

California is largest population in US (Florida 3rd largest population), but Florida is dominate in the dataset (CA = 19th rank in dataset)
`df['X_STATE'].value_counts()`

- States that only sleep for 1 to 2 hours per day:

```
sleepDf = df[(df['SLEPTIM1'] == 1.0) | (df['SLEPTIM1'] == 2.0)]
```

```
41    8
32    11
11    13
38    13
47    14
4     22
27    34
```

- 27/33 of people who sleeps for 1 to 2 hours are satisfied or very satisfied with their lives

- 0.486956522 of people who have 1 to 2 hours of sleep have been told they have depressive disorder (even if 80% of entire population do not have depression)

```
sleepDf.groupby('ADDEPEV2').count()['IDATE']
```

```
ADDEPEV2
1.0     56
2.0     59
```

- 0.533333333 of people who have 1 to 2 hours of sleep “Have just enough money to make ends meet” (where in the entire dataset, “having just enough money” is only 34% of the data)

```
sleepDf.groupby('SDHMONEY').count()['IDATE']
```

- 0.4 of people who have 1 to 2 hours of sleep don't have cell phones (unlike 18% in entire dataset)

```
sleepDf.groupby('CPDEMO1A').count()['IDATE']
```

- 0.4375 of people who have 1 to 2 hours of sleep have not been on the Internet in the last 30 days (18% only in the entire dataset)

```
sleepDf.groupby('INTERNET').count()['IDATE']
```

- 0.339130435 of people who have 1 to 2 hours of sleep are those whose employment status is “unable to work” (in the entire dataset, this is 7%)
sleepDf.groupby('EMPLOY1').count()['IDATE']

State analysis:

Code:

```
resList = []
topic = 'ADDEPEV2'
for state in df['X_STATE'].unique():
    tmpDf = df[(df['X_STATE'] == state)]
    ratioAr = tmpDf.groupby(topic).count()['IDATE'].values
    rat = (ratioAr[0])/(ratioAr[0]+ratioAr[1])
    resList.append([rat,state])
resList.sort()
resList
```

- Guam and Louisiana has the lowest rate to have Arthritis.

```
'HAVARTH3'
[[0.1899070385126162, 66],
 [0.22605898123324397, 6],
 [0.23732778788666492, 11],
 [0.25009819324430477, 49],
 [0.25307371021824815, 27]
```

- Guam, Hawaii, and District of Columbia is the least depressed. Rhode Island is ranked 7th (0.25) for highest depression ratio (West Virginia is highest).

- ADDEPEV2

- rat = (ratioAr[0])/(ratioAr[0]+ratioAr[1])

```
[0.09700996677740864, 66],
 [0.13619330662876342, 15],
 [0.1375908618899273, 11],
 [0.16213508872352606, 46],
 [0.16360672614962252, 34],
 [0.1641041732036398, 2],
 [0.1649209283220374, 8]
```

....

```
[0.2281178190933779, 22],
 [0.2284521324187295, 1],
 [0.22990412979351033, 26],
```

[0.2347375937165298, 44],
 [0.23759153783563874, 21],
 [0.23839939883524328, 41],
 [0.23850663375501388, 50],
 [0.2418188081295212, 47],
 [0.2494570276140242, 23],
 [0.25973548861131524, 54]]

- Utah (as well as Puerto Rico) has the highest alcohol abstinence rate (0.72 don't drink alcohol in past 30 days) (Drink any alcoholic beverages in past 30 days). Wisconsin has highest ratio of people drinking alcoholic beverages in past 30 days. Rhode Island is ranked 10 in the list

- 'DRNKANY5'
- $rat = (ratioAr[1]) / (ratioAr[0] + ratioAr[1])$
- [[0.33821912788567243, 55],
 [0.3638936621853836, 9],
 [0.3693413956528845, 50],
 [0.37263286499694565, 25],
 [0.37423878944454697, 33],
 [0.3806010928961749, 11],
 [0.38225129597630214, 27],
 [0.38284090909090907, 8],
 [0.3854415274463007, 41],
 [0.39112370160528803, 44],

- West Virginia has the highest rate of asthma (17.5%) (Rhode Island is ranked 3rd with 16.8%) (Ever Told Had Asthma)

- ASTHMA3
- [0.15866884563957806, 1],
 [0.1608910891089109, 25],
 [0.16670971074380164, 15],
 [0.16865894658199476, 41],
 [0.16898395721925133, 44],
 [0.17507799596256193, 54],
 [0.18401071907101385, 72]]

-

- Colorado and California has the highest ratio of people who exercised in the past 30 days (18-19%). Puerto Rico has the largest ratio and a large spike in people not exercising (Kentucky: 36%)->(Puerto Rico: 54%)

- 'EXERANY2'
- [0.32308760962056093, 39],
[0.3265228906545592, 54],
[0.33230265551109495, 13],
[0.33618046430135784, 28],
[0.34127631793700014, 40],
[0.3510433386837881, 1],
[0.3625742980118877, 5],
[0.3629473161033797, 21],
[0.5358516172811582, 72]]
-
- New Hampshire (3%) has the highest ratio of "Never wearing a seatbelt" with California having the lowest ratio (0.3%)
 - 'SEATBELT'
 - [0.02602905569007264, 46],
[0.02652931630565159, 31],
[0.027804810232170164, 29],
[0.027974504249291786, 30],
[0.029796391325939413, 42],
[0.03769774486704813, 2],
[0.04437643458301454, 33]]
- Washington DC has the highest rate in participating in a test for HIV (70% tested). Nebraska has the lowest rate of participation to test for HIV (80% has not been tested).
 - HIVTST6
 - [0.7570416450665284, 40],
[0.7596636175863086, 19],
[0.7676132668545416, 5],
[0.7824561403508772, 49],
[0.790390956194065, 38],
[0.7998543866035676, 31]]
-
- Montana has the highest ratio of people not using the Internet in the last 30 days (30% don't use the Internet). 40% of people in Puerto Rico has not used Internet in 30 days
 - INTERNET
 - [0.22875485723939853, 13],
[0.24980556851765437, 35],
[0.25027583670467085, 54],
[0.2556934893274984, 45],
[0.26914397472544005, 1],
[0.29645304312777104, 28],

[0.3036291088732662, 5],
[0.40371281592484903, 72]]

-

- 15% of West Virginia population is deaf, the highest deaf ratio.
 - DEAF
 - [0.1134656652360515, 29],
[0.12302047122441097, 5],
[0.12332195676905575, 56],
[0.12887746358183377, 30],
[0.1301189924277546, 40],
[0.15512465373961218, 54]]
- West Virginia has the highest rate of people who have difficult concentrating / remembering
 - DECIDE
 - [0.1321009310484415, 29],
[0.14003022574855956, 12],
[0.1405371642723298, 40],
[0.14415700267618198, 47],
[0.1443253881478242, 22],
[0.1498093742991702, 72],
[0.15363300492610837, 1],
[0.15812678741658723, 21],
[0.1777901371894698, 54]]
- Utah has the lowest smoking ratio (Smoked at least 100 cigarettes) (75% of population smoke less)
 - SMOKE100
 - [[0.24819566960705694, 49],
[0.28664568345323743, 72],
[0.35350063022802797, 6],
[0.3736673773987207, 11],
[0.3787305590584279, 16],
[0.3825789923142613, 48],
[0.39455307262569833, 13],
[0.39554890592855807, 17],
- Mississippi has highest rate of “Not doing anything to prevent pregnancy” (65%)
 - PFPVRN2
 - [[0.34034034034034033, 72],
[0.3454810495626822, 28],
[0.354014598540146, 54],
[0.3834510595358224, 1],

```
[0.3862530413625304, 48],  
[0.3896551724137931, 13],  
[0.41029723991507433, 18],  
[0.41074737492279184, 39],
```

- Texas has the highest number of people who report their weight in kg (use the metric system)

- WEIGHT2
- `df[(df['WEIGHT2']>=9000) & (df['WEIGHT2'] <= 9998)].groupby('X_STATE').count()['IDATE'].sort_values()`
- | | |
|----|-----|
| 25 | 35 |
| 9 | 35 |
| 4 | 48 |
| 27 | 50 |
| 18 | 52 |
| 53 | 61 |
| 34 | 76 |
| 12 | 96 |
| 36 | 106 |
| 48 | 111 |

-

- New York has the highest number of people who report their height in meters (use the metric system)

- HEIGHT3
- | | |
|----|-----|
| 17 | 62 |
| 25 | 62 |
| 4 | 81 |
| 8 | 84 |
| 53 | 84 |
| 31 | 94 |
| 18 | 112 |
| 48 | 224 |
| 36 | 228 |

-

-

Value				
Value	Label			
Frequency	Percentage	Weighted		
Percentage				
1	Alabama	6,754	1.50	1.48
2	Alaska	3,203	0.71	0.22
4	Arizona	15,499	3.44	2.12

5 Arkansas 5,322 1.18 0.90
6 California 9,358 2.08 11.95
8 Colorado 9,802 2.18 1.72
9 Connecticut 10,588 2.35 1.11
10 Delaware 4,139 0.92 0.30
11 District of Columbia 3,868 0.86 0.22
12 Florida 22,059 4.90 6.61
13 Georgia 6,056 1.35 3.11
15 Hawaii 7,754 1.72 0.44
16 Idaho 4,894 1.09 0.50
17 Illinois 5,545 1.23 3.87
18 Indiana 13,829 3.07 2.00
19 Iowa 7,699 1.71 0.95
20 Kansas 21,843 4.85 0.86
21 Kentucky 8,642 1.92 1.35
22 Louisiana 4,809 1.07 1.41
23 Maine 9,692 2.15 0.42
24 Maryland 13,588 3.02 1.85
25 Massachusetts 6,912 1.54 2.15
26 Michigan 10,889 2.42 3.04
27 Minnesota 17,095 3.80 1.67
28 Mississippi 5,076 1.13 0.89
29 Missouri 7,601 1.69 1.85
30 Montana 5,915 1.31 0.32
31 Nebraska 15,350 3.41 0.57
32 Nevada 3,764 0.84 0.91
33 New Hampshire
34 New Jersey 11,719 2.60 2.74
35 New Mexico 6,538 1.45 0.62
36 New York 12,249 2.72 6.11
37 North Carolina 4,916 1.09 3.13
38 North Dakota 6,992 1.55 0.23
39 Ohio 12,289 2.73 3.54
40 Oklahoma 6,638 1.48 1.17
41 Oregon 5,370 1.19 1.29
42 Pennsylvania 6,542 1.45 3.97
44 Rhode Island 5,632 1.25 0.33
45 South Carolina 11,311 2.51 1.54
46 South Dakota 7,012 1.56 0.26

47 Tennessee 5,843 1.30 2.04
48 Texas 12,255 2.72 8.27
49 Utah 10,251 2.28 0.86
50 Vermont 6,516 1.45 0.20
51 Virginia 9,630 2.14 2.59
53 Washington 13,279 2.95 2.26
54 West Virginia 5,472 1.22 0.57
55 Wisconsin 5,810 1.29 1.77
56 Wyoming 4,463 0.99 0.18
66 Guam 1,508 0.34 0.04
72 Puerto Rico 4,485 1.00 1.08