

I decided to investigate FBI Gun Data. The data comes from the FBI's National Instant Criminal Background Check System. The NICS is used by to determine whether a prospective buyer is eligible to buy firearms or explosives. Gun shops call into this system to ensure that each customer does not have a criminal record or isn't otherwise ineligible to make a purchase. I took the latest data with the entire 2017.

The second data is a census data 2010-2017. Questions I want to find out:

Which the ratio between background checks and population of each state?

Which states have had the highest growth in gun registrations?

Which states have had the highest growth in handgun purchases?

Which states have had the highest growth in longun purchases?

What is the trend of handgun and longgun purchases?

What is the overall trend of gun purchases?

In [1]:

```
%matplotlib inline
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In this project I divided the data into 2 groups: handgun and long gun background checks. I am strognly belive that the reason why people buy handguns and long guns are not the same at all.

In [2]:

```
months_activity=pd.read_csv('gun_data_new.csv')
```

In [3]:

```
census_data=pd.read_csv('census_data_2010-2017.csv',thousands=",")
```

Check all states in the census data.

In [4]:

```
print len(census_data['State']) #the ouput should be 51. 50 states+District of Columbia  
states_in_census_data=np.array(census_data['State']) #all states together from census data  
states_in_census_data
```

52

Out[4]:

```
array(['Alabama', 'Alaska', 'Arizona', 'Arkansas', 'California',  
      'Colorado', 'Connecticut', 'Delaware', 'District of Columbia',  
      'Florida', 'Georgia', 'Hawaii', 'Idaho', 'Illinois', 'Indiana',  
      'Iowa', 'Kansas', 'Kentucky', 'Louisiana', 'Maine', 'Maryland',  
      'Massachusetts', 'Michigan', 'Minnesota', 'Mississippi', 'Missouri',  
      'Montana', 'Nebraska', 'Nevada', 'New Hampshire', 'New Jersey',  
      'New Mexico', 'New York', 'North Carolina', 'North Dakota',  
      'Ohio', 'Oklahoma', 'Oregon', 'Pennsylvania', 'Rhode Island',  
      'South Carolina', 'South Dakota', 'Tennessee', 'Texas', 'Utah',  
      'Vermont', 'Virginia', 'Washington', 'West Virginia', 'Wisconsin',  
      'Wyoming', 'Puerto Rico'], dtype=object)
```

Cleaning data. I decided to delete Puerto Rico because we want to investigate the data only within the USA.

In [5]:

```
census_data=census_data[census_data.State != 'Puerto Rico']  
len(census_data)
```

Out[5]:

51

50 states + District of Columbia as expected.

Let's check if the number of states same as in the census data.

In [6]:

```
len(months_activity['state'].unique())
```

Out[6]:

55

As we can see there are 55 states instead of 51, let's find the difference

In [7]:

```
states_in_census_data=np.array(census_data['State']) #array of states in census_data_2010-2017.csv
states_in_gun_data=months_activity['state'].unique() #array of states in gun_data_new.csv
np.setdiff1d(states_in_gun_data,states_in_census_data) #check different states
```

Out[7]:

```
array(['Guam', 'Mariana Islands', 'Puerto Rico', 'Virgin Islands'],
      dtype=object)
```

Cleaning data. That's mean this states: 'Guam', 'Mariana Islands', 'Puerto Rico', 'Virgin Islands' are not included in 'census_data_2010-2017.csv'. I decided to delete them from the data.

In [8]:

```
states_for_investigation=census_data['State'].values #We have got this states from census data.
months_activity=months_activity[months_activity['state'].isin(states_for_investigation)]
len(months_activity['state'].unique())
```

Out[8]:

51

50 states + District of Columbia as expected.

In [9]:

```
census_data.head()
```

Out[9]:

	State	2010	2011	2012	2013	2014	2015	
0	Alabama	4785579	4798649	4813946	4827660	4840037	4850858	4860
1	Alaska	714015	722259	730825	736760	736759	737979	7415
2	Arizona	6407002	6465488	6544211	6616124	6706435	6802262	6908
3	Arkansas	2921737	2938640	2949208	2956780	2964800	2975626	2988
4	California	37327690	37672654	38019006	38347383	38701278	39032444	3929

I've decided to change indexes for both data frame. By default it was 0,1,2,3,4, etc and now we're going to turn into states.

In [10]:

```
census_data=census_data.rename(index=str, columns={'State': 'state'}) #changing column name from 'State' to 'state'.
census_data=census_data.set_index('state')
census_data.columns = [year for year in range(2010,2018)] #Changing type of columns name. From str '2010' to int 2010
census_data
```

Out[10]:

	2010	2011	2012	2013	2014	2015	
state							
Alabama	4785579	4798649	4813946	4827660	4840037	4850858	486
Alaska	714015	722259	730825	736760	736759	737979	741
Arizona	6407002	6465488	6544211	6616124	6706435	6802262	690
Arkansas	2921737	2938640	2949208	2956780	2964800	2975626	298
California	37327690	37672654	38019006	38347383	38701278	39032444	392
Colorado	5048029	5116411	5186330	5262556	5342311	5440445	553
Connecticut	3580171	3591927	3597705	3602470	3600188	3593862	358
Delaware	899712	907884	916868	925114	934805	944107	952
District of Columbia	605040	620336	635630	650114	660797	672736	684
Florida	18846461	19097369	19341327	19584927	19897747	20268567	206
Georgia	9712696	9810595	9911171	9981773	10083850	10199533	103
Hawaii	1363817	1378323	1392772	1408038	1417710	1426320	142
Idaho	1570912	1583180	1594673	1610187	1630391	1649324	168
Illinois	12841196	12862298	12878494	12890403	12882438	12862051	128
Indiana	6490029	6515358	6535665	6567484	6593182	6610596	663
Iowa	3050223	3063690	3074386	3089876	3105563	3118473	313
Kansas	2858403	2868756	2885316	2892900	2899553	2905789	290
Kentucky	4347948	4368505	4383673	4399121	4410415	4422057	443
Louisiana	4544871	4574388	4602681	4626795	4648797	4671211	468
Maine	1327568	1327968	1328101	1327975	1328903	1327787	133
Maryland	5788099	5843115	5891680	5932654	5970245	6000561	602
Massachusetts	6564943	6612178	6659627	6711138	6757925	6794002	682
Michigan	9876731	9876199	9886610	9899219	9914675	9918170	993

Minnesota	5310711	5345967	5377695	5416074	5452649	5483238	552
Mississippi	2970437	2977452	2982963	2987721	2988578	2985297	298
Missouri	5995681	6010280	6023267	6041142	6058014	6072640	609
Montana	990507	996866	1003522	1011921	1019931	1028317	103
Nebraska	1829956	1841641	1854862	1867414	1880920	1893564	190
Nevada	2702797	2718170	2752410	2786547	2831730	2883057	293
New Hampshire	1316700	1318345	1320923	1322622	1328684	1330134	133
New Jersey	8803708	8844694	8882095	8913735	8943010	8960001	897
New Mexico	2064607	2077744	2083590	2085161	2083207	2082264	208
New York	19405185	19526372	19625409	19712514	19773580	19819347	198
North Carolina	9574247	9662940	9755299	9849812	9941160	10041769	101
North Dakota	674518	684830	701380	722908	738658	754859	755
Ohio	11539282	11543332	11546969	11567845	11593741	11606027	116
Oklahoma	3759529	3785232	3815298	3849840	3875008	3904353	392
Oregon	3837073	3865845	3893920	3919664	3960673	4016537	408
Pennsylvania	12711063	12742811	12768034	12778450	12790341	12791124	127
Rhode Island	1053169	1052154	1052761	1052784	1054782	1055916	105
South Carolina	4635834	4672744	4719009	4765862	4824758	4892423	495
South Dakota	816227	823338	832576	842513	849455	854036	861
Tennessee	6355882	6396281	6450632	6490795	6540007	6590726	664
Texas	25241648	25644424	26078327	26479279	26954436	27454880	279
Utah	2775260	2815430	2854222	2899961	2938671	2984917	304
Vermont	625842	626210	625606	626044	625665	624455	623
Virginia	8025206	8107548	8188656	8261689	8316902	8366767	841
Washington	6741386	6819155	6890899	6963410	7046931	7152818	728
West Virginia	1854315	1854891	1855360	1852333	1847624	1839767	182
Wisconsin	5690403	5705812	5721075	5736673	5751272	5759744	577
Wyoming	564376	567602	576608	582341	583334	586102	584

Adding column "year". It will be helpful as we want to calculate sum of handgun and long_gun background checks each year for each state

In [11]:

```
months_activity['month'] = pd.to_datetime(months_activity['month'])
months_activity['year'] = months_activity['month'].dt.year
```

In [12]:

```
months_activity.head()
```

Out[12]:

	month	state	permit	permit_recheck	handgun	long_gun	other	multipl
0	2017-12-01	Alabama	24496.0	0.0	12895.0	15765.0	390.0	513
1	2017-12-01	Alaska	249.0	3.0	3675.0	3524.0	287.0	174
2	2017-12-01	Arizona	6243.0	415.0	17566.0	12840.0	1414.0	784
3	2017-12-01	Arkansas	3339.0	590.0	9189.0	13104.0	347.0	440
4	2017-12-01	California	53885.0	0.0	47613.0	35240.0	4885.0	0

5 rows x 9 columns

Let's calculate sum of handgun and long_gun background checks per year (1998-2017) into 2 series.

In [13]:

```
hand_gun_total_checks=months_activity.groupby(['year','state']).sum()['handgun']
long_gun_total_checks=months_activity.groupby(['year','state']).sum()['long_gun']
multiple_total_checks=months_activity.groupby(['year','state']).sum()['multiple']
print hand_gun_total_checks.head()
print long_gun_total_checks.head()
print multiple_total_checks.head()
```

```
year  state
1998  Alabama      10567.0
      Alaska       1365.0
      Arizona      7157.0
      Arkansas     4649.0
      California   28641.0
Name: handgun, dtype: float64
year  state
1998  Alabama      25723.0
      Alaska       2555.0
      Arizona      8275.0
      Arkansas    16914.0
      California   33438.0
Name: long_gun, dtype: float64
year  state
1998  Alabama        278
      Alaska         61
      Arizona        278
      Arkansas       189
      California      0
Name: multiple, dtype: int64
```

Creating 2 DataFrames for handgun and long_gun background checks which will be similar to census_data. We will take only the data from 2010 to 2017 because the census data available only for this years.

In [14]:

```
def creating_data_frame(total_checks,start_year,end_year):
    result=pd.DataFrame(index=months_activity['state'].unique())
    for year in range(start_year,end_year+1):
        result[year]=total_checks[year]
    return result
```

In [15]:

```
handgun_data=creating_data_frame(hand_gun_total_checks,2010,2017)
long_gun_data=creating_data_frame(long_gun_total_checks,2010,2017)

print handgun_data.head()
print long_gun_data.head()
```

	2010	2011	2012	2013	2014	
2015 \						
Alabama	116963.0	130490.0	177239.0	195052.0	178355.0	2437
34.0						
Alaska	25720.0	28291.0	35851.0	36678.0	35694.0	370
86.0						
Arizona	89162.0	108245.0	139050.0	138744.0	123375.0	1484
00.0						
Arkansas	49992.0	54659.0	69481.0	74884.0	60520.0	773
94.0						
California	245159.0	303169.0	375569.0	407355.0	495962.0	4778
22.0						

	2016	2017				
Alabama	153123.0	97751.0				
Alaska	37491.0	34556.0				
Arizona	166784.0	153522.0				
Arkansas	80244.0	72100.0				
California	560355.0	512465.0				

	2010	2011	2012	2013	2014	
2015 \						
Alabama	136519.0	146290.0	184072.0	203530.0	170815.0	1760
75.0						
Alaska	33177.0	34607.0	40505.0	44001.0	39625.0	363
72.0						
Arizona	69067.0	81285.0	107079.0	110309.0	91086.0	931
55.0						
Arkansas	72206.0	78806.0	94539.0	94837.0	82670.0	830
97.0						
California	275458.0	320988.0	423207.0	526191.0	369009.0	3431
27.0						

	2016	2017
Alabama	121963.0	86210.0
Alaska	36887.0	32648.0
Arizona	108988.0	99248.0
Arkansas	82120.0	76765.0
California	554550.0	318133.0

Now we want to know the ratio of background checks to the population of each state. We are going to divide background checks by a population of a state. We have two exactly the same data frame structure for both: background checks and census data.

In [16]:

```
ratio_background_checks_handgun=handgun_data.div(census_data)
ratio_background_checks_handgun # handgun
ratio_background_checks_long_gun=long_gun_data.div(census_data)
ratio_background_checks_long_gun #long_gun
```

Out[16]:

	2010	2011	2012	2013	2014	2015	20
Alabama	0.028527	0.030486	0.038237	0.042159	0.035292	0.036298	0.0250
Alaska	0.046465	0.047915	0.055424	0.059722	0.053783	0.049286	0.0497
Arizona	0.010780	0.012572	0.016362	0.016673	0.013582	0.013695	0.0157
Arkansas	0.024713	0.026817	0.032056	0.032074	0.027884	0.027926	0.0274
California	0.007379	0.008520	0.011131	0.013722	0.009535	0.008791	0.0141
Colorado	0.025682	0.028568	0.032965	0.035727	0.029212	0.029984	0.0325
Connecticut	0.011325	0.011536	0.015007	0.015949	0.016519	0.013453	0.0122
Delaware	0.010521	0.011755	0.015131	0.017753	0.020615	0.021230	0.0228
District of Columbia	0.000018	0.000015	0.000028	0.000057	0.000045	0.000048	0.0000
Florida	0.009905	0.011123	0.015345	0.016670	0.013807	0.013704	0.0153
Georgia	0.009837	0.011426	0.013942	0.014019	0.012318	0.012404	0.0124
Hawaii	0.000000	0.000000	0.000001	0.000000	0.000000	0.000000	0.0000
Idaho	0.026783	0.030292	0.035782	0.035415	0.031207	0.032407	0.0330
Illinois	0.009092	0.010413	0.013542	0.014327	0.012048	0.011419	0.0138
Indiana	0.017635	0.019955	0.026670	0.027967	0.024147	0.024261	0.0267
Iowa	0.013245	0.013656	0.014872	0.013389	0.011371	0.010964	0.0112
Kansas	0.023636	0.025853	0.031345	0.031257	0.024542	0.024059	0.0253
Kentucky	0.022979	0.025528	0.029562	0.029008	0.024730	0.024794	0.0245
Louisiana	0.026239	0.028446	0.034278	0.036233	0.031538	0.030110	0.0296
Maine	0.028600	0.030752	0.036619	0.038373	0.032556	0.031028	0.0351
Maryland	0.007871	0.009187	0.011192	0.014672	0.010731	0.010849	0.0119
Massachusetts	0.003336	0.003751	0.005261	0.005969	0.004942	0.005026	0.0061
Michigan	0.013062	0.013610	0.016138	0.015694	0.012685	0.012797	0.0129
Minnesota	0.022122	0.023825	0.027637	0.028073	0.024384	0.024954	0.0259
Mississippi	0.027247	0.029189	0.034986	0.033561	0.029509	0.030202	0.0310
Missouri	0.023677	0.026645	0.033750	0.035193	0.031164	0.030747	0.0332

Montana	0.048625	0.050535	0.059243	0.059230	0.054242	0.055965	0.0547
Nebraska	0.015378	0.016727	0.018739	0.017022	0.014012	0.014258	0.0138
Nevada	0.013617	0.015504	0.018763	0.017272	0.012645	0.011903	0.0129
New Hampshire	0.021235	0.023369	0.033009	0.035160	0.029282	0.028751	0.0334
New Jersey	0.002730	0.003166	0.004335	0.005642	0.004547	0.004524	0.0054
New Mexico	0.020485	0.021748	0.027623	0.029408	0.025394	0.025086	0.0266
New York	0.007603	0.008376	0.010928	0.010249	0.009132	0.009295	0.0099
North Carolina	0.013463	0.015385	0.019593	0.020253	0.016611	0.016630	0.0174
North Dakota	0.044793	0.053787	0.063483	0.062058	0.052693	0.048989	0.0450
Ohio	0.013843	0.016139	0.021591	0.023016	0.020321	0.020144	0.0209
Oklahoma	0.024221	0.029016	0.038410	0.040029	0.034950	0.032598	0.0332
Oregon	0.025613	0.027597	0.034377	0.033983	0.030017	0.031215	0.0350
Pennsylvania	0.048963	0.056075	0.068416	0.059919	0.011850	0.009243	0.0123
Rhode Island	0.005756	0.006882	0.009981	0.010570	0.007982	0.007957	0.0096
South Carolina	0.013229	0.015030	0.018257	0.019045	0.016402	0.017375	0.0176
South Dakota	0.050725	0.052958	0.063419	0.063649	0.057241	0.058924	0.0574
Tennessee	0.021327	0.024732	0.031945	0.033347	0.029028	0.030101	0.0316
Texas	0.014317	0.016514	0.020532	0.021470	0.018707	0.017731	0.0172
Utah	0.014952	0.016588	0.018912	0.018791	0.016105	0.016489	0.0177
Vermont	0.022675	0.024562	0.029715	0.030194	0.026126	0.026734	0.0302
Virginia	0.018126	0.018428	0.024799	0.027785	0.022563	0.021995	0.0249
Washington	0.015405	0.017123	0.022472	0.022889	0.018456	0.018128	0.0207
West Virginia	0.038886	0.044309	0.053224	0.057125	0.048482	0.046681	0.0443
Wisconsin	0.022336	0.022939	0.027628	0.028867	0.025197	0.025915	0.0276
Wyoming	0.038077	0.041504	0.049016	0.051128	0.046575	0.044811	0.0441

I have decided to delete 'Hawaii' because according to the data, this state doesn't give us any information.

In [17]:

```
ratio_background_checks_handgun=ratio_background_checks_handgun.drop(['Hawaii'])
ratio_background_checks_long_gun=ratio_background_checks_long_gun.drop(['Hawaii'])
```

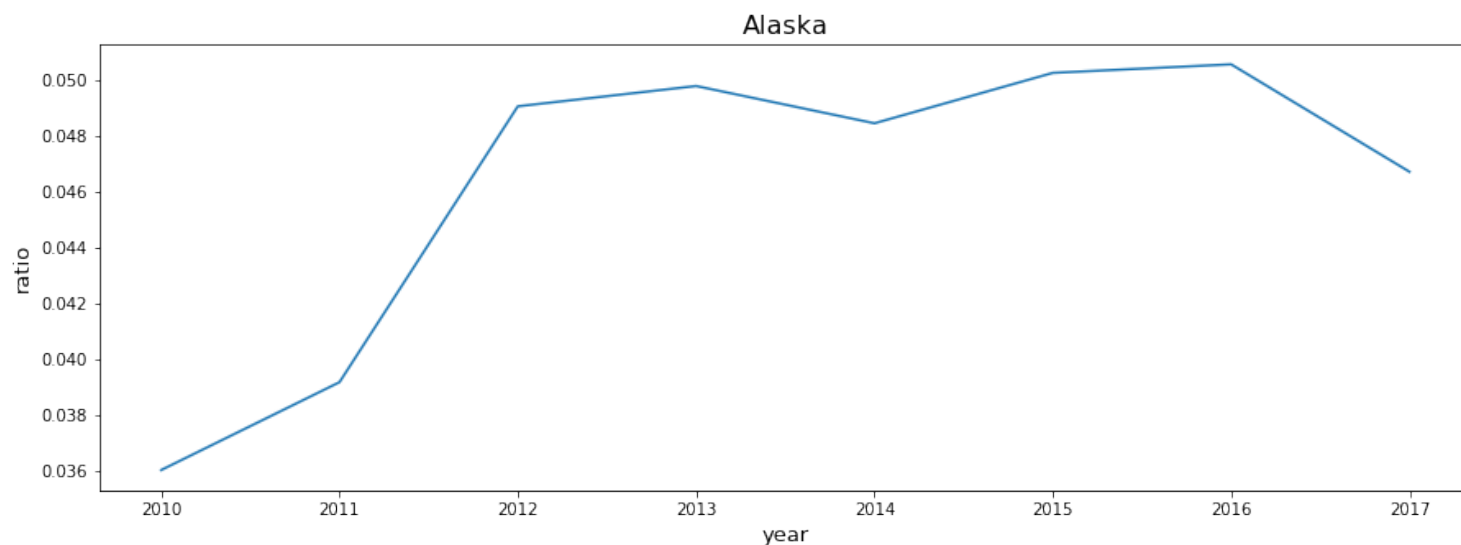
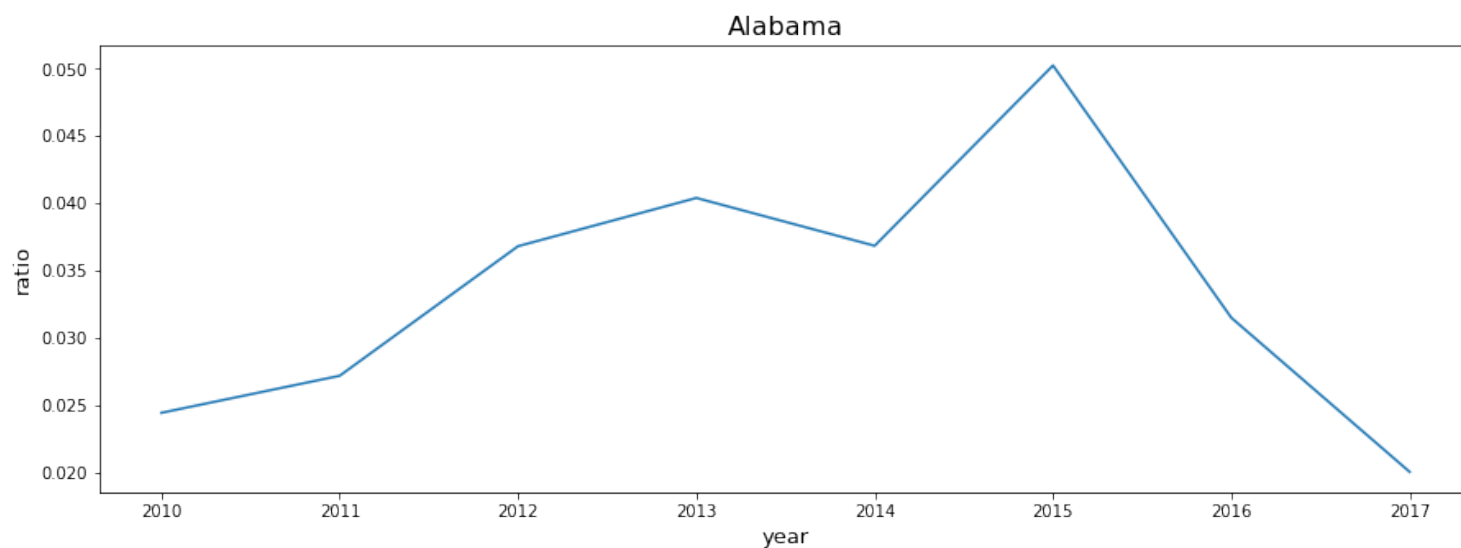
Creating a new function "plotting_state" which creates a plot for the state. And now we are ready to show trends in gun background checks(handgun) in each state from 2010 to 2017. Let's check what we can find.

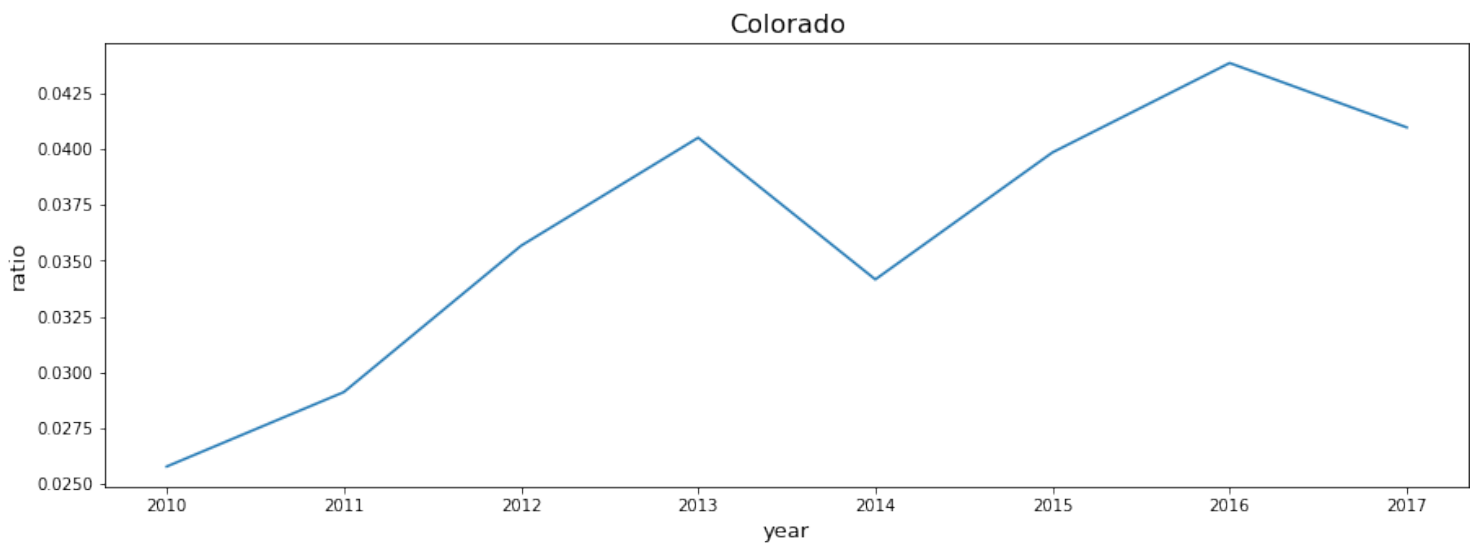
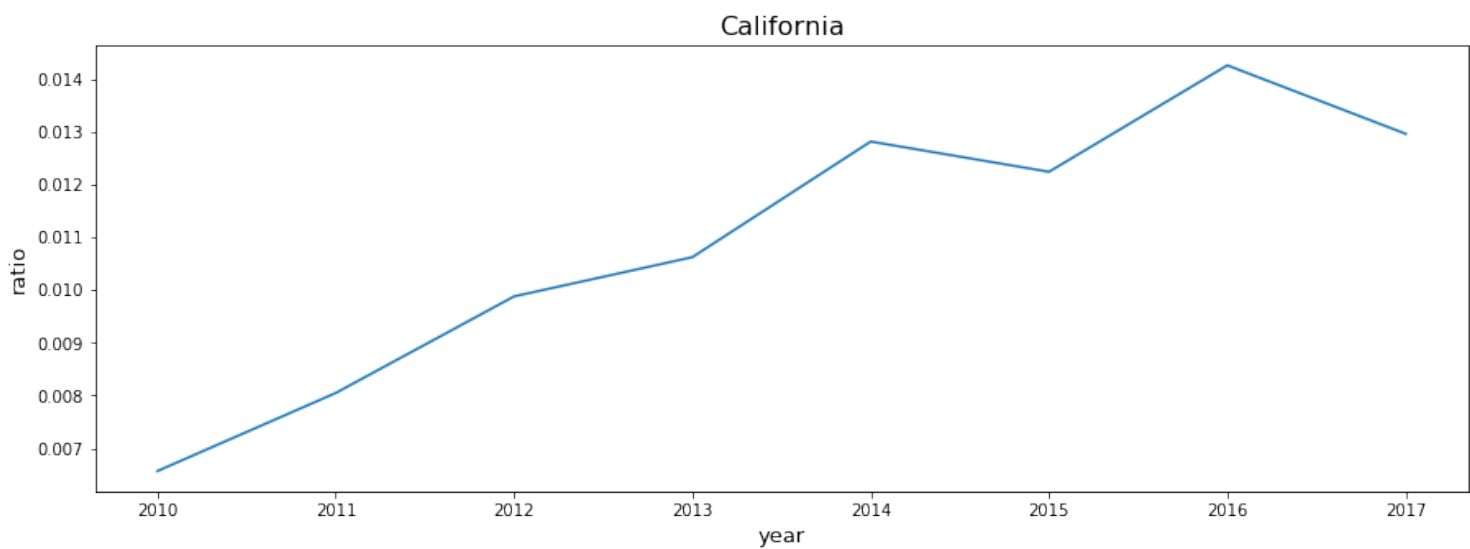
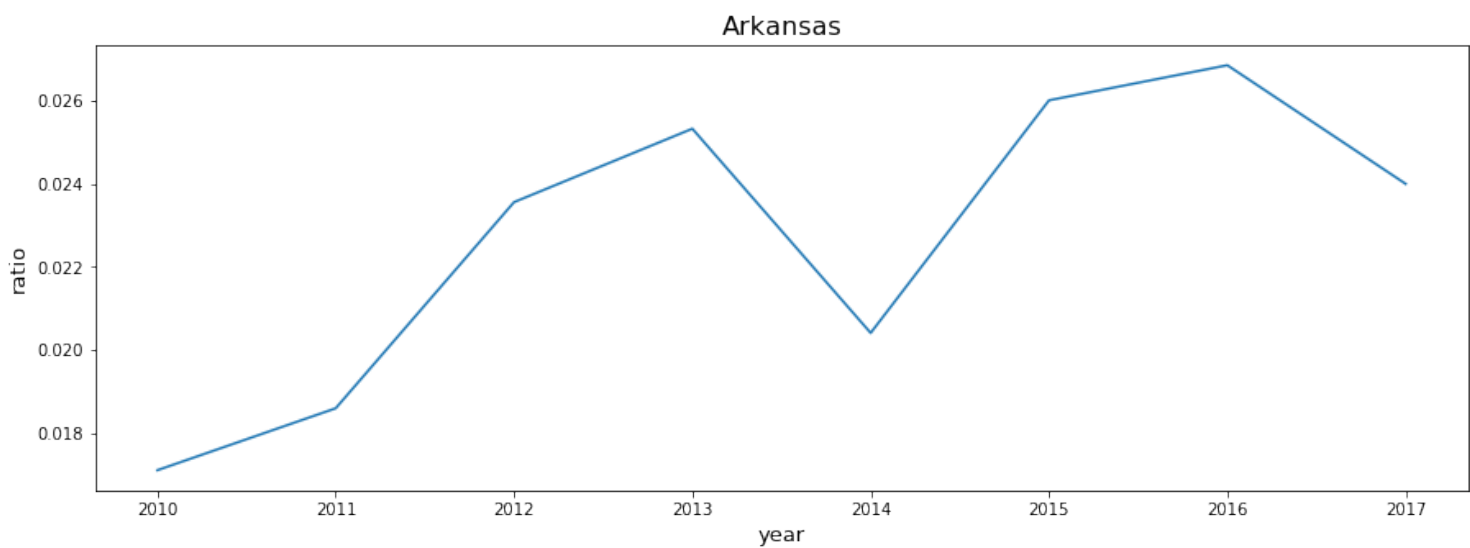
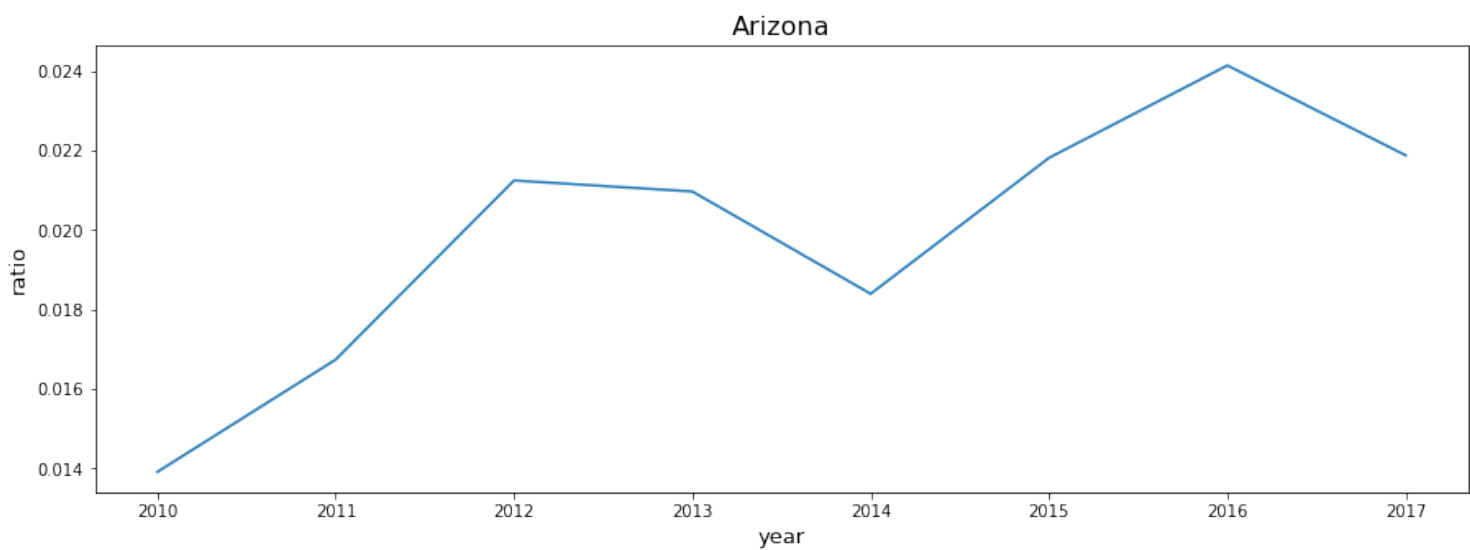
In [18]:

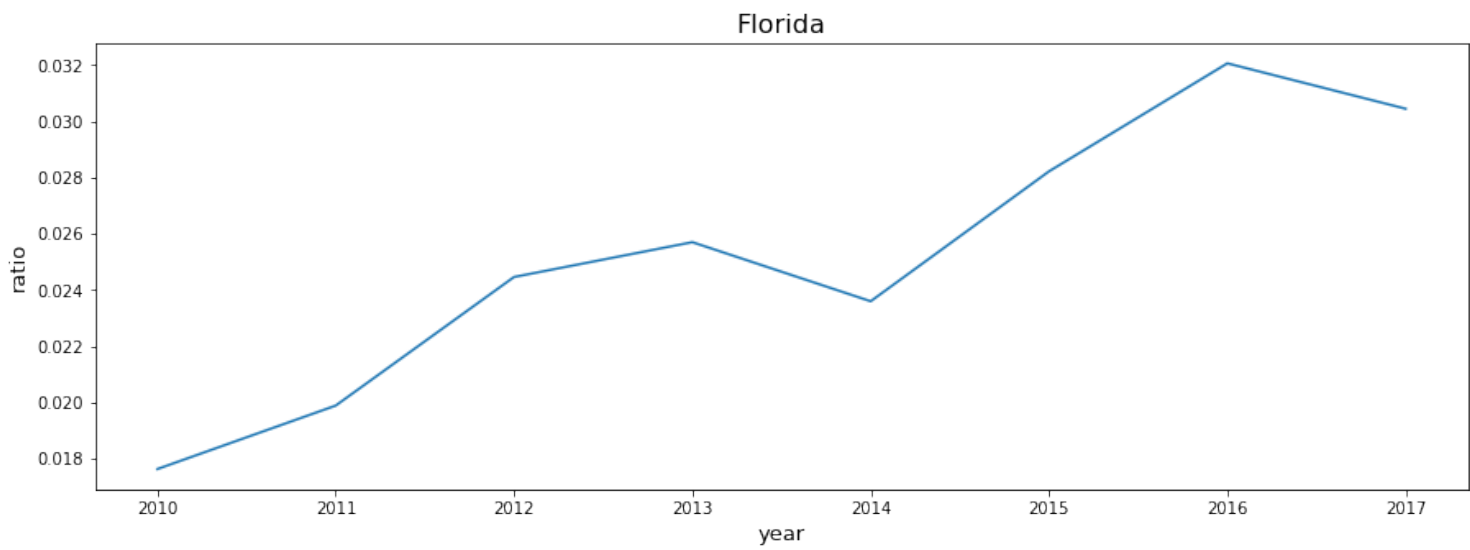
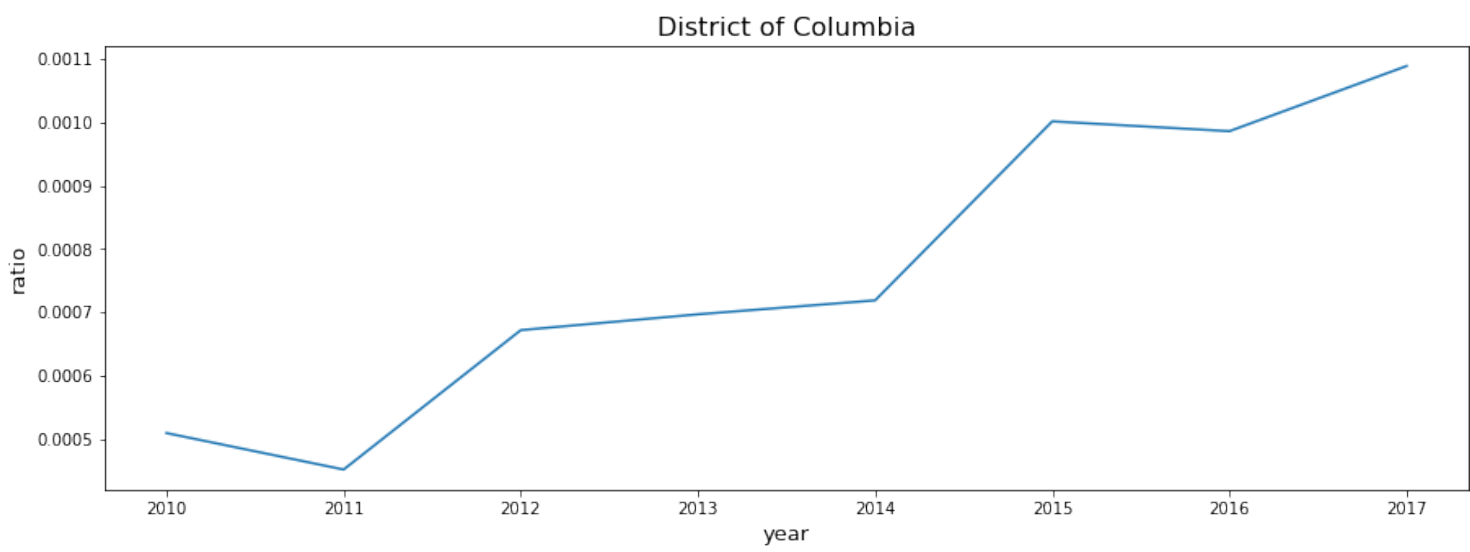
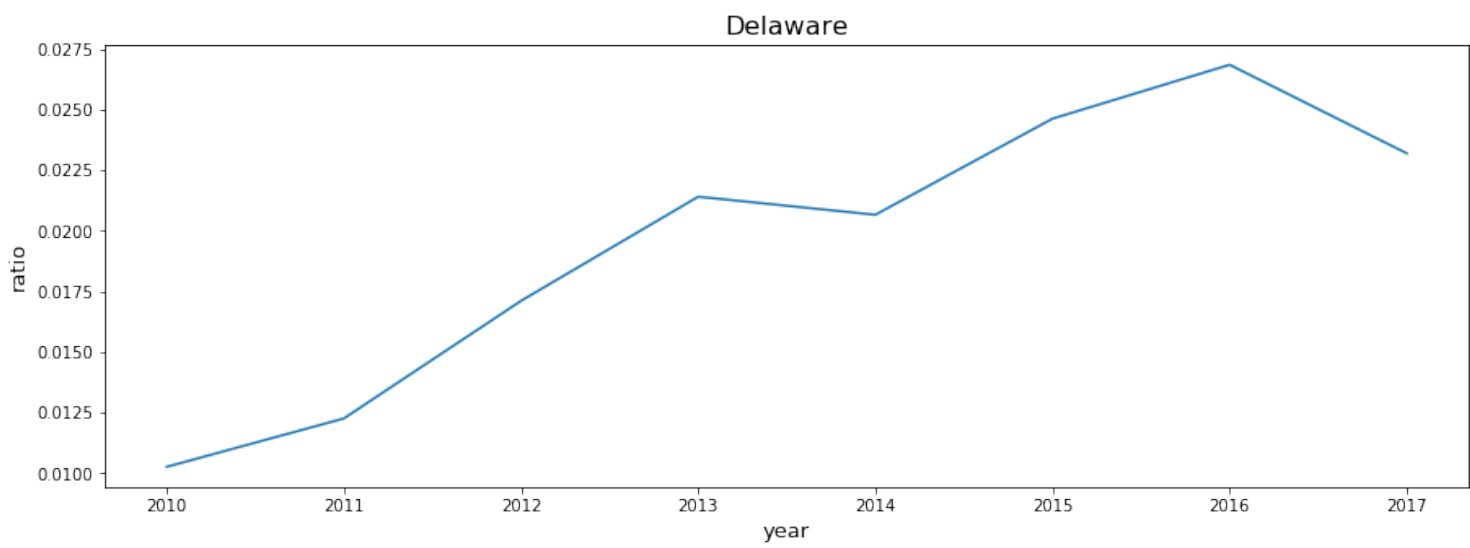
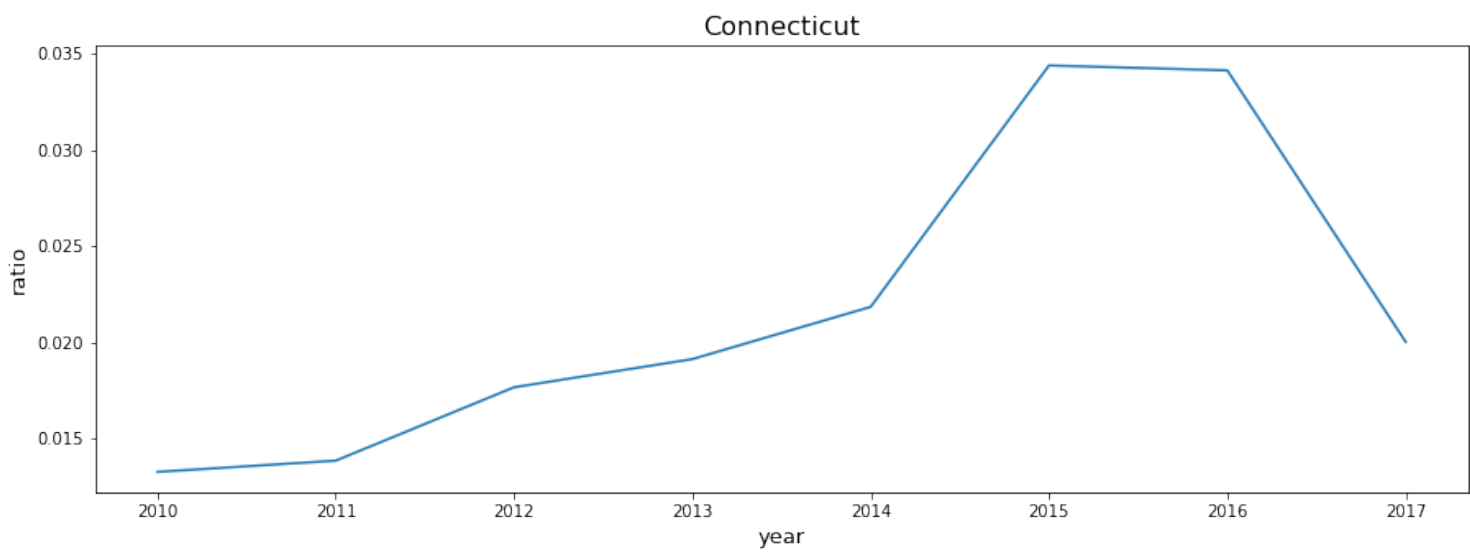
```
def plotting_state(state):  
    plt.figure(figsize=(15,5))  
    state.plot()  
    plt.title(state.name,fontsize=16)  
    plt.xlabel('year',fontsize=13)  
    plt.ylabel('ratio',fontsize=13)  
    for e in range(len(ratio_background_checks_handgun)):  
        plotting_state(ratio_background_checks_handgun.iloc[e])
```

/anaconda2/lib/python2.7/site-packages/matplotlib/pyplot.py:523: RuntimeWarning: More than 20 figures have been opened. Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly closed and may consume too much memory. (To control this warning, see the rcParam `figure.max_open_warning`).

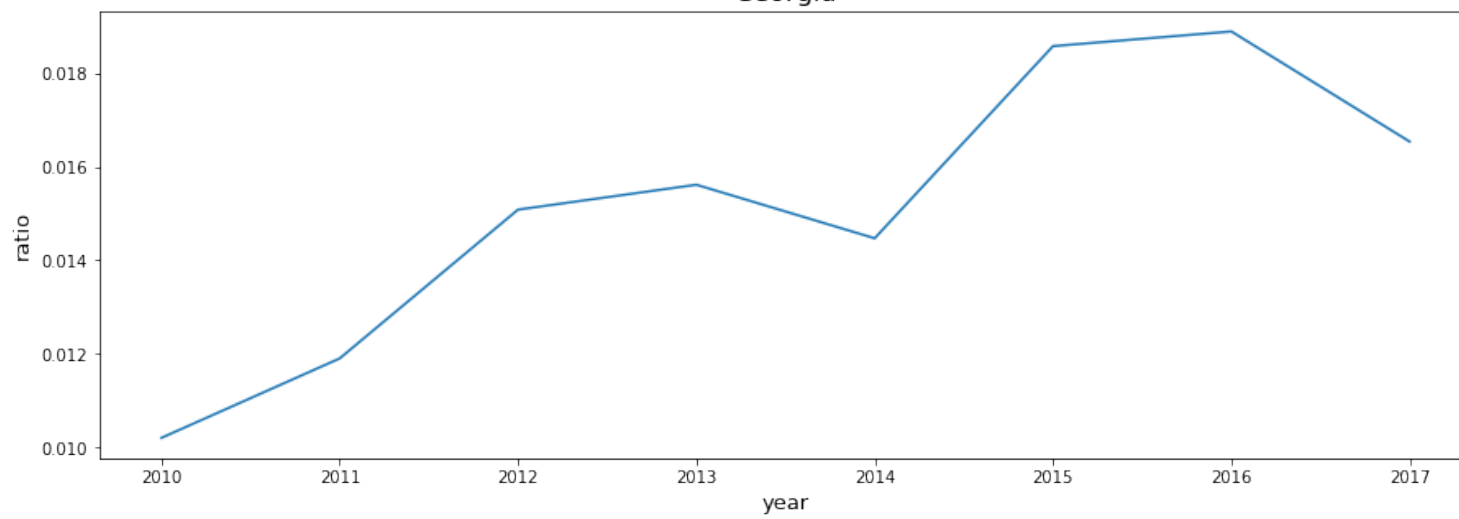
max_open_warning, RuntimeWarning)



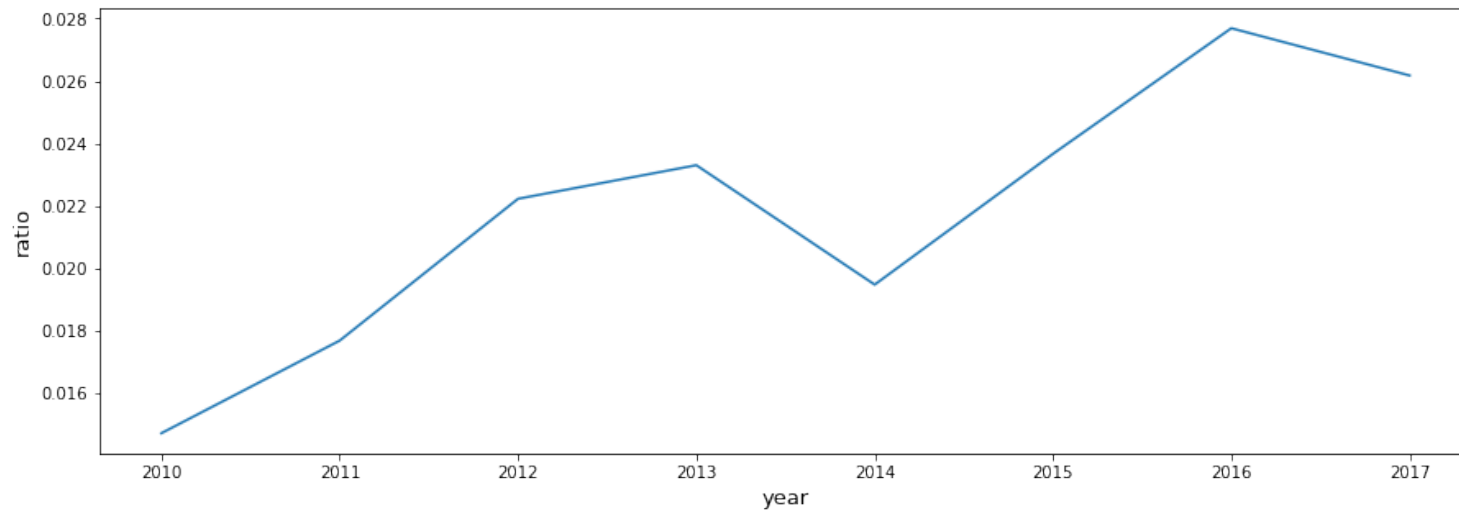




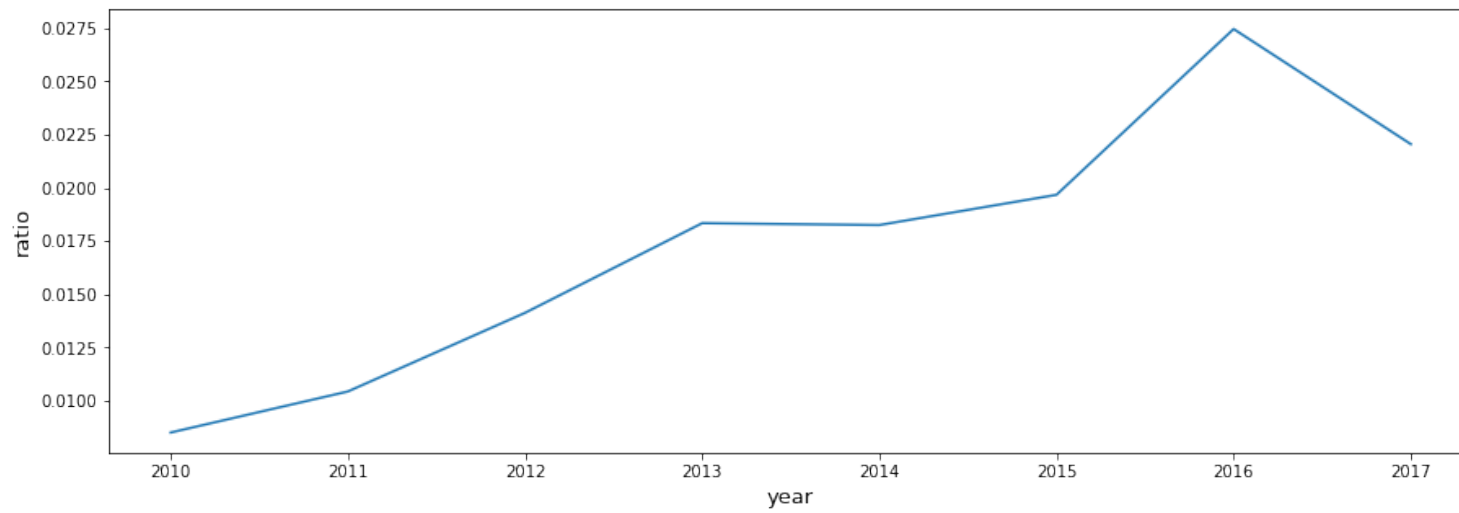
Georgia



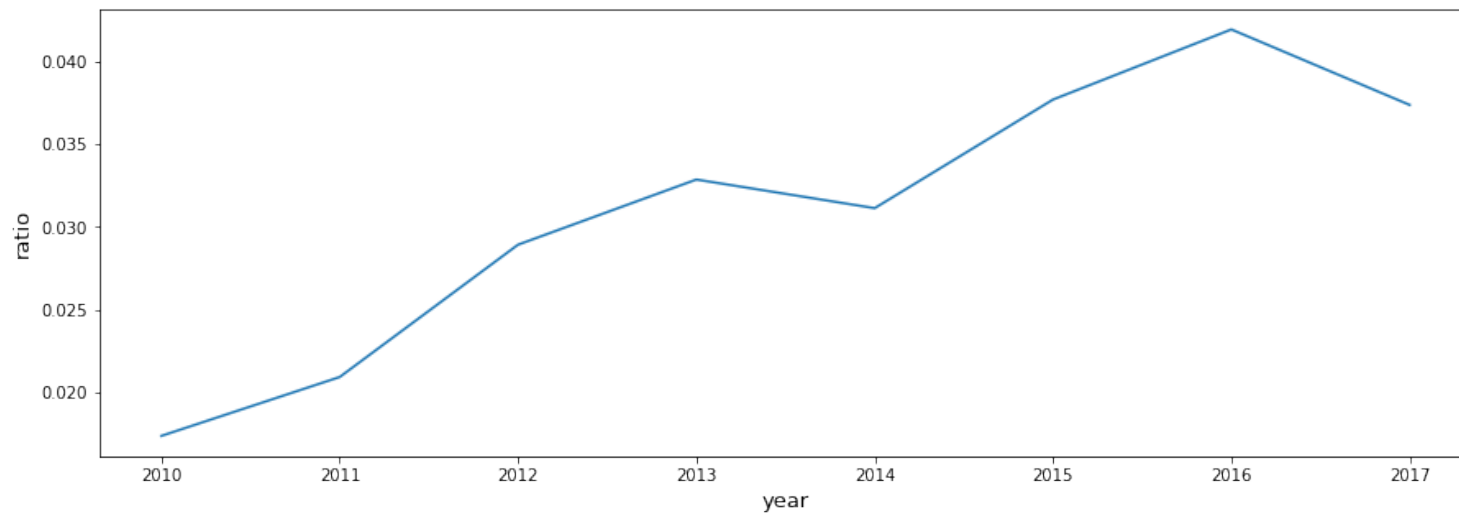
Idaho

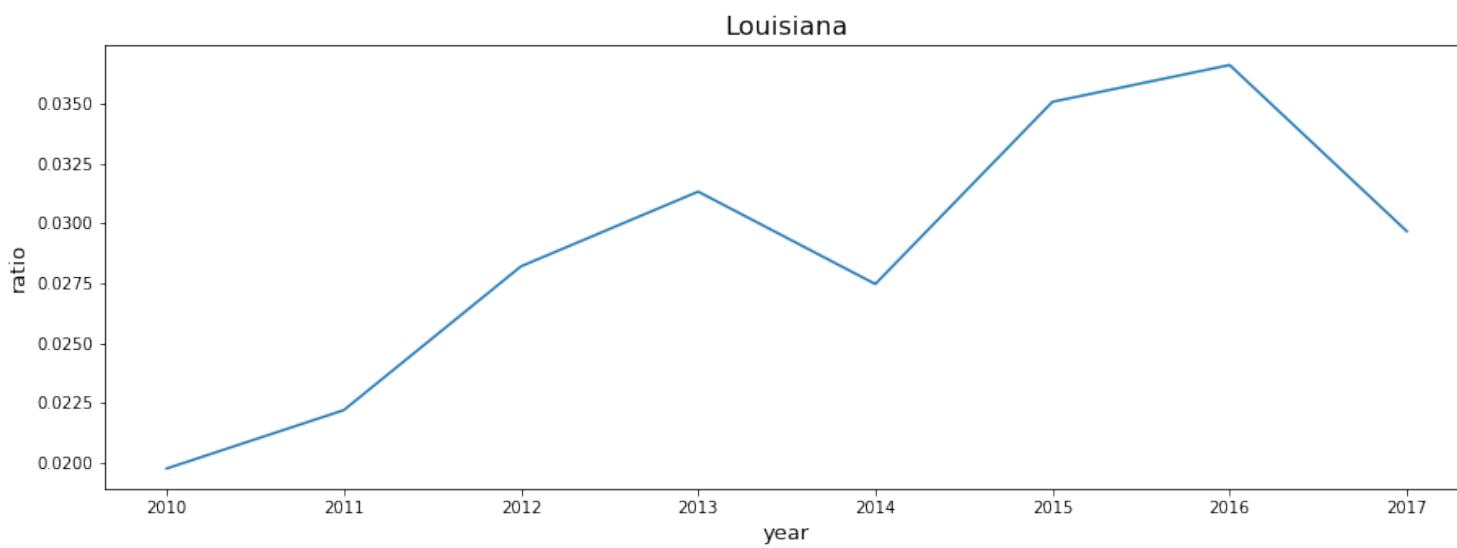
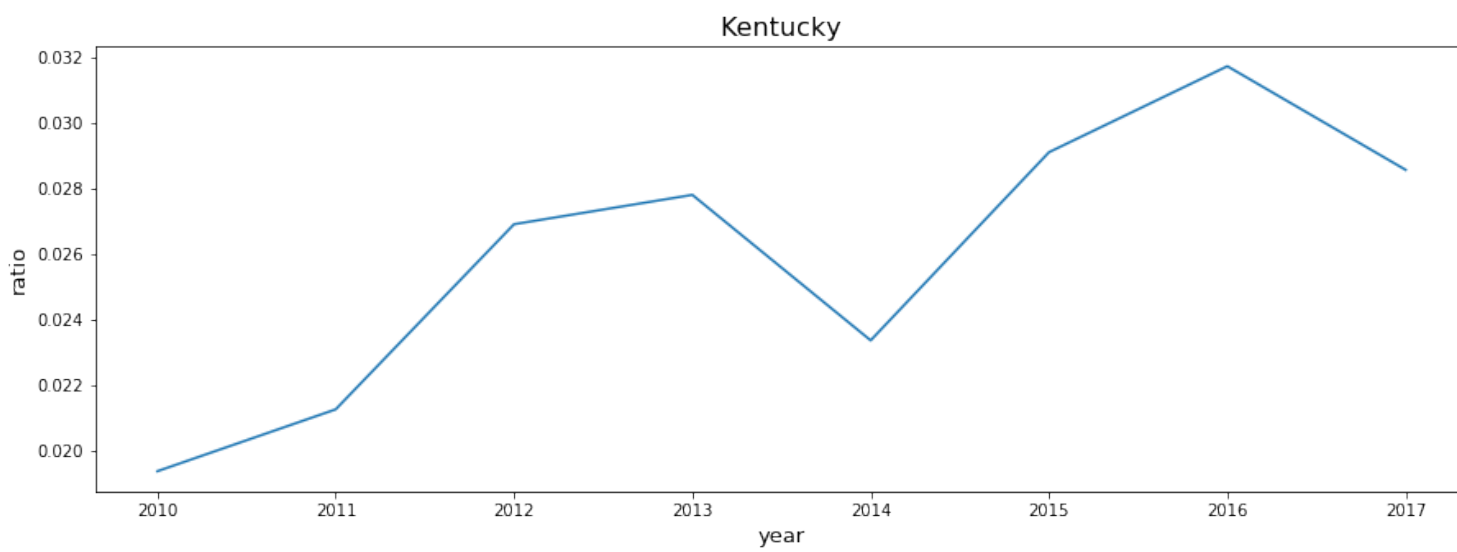
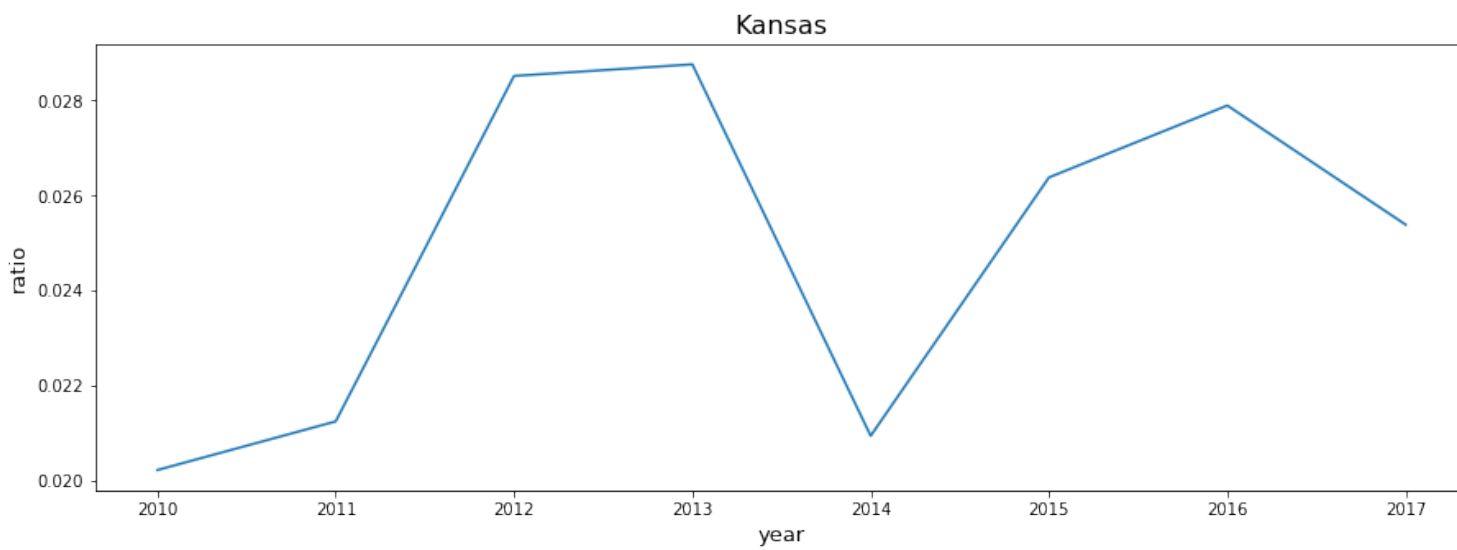
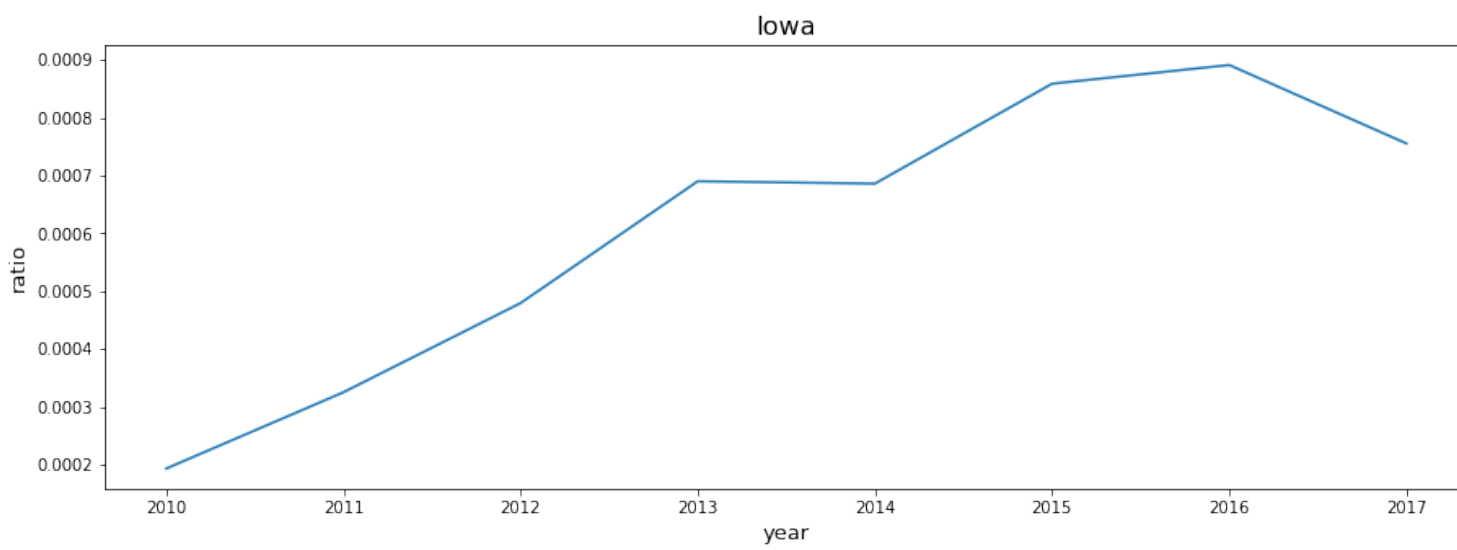


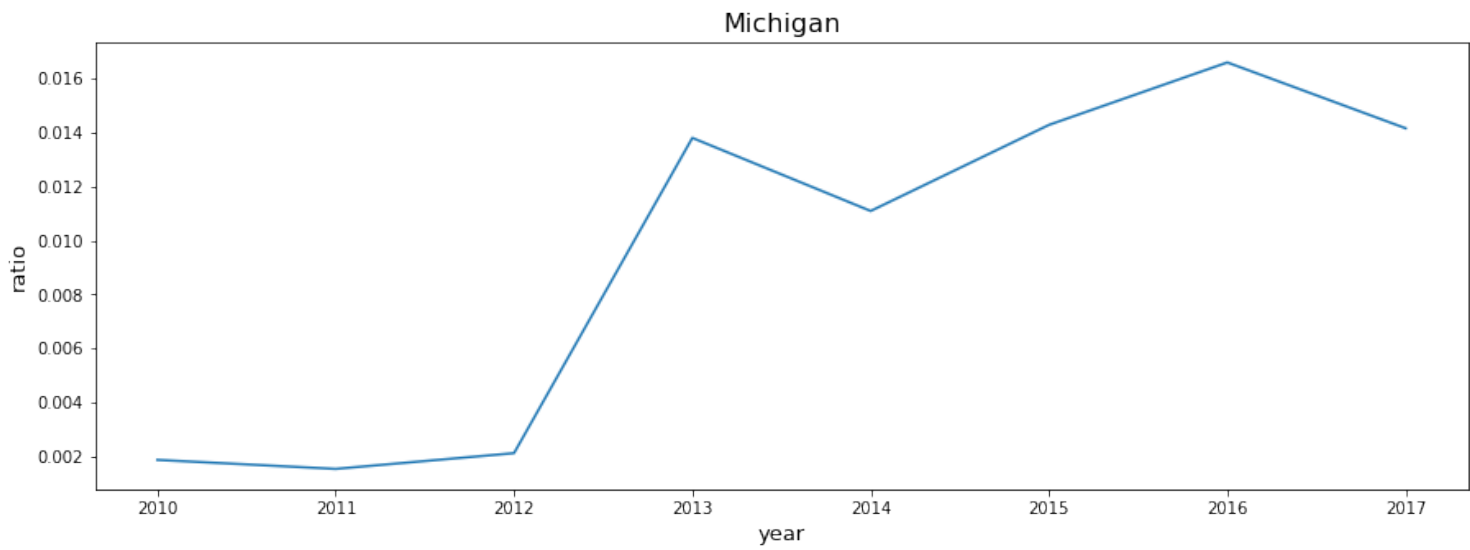
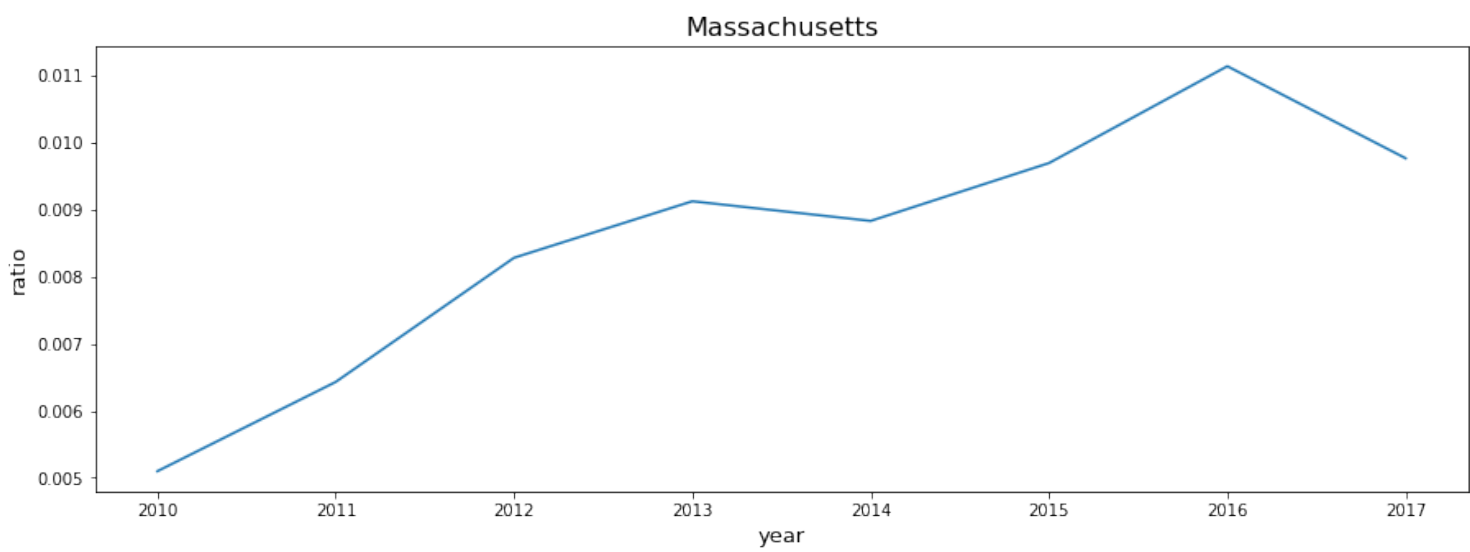
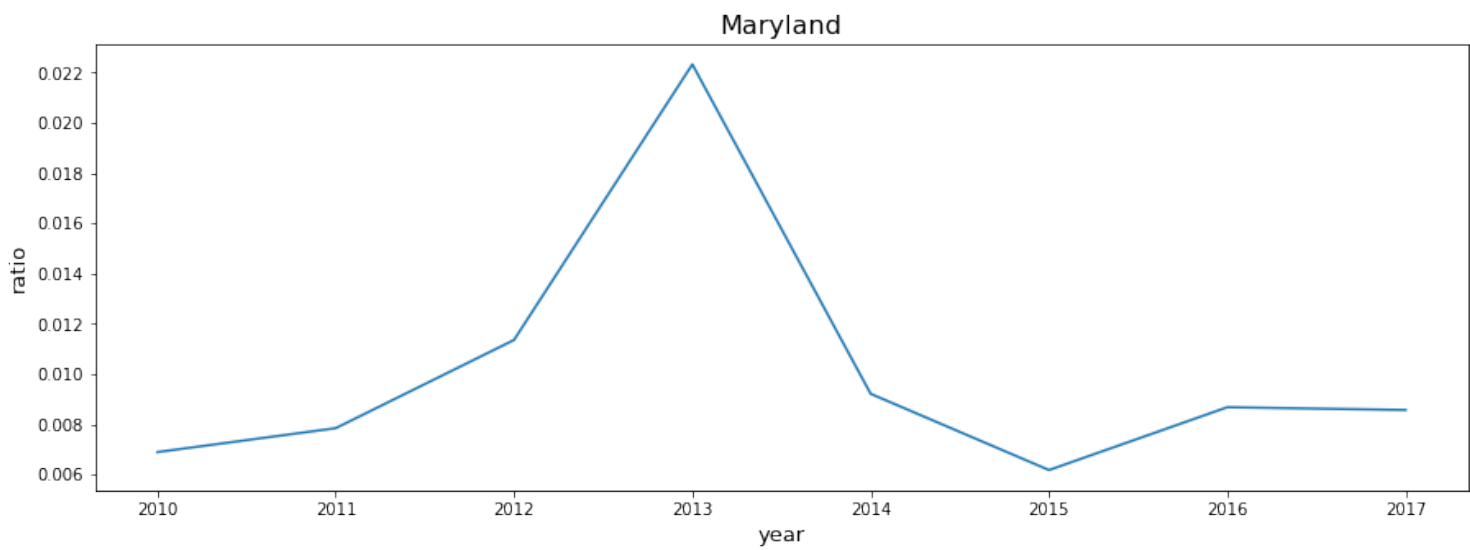
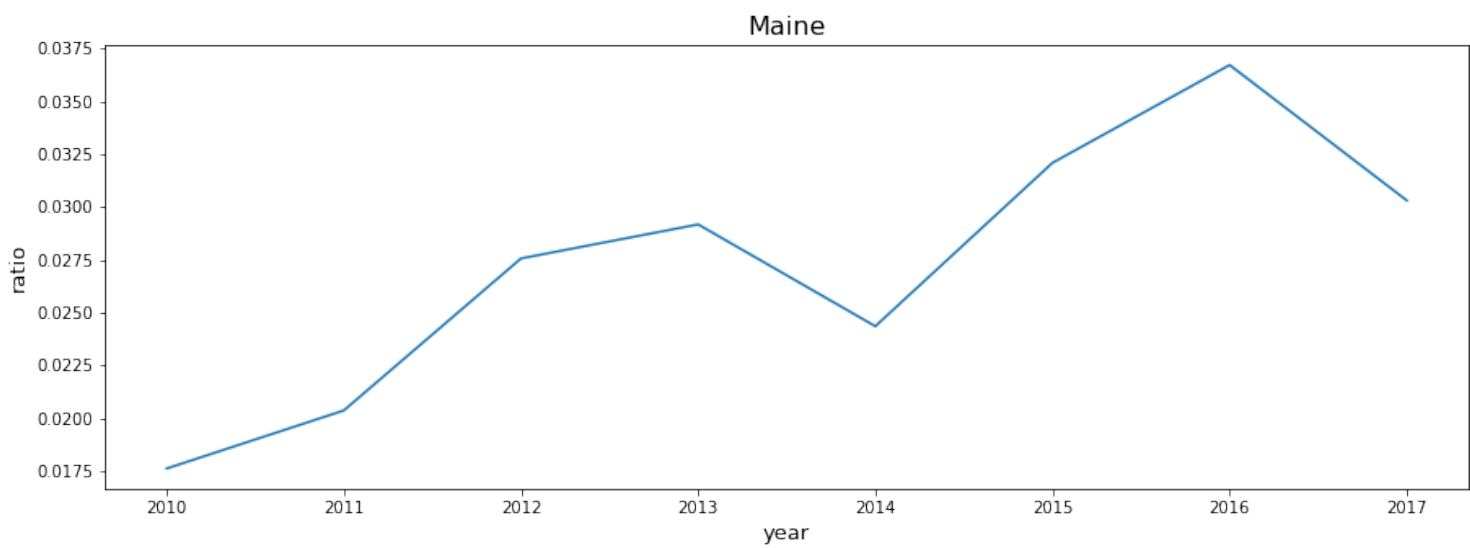
Illinois

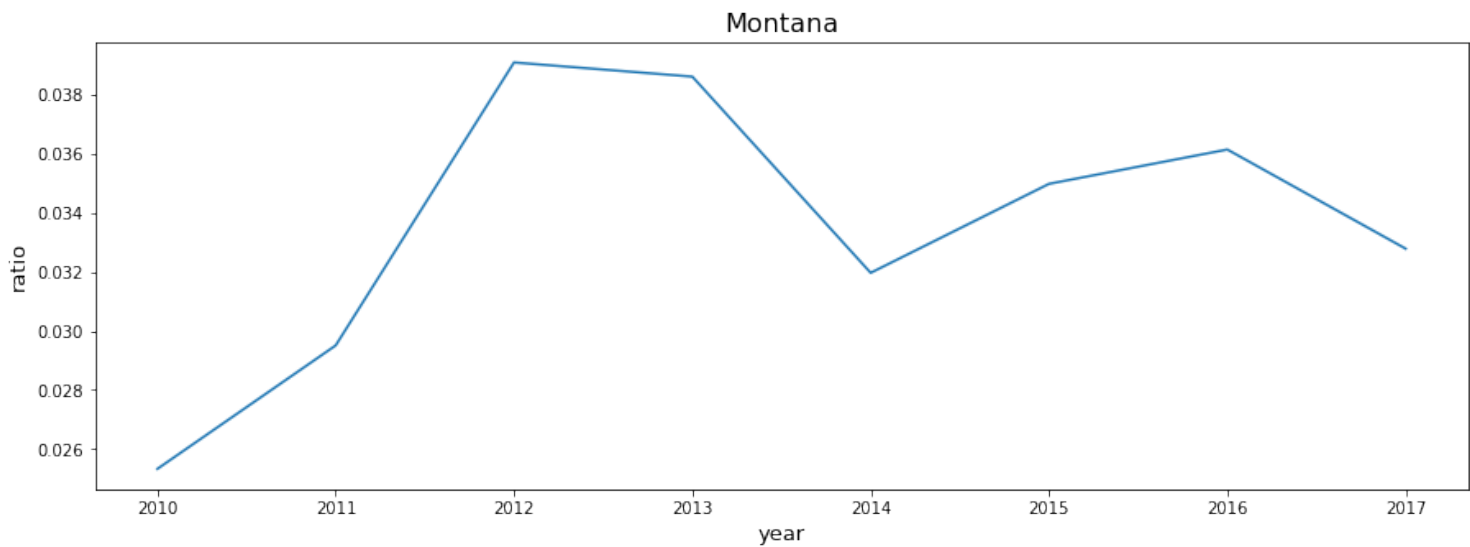
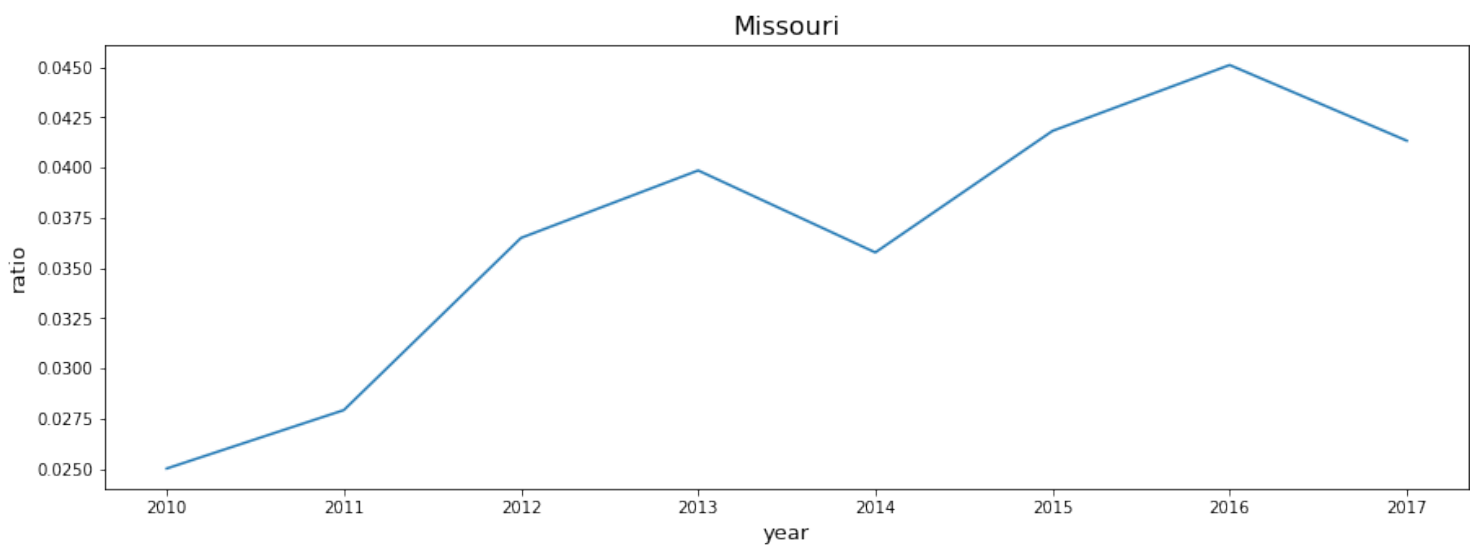
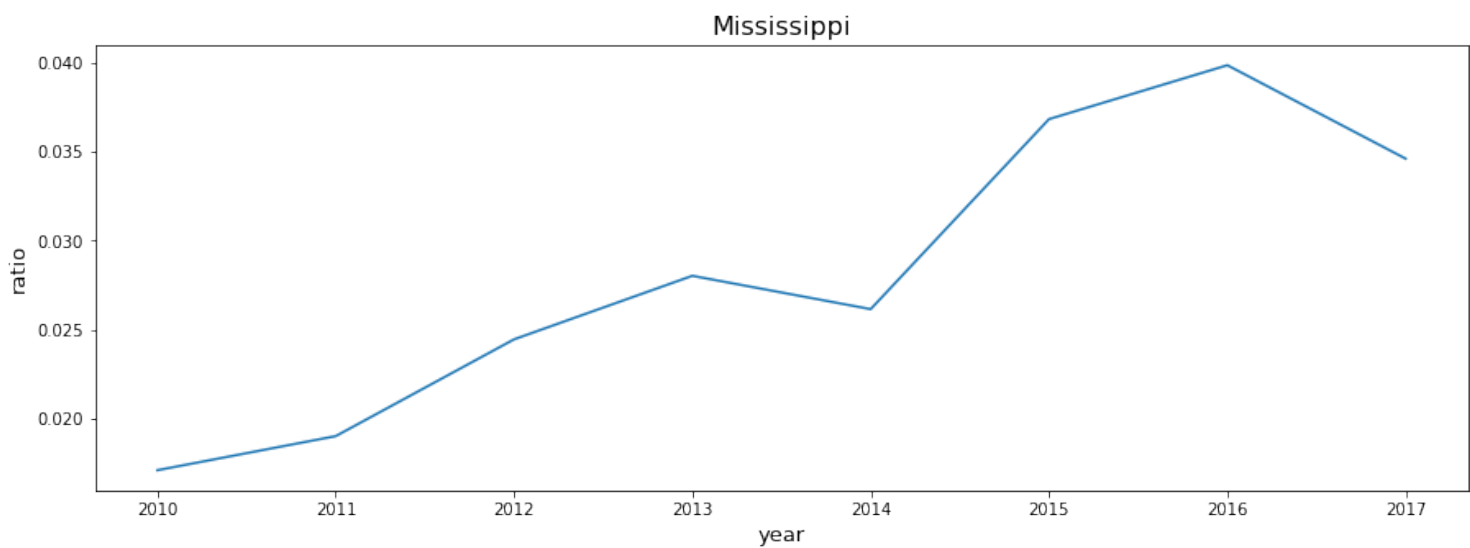
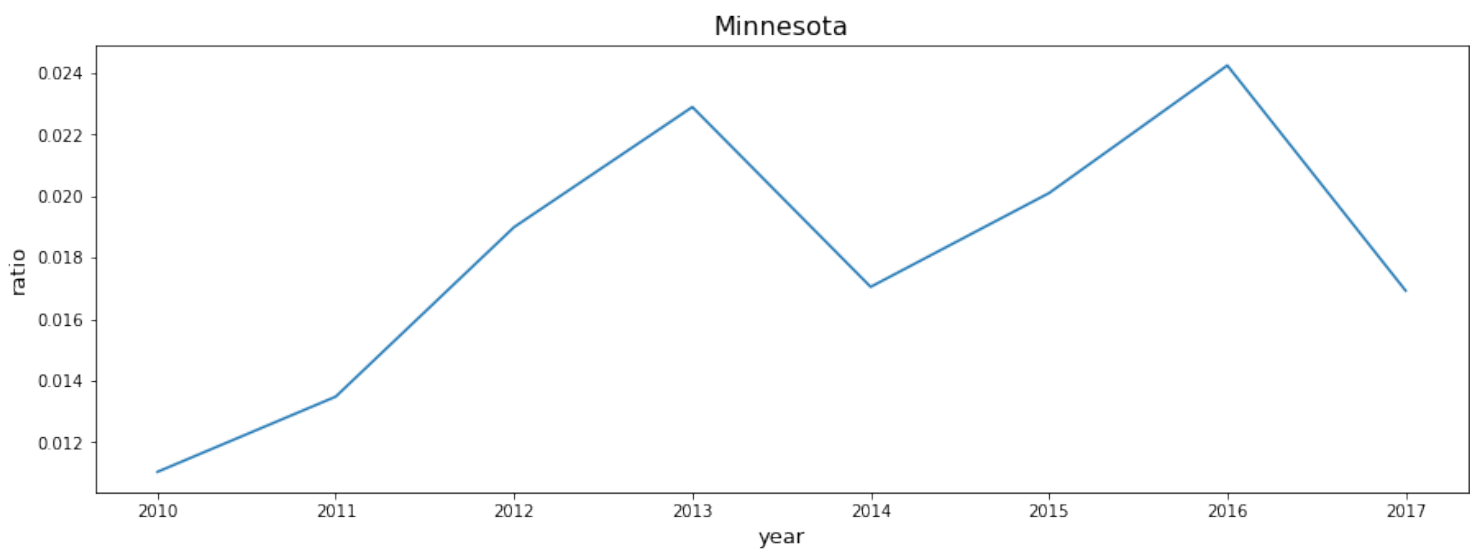


Indiana

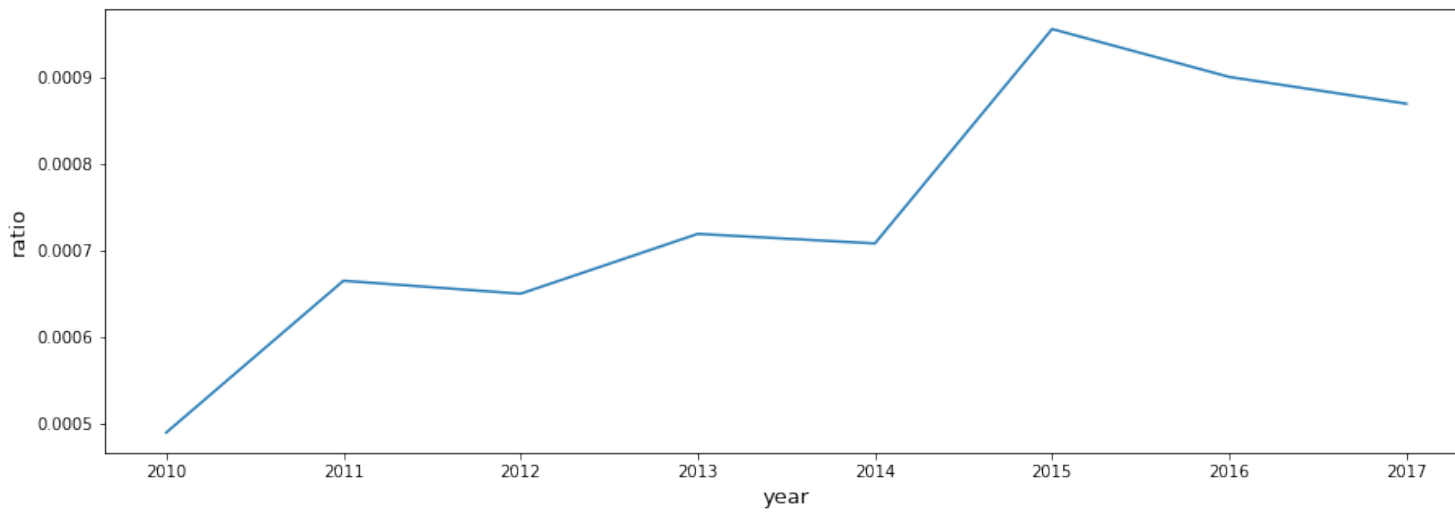




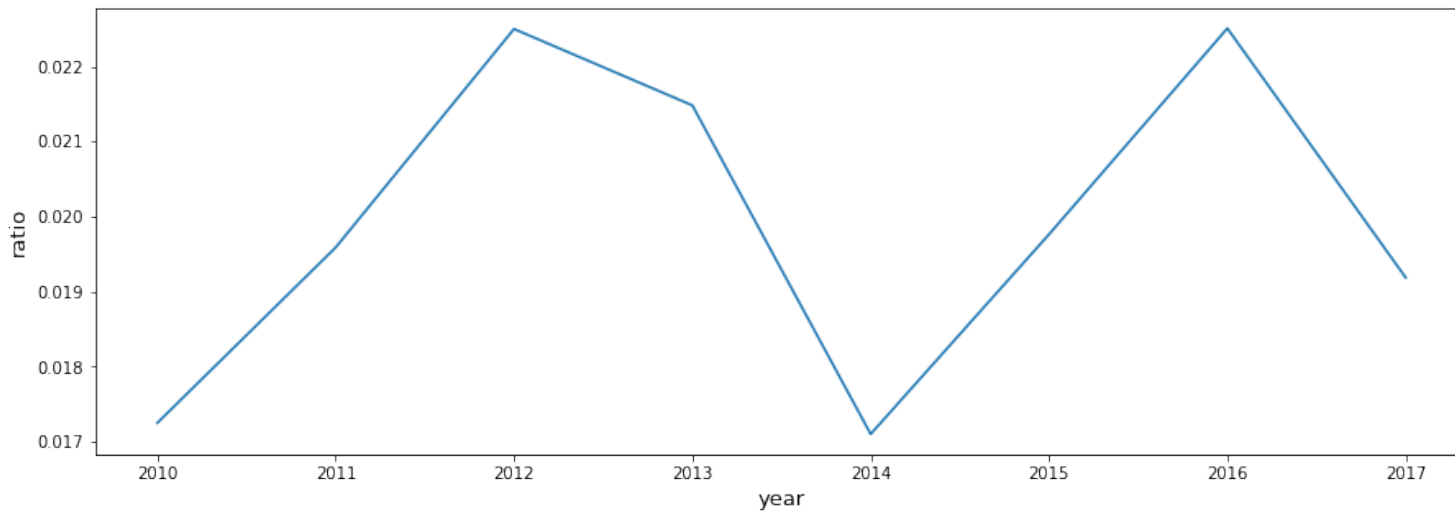




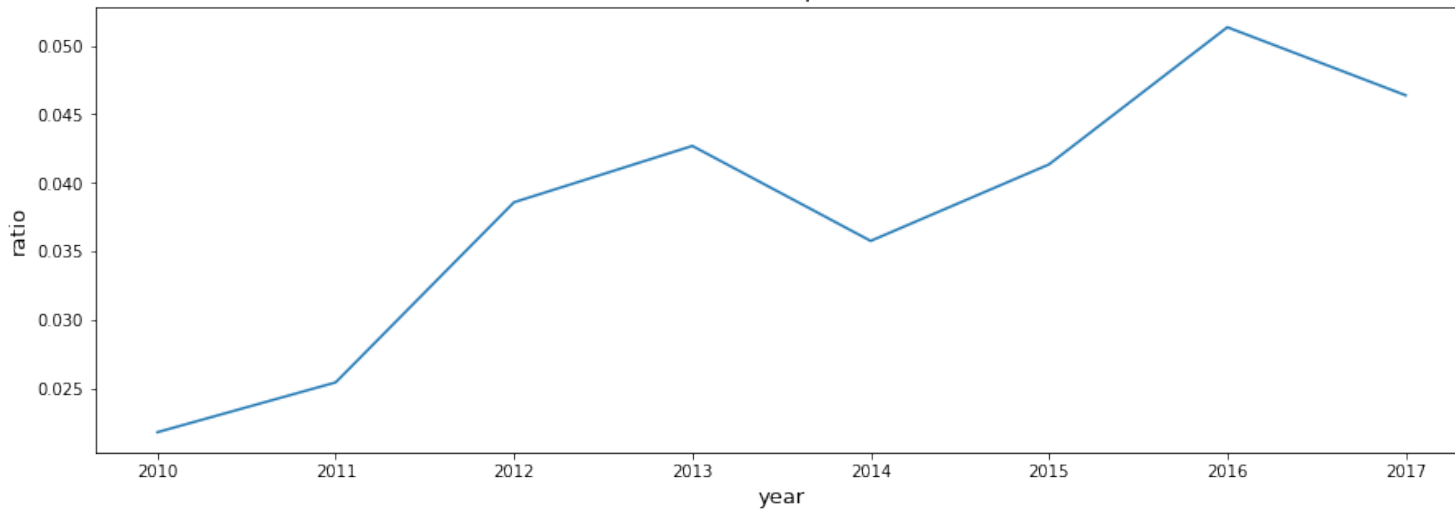
Nebraska



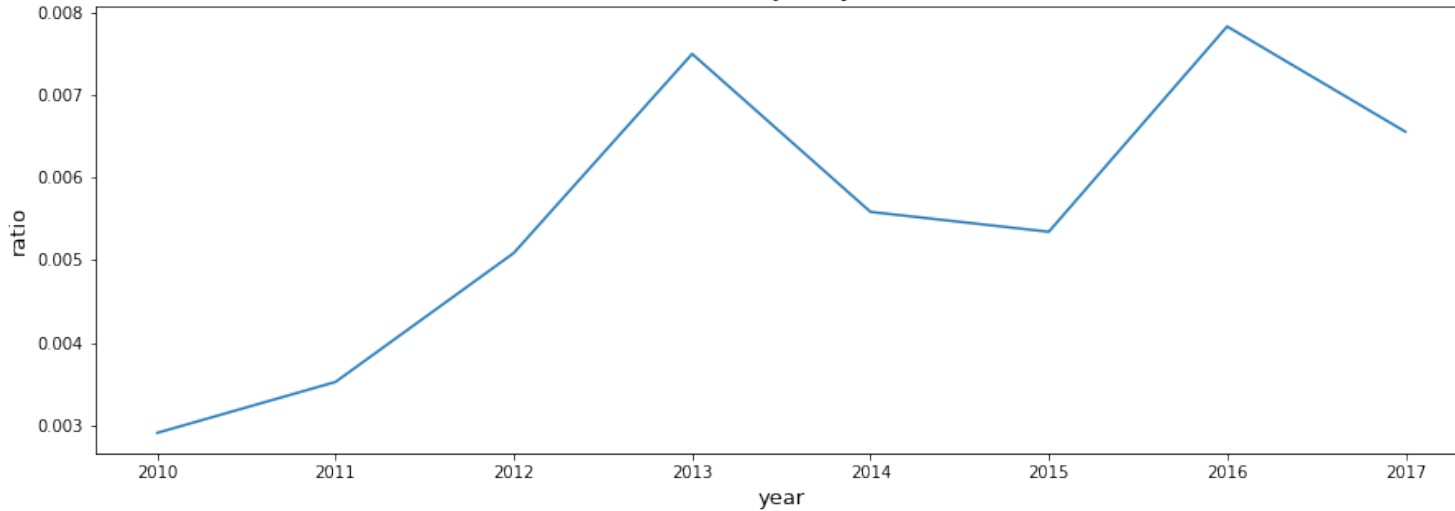
Nevada



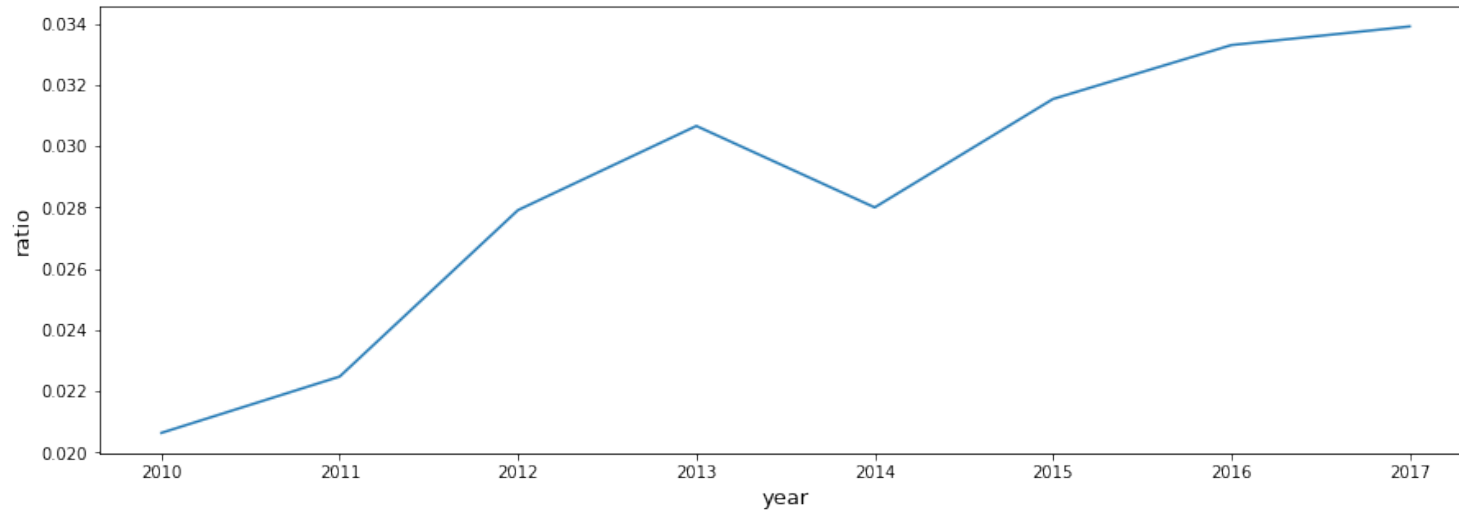
New Hampshire



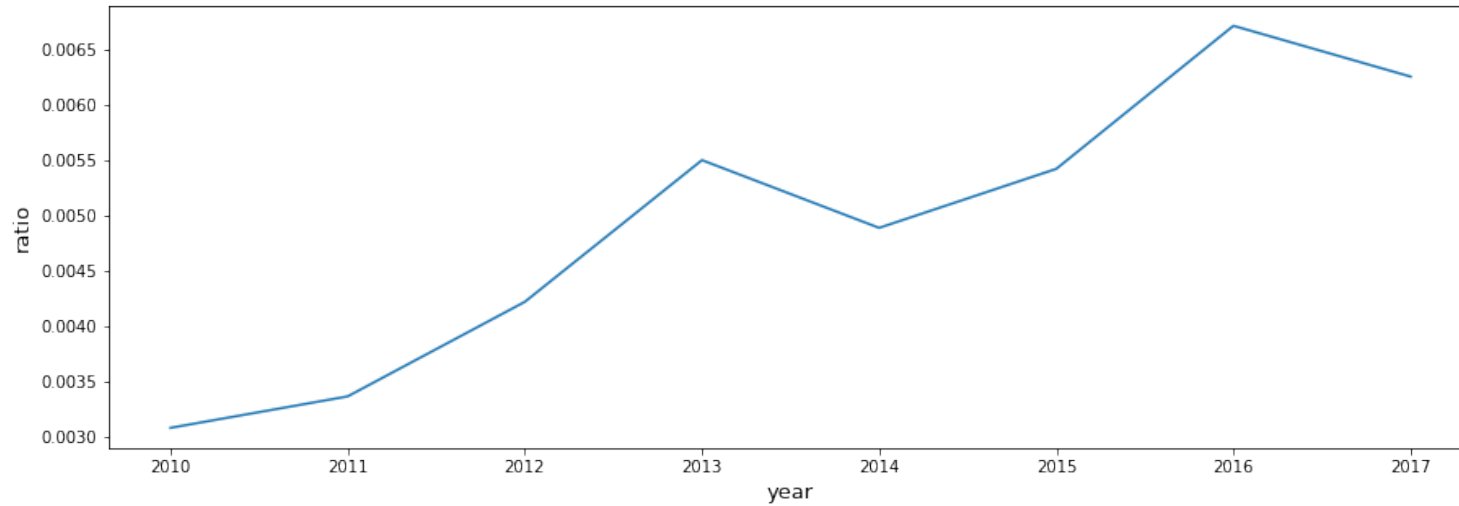
New Jersey



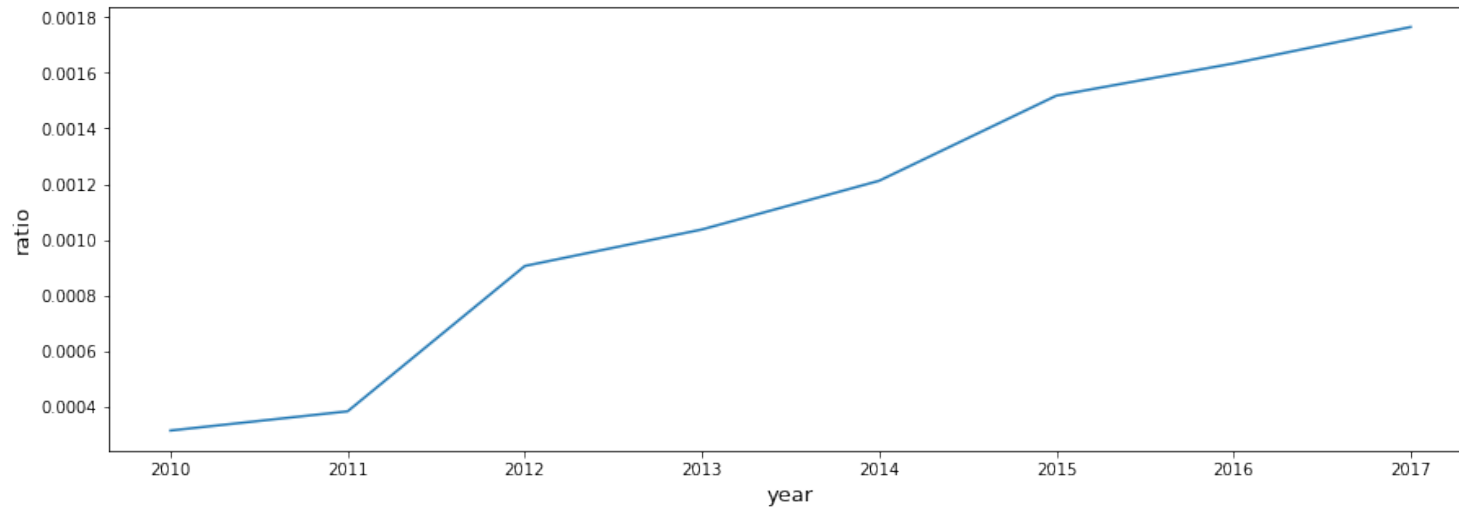
New Mexico



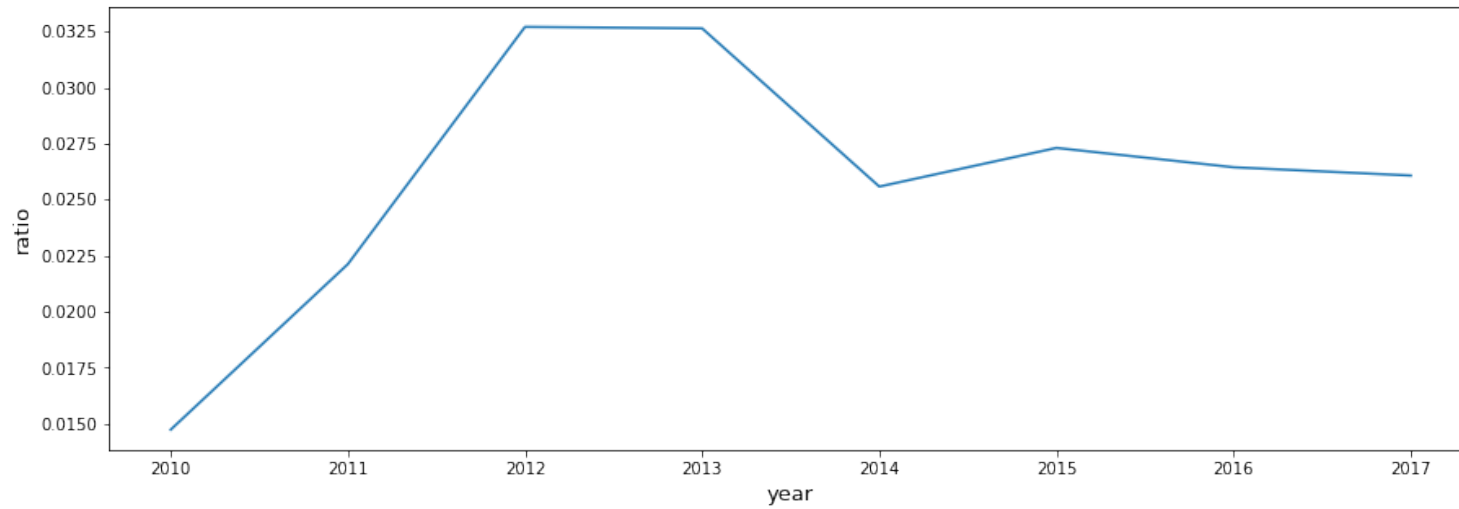
New York



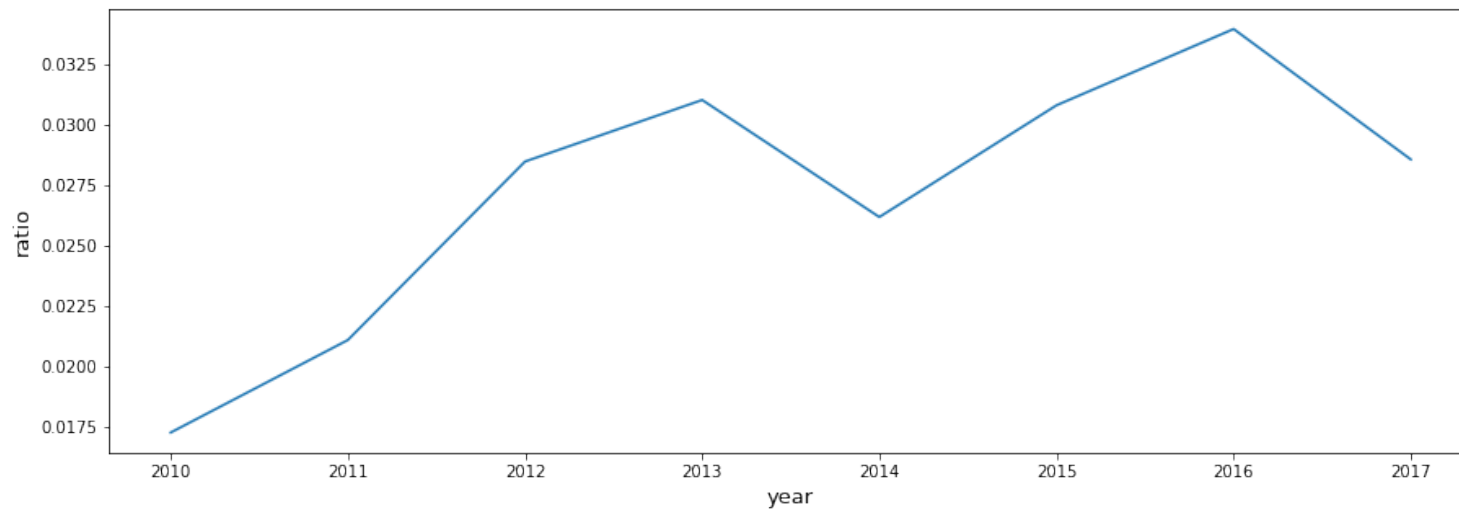
North Carolina



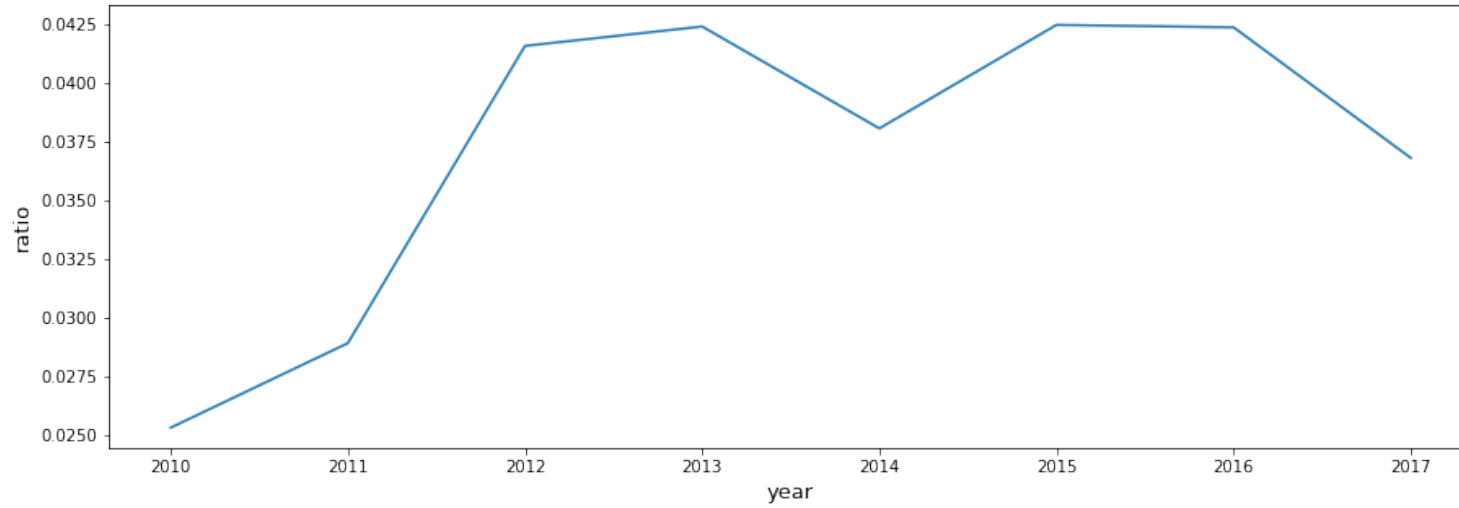
North Dakota



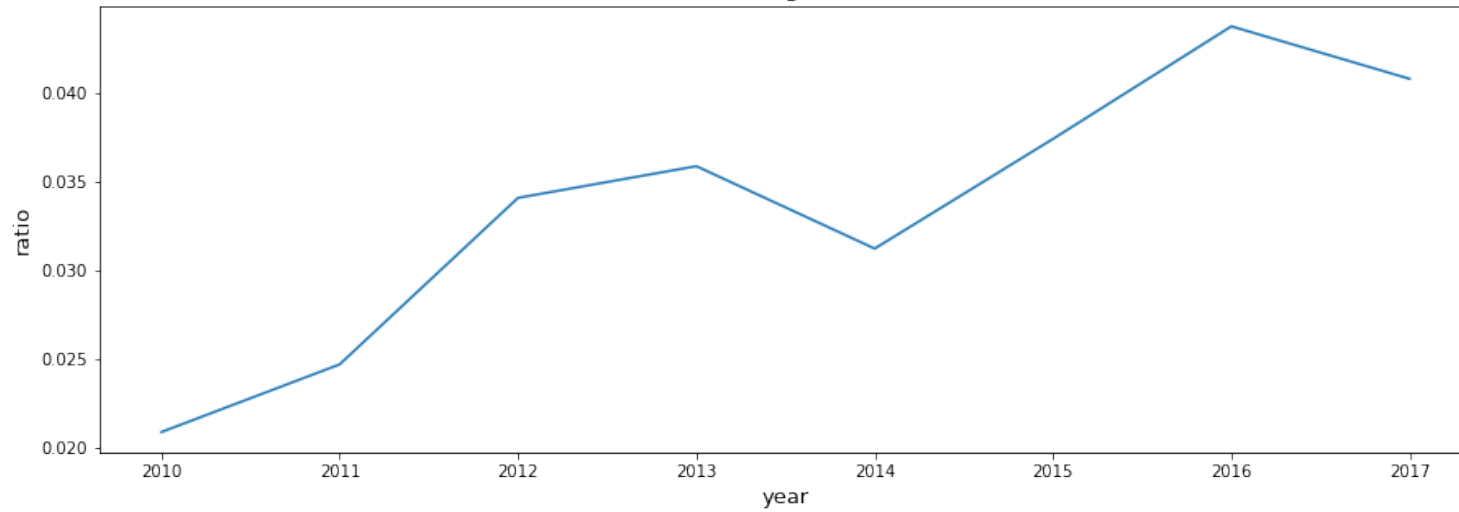
Ohio



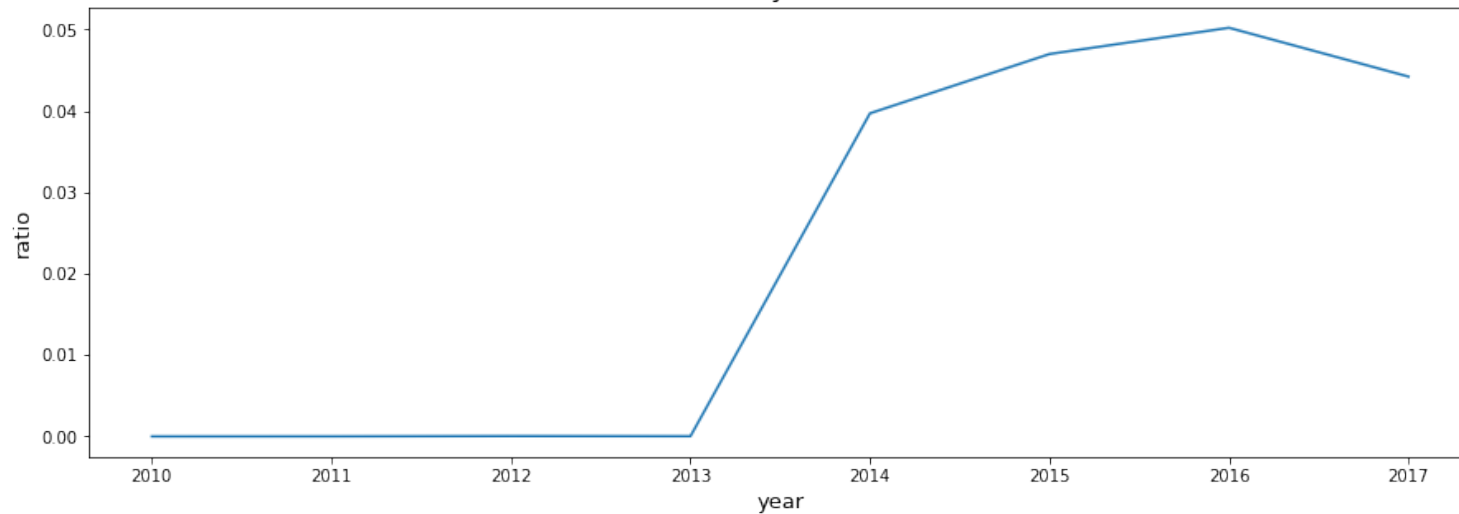
Oklahoma



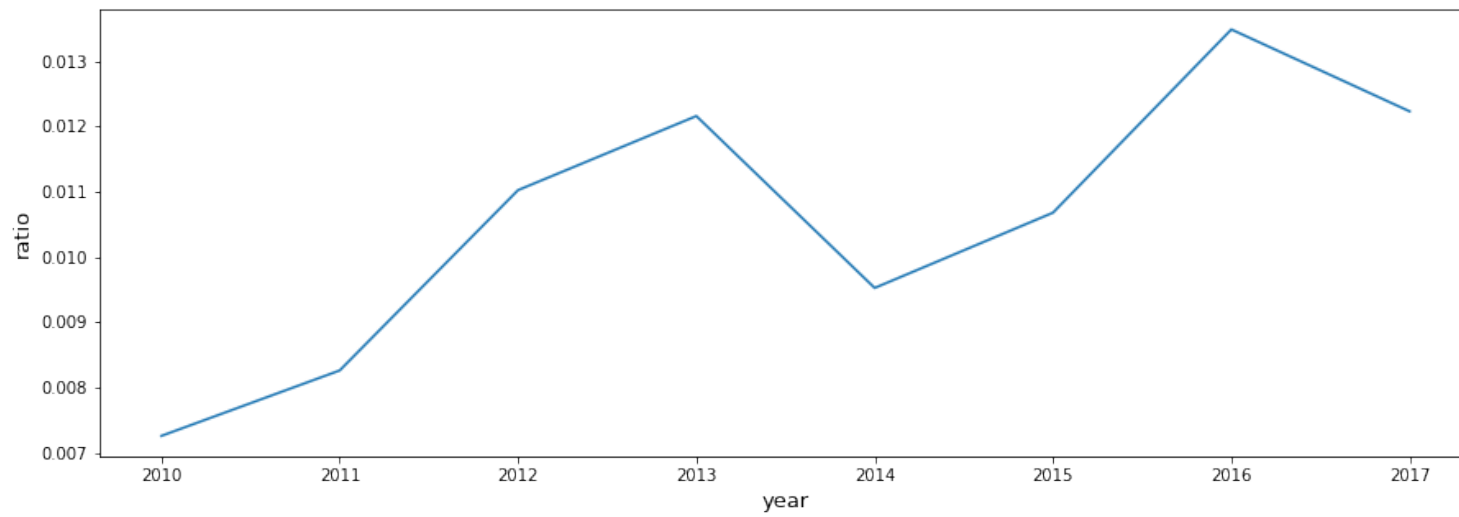
Oregon



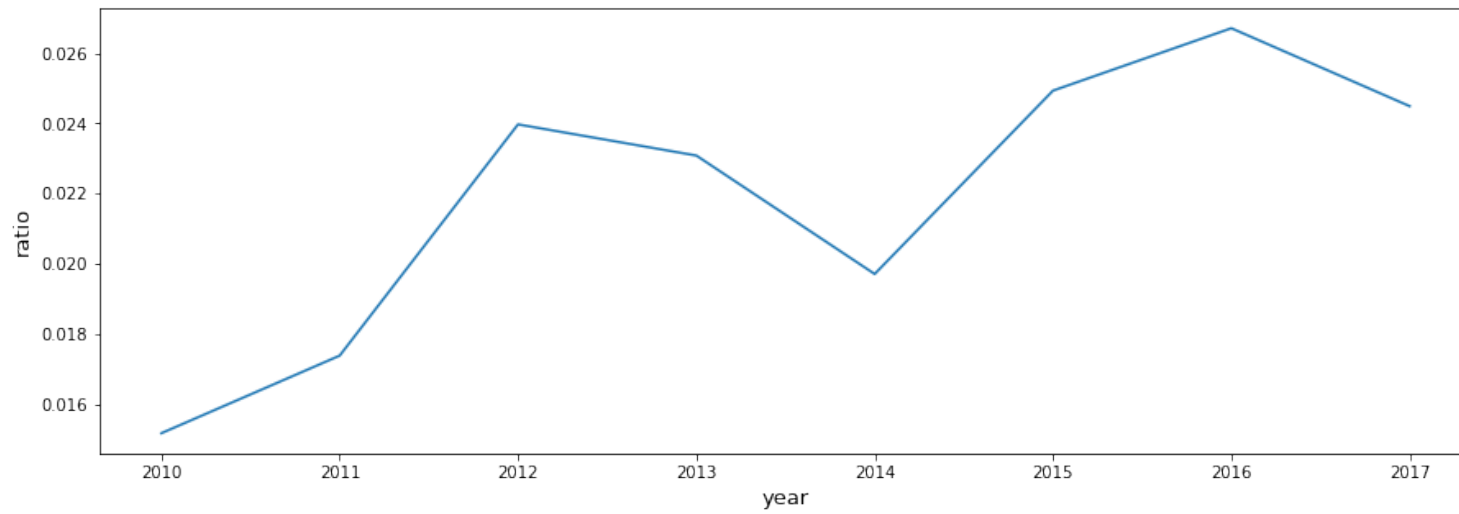
Pennsylvania



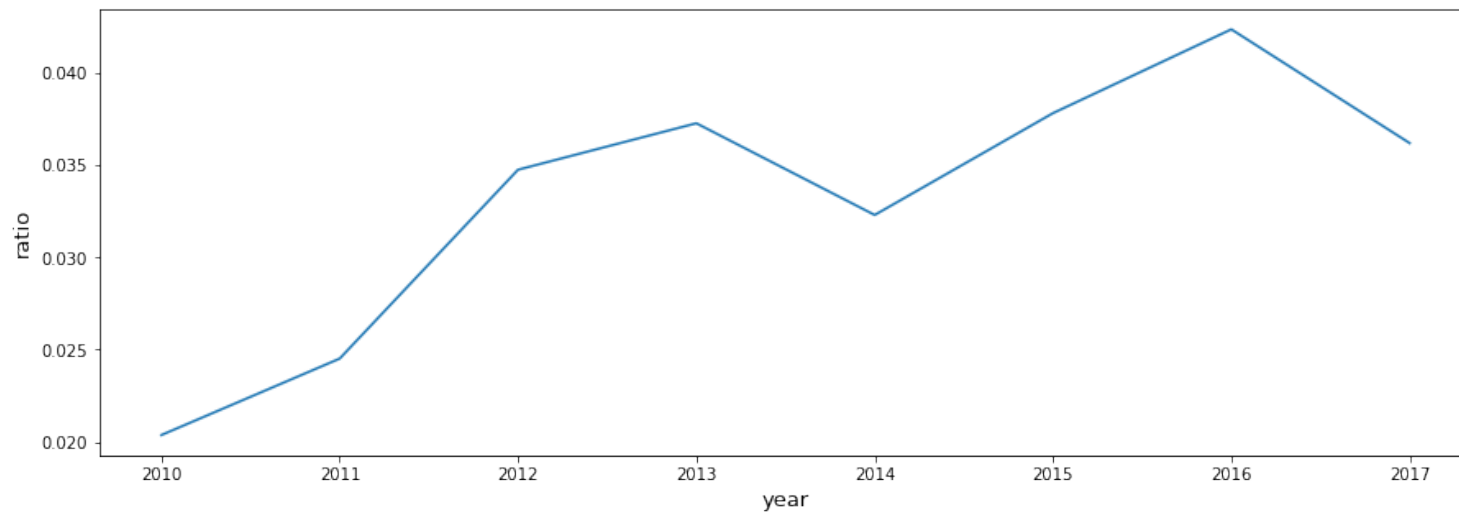
Rhode Island



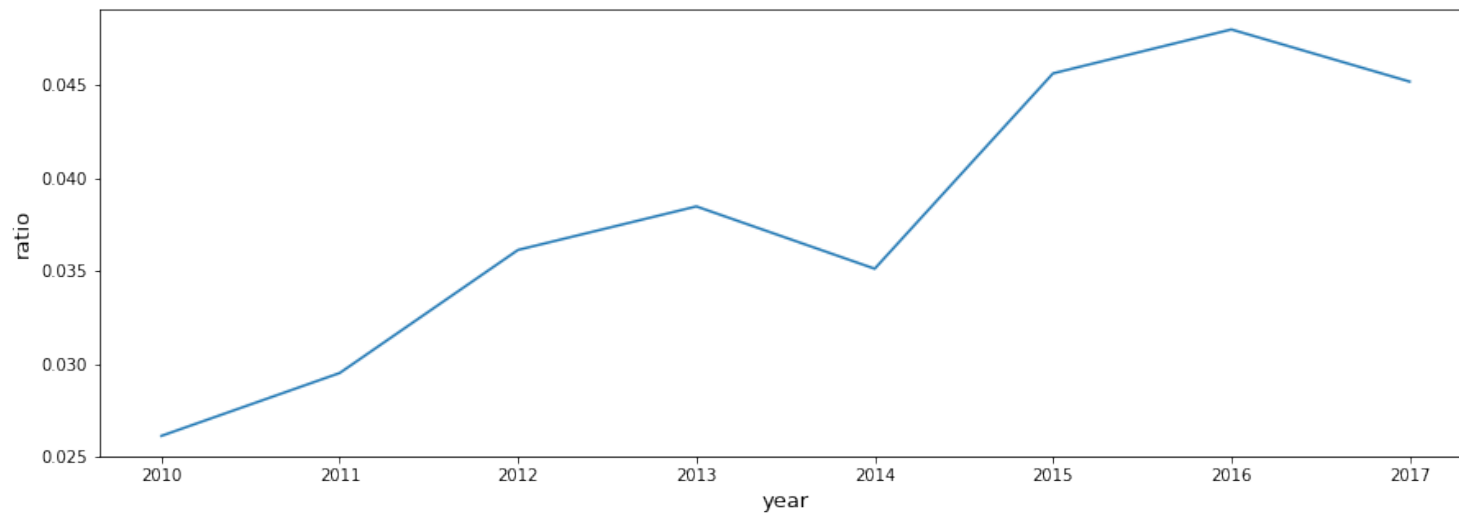
South Carolina



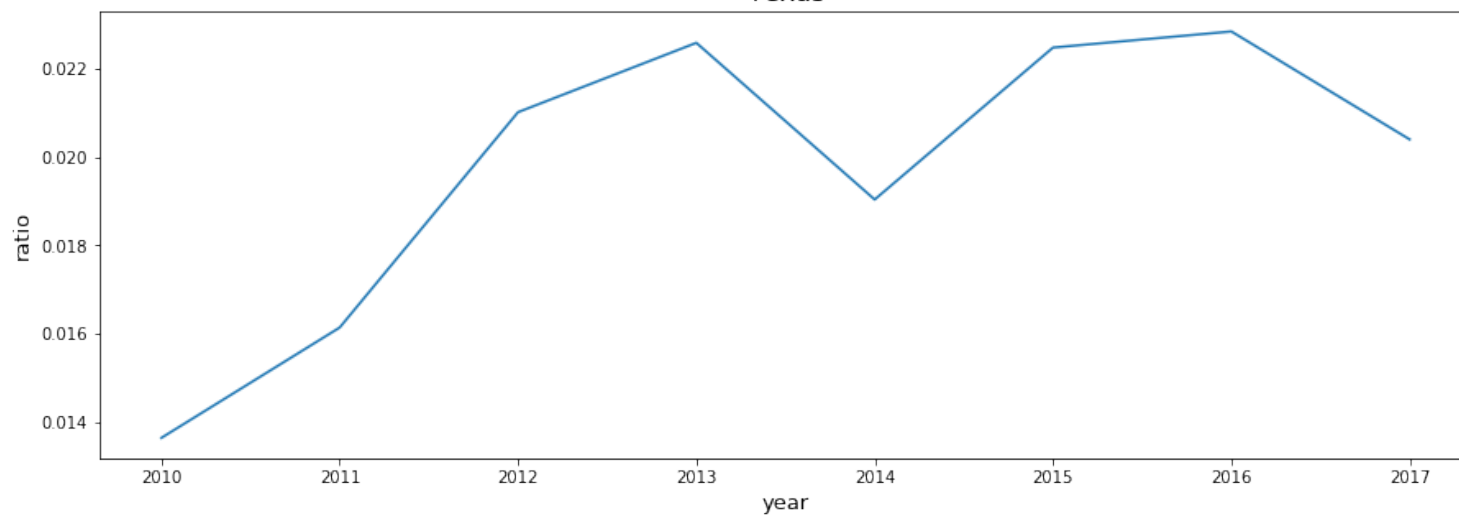
South Dakota



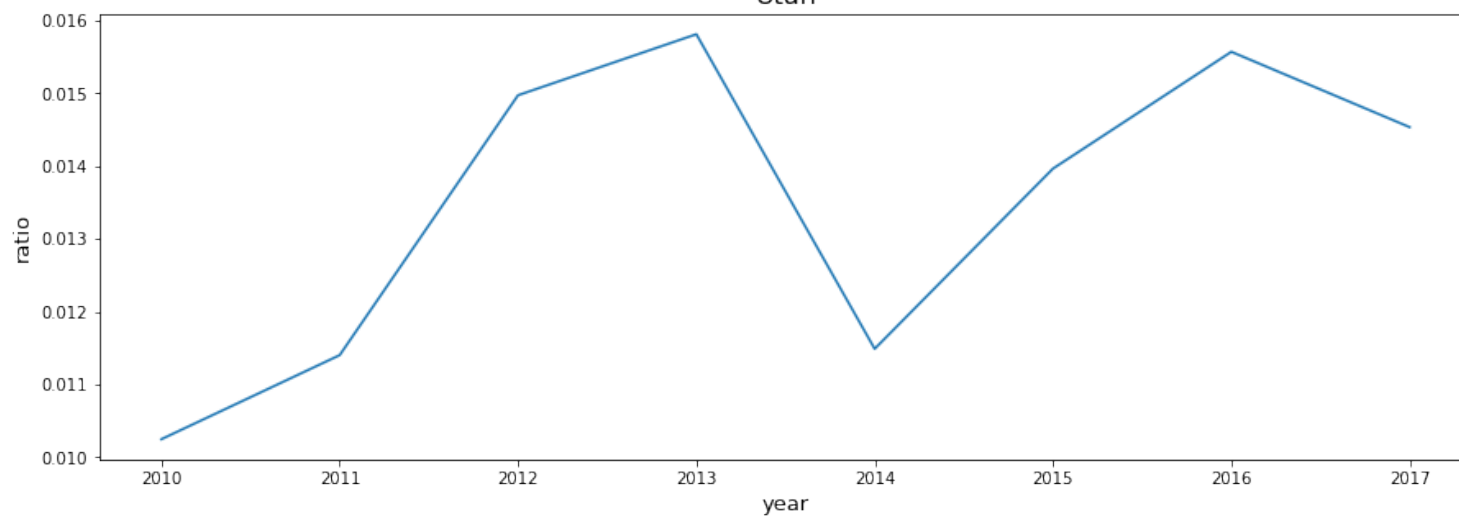
Tennessee



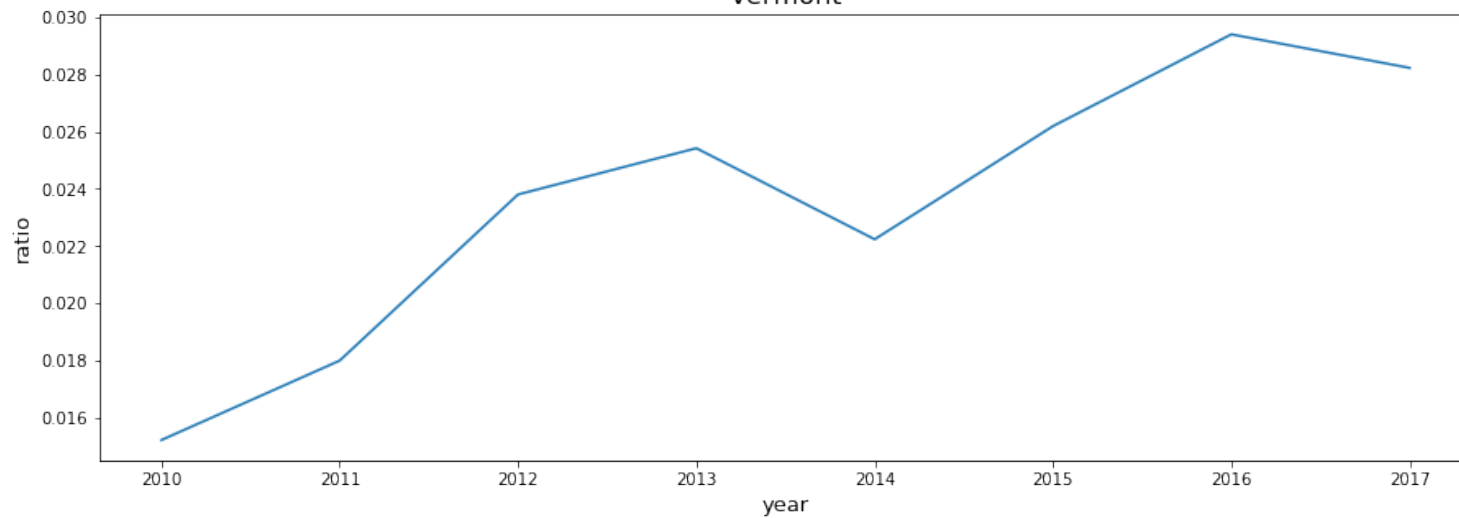
Texas



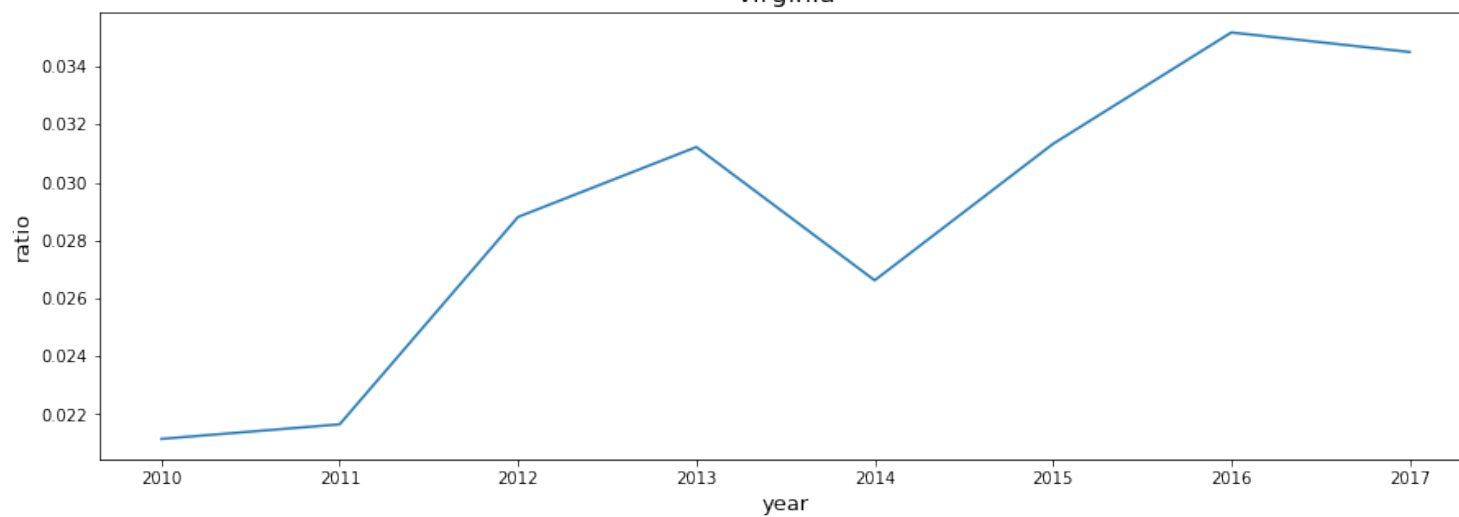
Utah

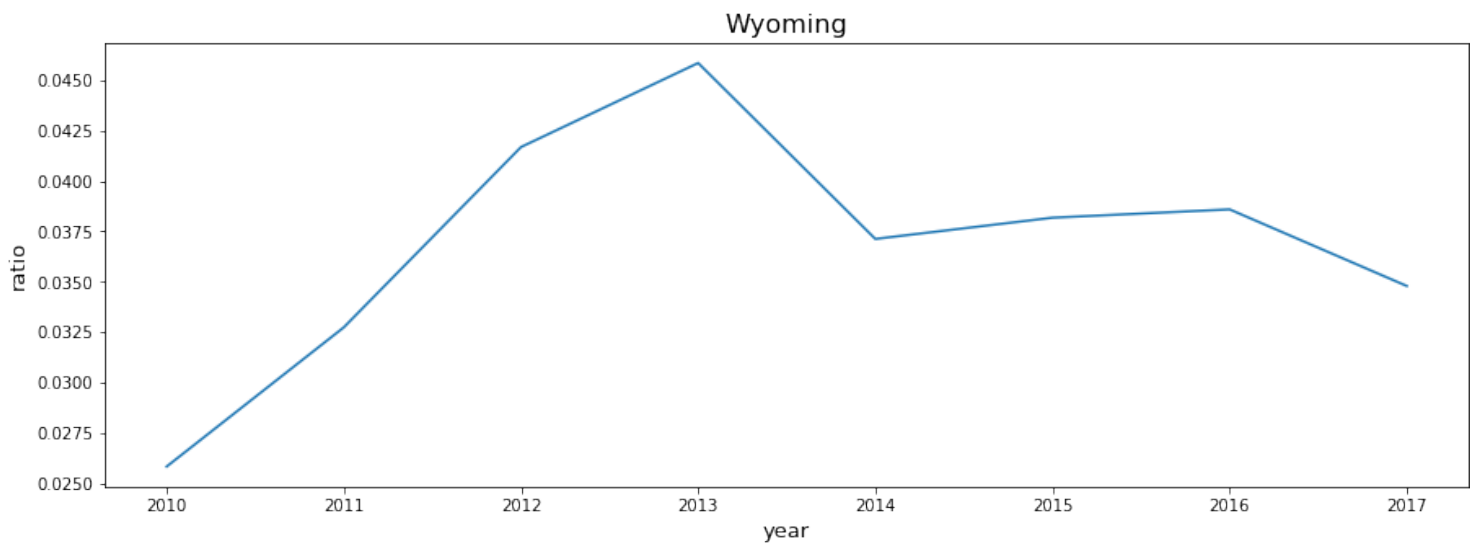
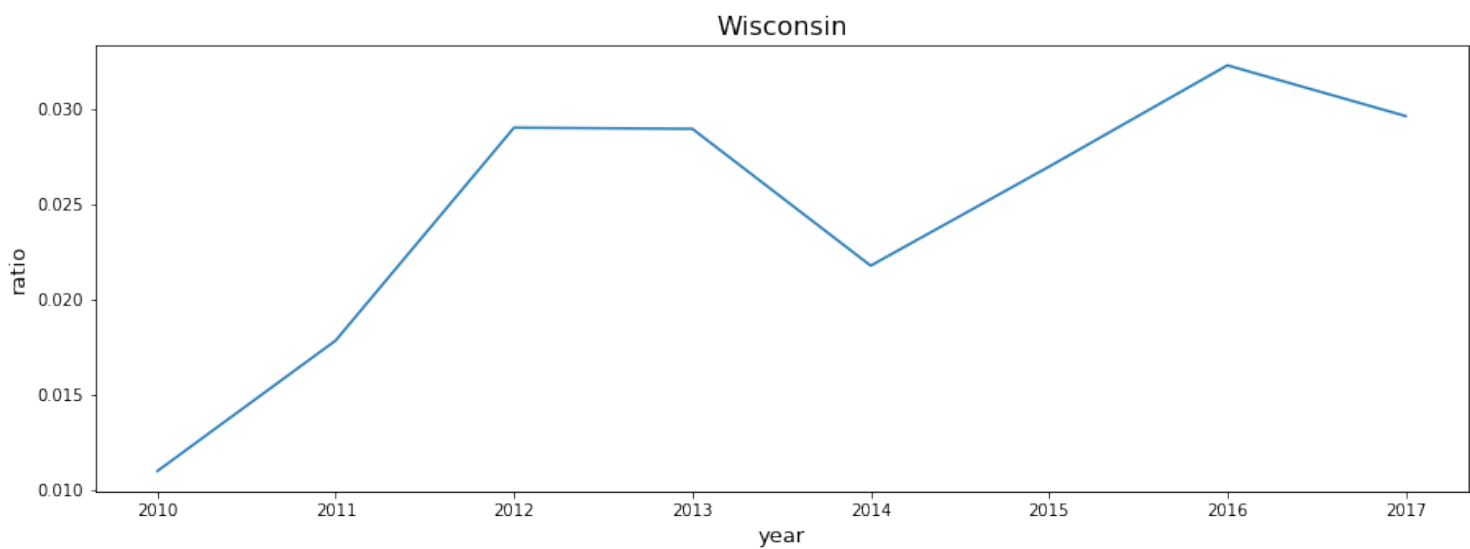
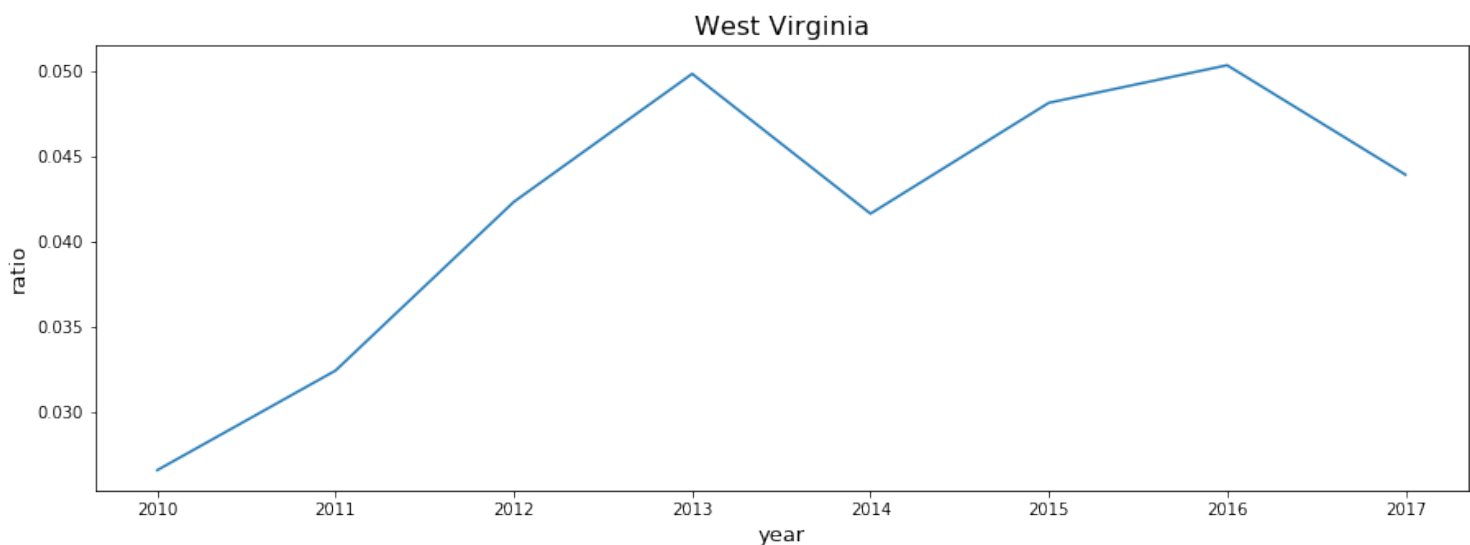
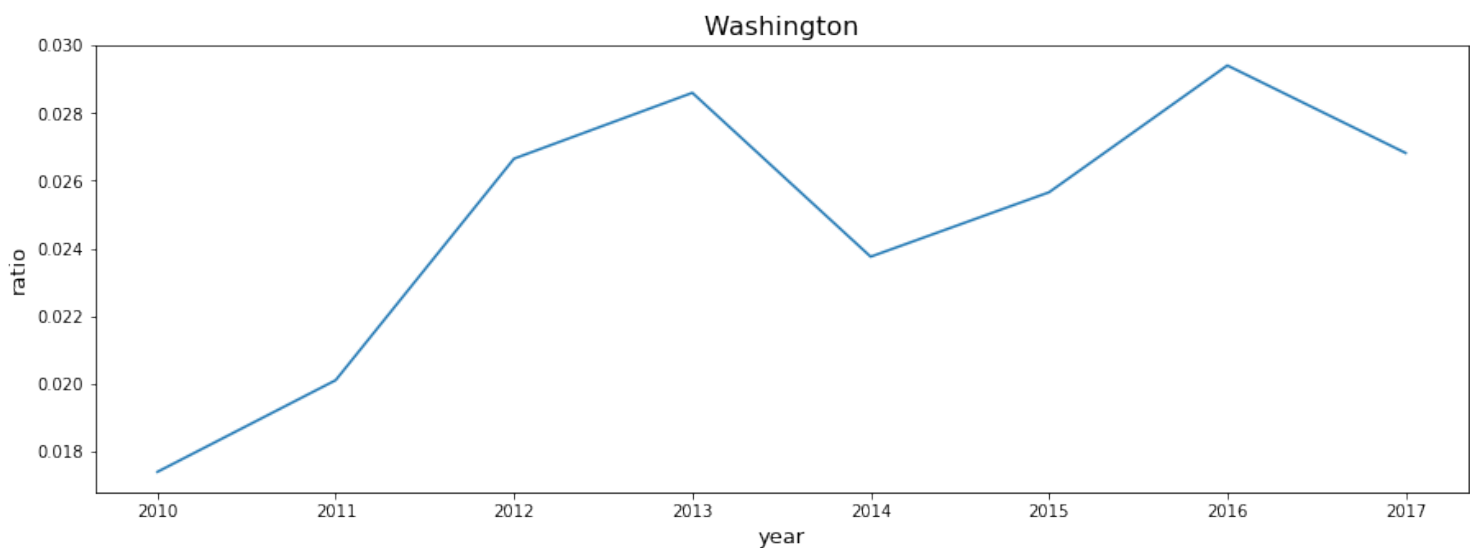


Vermont



Virginia





As we can see handgun background checks throughout the USA are rising. The overall trend of background checks is positive. And the very first thing I noticed that sometimes the high increase or decrease appears. I read a lot of information about what's going on with gun background checks in particular years and I found out that this bursts of background checks related to new gun laws or restrictions. I found it interesting and I wondered what happened. Everything has a causation. The gun laws in each state have differences, we can't compare a state to another state. The only thing that makes sense is to compare between years the same state. For example, when Maryland approved one of the nation's strictest gun-control measures in May 2013, gun sales jumped as buyers tried to beat the October deadline specified in the measure, which banned most semiautomatic rifles.

Now we are ready to answer some questions:

-Which states have had the highest growth in gun registrations?

-What is the overall trend of gun purchases?

Firstly we want to know top 10 states with the highest growth 2010-2017. We'll create a new function which should calculate % growth. For example, 2011/2010 produce 14%, 2012/2011 produce -3% and so on. After we'll apply the function to data frame and finally check the highest growth in gun registrations.

In [19]:

```
def percentage_func(series):  
    return series/series.shift()*100-100
```

In [20]:

```
percentage_data_handgun=ratio_background_checks_handgun.apply(percentage_func,  
axis=1) #handgun  
percentage_data_long_gun=ratio_background_checks_long_gun.apply(percentage_func,  
axis=1) #long_gun
```

In [21]:

```
print percentage_data_handgun.head(10)  
print percentage_data_long_gun.head(10)
```

	2010	2011	2012	2013	
2014 \					
Alabama	NaN	11.261328	35.394128	9.737650	-8.79
4112					
Alaska	NaN	8.740596	25.236967	1.482633	-2.68
2675					
Arizona	NaN	20.304423	26.913308	-1.304608	-12.27
4698					
Arkansas	NaN	8.706598	26.661714	7.500223	-19.40
0287					
California	NaN	22.529835	22.752517	7.534629	20.63
8458					
Colorado	NaN	12.929903	22.500479	13.567959	-15.66
8978					

Connecticut 1278	NaN	4.385388	27.463509	8.272517	14.17
Delaware 0904	NaN	19.456698	39.585827	25.111717	-3.48
District of Columbia 1314	NaN	-11.332510	48.830672	3.725418	3.16
Florida 4065	NaN	12.795004	22.987656	5.062424	-8.17

		2015	2016	2017	
Alabama	36.351822	-37.301391	-36.347764		
Alaska	3.728052	0.609037	-7.613378		
Arizona	18.589189	10.657576	-9.363604		
Arkansas	27.416430	3.245101	-10.629004		
California	-4.474944	16.484798	-9.101930		
Colorado	16.648157	10.046218	-6.577456		
Connecticut	57.336114	-0.745953	-41.292742		
Delaware	19.191994	9.027254	-13.608762		
District of Columbia	39.376541	-1.549221	10.444848		
Florida	19.572330	13.609127	-5.035613		
	2010	2011	2012	2013	

2014 \					
Alabama 88415	NaN	6.865383	25.426950	10.256763	-16.2
Alaska 45106	NaN	3.119599	15.670936	7.755952	-9.9
Arizona 38463	NaN	16.625459	30.148124	1.896742	-18.5
Arkansas 65184	NaN	8.512741	19.534344	0.058317	-13.0
California 12935	NaN	15.461795	30.644011	23.269490	-30.5
Colorado 34968	NaN	11.238995	15.392156	8.378507	-18.2
Connecticut 72701	NaN	1.858106	30.096024	6.273159	3.5
Delaware 17999	NaN	11.725541	28.720618	17.332983	16.1
District of Columbia 29744	NaN	-20.199259	95.187766	100.975948	-20.2
Florida 71661	NaN	12.295789	37.958215	8.632703	-17.1

		2015	2016	2017	
Alabama	2.849412	-30.870408	-29.520562		
Alaska	-8.361208	0.931358	-11.285237		
Arizona	0.830728	15.194881	-10.333653		
Arkansas	0.150810	-1.592596	-7.020283		
California	-7.802850	60.530647	-42.980722		
Colorado	2.642232	8.569482	-8.513602		
Connecticut	-18.555395	-9.227645	-40.970938		
Delaware	2.981280	7.518358	-16.131271		
District of Columbia	4.773661	47.457389	-7.551746		
Florida	-0.745034	11.708968	-13.270786		

Column '2010' is NaN as expected. The column doesn't have the previous column to divide.

I decided to find top 10 states which have the highest growth from 2010-2017.

In [22]:

```
print percentage_data_handgun.sum(axis=1).sort_values(ascending=False)
print percentage_data_long_gun.sum(axis=1).sort_values(ascending=False)
```

Pennsylvania	161872.302777
Michigan	580.348123
North Carolina	230.620915
Iowa	173.150021
Wisconsin	135.002216
Illinois	115.086107
New Jersey	113.273532
Maryland	102.660955
Delaware	95.283824
District of Columbia	92.657062
New Hampshire	92.483994
Indiana	88.611810
Mississippi	83.670982
New York	81.718180
Oregon	78.630341
North Dakota	78.424852
California	76.363362
Massachusetts	74.187889
Vermont	70.655252
South Dakota	70.653497
Connecticut	69.590111
Maine	69.033032
Nebraska	68.219938
Idaho	67.525303
Minnesota	66.366453
Rhode Island	65.035209
Ohio	62.566160
Tennessee	62.410584
West Virginia	61.186156
Florida	60.816863
South Carolina	59.504830
Missouri	57.741602
Georgia	57.105955
Virginia	57.096196
New Mexico	54.265187
Arizona	53.521585
Colorado	53.446283
Washington	52.260112
Louisiana	51.323478
Texas	49.271845
Oklahoma	47.911453
Utah	47.229793
Kentucky	47.155118
Arkansas	43.500775
Wyoming	38.995220
Kansas	35.668747
Montana	34.047488
Alaska	29.501233
Nevada	18.198054

Alabama	10.301661
dtype: float64	
District of Columbia	200.414017
New Jersey	70.174976
Delaware	68.265507
Massachusetts	54.618980
California	48.609436
Rhode Island	46.711510
Maryland	45.293411
Florida	39.408193
New Hampshire	37.939460
Indiana	36.591083
Ohio	35.936591
Arizona	35.823818
Missouri	35.722586
Virginia	34.098989
Tennessee	33.821302
Oklahoma	32.192343
Illinois	30.820115
Washington	27.421122
Vermont	27.400448
New Mexico	27.122074
Oregon	26.355334
New York	24.374430
South Carolina	23.654676
Wisconsin	21.813998
Idaho	21.109484
North Carolina	20.495839
Colorado	19.472802
Georgia	15.357404
Texas	14.211405
West Virginia	12.742516
Utah	10.116643
Maine	9.428207
Mississippi	6.886256
Wyoming	6.684718
Arkansas	6.578148
Minnesota	5.214394
Montana	3.718190
South Dakota	3.254906
Kansas	2.788172
Louisiana	0.934855
Kentucky	0.297604
Michigan	0.150982
Alaska	-2.113708
North Dakota	-4.115218
Nevada	-12.152115
Nebraska	-15.108639
Iowa	-22.646966
Connecticut	-26.953987
Alabama	-31.280876
Pennsylvania	-37.929629
dtype: float64	

Pennsylvania has increased by 161872.3% for handguns but has the lowest growth in long_guns -37.9% since 2010. I wondered why it happened and found out that the state has the strictest gun laws, they also support their own background check system. The gun laws related to concealed carry, permit to purchase - carry a lot of weight.

-What is the overall trend of gun purchases?

To answer this question we have to convert background checks into estimated sales, we relied on a method suggested in the Small Arms Survey by Jurgen Brauer, a professor at Georgia Regents University. Each long gun and handgun check was counted as 1.1 sales. Each multiple-gun check was counted as two sales. Permit checks and other types of checks were omitted. The multiplier is an estimate based on Mr. Brauer's interviews with gun shop owners.

Creating function calculate_purchases(total_checks,start_year,end_year), total_checks can be hand_gun_total_checks,long_gun_total_checks, multiple_total_checks.

In [23]:

```
def calculate_purchases(total_checks,start_year,end_year):  
    result=pd.DataFrame(index=months_activity['state'].unique())  
    for year in range(start_year,end_year+1):  
        result[year]=total_checks[year]  
    return result
```

In [24]:

```
purchases_handgun=calculate_purchases(hand_gun_total_checks,1998,2017)
purchases_handgun=purchases_handgun.mul(1.1)#Each handgun check was counted as 1.1 sales
purchases_handgun.head(10)
```

Out[24]:

	1998	1999	2000	2001	2002	2003	2004	
Alabama	11623.7	103998.4	90181.3	92273.5	71823.4	74783.5	76925.2	8
Alaska	1501.5	15772.9	13985.4	14342.9	11554.4	12255.1	14328.6	7
Arizona	7872.7	85913.3	71793.7	69515.6	59216.3	54597.4	61859.6	7
Arkansas	5113.9	55575.3	47803.8	44140.8	33492.8	34573.0	34193.5	3
California	31505.1	409082.3	361476.5	308048.4	258633.1	174692.1	194942.0	2
Colorado	8717.5	112604.8	129800.0	144404.7	73455.8	74207.1	76463.2	8
Connecticut	1673.1	24612.5	8.8	10355.4	27148.0	28146.8	25044.8	2
Delaware	557.7	6576.9	5442.8	5724.4	5243.7	5498.9	5474.7	6
District of Columbia	0.0	20.9	18.7	13.2	6.6	12.1	15.4	2
Florida	12657.7	141650.3	134818.2	155269.4	142692.0	147491.3	159230.5	2



Same for long_gun background checks

In [25]:

```
purchases_long_gun=calculate_purchases(long_gun_total_checks,1998,2017)
purchases_long_gun=purchases_long_gun.mul(1.1)#Each long_gun check was counted
as 1.1 sales
purchases_long_gun.head(10)
```

Out[25]:

	1998	1999	2000	2001	2002	2003	2004	
Alabama	28295.3	163918.7	150443.7	148405.4	122472.9	125514.4	129461.2	-
Alaska	2810.5	30569.0	28942.1	27860.8	22779.9	22423.5	24154.9	2
Arizona	9102.5	78501.5	71310.8	74222.5	65896.6	62702.2	68735.7	7
Arkansas	18605.4	139562.5	128573.5	122017.5	90409.0	90757.7	94023.6	9
California	36781.8	451130.9	333020.6	372076.1	283055.3	230566.6	241462.1	2
Colorado	16756.3	161978.3	196777.9	228780.2	121770.0	109604.0	120486.3	-
Connecticut	2734.6	25146.0	4.4	14468.3	28689.1	30426.0	27549.5	2
Delaware	1779.8	12558.7	11286.0	12329.9	12479.5	12245.2	12316.7	-
District of Columbia	0.0	2.2	0.0	5.5	1.1	0.0	0.0	-
Florida	21211.3	139948.6	135213.1	156900.7	150885.9	159220.6	170866.3	-



Same for multiple background checks

In [26]:

```
purchases_multiple=calculate_purchases(multiple_total_checks,1998,2017)
purchases_multiple=purchases_multiple.mul(2)#Each multiple check was counted as 2 sales
purchases_multiple.head(10)
```

Out[26]:

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Alabama	556	6390	6298	6980	7096	7240	7264	7952	7730	8506
Alaska	122	1884	1868	2094	1700	1696	1764	2104	2002	2248
Arizona	556	6906	5920	5702	4920	5252	6084	6534	6812	6472
Arkansas	378	5892	5072	5264	5090	4958	5090	5358	5298	5580
California	0	0	0	0	0	0	78	136	33248	382266
Colorado	1120	13410	21062	18840	9126	9140	10264	10772	10778	43430
Connecticut	12	136	0	948	2778	2110	2226	1946	1438	1328
Delaware	20	396	278	380	408	338	416	416	440	592
District of Columbia	0	0	0	0	0	0	0	0	4	4
Florida	694	9650	9058	11190	10334	11256	12132	13972	15784	17478

Now we are ready to calculate the total gun purchases per state.

In [27]:

```
total_gun_purchases=purchases_handgun+purchases_long_gun+purchases_multiple
total_gun_purchases
```

Out[27]:

	1998	1999	2000	2001	2002	2003	200
Alabama	40475.0	274307.1	246923.0	247658.9	201392.3	207537.9	213650.
Alaska	4434.0	48225.9	44795.5	44297.7	36034.3	36374.6	40247.5
Arizona	17531.2	171320.8	149024.5	149440.1	130032.9	122551.6	136679.
Arkansas	24097.3	201029.8	181449.3	171422.3	128991.8	130288.7	133307.
California	68286.9	860213.2	694497.1	680124.5	541688.4	405258.7	436482.
Colorado	26593.8	287993.1	347639.9	392024.9	204351.8	192951.1	207213.
Connecticut	4419.7	49894.5	13.2	25771.7	58615.1	60682.8	54820.3
Delaware	2357.5	19531.6	17006.8	18434.3	18131.2	18082.1	18207.4

District of Columbia	0.0	23.1	18.7	18.7	7.7	12.1	15.4
Florida	34563.0	291248.9	279089.3	323360.1	303911.9	317967.9	342228.
Georgia	3310.3	275940.4	241064.2	237078.2	214472.3	217583.6	218886.
Hawaii	33.0	4.4	0.0	0.0	0.0	0.0	0.0
Idaho	7778.6	81170.2	74579.1	71664.3	56792.7	56726.7	58685.9
Illinois	19370.7	209167.5	170580.9	172620.0	164949.5	170016.3	174857.
Indiana	24786.0	223574.9	126456.7	100839.7	88593.5	116808.7	189553.
Iowa	7498.4	50704.0	46747.6	45627.8	42443.1	46611.2	49937.4
Kansas	13759.6	118434.2	108049.9	104587.1	94249.3	97620.1	99835.0
Kentucky	35433.9	277039.1	256859.1	259099.9	213521.3	211760.0	221122.
Louisiana	30595.2	217866.5	194349.7	191827.6	157197.9	165057.8	162333.
Maine	4559.8	52511.4	50080.7	55396.3	52738.9	54043.0	55890.3
Maryland	8115.8	59319.6	53282.9	58277.8	53738.9	54317.7	57816.7
Massachusetts	1592.7	21606.0	28020.4	32734.5	37383.1	43296.8	42275.0
Michigan	22570.0	202472.0	196615.7	205488.9	217696.5	225185.6	224028.
Minnesota	11169.2	163722.0	168500.5	168220.7	161260.2	182502.3	186886.
Mississippi	28888.5	202189.3	182250.1	169683.6	129521.2	131065.0	133202.
Missouri	26605.5	228979.8	212080.9	216401.4	195330.1	201123.3	208716.
Montana	7124.0	81489.8	79321.1	79540.5	64747.6	68097.5	70538.3
Nebraska	5208.4	34749.7	32529.4	31553.5	28717.5	30312.3	32872.9
Nevada	7218.8	70360.3	56034.2	56800.1	53040.3	55590.3	59293.4
New Hampshire	2697.0	41573.8	43630.3	46232.0	43686.9	43950.0	43808.0
New Jersey	4178.8	42915.8	41293.6	43963.3	48101.3	46438.0	41593.3
New Mexico	10058.7	105575.5	93881.0	91292.9	70516.4	70152.7	73920.6
New York	14991.7	131160.3	120728.1	130462.0	122562.2	123814.2	133922.
North Carolina	30064.3	222757.7	207254.1	207078.3	157794.3	161607.0	161610.
North Dakota	2696.9	32860.0	33286.8	32946.5	30070.2	31455.0	33367.0
Ohio	36123.3	332259.0	306644.3	322076.6	310020.6	312191.6	342951.
Oklahoma	22325.5	189463.8	175652.1	176324.8	140698.0	149300.9	156511.
Oregon	15516.4	178115.5	150656.6	156977.3	152706.3	159048.7	166905.
Pennsylvania	53918.8	599202.2	532995.4	561553.6	623318.5	567271.5	553883.
Rhode Island	1181.1	12404.4	11234.7	12130.9	11233.8	11445.9	11919.7

South Carolina	10581.8	155137.9	134392.6	128553.9	104910.2	107219.9	110474.
South Dakota	3043.2	34650.5	40649.1	44731.0	45049.2	48453.6	51653.4
Tennessee	27431.0	293897.9	261221.8	256955.1	248262.1	245121.3	243512.
Texas	85647.0	723811.7	652845.1	647641.9	522338.8	533291.1	542628.
Utah	11857.3	86544.6	74806.4	75096.0	68591.1	70586.7	72721.7
Vermont	2339.1	24623.4	22481.0	22900.7	22757.3	22221.7	21685.4
Virginia	27905.3	230886.2	209065.2	217496.3	212645.5	222263.3	236641.
Washington	12488.3	133547.7	126973.7	139635.9	135184.7	152693.5	162735.
West Virginia	15758.2	149723.2	142941.2	148163.7	117926.6	114577.7	115216.
Wisconsin	17006.0	197362.2	198637.0	199638.1	188674.9	204606.0	202858.
Wyoming	3740.2	37512.8	35227.5	35603.6	29907.0	29582.6	31712.2

Finally we have the data with total gun purchases for each state for each year. I want to see a trend of gun purchases for each state.

According to documentation:

```
df.plot()
```

Line plot with one line per column.

So, I'm going to change columns to rows

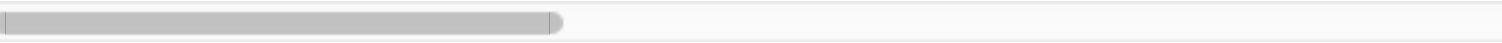
In [28]:

```
inversed_purchases = total_gun_purchases.T
inversed_purchases.head(10)
```

Out[28]:

	Alabama	Alaska	Arizona	Arkansas	California	Colorado	Connecticut	De
1998	40475.0	4434.0	17531.2	24097.3	68286.9	26593.8	4419.7	23
1999	274307.1	48225.9	171320.8	201029.8	860213.2	287993.1	49894.5	19
2000	246923.0	44795.5	149024.5	181449.3	694497.1	347639.9	13.2	17
2001	247658.9	44297.7	149440.1	171422.3	680124.5	392024.9	25771.7	18
2002	201392.3	36034.3	130032.9	128991.8	541688.4	204351.8	58615.1	18
2003	207537.9	36374.6	122551.6	130288.7	405258.7	192951.1	60682.8	18
2004	213650.4	40247.5	136679.3	133307.1	436482.1	207213.5	54820.3	18
2005	223808.3	45034.8	160222.7	137879.4	504942.5	221092.0	56325.6	18
2006	232324.7	48404.4	177284.5	142397.6	504970.9	227299.8	58656.7	18
2007	235261.1	51614.9	163358.4	135615.4	807970.4	260062.9	65807.8	17

10 rows × 51 columns



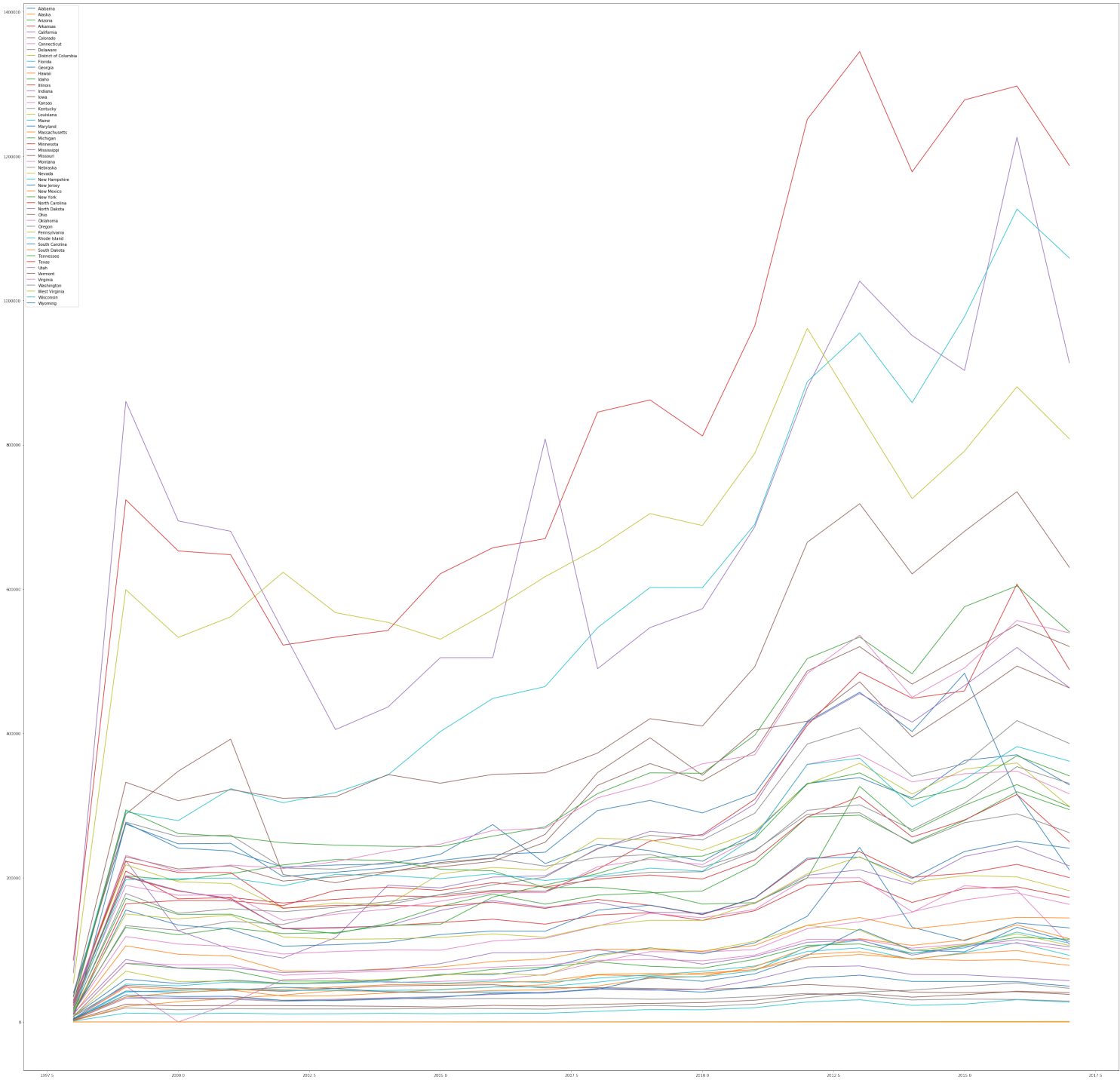
As expected. Now we can plot the data.

In [29]:

```
inversed_purchases.plot(figsize=(50,50))
```

Out[29]:

<matplotlib.axes._subplots.AxesSubplot at 0x1a23774d50>



15 states have a positive trend, the other states have a stable trend. I've noticed that gun purchases in 2017 fell down for all states.

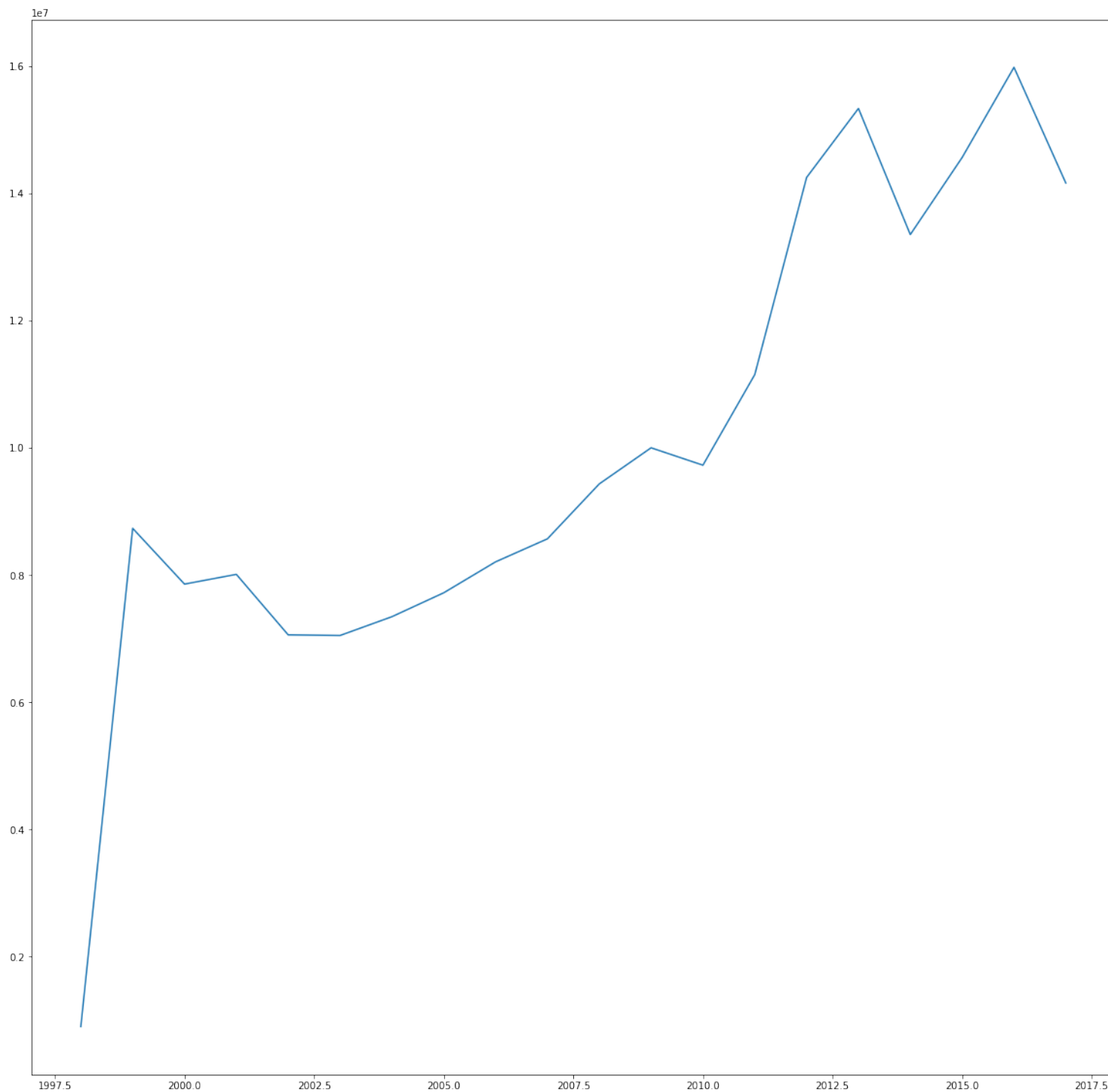
What is the overall trend of gun purchases?

In [30]:

```
total_gun_purchases.sum(axis=0).plot(figsize=(20,20)) #sum of gun purchases each year
```

Out[30]:

<matplotlib.axes._subplots.AxesSubplot at 0x1a27a6fd90>

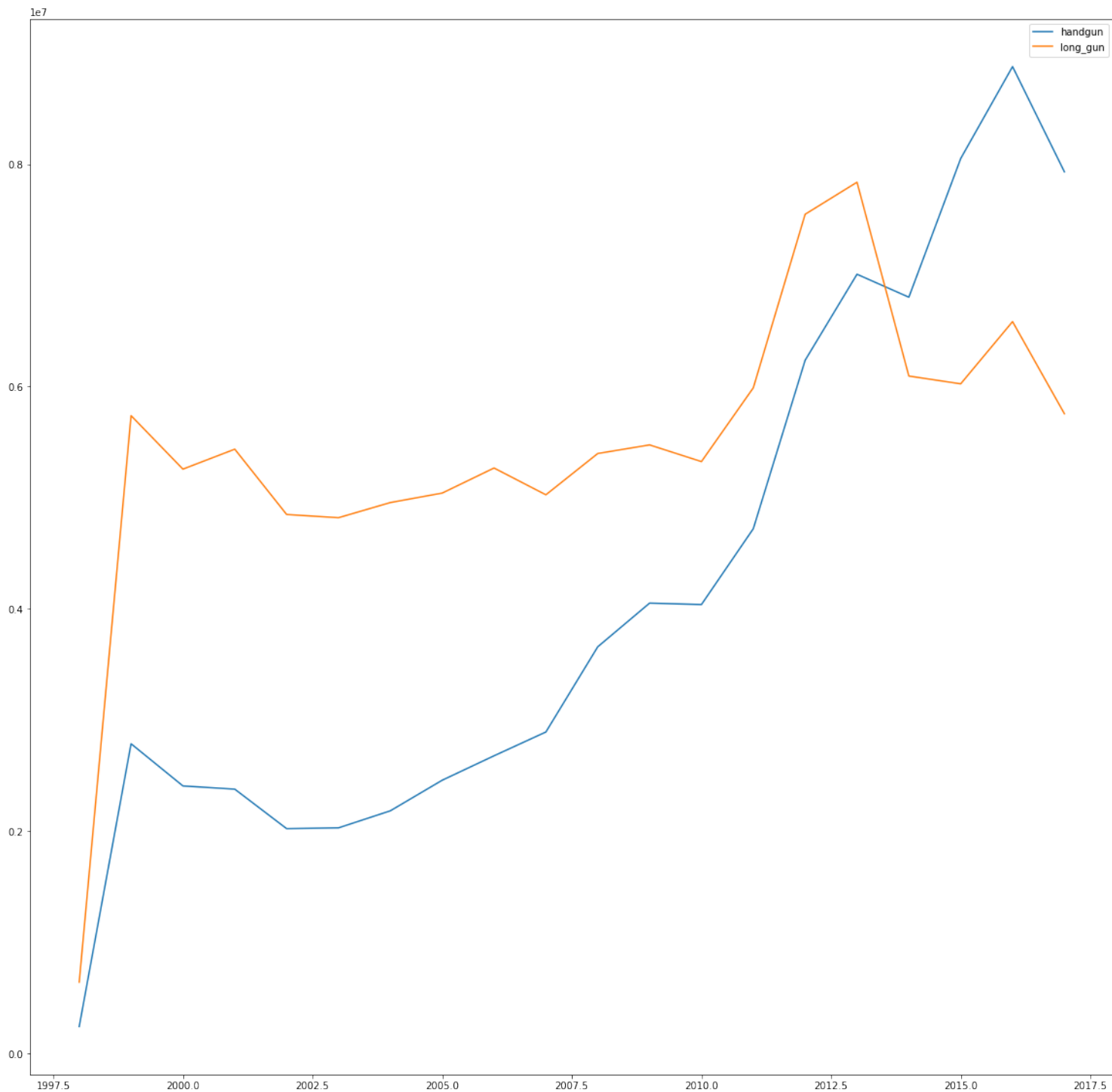


The overall trend of gun purchases is positive. I've noticed that in 2012 gun purchases rised, I wondered why it happened and discovered that Mr. Obama called for new buying restrictions after the mass shooting at Sandy Hook Elementary School. And the same pattern occured in 2015 after a call from President Obama to make it harder to buy assault weapons after the terrorist attack in San Bernardino, Calif.

Now I decided to compare handgun to long_gun sales

In [31]:

```
purchases_handgun.sum(axis=0).plot(figsize=(20,20))
purchases_long_gun.sum(axis=0).plot(figsize=(20,20))
plt.legend(['handgun', 'long_gun']);
```



The increase is mostly due to higher sales of handguns, which are typically bought for self-defense.

I decided to check states with the highest gun purchases(1998-2017)

In [32]:

```
total_gun_purchases.sum(axis=1).sort_values(ascending=False)
```

Out[32]:

Texas	16677859.7
California	13700154.1
Pennsylvania	13058504.3
Florida	11512704.7
Ohio	8727028.3
Tennessee	6992322.6
Virginia	6561852.7
Colorado	6535912.1
Missouri	6426532.6
Alabama	5615744.1
Illinois	5547024.6
Indiana	5258123.7
Georgia	5143490.7
Wisconsin	4790482.1
Louisiana	4770008.0
Washington	4767332.9
Kentucky	4701135.1
Oklahoma	4520153.1
Michigan	4315080.4
Oregon	4221735.6
New York	4212272.1
Minnesota	4127792.8
Arizona	3843535.0
North Carolina	3570492.7
Mississippi	3386331.1
South Carolina	3145107.7
Arkansas	3076705.1
West Virginia	2955760.0
Kansas	2611944.9
New Mexico	2001982.1
Maryland	1851202.4
Utah	1735533.7
Connecticut	1712734.2
Montana	1676576.7
Nevada	1653736.1
Idaho	1590113.9
Maine	1349709.9
New Hampshire	1308445.9
New Jersey	1282095.0
Massachusetts	1274694.2
South Dakota	1262940.4
Alaska	1188756.3
North Dakota	908393.3
Iowa	888057.9
Wyoming	841029.3
Nebraska	623450.6
Vermont	545671.7
Delaware	525621.7
Rhode Island	340353.3
District of Columbia	6006.4
Hawaii	45.8

dtype: float64

Texas, California, Pennsylvania, Florida, Ohio are leaders in gun purchases.

What is the growth of gun purchases from 1998 to 2017?

Let's check it.

In [33]:

```
total_gun_purchases.sum(axis=0)
```

Out[33]:

```
1998      899926.7
1999      8731075.2
2000      7854358.0
2001      8007449.5
2002      7056507.2
2003      7046718.6
2004      7341818.4
2005      7717761.7
2006      8203966.1
2007      8565683.3
2008      9429020.5
2009      9996644.3
2010      9723535.5
2011     11147586.2
2012     14245912.9
2013     15330213.7
2014     13347932.3
2015     14559585.4
2016     15976974.6
2017     14157603.5
dtype: float64
```

In [34]:

```
growth=total_gun_purchases.sum(axis=0)[2017]/total_gun_purchases.sum(axis=0)[1998]*100
growth
```

Out[34]:

```
1573.1951835632835
```

Amazing! The growth is 1573%!

Gun sales have increased almost 16 times in a decade, to about 14 million in 2017 from about 900,000 in 1998. More firearms are sold to residents in the United States than in any other country. I'm impressed. I've also noticed is that after a call for gun restrictions, sales go up.