:
                1
                   data
    Out[76]: {'rollnumber': ['17je5a0501',
                 '17je5a0502',
                '17je5a0503',
                '17je5a0504',
                '17je5a0505',
                '17je5a0506',
                '17je5a0507',
                '17je5a0508',
                '17je5a0509',
                '17je5a05010'
                '17je5a05011',
                '17je5a05012',
                '17je5a05013',
                '17je5a05014',
                '17je5a05015',
                '17je5a05016',
                '17je5a05017',
                '17je5a05018',
                '17je5a05019',
                '17je5a05020',
                '17je5a05021',
                '17je5a05022',
                '17je5a05023',
                '17je5a05024',
                '17je5a05025',
                '17je5a05026',
                '17je5a05027',
                '17je5a05028',
                '17je5a05029'
                '17je5a05030',
                '17je5a05031',
                '17je5a05032',
                '17je5a05033',
                '17je5a05034',
                '17je5a05035',
                '17je5a05036',
                '17je5a05037',
                '17je5a05038'
                '17je5a05039',
                '17je5a05040',
                '17je5a05041',
                '17je5a05042',
                '17je5a05043',
                '17je5a05044',
                '17je5a05045',
                '17je5a05046',
                '17je5a05047',
                '17je5a05048'
                '17je5a05049',
                '17je5a05050',
                '17je5a05051',
                '17je5a05052',
                '17je5a05053',
                '17je5a05054',
```

```
'17je5a05055',
  '17je5a05056',
  '17je5a05057',
  '17je5a05058',
  '17je5a05059',
  '17je5a05060'],
 'python': array([92, 98, 41, 54, 16, 9, 26, 26, 98, 29, 15, 2, 56, 30, 5
7, 75, 35,
       92, 39, 69, 23, 22, 7, 37, 98, 18, 65, 1, 87, 30, 21, 47, 85, 8,
       74, 17, 71, 39, 44, 37, 48, 8, 59, 82, 74, 26, 58, 58, 96, 42, 35,
       82, 55, 2, 93, 87, 90, 96, 48, 18]),
 'machine learning': array([86, 68, 28, 17, 2, 34, 60, 71, 87, 65, 75, 46,
10, 27, 67, 5, 46,
        31, 6, 75, 26, 22, 15, 40, 74, 58, 27, 86, 58, 71, 41, 58, 11, 99,
        3, 60, 90, 69, 82, 4, 40, 92, 60, 15, 42, 79, 86, 11, 60, 51, 57,
        89, 71, 84, 37, 74, 73, 41, 65, 40]),
 'deep learning': array([33, 27, 35, 25, 71, 76, 55, 81, 98, 79, 18, 13, 9
7, 83, 47, 93, 46,
        51, 59, 51, 73, 67, 17, 94, 51, 61, 42, 42, 10, 60, 21, 87, 92, 82,
       37, 45, 62, 88, 27, 6, 76, 57, 36, 39, 78, 12, 92, 78, 74, 31, 28,
        71, 57, 54, 61, 88, 76, 80, 60, 52]),
 'data analysis': array([66, 38, 94, 98, 20, 57, 12, 91, 47, 19, 56, 50, 6
7, 51, 91, 42, 56,
       53, 39, 61, 27, 83, 53, 2, 18, 56, 54, 4, 13, 62, 68, 73, 32, 36,
       11, 59, 91, 21, 9, 13, 56, 14, 62, 48, 68, 34, 21, 19, 90, 77, 99,
        3, 64, 4, 67, 61, 93, 18, 9, 91])}
```

1 | df=pd.DataFrame(data,index = [np.arange(1,61)])

In [78]: № 1 df

Out[78]:

	rollnumber	python	machine learning	deep learning	data analysis
1	17je5a0501	92	86	33	66
2	17je5a0502	98	68	27	38
3	17je5a0503	41	28	35	94
4	17je5a0504	54	17	25	98
5	17je5a0505	16	2	71	20
6	17je5a0506	9	34	76	57
7	17je5a0507	26	60	55	12
8	17je5a0508	26	71	81	91
9	17je5a0509	98	87	98	47
10	17je5a05010	29	65	79	19
11	17je5a05011	15	75	18	56
12	17je5a05012	2	46	13	50
13	17je5a05013	56	10	97	67
14	17je5a05014	30	27	83	51
15	17je5a05015	57	67	47	91
16	17je5a05016	75	5	93	42
17	17je5a05017	35	46	46	56
18	17je5a05018	92	31	51	53
19	17je5a05019	39	6	59	39
20	17je5a05020	69	75	51	61
21	17je5a05021	23	26	73	27
22	17je5a05022	22	22	67	83
23	17je5a05023	7	15	17	53
24	17je5a05024	37	40	94	2
25	17je5a05025	98	74	51	18
26	17je5a05026	18	58	61	56
27	17je5a05027	65	27	42	54
28	17je5a05028	1	86	42	4
29	17je5a05029	87	58	10	13
30	17je5a05030	30	71	60	62
31	17je5a05031	21	41	21	68
32	17je5a05032	47	58	87	73
33	17je5a05033	85	11	92	32
34	17je5a05034	8	99	82	36

	rollnumber	python	machine learning	deep learning	data analysis
35	17je5a05035	74	3	37	11
36	17je5a05036	17	60	45	59
37	17je5a05037	71	90	62	91
38	17je5a05038	39	69	88	21
39	17je5a05039	44	82	27	9
40	17je5a05040	37	4	6	13
41	17je5a05041	48	40	76	56
42	17je5a05042	8	92	57	14
43	17je5a05043	59	60	36	62
44	17je5a05044	82	15	39	48
45	17je5a05045	74	42	78	68
46	17je5a05046	26	79	12	34
47	17je5a05047	58	86	92	21
48	17je5a05048	58	11	78	19
49	17je5a05049	96	60	74	90
50	17je5a05050	42	51	31	77
51	17je5a05051	35	57	28	99
52	17je5a05052	82	89	71	3
53	17je5a05053	55	71	57	64
54	17je5a05054	2	84	54	4
55	17je5a05055	93	37	61	67
56	17je5a05056	87	74	88	61
57	17je5a05057	90	73	76	93
58	17je5a05058	96	41	80	18
59	17je5a05059	48	65	60	9
60	17je5a05060	18	40	52	91

Out[79]:

	rollnumber	python	machine learning	deep learning	data analysis
1	17je5a0501	92	86	33	66
2	17je5a0502	98	68	27	38
3	17je5a0503	41	28	35	94
4	17je5a0504	54	17	25	98
5	17je5a0505	16	2	71	20

In [80]: ► df.tail()

Out[80]: rollnumber python machine learning deep learning data analysis 17je5a05056 17je5a05057 17je5a05058 17je5a05059 17je5a05060

Out[81]:		rollnumber	python	machine learning	deep learning	data analysis
	51	17je5a05051	35	57	28	99
	5	17je5a0505	16	2	71	20
	58	17je5a05058	96	41	80	18
	41	17je5a05041	48	40	76	56
	27	17je5a05027	65	27	42	54

```
df['machine learning']
In [82]:
    Out[82]: 1
                      86
               2
                      68
               3
                      28
               4
                      17
               5
                       2
               6
                      34
               7
                      60
               8
                      71
               9
                      87
               10
                      65
               11
                      75
               12
                      46
               13
                      10
               14
                      27
               15
                      67
               16
                       5
               17
                      46
               18
                      31
               19
                       6
               20
                      75
               21
                      26
               22
                      22
                      15
               23
               24
                      40
               25
                      74
               26
                      58
               27
                      27
               28
                      86
               29
                      58
               30
                      71
               31
                      41
               32
                      58
               33
                      11
               34
                      99
               35
                       3
               36
                      60
               37
                      90
               38
                      69
               39
                      82
               40
                       4
               41
                      40
               42
                      92
               43
                      60
               44
                      15
               45
                      42
                      79
               46
               47
                      86
               48
                      11
               49
                      60
               50
                      51
               51
                      57
               52
                      89
               53
                      71
               54
                      84
```

```
55
                    37
              56
                    74
              57
                    73
              58
                    41
              59
                    65
              60
                    40
             Name: machine learning, dtype: int32
In [83]:
                  df['machine learning'].head()
    Out[83]: 1
                   86
              2
                   68
              3
                   28
              4
                   17
              5
                    2
              Name: machine learning, dtype: int32
In [84]:
                  df['machine learning'].sample(5)
   Out[84]:
             32
                    58
              18
                    31
                    87
              10
                    65
              19
              Name: machine learning, dtype: int32
```

In [85]: ▶ 1 df

Out[85]:

	rollnumber	python	machine learning	deep learning	data analysis
1	17je5a0501	92	86	33	66
2	17je5a0502	98	68	27	38
3	17je5a0503	41	28	35	94
4	17je5a0504	54	17	25	98
5	17je5a0505	16	2	71	20
6	17je5a0506	9	34	76	57
7	17je5a0507	26	60	55	12
8	17je5a0508	26	71	81	91
9	17je5a0509	98	87	98	47
10	17je5a05010	29	65	79	19
11	17je5a05011	15	75	18	56
12	17je5a05012	2	46	13	50
13	17je5a05013	56	10	97	67
14	17je5a05014	30	27	83	51
15	17je5a05015	57	67	47	91
16	17je5a05016	75	5	93	42
17	17je5a05017	35	46	46	56
18	17je5a05018	92	31	51	53
19	17je5a05019	39	6	59	39
20	17je5a05020	69	75	51	61
21	17je5a05021	23	26	73	27
22	17je5a05022	22	22	67	83
23	17je5a05023	7	15	17	53
24	17je5a05024	37	40	94	2
25	17je5a05025	98	74	51	18
26	17je5a05026	18	58	61	56
27	17je5a05027	65	27	42	54
28	17je5a05028	1	86	42	4
29	17je5a05029	87	58	10	13
30	17je5a05030	30	71	60	62
31	17je5a05031	21	41	21	68
32	17je5a05032	47	58	87	73
33	17je5a05033	85	11	92	32
34	17je5a05034	8	99	82	36

	rollnumber	python	machine learning	deep learning	data analysis
35	17je5a05035	74	3	37	11
36	17je5a05036	17	60	45	59
37	17je5a05037	71	90	62	91
38	17je5a05038	39	69	88	21
39	17je5a05039	44	82	27	9
40	17je5a05040	37	4	6	13
41	17je5a05041	48	40	76	56
42	17je5a05042	8	92	57	14
43	17je5a05043	59	60	36	62
44	17je5a05044	82	15	39	48
45	17je5a05045	74	42	78	68
46	17je5a05046	26	79	12	34
47	17je5a05047	58	86	92	21
48	17je5a05048	58	11	78	19
49	17je5a05049	96	60	74	90
50	17je5a05050	42	51	31	77
51	17je5a05051	35	57	28	99
52	17je5a05052	82	89	71	3
53	17je5a05053	55	71	57	64
54	17je5a05054	2	84	54	4
55	17je5a05055	93	37	61	67
56	17je5a05056	87	74	88	61
57	17je5a05057	90	73	76	93
58	17je5a05058	96	41	80	18
59	17je5a05059	48	65	60	9
60	17je5a05060	18	40	52	91

machine learning 57
deep learning 28
data analysis 99

Name: (51,), dtype: object

5/26/2021

```
Day-3
                    df.loc[50:61][['rollnumber','python','machine learning']]
In [89]:
    Out[89]:
                     rollnumber python machine learning
                50
                    17je5a05050
                                                     51
                                    42
                    17je5a05051
                51
                                    35
                                                     57
                52 17je5a05052
                                    82
                                                     89
                   17je5a05053
                                                     71
                                    55
                   17je5a05054
                                     2
                                                     84
                55 17je5a05055
                                    93
                                                     37
                56
                   17je5a05056
                                    87
                                                     74
                   17je5a05057
                                    90
                                                     73
                58 17je5a05058
                                                     41
                                    96
                    17je5a05059
                                    48
                                                     65
                60 17je5a05060
                                    18
                                                     40
In [90]:
                    df.isnull().sum()
```

```
Out[90]:
             rollnumber
                                  0
              python
                                  0
             machine learning
                                  0
              deep learning
                                  0
              data analysis
                                  0
              dtype: int64
                  data = {'ravi':[40,50,np.nan,60,70],'anil':[70,np.nan,60,np.nan,80]}
In [91]:
In [92]:
               1
                  data
    Out[92]: {'ravi': [40, 50, nan, 60, 70], 'anil': [70, nan, 60, nan, 80]}
                  df = pd.DataFrame(data,index = [1,2,3,4,5])
In [93]:
In [94]:
                  df
   Out[94]:
                 ravi
                       anil
                 40.0
                      70.0
              1
              2
                 50.0 NaN
                 NaN
                      60.0
                 60.0 NaN
              5 70.0 80.0
```

```
df.isnull().sum()
In [95]:
    Out[95]: ravi
                       1
              anil
                       2
              dtype: int64
In [96]:
                   df.fillna(50)
    Out[96]:
                  ravi
                       anil
               1 40.0 70.0
                  50.0
                       50.0
               3 50.0 60.0
                 60.0 50.0
               5 70.0 80.0
In [97]:
                1
                   df
    Out[97]:
                  ravi
                        anil
                  40.0
                      70.0
                  50.0 NaN
                  NaN
                      60.0
                  60.0
                      NaN
               5 70.0 80.0
In [98]:
                   df.dropna()
    Out[98]:
                  ravi
                      anil
                 40.0 70.0
               5 70.0 80.0
                   df.fillna(method = 'ffill')
In [99]:
    Out[99]:
                  ravi
                      anil
               1 40.0 70.0
               2 50.0 70.0
                 50.0 60.0
                 60.0 60.0
               5 70.0 80.0
```

In [100]:

df.fillna(method = 'bfill')

```
Out[100]:
                    ravi
                         anil
                    40.0
                         70.0
                    50.0
                         60.0
                    60.0
                         60.0
                    60.0
                         80.0
                    70.0 80.0
In [102]:
                     df.fillna(method='ffill',inplace = True)
In [103]:
                  1
                     df
    Out[103]:
                    ravi
                         anil
                        70.0
                 1
                    40.0
                    50.0 70.0
                    50.0
                         60.0
                         60.0
                    60.0
                   70.0 80.0
In [104]:
                     d=pd.read_csv('https://raw.githubusercontent.com/AP-State-Skill-Developm
             H
                  1
                  2
                     d
    Out[104]:
                       GEOID
                                                2006
                                                              2008
                                                                     2009
                                  State
                                         2005
                                                       2007
                                                                            2010
                                                                                   2011
                                                                                          2012
                                                                                                2013
                    04000US01
                               Alabama
                                        37150
                                               37952
                                                      42212
                                                             44476
                                                                    39980
                                                                           40933
                                                                                  42590
                                                                                        43464
                                                                                               41381
                    04000US02
                                 Alaska
                                        55891
                                               56418 62993
                                                             63989
                                                                    61604
                                                                           57848
                                                                                 57431
                                                                                        63648
                                                                                                61137
                    04000US04
                                                             46914
                                                                    45739
                                                                           46896
                                                                                         47044
                                 Arizona
                                        45245
                                               46657
                                                      47215
                                                                                  48621
                                                                                                50602
                    04000US05
                               Arkansas
                                        36658
                                               37057
                                                      40795
                                                             39586
                                                                    36538
                                                                           38587
                                                                                  41302
                                                                                        39018
                                                                                               39919
                    04000US06 California 51755 55319 55734 57014
                                                                    56134 54283 53367
                                                                                        57020 57528
In [109]:
                     df = pd.read_csv("income.csv")
```

```
In [107]:
                   1
                      df
    Out[107]:
                        GEOID
                                           2005
                                                   2006
                                                          2007
                                                                 2008
                                                                        2009
                                                                               2010
                                                                                       2011
                                                                                                     2013
                                    State
                                                                                              2012
                    04000US01
                                          37150
                                                  37952
                                                         42212
                                                                44476
                                                                       39980
                                                                              40933
                                                                                     42590
                                                                                             43464
                                                                                                    41381
                  0
                                 Alabama
                    04000US02
                                          55891
                                                        62993
                                                                63989
                                                                       61604
                                                                              57848
                                                                                            63648
                                   Alaska
                                                 56418
                                                                                     57431
                                                                                                    61137
                     04000US04
                                  Arizona
                                          45245
                                                 46657
                                                         47215
                                                                46914
                                                                       45739
                                                                              46896
                                                                                     48621
                                                                                             47044
                                                                                                    50602
                  3
                     04000US05
                                          36658
                                                  37057
                                                                39586
                                                                              38587
                                                                                     41302
                                                                                            39018
                                                                                                    39919
                                Arkansas
                                                         40795
                                                                       36538
                     04000US06
                                          51755
                                                 55319
                                                         55734
                                                                57014
                                                                       56134
                                                                              54283
                                                                                            57020
                                California
                                                                                     53367
                                                                                                    57528
In [110]:
             H
                      df.shape
                 (5, 11)
    Out[110]:
In [111]:
                   1
                      df.isnull().sum()
    Out[111]:
                GEOID
                           0
                 State
                           0
                 2005
                           0
                 2006
                           0
                 2007
                           0
                 2008
                           0
                 2009
                           0
                 2010
                           0
                 2011
                           0
                 2012
                            0
                 2013
                            0
                 dtype: int64
In [115]:
                      df = pd.read_csv('income.csv')
             H
                   1
                      df
In [116]:
                   1
    Out[116]:
                         GEOID
                                    State
                                           2005
                                                    2006
                                                           2007
                                                                    2008
                                                                           2009
                                                                                    2010
                                                                                           2011
                                                                                                  2012
                                                                                                          201
                    04000US01
                                 Alabama
                                          37150
                                                 37952.0
                                                          42212
                                                                 44476.0
                                                                          39980
                                                                                 40933.0
                                                                                          42590
                                                                                                 43464
                                                                                                        4138
                     04000US02
                                          55891
                                                 56418.0
                                                          62993
                                                                          61604
                                                                                 57848.0
                                                                                          57431
                                                                                                 63648
                                                                                                         6113
                                   Alaska
                                                                     NaN
                     04000US04
                                          45245
                                                     NaN
                                                          47215
                                                                 46914.0
                                                                          45739
                                                                                          48621
                                                                                                 47044
                                                                                                        5060
                                  Arizona
                                                                                    NaN
                     04000US05
                                                 37057.0
                                                                                                 39018
                                 Arkansas
                                          36658
                                                          40795
                                                                     NaN
                                                                          36538
                                                                                 38587.0
                                                                                          41302
                                                                                                        3991
                     04000US06
                                California
                                          51755
                                                 55319.0
                                                          55734
                                                                 57014.0
                                                                          56134
                                                                                 54283.0
                                                                                          53367
                                                                                                 57020
                                                                                                        5752
```

```
In [117]:
                    df.isnull().sum()
    Out[117]: GEOID
                          0
                State
                          0
                2005
                          0
                2006
                          1
                2007
                          0
                2008
                          2
                2009
                          0
                2010
                          1
                2011
                          0
                2012
                2013
                dtype: int64
In [124]:
                     df.fillna(67189,inplace = True)
In [125]:
                  1
                     df
    Out[125]:
                       GEOID
                                  State
                                         2005
                                                 2006
                                                        2007
                                                                2008
                                                                       2009
                                                                               2010
                                                                                      2011
                                                                                            2012
                                                                                                   201:
                   04000US01
                               Alabama
                                        37150
                                              37952.0
                                                       42212
                                                             44476.0
                                                                      39980
                                                                            40933.0
                                                                                     42590
                                                                                           43464
                                                                                                  4138
                   04000US02
                                        55891
                                              56418.0
                                                      62993
                                                             67189.0
                                                                      61604
                                                                            57848.0
                                                                                    57431
                                                                                           63648
                                Alaska
                                                                                                  6113
                   04000US04
                                        45245
                                              67189.0
                                                      47215
                                                             46914.0
                                                                      45739
                                                                            67189.0
                                                                                    48621
                                                                                           47044
                                                                                                  5060
                                Arizona
                   04000US05
                                              37057.0
                               Arkansas
                                        36658
                                                       40795
                                                             67189.0
                                                                      36538
                                                                            38587.0
                                                                                    41302
                                                                                           39018
                                                                                                  3991
                                                                                          57020 5752
                   04000US06
                              California
                                       51755 55319.0 55734 57014.0 56134
                                                                            54283.0 53367
In [126]:
                     df.isnull().sum()
    Out[126]:
               GEOID
                          0
                State
                          0
                2005
                          0
                2006
                          0
                2007
                          0
                2008
                          0
                2009
                2010
                          0
                2011
                          0
                2012
                          0
                2013
                dtype: int64
In [127]:
                     data = {'ravi':[40,50,np.nan,60,70],'anil':[70,np.nan,60,np.nan,80]}
In [128]:
                     data
    Out[128]: {'ravi': [40, 50, nan, 60, 70], 'anil': [70, nan, 60, nan, 80]}
                     df = pd.DataFrame(data,index=[1,2,3,4,5])
In [129]:
```

```
df['anil'].fillna(60,inplace=True)
In [133]:
In [134]:
                   df
   Out[134]:
                       anil
                  ravi
               1 40.0 70.0
               2 50.0 60.0
               3 60.0 60.0
               4 60.0 60.0
               5 70.0 80.0
In [135]:
                   df.sum()
   Out[135]: ravi
                       280.0
              anil
                       330.0
              dtype: float64
In [136]:
                   df.min()
   Out[136]: ravi
                       40.0
              anil
                       60.0
              dtype: float64
In [137]:
                   df.max()
   Out[137]: ravi
                       70.0
              anil
                       80.0
              dtype: float64
                   df['ravi'].mean()
In [138]:
   Out[138]: 56.0
                  df['anil'].mean()
In [139]:
   Out[139]: 66.0
  In [ ]:
                1
```