

Final proposal: Opioid prescription ¶

Shihui Wang sw2804@nyu.edu

This project will investigate how opioid prescriptions are related to the practitioners' field, geography and local income in the US. The key dataset is a Socrata API. It contains information about prescribers name, practice address and type, as well as opioid prescribing rate. I combined it with census api on income per capita in 2016 by zipcode. I anticipate the project to have the following parts:

- Basic statistics and graphs on prescription by professions, state income and opioid death rate by state.
- Visualization of prescription on overdose death and local per capita income.
- A national map of opioid prescription and opioid overdose death by state.

```
In [1]: import pandas as pd
        from sodapy import Socrata
        import numpy as np
        import matplotlib.pyplot as plt
        from census import Census
        from us import states
        import requests
        import json
```

```
In [2]: client = Socrata("data.cms.gov", 'jwf3Ur2ya3N6XrQBYjgV4vdXw', username=
        "sw2804@nyu.edu",
        password="mW@ng4*0o!")
```

```
In [3]: results = client.get("3hwr-y534", limit=1072997)
```

```
In [4]: results_df = pd.DataFrame.from_records(results)
```

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

```
Out[5]:
```

	npi	nppes_provider_first_name	nppes_provider_last_org_name	nppes_pro
0	1831127620	FRED	BRESSLER	TX
1	1588095855	REID	BREWSTER	CA
2	1144293044	BRUCE	BRIDEWELL	FL
3	1255428645	THOMAS	BRIDGMAN	WA
4	1619082526	MARTIN	BRIGGS	LA

```
In [6]: opioid = results_df.dropna()
        opioid.head()
```

```
Out[6]:
```

	npi	nppes_provider_first_name	nppes_provider_last_org_name	nppes_pr
7	1326336751	LOUIS	BRITT	TN
14	1841342003	CHARLES	BROWN	CA
15	1811325749	KARA	BROWN	OR
26	1194710434	JILL	BUCK	IN
31	1417264565	RYAN	BURKE	NM

```
In [7]: opioid.rename(columns={'nppes_provider_zip5':'zipcode'}, inplace=True)
        opioid
```

/Applications/anaconda3/lib/python3.6/site-packages/pandas/core/frame.py:3027: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: <http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy>

```
return super(DataFrame, self).rename(**kwargs)
```

```
Out[7]:
```

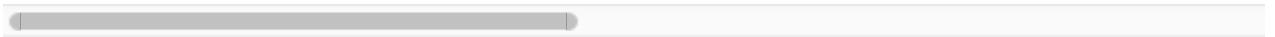
	npi	nppes_provider_first_name	nppes_provider_last_org_name	npp
7	1326336751	LOUIS	BRITT	TN
14	1841342003	CHARLES	BROWN	CA

15	1811325749	KARA	BROWN	OR
26	1194710434	JILL	BUCK	IN
31	1417264565	RYAN	BURKE	NM
34	1235186511	MARY	BURTON	UT
54	1265506067	LARISA	CANTER	CA
60	1164418778	BRENDA	CARRELL	MO
68	1649243320	RAYMOND	CAST	FL
72	1639395957	KIMBERLY	CERTA	VA
76	1982036091	JAMMIE	CHANG	TX
81	1346370509	HAI-SOU	CHEN	CA
86	1528247590	SUMA	CHERUKURI	MI
88	1013024538	SIV	CHHAY	TX
102	1629278502	PIER	CIPRIANI	PA
118	1134141542	SARAH	FRYE	MO
154	1033220991	PAUL	GOLDEN	CA
166	1508970039	AZIZ	GOPALANI	PA
182	1639368749	ELLEN	GREENE	NY
183	1346351533	JAMES	GREENE	GA
202	1518933571	STEVEN	KIRKHAM	OH
225	1578614061	KIMBERLY	KRON	OH
227	1174810360	LAURA	KRUTER	NY
235	1861649527	SUMEET	KUMAR	TX
238	1295929586	PRASAD	KURELLA	OH
245	1700871191	CHRISTINA	LANNOM	TN
251	1235290594	JULIE	LARSON	CO
257	1225041262	ALEX	LAU	CA
269	1619984705	SARA	MOBASSERI	GA
270	1962437491	ELLEN	MODELL	WA
...

1072967	1326013780	ARMANDO	WYATT	VA
1072968	1629041298	DAVID	GIRDANY	PA
1072969	1639313158	MARGARET	RODENBECK	IN
1072970	1679534564	TIMOTHY	MARTIN	PA
1072971	1124149190	ELIZABETH	SHARP	AR
1072972	1235225780	CLAUDE	FORTIN	IL
1072973	1891892626	DAVID	BENDER	WV
1072974	1801850995	SHELBY	BAILEY	AL
1072975	1134221062	MICHAEL	MAROULES	NJ
1072976	1669676714	MEERA	REDDY	CA
1072977	1194739508	STEVEN	GUTTENBERG	DC
1072978	1639126824	REGULO	TOBIAS	KY
1072979	1780600338	JAY	STIEFEL	NJ
1072980	1700996659	WILLIAM	MOSS	LA
1072981	1275700528	MEENA	SINGH	KS
1072982	1154765816	ALEXANDRA	DRESSEN	MA
1072983	1073574497	DARIUS	CLARKE	TX
1072984	1639102031	PERRY	WYNER	NY
1072985	1538276472	THOMAS	KRUPITZER	OH
1072986	1023178506	STEVEN	CHANDLER	AL
1072987	1285613653	MICHAEL	HOUSE	TX
1072988	1932272887	CHARLES	WHITEHILL	CA
1072989	1083601033	MALCOLM	SCHWARTZ	NJ
1072990	1497774723	DAVID	FRANK	WA
1072991	1497735096	HOWARD	SANDER	NY
1072992	1346419918	THANH	NGUYEN	CA
1072993	1841356219	CHRISTINE	PREBLICK	NY
1072994	1982609103	GORDON	SIEGEL	IL

1072995	1891880902	NORA	MEANEY-ELMAN	NY
1072996	1538507330	NUJUD	DAHAM	MA

753998 rows × 9 columns



Now getting income per capita at state level and zipcode level in 2016 data from census

```
In [8]: my_api_key = '34caee0abc641aead3785ea57203c35b70def3ee'
c = Census(my_api_key)
code = ("NAME", "B19301_001E")    #PER CAPITA INCOME IN THE PAST 12 MONTHS (IN 2016 INFLATION-ADJUSTED DOLLARS)
state_wealth = c.acs5.get(code, {'for': 'state:* '})
state_wealth = pd.DataFrame(state_wealth)
state_wealth.rename(columns={'NAME': 'State', 'B19301_001E': 'state_per_capita_income'}, inplace=True)
state_wealth.drop('state', axis=1, inplace=True)
state_wealth
```

Out[8]:

	state_per_capita_income	State
0	24736.0	Alabama
1	34191.0	Alaska
2	26686.0	Arizona
3	23401.0	Arkansas
4	31458.0	California
5	33230.0	Colorado
6	39906.0	Connecticut
7	31118.0	Delaware
8	48781.0	District of Columbia
9	27598.0	Florida
10	26678.0	Georgia
11	30970.0	Hawaii
12	24280.0	Idaho
13	31502.0	Illinois
14	26117.0	Indiana

15	28872.0	Iowa
16	28478.0	Kansas
17	24802.0	Kentucky
18	25515.0	Louisiana
19	28473.0	Maine
20	37756.0	Maryland
21	38069.0	Massachusetts
22	27549.0	Michigan
23	33225.0	Minnesota
24	21651.0	Mississippi
25	27044.0	Missouri
26	27309.0	Montana
27	28596.0	Nebraska
28	27253.0	Nevada
29	35264.0	New Hampshire
30	37538.0	New Jersey
31	24459.0	New Mexico
32	34212.0	New York
33	26779.0	North Carolina
34	33107.0	North Dakota
35	27800.0	Ohio
36	25628.0	Oklahoma
37	28822.0	Oregon
38	30137.0	Pennsylvania
39	31904.0	Rhode Island
40	25521.0	South Carolina
41	27516.0	South Dakota
42	26019.0	Tennessee
43	27828.0	Texas

44	25600.0	Utah
45	30663.0	Vermont
46	34967.0	Virginia
47	32999.0	Washington
48	24002.0	West Virginia
49	29253.0	Wisconsin
50	30139.0	Wyoming
51	11688.0	Puerto Rico

```
In [9]: zipcode = pd.DataFrame(c.acs5.get(code,
                                     {'for': 'zip code tabulation
                                     area: *'}))

        zipcode.head()
```

Out[9]:

	B19301_001E	NAME	zip code tabulation area
0	30430.0	ZCTA5 01001	01001
1	26072.0	ZCTA5 01002	01002
2	3829.0	ZCTA5 01003	01003
3	32169.0	ZCTA5 01005	01005
4	36359.0	ZCTA5 01007	01007

```
In [10]: zipcode.rename(columns={'zip code tabulation area':'zipcode', 'B19301_0
                                01E':'per_capita_income'}, inplace=True)
```

Merging these two datasets together by zipcode

```

In [11]: combo = pd.merge(opioid, zipcode, # left df, right df
                           on='zipcode',    # link with cntry
                           how='left',      # add to left...does this matter here?
                           inner: inner section of the keys,
                           # outter: takes all the keys on both dataframes
                           indicator=True) # Tells us what happened

combo.rename(columns={'nppes_provider_state': 'Abbreviation', 'nppes_provider_first_name': 'first_name', 'nppes_provider_last_org_name': 'last_name'}, inplace=True)
combo[combo["_merge"] == "left_only"]

```

Out[11]:

	npi	first_name	last_name	Abbreviation	zipcode	opiod_claim
80	1720139348	SCOTT	ZAMVIL	CA	94143	0
120	1306067921	ROCHELLE	AASER	TN	37933	0
175	1205096609	MOHAMED	ABAZEED	OH	44195	0
182	1538509062	ELHAM	ABBAS	KS	66160	0
192	1790756179	UME	ABBAS	OH	44195	0
194	1689881781	FARHA	ABBASI	MI	48824	0
258	1346276573	RICHARD	ABBOTT	CA	94143	0
306	1164693008	KHALED	ABDEL-KADER	TN	37232	0
308	1518181239	AHMED	ABDEL-LATIF	KY	40536	0
312	1982620340	MAY	ABDEL-WAHAB	OH	44195	0
392	1366509804	IBRAHIM	ABDULLAH	NE	68198	0
393	1558504142	LUBNA	ABDULLAH	MD	21287	0
446	1659690337	OLIVER	ABELA	MI	48824	0
469	1942263975	TOM	ABELSON	OH	44195	0
487	1093009219	MAC	ABERNATHY	SC	29425	0
555	1558415067	THOMAS	ABLEMAN	VA	22908	0

558	1770528192	ARTHUR	ABLIN	CA	94143	0
592	1144427766	NUHAD	ABOU ZEID	NC	27157	0
600	1063617470	HIBA	ABOUASSI	NC	27710	0
601	1265450233	NABIL	ABOUCHALA	ND	58122	0
608	1164715272	NAEL	ABOUL-HOSN	KY	40536	0
629	1720242464	ANU	ABRAHAM	UT	84132	0
647	1073660916	FENOTE	ABRAHAM	MD	20889	0
660	1831486182	JULIE	ABRAHAM	OH	44195	0
683	1740456144	SHINY	ABRAHAM	OH	44195	0
687	1467427096	THEODORE	ABRAHAM	MD	21287	0
814	1912081308	PAUL	ABRINKO	CA	94143	0
834	1366784902	DENEEN	ABSTON	OH	44195	0
838	1073546016	CHRISTINE	ABT	IL	60122	0
846	1720375785	AMAL	ABU LIBDEH	VA	22908	0
...
753691	1366436214	ANITHA	KUCHIPUDI	NH	3076	28
753708	1033313747	STEFAN	HOLUBAR	NH	3756	30
753721	1225384951	KRISTIN	MIKKELSEN	ME	4005	12
753726	1538143052	LYNN	TANOUE	CT	6519	15
753731	1538326640	MELISSA	MONAGHAN	NJ	7740	40
753748	1003840836	GAIL	LEE	MA	2111	25
753754	1841220829	YAN	LUPYAN	NJ	8816	119
753777	1023349503	SWATHI	ELURI	MD	21287	23
753780	1407859069	ANTONELLA	MAIETTA	NJ	7801	15

753792	1467476044	RAMON	MENDEZ-SEXTO	PR	966	84
753806	1477598928	NANCY	SCANGARELLO	NJ	7102	32
753810	1578990388	BERNADENE	LAWRENCE-PHILLIP	CT	6810	19
753811	1568590487	DOUGLASS	BIBULD	MA	2301	27
753820	1093857476	GERARD	GRAHAM	CT	6001	11
753830	1790787323	PAUL	BARLOW	NH	3038	119
753832	1245280270	GRISELLE	GERENA RAMIREZ	PR	754	338
753838	1861619967	DEBORAH	APPLEYARD	VI	820	77
753856	1124133657	DANIEL	BROWN	MA	1960	37
753861	1124002365	DAWN	HOGAN	RI	2863	58
753876	1407881915	MICHAEL	VIRATA	CT	6511	73
753897	1225041965	PAUL	FONTANAZZA	NJ	7029	11
753910	1992080212	EDUARDO	SALGUERO	PR	727	26
753931	1871852954	ANDREA	WHITE	TX	77555	18
753934	1417212192	JASON	POCO	NJ	8873	169
753940	1851310726	RAASHAN	WILLIAMS	NJ	7087	22
753976	1134221062	MICHAEL	MAROULES	NJ	7013	117
753980	1780600338	JAY	STIEFEL	NJ	8096	58
753983	1154765816	ALEXANDRA	DRESSEN	MA	2148	12
753990	1083601033	MALCOLM	SCHWARTZ	NJ	7090	28
753997	1538507330	NUJUD	DAHAM	MA	2111	12

94195 rows × 12 columns



```
In [12]: combo.dropna(inplace=True)
         combo.drop('_merge',axis=1,inplace=True)
         combo
```

Out[12]:

--	--	--	--	--	--	--

	npi	first_name	last_name	Abbreviation	zipcode	opioid_clai
0	1326336751	LOUIS	BRITT	TN	38018	0
1	1841342003	CHARLES	BROWN	CA	92618	0
2	1811325749	KARA	BROWN	OR	97015	0
3	1194710434	JILL	BUCK	IN	46143	0
4	1417264565	RYAN	BURKE	NM	87106	0
5	1235186511	MARY	BURTON	UT	84105	0
6	1265506067	LARISA	CANTER	CA	92660	0
7	1164418778	BRENDA	CARRELL	MO	65804	0
8	1649243320	RAYMOND	CAST	FL	32209	0
9	1639395957	KIMBERLY	CERTA	VA	22192	0
10	1982036091	JAMMIE	CHANG	TX	75034	0
11	1346370509	HAI-SOU	CHEN	CA	91776	0
12	1528247590	SUMA	CHERUKURI	MI	49525	0
13	1013024538	SIV	CHHAY	TX	75052	0
14	1629278502	PIER	CIPRIANI	PA	18940	0
15	1134141542	SARAH	FRYE	MO	65804	0
16	1033220991	PAUL	GOLDEN	CA	95350	0

17	1508970039	AZIZ	GOPALANI	PA	15017	0
18	1639368749	ELLEN	GREENE	NY	10016	0
19	1346351533	JAMES	GREENE	GA	30329	0
20	1518933571	STEVEN	KIRKHAM	OH	43302	0
21	1578614061	KIMBERLY	KRON	OH	44646	0
22	1174810360	LAURA	KRUTER	NY	11795	0
23	1861649527	SUMEET	KUMAR	TX	78229	0
24	1295929586	PRASAD	KURELLA	OH	44718	0
25	1700871191	CHRISTINA	LANNOM	TN	37067	0
26	1235290594	JULIE	LARSON	CO	81505	0
27	1225041262	ALEX	LAU	CA	94063	0
28	1619984705	SARA	MOBASSERI	GA	30309	0
29	1962437491	ELLEN	MODELL	WA	98026	0
...
753963	1033181821	CHRISTOPHER	THOMAS	FL	33484	329
753964	1700852092	ARTHUR	OZOLIN	WA	98405	88
753965	1376503623	SUSANNE	JOHNSON	TX	75093	38
753966	1538176011	ROBERT	PADILLA	CA	95350	11

753967	1366456584	DOROTHY	TANNER	IN	47274	1539
753968	1326013780	ARMANDO	WYATT	VA	23510	91
753969	1629041298	DAVID	GIRDANY	PA	15501	70
753970	1639313158	MARGARET	RODENBECK	IN	46845	23
753971	1679534564	TIMOTHY	MARTIN	PA	15024	268
753972	1124149190	ELIZABETH	SHARP	AR	72703	111
753973	1235225780	CLAUDE	FORTIN	IL	62702	523
753974	1891892626	DAVID	BENDER	WV	26354	1345
753975	1801850995	SHELBY	BAILEY	AL	35661	42
753977	1669676714	MEERA	REDDY	CA	91367	198
753978	1194739508	STEVEN	GUTTENBERG	DC	20006	26
753979	1639126824	REGULO	TOBIAS	KY	40207	148
753981	1700996659	WILLIAM	MOSS	LA	70601	116
753982	1275700528	MEENA	SINGH	KS	66216	34
753984	1073574497	DARIUS	CLARKE	TX	78701	83
753985	1639102031	PERRY	WYNER	NY	11570	61

753986	1538276472	THOMAS	KRUPITZER	OH	44256	228
753987	1023178506	STEVEN	CHANDLER	AL	35601	450
753988	1285613653	MICHAEL	HOUSE	TX	76201	78
753989	1932272887	CHARLES	WHITEHILL	CA	94611	302
753991	1497774723	DAVID	FRANK	WA	98101	46
753992	1497735096	HOWARD	SANDER	NY	10016	15
753993	1346419918	THANH	NGUYEN	CA	95340	98
753994	1841356219	CHRISTINE	PREBLICK	NY	10003	14
753995	1982609103	GORDON	SIEGEL	IL	60611	21
753996	1891880902	NORA	MEANEY-ELMAN	NY	14221	52

659266 rows × 11 columns

```
In [14]: state=pd.read_csv('https://github.com/jasonong/List-of-US-States/raw/master/states.csv')
```

```
In [15]: combo2=pd.merge(combo,state,on='Abbreviation',how='left',indicator=True)
         combo2[combo2["_merge"] == "left_only"]
         combo2.dropna(inplace=True)
         combo2.drop('_merge',axis=1,inplace=True)
         combo2
```

Out[15]:

	npi	first_name	last_name	Abbreviation	zipcode	opioid_clai
0	1326336751	LOUIS	BRITT	TN	38018	0

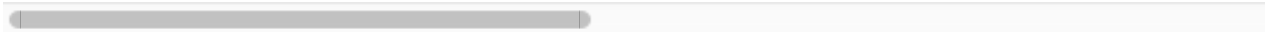
1	1841342003	CHARLES	BROWN	CA	92618	0
2	1811325749	KARA	BROWN	OR	97015	0
3	1194710434	JILL	BUCK	IN	46143	0
4	1417264565	RYAN	BURKE	NM	87106	0
5	1235186511	MARY	BURTON	UT	84105	0
6	1265506067	LARISA	CANTER	CA	92660	0
7	1164418778	BRENDA	CARRELL	MO	65804	0
8	1649243320	RAYMOND	CAST	FL	32209	0
9	1639395957	KIMBERLY	CERTA	VA	22192	0
10	1982036091	JAMMIE	CHANG	TX	75034	0
11	1346370509	HAI-SOU	CHEN	CA	91776	0
12	1528247590	SUMA	CHERUKURI	MI	49525	0
13	1013024538	SIV	CHHAY	TX	75052	0
14	1629278502	PIER	CIPRIANI	PA	18940	0
15	1134141542	SARAH	FRYE	MO	65804	0
16	1033220991	PAUL	GOLDEN	CA	95350	0
17	1508970039	AZIZ	GOPALANI	PA	15017	0

18	1639368749	ELLEN	GREENE	NY	10016	0
19	1346351533	JAMES	GREENE	GA	30329	0
20	1518933571	STEVEN	KIRKHAM	OH	43302	0
21	1578614061	KIMBERLY	KRON	OH	44646	0
22	1174810360	LAURA	KRUTER	NY	11795	0
23	1861649527	SUMEET	KUMAR	TX	78229	0
24	1295929586	PRASAD	KURELLA	OH	44718	0
25	1700871191	CHRISTINA	LANNOM	TN	37067	0
26	1235290594	JULIE	LARSON	CO	81505	0
27	1225041262	ALEX	LAU	CA	94063	0
28	1619984705	SARA	MOBASSERI	GA	30309	0
29	1962437491	ELLEN	MODELL	WA	98026	0
...
659236	1033181821	CHRISTOPHER	THOMAS	FL	33484	329
659237	1700852092	ARTHUR	OZOLIN	WA	98405	88
659238	1376503623	SUSANNE	JOHNSON	TX	75093	38
659239	1538176011	ROBERT	PADILLA	CA	95350	11
659240	1366456584	DOROTHY	TANNER	IN	47274	1539

659241	1326013780	ARMANDO	WYATT	VA	23510	91
659242	1629041298	DAVID	GIRDANY	PA	15501	70
659243	1639313158	MARGARET	RODENBECK	IN	46845	23
659244	1679534564	TIMOTHY	MARTIN	PA	15024	268
659245	1124149190	ELIZABETH	SHARP	AR	72703	111
659246	1235225780	CLAUDE	FORTIN	IL	62702	523
659247	1891892626	DAVID	BENDER	WV	26354	1345
659248	1801850995	SHELBY	BAILEY	AL	35661	42
659249	1669676714	MEERA	REDDY	CA	91367	198
659250	1194739508	STEVEN	GUTTENBERG	DC	20006	26
659251	1639126824	REGULO	TOBIAS	KY	40207	148
659252	1700996659	WILLIAM	MOSS	LA	70601	116
659253	1275700528	MEENA	SINGH	KS	66216	34
659254	1073574497	DARIUS	CLARKE	TX	78701	83
659255	1639102031	PERRY	WYNER	NY	11570	61
659256	1538276472	THOMAS	KRUPITZER	OH	44256	228
659257	1023178506	STEVEN	CHANDLER	AL	35601	450

659258	1285613653	MICHAEL	HOUSE	TX	76201	78
659259	1932272887	CHARLES	WHITEHILL	CA	94611	302
659260	1497774723	DAVID	FRANK	WA	98101	46
659261	1497735096	HOWARD	SANDER	NY	10016	15
659262	1346419918	THANH	NGUYEN	CA	95340	98
659263	1841356219	CHRISTINE	PREBLICK	NY	10003	14
659264	1982609103	GORDON	SIEGEL	IL	60611	21
659265	1891880902	NORA	MEANEY-ELMAN	NY	14221	52

659241 rows × 12 columns



I also want to get death by opioid overdose data

```
In [16]: url='https://data.opendatasoft.com/api/records/1.0/search/?dataset=prescription-opioid-overdose-deaths-and-death-rate-per-100000-population-age-adj%40public&rows=867&facet=prescription_opioid_overdose_deaths_missing_reason&facet=footnotes&facet=location&facet=year'
```

```
In [17]: r=requests.get(url)
```

```
In [18]: print(r.headers["content-type"])
application/json; charset=utf-8
```

```
In [19]: data= json.loads(r.content.decode("utf-8"))
```

```
In [20]: df=pd.DataFrame(data['records'])
dff=pd.DataFrame(df['fields'].tolist())
dff.drop('prescription_opioid_overdose_deaths_missing_reason',axis=1,inplace=True)
death=dff.set_index('year').loc['2015']
death.rename(columns={'location':'State'}, inplace=True)
death
```

Out[20]:

	State	prescription_opioid_overdose_deaths
year		
2015	Maryland	747.0
2015	Minnesota	221.0
2015	West Virginia	512.0
2015	Washington	400.0
2015	Montana	40.0
2015	Idaho	73.0
2015	Alabama	174.0
2015	Arkansas	186.0
2015	Nebraska	48.0
2015	North Dakota	22.0
2015	Vermont	57.0
2015	Louisiana	149.0
2015	Oregon	228.0
2015	Rhode Island	216.0
2015	South Dakota	21.0
2015	New Mexico	211.0
2015	Ohio	1800.0
2015	California	1414.0
2015	Pennsylvania	853.0
2015	Texas	724.0
2015	Alaska	67.0

2015	Delaware	77.0
2015	Kansas	127.0
2015	New Hampshire	340.0
2015	Virginia	548.0
2015	Tennessee	853.0
2015	Wisconsin	385.0
2015	District of Columbia	49.0
2015	Iowa	122.0
2015	Massachusetts	1132.0
2015	Wyoming	36.0
2015	Connecticut	412.0
2015	New Jersey	473.0
2015	Florida	1417.0
2015	Kentucky	644.0
2015	Indiana	279.0
2015	Michigan	873.0
2015	Missouri	439.0
2015	Nevada	316.0
2015	Arizona	406.0
2015	Mississippi	110.0
2015	South Carolina	468.0
2015	Utah	423.0
2015	Illinois	588.0
2015	Hawaii	50.0
2015	Maine	203.0
2015	New York	1390.0
2015	Oklahoma	389.0
2015	Colorado	328.0
2015	North Carolina	864.0

2015	Georgia	694.0
-------------	---------	-------

```
In [21]: combo3=pd.merge(combo2,death,on='State',how='left',indicator=True)
        combo3.drop(['NAME','_merge'],axis=1,inplace=True)
        combo3
```

Out[21]:

	npi	first_name	last_name	Abbreviation	zipcode	opioid_clai
0	1326336751	LOUIS	BRITT	TN	38018	0
1	1841342003	CHARLES	BROWN	CA	92618	0
2	1811325749	KARA	BROWN	OR	97015	0
3	1194710434	JILL	BUCK	IN	46143	0
4	1417264565	RYAN	BURKE	NM	87106	0
5	1235186511	MARY	BURTON	UT	84105	0
6	1265506067	LARISA	CANTER	CA	92660	0
7	1164418778	BRENDA	CARRELL	MO	65804	0
8	1649243320	RAYMOND	CAST	FL	32209	0
9	1639395957	KIMBERLY	CERTA	VA	22192	0
10	1982036091	JAMMIE	CHANG	TX	75034	0
11	1346370509	HAI-SOU	CHEN	CA	91776	0
12	1528247590	SUMA	CHERUKURI	MI	49525	0
13	1013024538	SIV	CHHAY	TX	75052	0
14	1629278502	PIER	CIPRIANI	PA	18940	0
15	1134141542	SARAH	FRYE	MO	65804	0
16	1033220991	PAUL	GOLDEN	CA	95350	0
17	1508970039	AZIZ	GOPALANI	PA	15017	0
18	1639368749	ELLEN	GREENE	NY	10016	0
19	1346351533	JAMES	GREENE	GA	30329	0
20	1518933571	STEVEN	KIRKHAM	OH	43302	0

21	1578614061	KIMBERLY	KRON	OH	44646	0
22	1174810360	LAURA	KRUTER	NY	11795	0
23	1861649527	SUMEET	KUMAR	TX	78229	0
24	1295929586	PRASAD	KURELLA	OH	44718	0
25	1700871191	CHRISTINA	LANNOM	TN	37067	0
26	1235290594	JULIE	LARSON	CO	81505	0
27	1225041262	ALEX	LAU	CA	94063	0
28	1619984705	SARA	MOBASSERI	GA	30309	0
29	1962437491	ELLEN	MODELL	WA	98026	0
...
659211	1033181821	CHRISTOPHER	THOMAS	FL	33484	329
659212	1700852092	ARTHUR	OZOLIN	WA	98405	88
659213	1376503623	SUSANNE	JOHNSON	TX	75093	38
659214	1538176011	ROBERT	PADILLA	CA	95350	11
659215	1366456584	DOROTHY	TANNER	IN	47274	1539
659216	1326013780	ARMANDO	WYATT	VA	23510	91
659217	1629041298	DAVID	GIRDANY	PA	15501	70
659218	1639313158	MARGARET	RODENBECK	IN	46845	23
659219	1679534564	TIMOTHY	MARTIN	PA	15024	268
659220	1124149190	ELIZABETH	SHARP	AR	72703	111
659221	1235225780	CLAUDE	FORTIN	IL	62702	523
659222	1891892626	DAVID	BENDER	WV	26354	1345
659223	1801850995	SHELBY	BAILEY	AL	35661	42

659224	1669676714	MEERA	REDDY	CA	91367	198
659225	1194739508	STEVEN	GUTTENBERG	DC	20006	26
659226	1639126824	REGULO	TOBIAS	KY	40207	148
659227	1700996659	WILLIAM	MOSS	LA	70601	116
659228	1275700528	MEENA	SINGH	KS	66216	34
659229	1073574497	DARIUS	CLARKE	TX	78701	83
659230	1639102031	PERRY	WYNER	NY	11570	61
659231	1538276472	THOMAS	KRUPITZER	OH	44256	228
659232	1023178506	STEVEN	CHANDLER	AL	35601	450
659233	1285613653	MICHAEL	HOUSE	TX	76201	78
659234	1932272887	CHARLES	WHITEHILL	CA	94611	302
659235	1497774723	DAVID	FRANK	WA	98101	46
659236	1497735096	HOWARD	SANDER	NY	10016	15
659237	1346419918	THANH	NGUYEN	CA	95340	98
659238	1841356219	CHRISTINE	PREBLICK	NY	10003	14
659239	1982609103	GORDON	SIEGEL	IL	60611	21
659240	1891880902	NORA	MEANEY-ELMAN	NY	14221	52

659241 rows × 12 columns



```
In [22]: final_df=pd.merge(combo3,state_wealth,on='State',how='left')
         final_df
```

Out[22]:

Out[44]:

	npi	first_name	last_name	Abbreviation	zipcode	opioid_clai
0	1326336751	LOUIS	BRITT	TN	38018	0
1	1841342003	CHARLES	BROWN	CA	92618	0
2	1811325749	KARA	BROWN	OR	97015	0
3	1194710434	JILL	BUCK	IN	46143	0
4	1417264565	RYAN	BURKE	NM	87106	0
5	1235186511	MARY	BURTON	UT	84105	0
6	1265506067	LARISA	CANTER	CA	92660	0
7	1164418778	BRENDA	CARRELL	MO	65804	0
8	1649243320	RAYMOND	CAST	FL	32209	0
9	1639395957	KIMBERLY	CERTA	VA	22192	0
10	1982036091	JAMMIE	CHANG	TX	75034	0
11	1346370509	HAI-SOU	CHEN	CA	91776	0
12	1528247590	SUMA	CHERUKURI	MI	49525	0
13	1013024538	SIV	CHHAY	TX	75052	0
14	1629278502	PIER	CIPRIANI	PA	18940	0
15	1134141542	SARAH	FRYE	MO	65804	0
16	1033220991	PAUL	GOLDEN	CA	95350	0
17	1508970039	AZIZ	GOPALANI	PA	15017	0
18	1639368749	ELLEN	GREENE	NY	10016	0
19	1346351533	JAMES	GREENE	GA	30329	0
20	1518933571	STEVEN	KIRKHAM	OH	43302	0
21	1578614061	KIMBERLY	KRON	OH	44646	0
22	1174810360	LAURA	KRUTER	NY	11795	0
23	1861649527	SUMEET	KUMAR	TX	78229	0
24	1295929586	PRASAD	KURELLA	OH	44718	0

25	1700871191	CHRISTINA	LANNOM	TN	37067	0
26	1235290594	JULIE	LARSON	CO	81505	0
27	1225041262	ALEX	LAU	CA	94063	0
28	1619984705	SARA	MOBASSERI	GA	30309	0
29	1962437491	ELLEN	MODELL	WA	98026	0
...
659211	1033181821	CHRISTOPHER	THOMAS	FL	33484	329
659212	1700852092	ARTHUR	OZOLIN	WA	98405	88
659213	1376503623	SUSANNE	JOHNSON	TX	75093	38
659214	1538176011	ROBERT	PADILLA	CA	95350	11
659215	1366456584	DOROTHY	TANNER	IN	47274	1539
659216	1326013780	ARMANDO	WYATT	VA	23510	91
659217	1629041298	DAVID	GIRDANY	PA	15501	70
659218	1639313158	MARGARET	RODENBECK	IN	46845	23
659219	1679534564	TIMOTHY	MARTIN	PA	15024	268
659220	1124149190	ELIZABETH	SHARP	AR	72703	111
659221	1235225780	CLAUDE	FORTIN	IL	62702	523
659222	1891892626	DAVID	BENDER	WV	26354	1345
659223	1801850995	SHELBY	BAILEY	AL	35661	42
659224	1669676714	MEERA	REDDY	CA	91367	198
659225	1194739508	STEVEN	GUTTENBERG	DC	20006	26
659226	1639126824	REGULO	TOBIAS	KY	40207	148

659227	1700996659	WILLIAM	MOSS	LA	70601	116
659228	1275700528	MEENA	SINGH	KS	66216	34
659229	1073574497	DARIUS	CLARKE	TX	78701	83
659230	1639102031	PERRY	WYNER	NY	11570	61
659231	1538276472	THOMAS	KRUPITZER	OH	44256	228
659232	1023178506	STEVEN	CHANDLER	AL	35601	450
659233	1285613653	MICHAEL	HOUSE	TX	76201	78
659234	1932272887	CHARLES	WHITEHILL	CA	94611	302
659235	1497774723	DAVID	FRANK	WA	98101	46
659236	1497735096	HOWARD	SANDER	NY	10016	15
659237	1346419918	THANH	NGUYEN	CA	95340	98
659238	1841356219	CHRISTINE	PREBLICK	NY	10003	14
659239	1982609103	GORDON	SIEGEL	IL	60611	21
659240	1891880902	NORA	MEANEY-ELMAN	NY	14221	52

659241 rows × 13 columns

