In [ ]:
```
'''
CS 418: Project 1
Authors: Anusha Sagi, Fatima Kahack, Lydia Tse
Description: The following is code that merges
the data sets regarding the election results
and the 2018 census information.
'''
```

In [1]:
```
# Load libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
```

In [2]:
```
# Load Election dataset
data_election = pd.read_csv('election_train.csv')
print(data_election)
```

```
      Year State             County      Office       Party     Votes
0     2018    AZ      Apache County  US Senator  Democratic   16298.0
1     2018    AZ      Apache County  US Senator  Republican    7810.0
2     2018    AZ     Cochise County  US Senator  Democratic   17383.0
3     2018    AZ     Cochise County  US Senator  Republican   26929.0
4     2018    AZ    Coconino County  US Senator  Democratic   34240.0
5     2018    AZ    Coconino County  US Senator  Republican   19249.0
6     2018    AZ        Gila County  US Senator  Democratic    7643.0
7     2018    AZ        Gila County  US Senator  Republican   12180.0
8     2018    AZ      Graham County  US Senator  Democratic    3368.0
9     2018    AZ      Graham County  US Senator  Republican    6870.0
10    2018    AZ       La Paz County  US Senator  Democratic    1609.0
11    2018    AZ       La Paz County  US Senator  Republican    3265.0
12    2018    AZ    Maricopa County  US Senator  Democratic  732671.0
13    2018    AZ    Maricopa County  US Senator  Republican  672505.0
14    2018    AZ      Mohave County  US Senator  Democratic   19214.0
15    2018    AZ      Mohave County  US Senator  Republican   50209.0
16    2018    AZ      Navajo County  US Senator  Democratic   16624.0
17    2018    AZ      Navajo County  US Senator  Republican   18767.0
18    2018    AZ        Pima County  US Senator  Democratic  221242.0
19    2018    AZ        Pima County  US Senator  Republican  160550.0
20    2018    AZ  Santa Cruz County  US Senator  Democratic    9241.0
21    2018    AZ  Santa Cruz County  US Senator  Republican    3828.0
22    2018    AZ     Yavapai County  US Senator  Democratic   40160.0
23    2018    AZ     Yavapai County  US Senator  Republican   65308.0
24    2018    CT   Fairfield County  US Senator  Democratic  210899.0
25    2018    CT   Fairfield County  US Senator  Republican  131321.0
26    2018    CT    Hartford County  US Senator  Democratic  203591.0
27    2018    CT    Hartford County  US Senator  Republican  123864.0
28    2018    CT   Middlesex County  US Senator  Democratic   42383.0
29    2018    CT   Middlesex County  US Senator  Republican   32836.0
...    ...   ...               ...         ...         ...       ...
2380  2018    WY      Albany County  US Senator  Democratic    7576.0
2381  2018    WY      Albany County  US Senator  Republican    6366.0
2382  2018    WY    Campbell County  US Senator  Democratic    1628.0
2383  2018    WY    Campbell County  US Senator  Republican   11020.0
2384  2018    WY      Carbon County  US Senator  Democratic    1359.0
2385  2018    WY      Carbon County  US Senator  Republican    3673.0
2386  2018    WY    Converse County  US Senator  Democratic     834.0
2387  2018    WY    Converse County  US Senator  Republican    3959.0
2388  2018    WY     Fremont County  US Senator  Democratic    4734.0
2389  2018    WY     Fremont County  US Senator  Republican    9262.0
2390  2018    WY      Goshen County  US Senator  Democratic    1020.0
2391  2018    WY      Goshen County  US Senator  Republican    3658.0
2392  2018    WY     Johnson County  US Senator  Democratic     722.0
2393  2018    WY     Johnson County  US Senator  Republican    3085.0
2394  2018    WY     Lincoln County  US Senator  Democratic    1152.0
2395  2018    WY     Lincoln County  US Senator  Republican    5846.0
2396  2018    WY     Natrona County  US Senator  Democratic    7285.0
2397  2018    WY     Natrona County  US Senator  Republican   16359.0
2398  2018    WY    Niobrara County  US Senator  Democratic     144.0
2399  2018    WY    Niobrara County  US Senator  Republican     980.0
2400  2018    WY      Platte County  US Senator  Democratic     801.0
2401  2018    WY      Platte County  US Senator  Republican    2850.0
2402  2018    WY    Sublette County  US Senator  Democratic     668.0
2403  2018    WY    Sublette County  US Senator  Republican    2653.0
2404  2018    WY  Sweetwater County  US Senator  Democratic    3943.0
2405  2018    WY  Sweetwater County  US Senator  Republican    8577.0
2406  2018    WY       Uinta County  US Senator  Democratic    1371.0
2407  2018    WY       Uinta County  US Senator  Republican    4713.0
2408  2018    WY    Washakie County  US Senator  Democratic     588.0
2409  2018    WY    Washakie County  US Senator  Republican    2423.0

[2410 rows x 6 columns]
```

In [3]:
```
# 1. Reshape dataset from long format to wide format - pivot_table()
data_election = pd.pivot_table(data_election, index = ['Year','State','County','Office'], columns = 'Party', values = 'Votes', aggfunc = np.sum).reset_index()
print(data_election)
#data_a_tidy.to_excel("output.xlsx")
```

Party	Year	State	County	Office	Democratic	Republican
0	2018	AZ	Apache County	US Senator	16298.0	7810.0
1	2018	AZ	Cochise County	US Senator	17383.0	26929.0
2	2018	AZ	Coconino County	US Senator	34240.0	19249.0
3	2018	AZ	Gila County	US Senator	7643.0	12180.0
4	2018	AZ	Graham County	US Senator	3368.0	6870.0
5	2018	AZ	La Paz County	US Senator	1609.0	3265.0
6	2018	AZ	Maricopa County	US Senator	732671.0	672505.0
7	2018	AZ	Mohave County	US Senator	19214.0	50209.0
8	2018	AZ	Navajo County	US Senator	16624.0	18767.0
9	2018	AZ	Pima County	US Senator	221242.0	160550.0
10	2018	AZ	Santa Cruz County	US Senator	9241.0	3828.0
11	2018	AZ	Yavapai County	US Senator	40160.0	65308.0
12	2018	CT	Fairfield County	US Senator	210899.0	131321.0
13	2018	CT	Hartford County	US Senator	203591.0	123864.0
14	2018	CT	Middlesex County	US Senator	42383.0	32836.0
15	2018	CT	New Haven County	US Senator	179714.0	126004.0
16	2018	CT	Tolland County	US Senator	34732.0	28046.0
17	2018	CT	Windham County	US Senator	20490.0	19032.0
18	2018	DE	Sussex County	US Senator	40675.0	50391.0
19	2018	FL	Alachua County	US Senator	74493.0	40599.0
20	2018	FL	Baker County	US Senator	1945.0	8579.0
21	2018	FL	Bay County	US Senator	16723.0	46681.0
22	2018	FL	Bradford County	US Senator	2879.0	7576.0
23	2018	FL	Brevard County	US Senator	121112.0	160305.0
24	2018	FL	Broward County	US Senator	472239.0	211397.0
25	2018	FL	Charlotte County	US Senator	33525.0	52916.0
26	2018	FL	Citrus County	US Senator	22660.0	48008.0
27	2018	FL	Collier County	US Senator	54390.0	101266.0
28	2018	FL	Desoto County	US Senator	3328.0	5503.0
29	2018	FL	Dixie County	US Senator	1322.0	4442.0
...	...	...	...	...	...	...
1175	2018	WV	Pocahontas County	US Senator	1269.0	1411.0
1176	2018	WV	Preston County	US Senator	3686.0	5943.0
1177	2018	WV	Raleigh County	US Senator	10581.0	12620.0
1178	2018	WV	Randolph County	US Senator	4472.0	4017.0
1179	2018	WV	Ritchie County	US Senator	1082.0	1961.0
1180	2018	WV	Roane County	US Senator	2165.0	1899.0
1181	2018	WV	Summers County	US Senator	2069.0	1868.0
1182	2018	WV	Taylor County	US Senator	2376.0	2642.0
1183	2018	WV	Tucker County	US Senator	1469.0	1502.0
1184	2018	WV	Tyler County	US Senator	1065.0	1603.0
1185	2018	WV	Upshur County	US Senator	3102.0	4010.0
1186	2018	WV	Wayne County	US Senator	6395.0	5954.0
1187	2018	WV	Wetzel County	US Senator	2518.0	2135.0
1188	2018	WV	Wood County	US Senator	14189.0	13696.0
1189	2018	WV	Wyoming County	US Senator	2607.0	3096.0
1190	2018	WY	Albany County	US Senator	7576.0	6366.0
1191	2018	WY	Campbell County	US Senator	1628.0	11020.0
1192	2018	WY	Carbon County	US Senator	1359.0	3673.0
1193	2018	WY	Converse County	US Senator	834.0	3959.0
1194	2018	WY	Fremont County	US Senator	4734.0	9262.0
1195	2018	WY	Goshen County	US Senator	1020.0	3658.0
1196	2018	WY	Johnson County	US Senator	722.0	3085.0
1197	2018	WY	Lincoln County	US Senator	1152.0	5846.0
1198	2018	WY	Natrona County	US Senator	7285.0	16359.0
1199	2018	WY	Niobrara County	US Senator	144.0	980.0
1200	2018	WY	Platte County	US Senator	801.0	2850.0
1201	2018	WY	Sublette County	US Senator	668.0	2653.0
1202	2018	WY	Sweetwater County	US Senator	3943.0	8577.0
1203	2018	WY	Uinta County	US Senator	1371.0	4713.0
1204	2018	WY	Washakie County	US Senator	588.0	2423.0

[1205 rows x 6 columns]

In [4]:
```
# Load demographics dataset
data_demographics = pd.read_csv('demographics_train.csv')
print(data_demographics)
```

	State	County	FIPS	Total Population
0	Wisconsin	La Crosse	55063	117538
1	Virginia	Alleghany	51005	15919
2	Indiana	Fountain	18045	16741
3	Ohio	Geauga	39055	94020
4	Wisconsin	Jackson	55053	20566
5	Texas	Baylor	48023	3639
6	Nebraska	Madison	31119	35125
7	Hawaii	Hawaii	15001	193680
8	Tennessee	Henry	47079	32291
9	Michigan	Oceana	26127	26152
10	Nebraska	Pierce	31139	7179
11	Texas	Jack	48237	8866
12	Florida	Walton	12131	61528
13	Virginia	Washington	51191	54562
14	Florida	Escambia	12033	309574
15	Texas	Wheeler	48483	5642
16	Arizona	Yavapai	4025	218586
17	Nebraska	Loup	31115	542
18	Michigan	Antrim	26009	23215
19	Minnesota	Wabasha	27157	21327
20	Nevada	Elko	32007	52029
21	Wisconsin	Dodge	55027	88404
22	Minnesota	Lake of the Woods	27077	3901
23	West Virginia	Tucker	54093	6922
24	Virginia	Lexington City	51678	7036
25	Florida	Gilchrist	12041	17033
26	Tennessee	Claiborne	47025	31701
27	Connecticut	Middlesex	9007	164438
28	Utah	Sevier	49041	20913
29	Massachusetts	Hampshire	25015	161035
...	...	...	...	...
1186	West Virginia	Wood	54107	86262
1187	Florida	Orange	12095	1256055
1188	North Dakota	Walsh	38099	10995
1189	Indiana	White	18181	24265
1190	Montana	Carter	30011	1295
1191	Indiana	Pulaski	18131	12910
1192	Washington	Lewis	53041	75724
1193	Pennsylvania	Greene	42059	37669
1194	Texas	Yoakum	48501	8316
1195	Montana	Pondera	30073	6166
1196	North Dakota	Dickey	38021	5160
1197	Texas	Refugio	48391	7315
1198	Virginia	Smyth	51173	31513
1199	Virginia	Charlotte	51037	12232
1200	Michigan	Grand Traverse	26055	90715
1201	Indiana	Rush	18139	16873
1202	Montana	Deer Lodge	30023	9176
1203	Ohio	Washington	39167	61154
1204	Wisconsin	Portage	55097	70551
1205	New York	Monroe	36055	749236
1206	Florida	Broward	12011	1863780
1207	Washington	Wahkiakum	53069	4051
1208	New York	Tioga	36107	49649
1209	Montana	Carbon	30009	10340
1210	Texas	Johnson	48251	157544
1211	Montana	Lincoln	30053	19268
1212	Ohio	Tuscarawas	39157	92579
1213	Michigan	Newaygo	26123	47957
1214	Tennessee	Lauderdale	47097	27261
1215	Texas	Sabine	48403	10367

	Citizen Voting-Age Population	Percent White, not Hispanic or Latino
0	0	90.537528
1	12705	91.940449
2	12750	95.705155
3	0	95.837056
4	15835	86.662453
5	0	86.644683
6	24885	81.249822
7	0	30.401694
8	25285	87.662197
9	18930	82.486999
10	5385	96.893718
11	6535	78.411911
12	47490	84.447731
13	0	95.484037
14	0	65.219624
15	3785	68.397731
16	0	81.159361
17	435	97.970480
18	0	95.179841
19	16385	94.926619
20	34740	67.354744
21	0	90.629383
22	0	93.975904
23	5690	97.688529

24	6335	80.542922
25	0	87.295250
26	25350	95.495410
27	0	84.738929
28	0	92.244059
29	0	84.610799
...	...	...
1186	67640	95.478890
1187	854605	42.656970
1188	8280	85.657117
1189	0	90.245209
1190	0	96.602317
1191	9870	94.771495
1192	56900	84.489726
1193	0	92.598689
1194	0	34.271284
1195	0	80.960104
1196	0	93.720930
1197	0	42.761449
1198	24960	94.376924
1199	9540	67.282538
1200	70710	92.673758
1201	0	96.325490
1202	0	90.747602
1203	0	95.364163
1204	0	91.815850
1205	0	71.329327
1206	0	39.245351
1207	0	89.681560
1208	38525	95.454088
1209	0	94.787234
1210	0	74.343041
1211	15640	93.351671
1212	70485	95.155489
1213	0	90.716684
1214	0	60.456330
1215	0	86.341275

	Percent Black, not Hispanic or Latino	Percent Hispanic or Latino	\
0	1.214075	1.724549	
1	5.207614	1.432251	
2	0.400215	2.359477	
3	1.256116	1.294405	
4	1.983857	3.082758	
5	1.841165	8.353943	
6	1.155872	14.217794	
7	0.547811	12.405514	
8	8.599919	2.201852	
9	1.131845	14.419547	
10	0.222872	1.587965	
11	4.376269	15.880893	
12	4.950592	5.888376	
13	1.268282	1.414904	
14	21.532816	5.389988	
15	2.658632	26.462247	
16	0.518331	14.054880	
17	0.000000	0.000000	
18	0.323067	1.955632	
19	0.150045	2.874291	
20	1.060947	24.067347	
21	2.085878	4.485091	
22	0.102538	1.230454	
23	0.173360	0.650101	
24	8.996589	4.178511	
25	5.988376	5.336699	
26	1.050440	1.113530	
27	4.664980	5.641032	
28	0.344283	4.882131	
29	2.683268	5.253516	
...	...	...	
1186	1.123322	1.032900	
1187	19.981211	29.338126	
1188	0.300136	10.759436	
1189	0.461570	7.554090	
1190	0.000000	1.389961	
1191	0.209140	2.788536	
1192	0.662934	9.629708	
1193	4.149300	1.422921	
1194	0.000000	63.552189	
1195	0.389231	1.881284	
1196	1.472868	3.333333	
1197	4.743677	49.323308	
1198	2.068987	1.948402	
1199	30.657292	0.948332	
1200	1.234636	2.621397	
1201	1.149766	1.428317	
1202	0.348736	3.073235	
1203	1.100500	1.013834	

1204	0.687446	2.951057
1205	14.497969	8.120272
1206	27.214639	27.564841
1207	0.123426	4.640829
1208	0.713005	1.695905
1209	0.009671	2.379110
1210	2.487559	19.762733
1211	0.057089	2.678015
1212	0.804718	2.349345
1213	1.317847	5.728048
1214	34.789626	2.380690
1215	7.080158	3.839105

	Percent Foreign Born	Percent Female	Percent Age 29 and Under	\
0	2.976059	51.171536	43.241335	
1	1.300333	51.077329	31.660280	
2	1.547100	49.770026	35.899887	
3	2.578175	50.678579	36.281642	
4	1.376058	46.649810	36.292911	
5	2.473207	51.662545	30.090684	
6	6.784342	50.448399	41.432028	
7	11.003717	50.143019	36.008881	
8	1.560806	51.441578	33.238364	
9	5.578923	49.395840	36.643469	
10	0.780053	49.658727	36.634629	
11	5.549289	43.187458	38.732236	
12	5.759979	49.349889	33.165388	
13	1.611011	50.500348	32.005792	
14	4.747492	50.299444	41.364585	
15	9.517901	49.202410	39.188231	
16	6.456955	51.092476	28.717301	
17	0.000000	52.398524	30.996310	
18	2.015938	50.273530	29.450786	
19	1.355090	50.171145	34.594645	
20	8.602895	47.948260	44.928790	
21	1.884530	47.788562	34.642098	
22	1.563702	47.962061	28.479877	
23	0.592314	50.245594	29.731292	
24	3.936896	43.447982	67.367823	
25	2.906123	47.390360	38.307990	
26	1.164001	51.200278	35.673954	
27	7.652732	51.098894	33.241708	
28	1.989193	48.893033	43.877014	
29	7.846120	53.230664	44.243177	
...	...	...	...	
1186	1.001600	51.756277	34.726763	
1187	20.157318	50.827551	43.005123	
1188	3.919964	48.722146	34.588449	
1189	4.314857	50.294663	36.130229	
1190	0.077220	51.737452	30.810811	
1191	0.805577	49.310612	36.018590	
1192	4.512440	50.158470	35.731076	
1193	0.809684	48.084632	35.267727	
1194	25.637326	49.158249	46.284271	
1195	2.189426	49.724295	37.528381	
1196	1.763566	49.806202	36.531008	
1197	2.734108	50.799727	36.896787	
1198	1.256624	50.423635	33.316409	
1199	0.662198	49.959124	35.080118	
1200	2.421871	50.830623	35.075787	
1201	1.060866	50.974930	36.745096	
1202	1.874455	46.785092	29.980384	
1203	1.430814	50.596854	34.483435	
1204	3.119729	49.846211	42.641493	
1205	8.460752	51.694126	39.551223	
1206	32.715718	51.376504	36.701971	
1207	3.505307	51.246606	26.783510	
1208	2.229652	50.554895	34.465951	
1209	1.276596	49.990329	28.143133	
1210	6.138603	50.168842	41.296400	
1211	2.345858	49.974050	27.979033	
1212	1.650482	50.823621	36.475875	
1213	1.978856	49.656984	36.620306	
1214	1.757089	47.734859	39.528997	
1215	0.752387	50.506415	31.243368	

	Percent Age 65 and Older	Median Household Income	Percent Unemployed	\
0	14.702479	51477	4.796952	
1	23.902255	45538	4.560986	
2	18.941521	45924	7.978789	
3	18.028079	74165	4.036902	
4	17.587280	49608	5.569698	
5	24.402308	34382	1.377410	
6	15.404982	48673	3.094085	
7	17.580545	53936	7.437632	
8	21.476572	38378	7.912277	
9	19.088406	41952	8.275556	
10	18.540187	53125	3.981142	

11	15.677871	50390	4.680468
12	18.783318	46910	7.295474
13	20.301675	43835	6.425857
14	15.774581	46117	8.297930
15	17.795108	51082	5.424099
16	28.272625	46638	8.525986
17	24.538745	56750	0.335570
18	25.410295	48825	7.815784
19	18.807146	58865	3.382303
20	9.229468	74672	4.896851
21	16.304692	54111	4.849356
22	21.814919	45732	3.888099
23	22.594626	43529	8.190184
24	14.198408	34464	3.926590
25	19.567898	40881	9.732220
26	18.390587	33428	8.446716
27	17.703937	79837	5.257635
28	15.583608	48872	5.340291
29	14.755177	62608	7.146733
...	...	...	...
1186	18.559737	43944	7.625458
1187	10.741090	49391	8.072223
1188	20.627558	51181	3.314246
1189	18.569957	51547	5.421587
1190	24.942085	46985	2.656250
1191	18.086754	44884	7.046070
1192	19.609371	44526	9.684249
1193	17.104250	49116	6.993228
1194	11.423761	56655	9.136546
1195	19.656179	42193	8.497409
1196	21.705426	55882	2.549773
1197	21.052632	50145	7.469005
1198	19.994923	38906	7.343413
1199	20.372793	33837	6.997030
1200	17.003803	55597	5.848530
1201	17.412434	46380	5.658982
1202	20.891456	39212	3.727181
1203	19.151650	44763	4.689885
1204	14.644725	52411	6.178216
1205	15.491914	53568	7.306146
1206	15.371879	52954	8.710312
1207	31.325599	48116	6.971514
1208	17.925839	58115	6.720605
1209	22.746615	52869	3.743112
1210	13.209643	59095	6.337092
1211	24.828732	35461	11.570571
1212	17.836658	46992	5.734798
1213	17.728382	44049	8.885564
1214	13.891640	32353	11.427738
1215	28.995852	32500	8.259134

	Percent Less than High School Degree \
0	5.474767
1	15.537543
2	12.032155
3	8.928599
4	11.792912
5	13.066667
6	11.678672
7	8.362326
8	15.710841
9	14.335429
10	5.227552
11	21.130221
12	15.083974
13	15.913833
14	10.053723
15	19.678609
16	9.830672
17	2.736318
18	9.311741
19	6.811536
20	16.622189
21	10.956491
22	10.309278
23	12.702390
24	14.962495
25	17.294025
26	22.759203
27	6.153444
28	10.556637
29	5.741345
...	...
1186	10.370080
1187	12.319191
1188	13.114754
1189	10.394650
1190	7.991588

1191	12.931034
1192	12.929994
1193	13.277047
1194	28.571429
1195	10.801144
1196	12.319456
1197	20.225171
1198	17.449694
1199	20.028275
1200	5.543796
1201	15.398535
1202	10.536314
1203	9.795325
1204	6.784049
1205	9.616412
1206	11.688725
1207	6.868490
1208	9.753866
1209	6.142694
1210	16.183214
1211	11.891078
1212	14.068037
1213	14.194114
1214	24.485474
1215	14.643237

	Percent Less than Bachelor's Degree	Percent Rural
0	67.529757	16.827753
1	83.711604	52.393846
2	85.538940	65.951276
3	62.730824	63.968990
4	86.129256	72.238251
5	73.409524	100.000000
6	80.345221	27.904576
7	72.408979	37.990804
8	85.305261	66.848129
9	82.947912	89.890854
10	78.618286	100.000000
11	89.221949	53.217603
12	73.691263	63.125193
13	75.223085	71.676143
14	75.116986	8.302225
15	85.879874	100.000000
16	74.458362	33.197178
17	78.109453	100.000000
18	73.314054	100.000000
19	78.761297	64.541428
20	81.880924	37.850793
21	83.745359	48.641828
22	83.024055	100.000000
23	85.794140	100.000000
24	55.112515	0.000000
25	87.654646	83.918767
26	86.813089	71.554962
27	59.071541	24.540066
28	83.495524	64.714931
29	55.277464	27.443699
...	...	...
1186	79.738786	26.774461
1187	68.077150	2.042138
1188	83.281291	62.055940
1189	84.315482	67.966563
1190	82.860147	100.000000
1191	89.330497	80.943143
1192	84.490042	60.654695
1193	82.669606	66.832963
1194	82.183908	37.327072
1195	77.324750	58.101739
1196	72.047579	100.000000
1197	89.203860	60.679940
1198	85.209810	75.285643
1199	86.958058	100.000000
1200	68.328672	48.023820
1201	85.971564	61.160304
1202	84.100541	33.641643
1203	81.590788	56.562207
1204	69.694941	36.031649
1205	63.335285	6.449975
1206	68.984269	0.016933
1207	84.928385	100.000000
1208	74.963552	65.658680
1209	70.838103	100.000000
1210	82.255463	37.874170
1211	81.041325	79.793773
1212	85.020218	41.580437
1213	85.660708	83.840281
1214	91.425468	58.662592
1215	87.060703	100.000000

```
[1216 rows x 17 columns]
```

In [5]:
```
#2i-Merge elections and demographics
#Rename column names in election - State
data_state_codes = pd.read_csv('State_codes.csv')
data_election['State'] = data_election['State'].map(data_state_codes.set_index('State_code')['State_name'])

print(data_election)
```

Party	Year	State	County	Office	Democratic	Republican
0	2018	Arizona	Apache County	US Senator	16298.0	7810.0
1	2018	Arizona	Cochise County	US Senator	17383.0	26929.0
2	2018	Arizona	Coconino County	US Senator	34240.0	19249.0
3	2018	Arizona	Gila County	US Senator	7643.0	12180.0
4	2018	Arizona	Graham County	US Senator	3368.0	6870.0
5	2018	Arizona	La Paz County	US Senator	1609.0	3265.0
6	2018	Arizona	Maricopa County	US Senator	732671.0	672505.0
7	2018	Arizona	Mohave County	US Senator	19214.0	50209.0
8	2018	Arizona	Navajo County	US Senator	16624.0	18767.0
9	2018	Arizona	Pima County	US Senator	221242.0	160550.0
10	2018	Arizona	Santa Cruz County	US Senator	9241.0	3828.0
11	2018	Arizona	Yavapai County	US Senator	40160.0	65308.0
12	2018	Connecticut	Fairfield County	US Senator	210899.0	131321.0
13	2018	Connecticut	Hartford County	US Senator	203591.0	123864.0
14	2018	Connecticut	Middlesex County	US Senator	42383.0	32836.0
15	2018	Connecticut	New Haven County	US Senator	179714.0	126004.0
16	2018	Connecticut	Tolland County	US Senator	34732.0	28046.0
17	2018	Connecticut	Windham County	US Senator	20490.0	19032.0
18	2018	Delaware	Sussex County	US Senator	40675.0	50391.0
19	2018	Florida	Alachua County	US Senator	74493.0	40599.0
20	2018	Florida	Baker County	US Senator	1945.0	8579.0
21	2018	Florida	Bay County	US Senator	16723.0	46681.0
22	2018	Florida	Bradford County	US Senator	2879.0	7576.0
23	2018	Florida	Brevard County	US Senator	121112.0	160305.0
24	2018	Florida	Broward County	US Senator	472239.0	
25	2018	Florida	Charlotte County	US Senator	33525.0	
26	2018	Florida	Citrus County	US Senator	22660.0	
27	2018	Florida	Collier County	US Senator	54390.0	
28	2018	Florida	Desoto County	US Senator	3328.0	
29	2018	Florida	Dixie County	US Senator	1322.0	
...	...	...	...	...	...	...
1175	2018	West Virginia	Pocahontas County	US Senator	1269.0	
1176	2018	West Virginia	Preston County	US Senator	3686.0	
1177	2018	West Virginia	Raleigh County	US Senator	10581.0	
1178	2018	West Virginia	Randolph County	US Senator	4472.0	
1179	2018	West Virginia	Ritchie County	US Senator	1082.0	
1180	2018	West Virginia	Roane County	US Senator	2165.0	
1181	2018	West Virginia	Summers County	US Senator	2069.0	
1182	2018	West Virginia	Taylor County	US Senator	2376.0	
1183	2018	West Virginia	Tucker County	US Senator	1469.0	
1184	2018	West Virginia	Tyler County	US Senator	1065.0	
1185	2018	West Virginia	Upshur County	US Senator	3102.0	
1186	2018	West Virginia	Wayne County	US Senator	6395.0	
1187	2018	West Virginia	Wetzel County	US Senator	2518.0	
1188	2018	West Virginia	Wood County	US Senator	14189.0	
1189	2018	West Virginia	Wyoming County	US Senator	2607.0	
1190	2018	Wyoming	Albany County	US Senator	7576.0	
1191	2018	Wyoming	Campbell County	US Senator	1628.0	
1192	2018	Wyoming	Carbon County	US Senator	1359.0	
1193	2018	Wyoming	Converse County	US Senator	834.0	
1194	2018	Wyoming	Fremont County	US Senator	4734.0	
1195	2018	Wyoming	Goshen County	US Senator	1020.0	
1196	2018	Wyoming	Johnson County	US Senator	722.0	
1197	2018	Wyoming	Lincoln County	US Senator	1152.0	
1198	2018	Wyoming	Natrona County	US Senator	7285.0	
1199	2018	Wyoming	Niobrara County	US Senator	144.0	
1200	2018	Wyoming	Platte County	US Senator	801.0	
1201	2018	Wyoming	Sublette County	US Senator	668.0	
1202	2018	Wyoming	Sweetwater County	US Senator	3943.0	
1203	2018	Wyoming	Uinta County	US Senator	1371.0	
1204	2018	Wyoming	Washakie County	US Senator	588.0	

```
24       211397.0
25        52916.0
26        48008.0
27       101266.0
28         5503.0
29         4442.0
...           ...
1175       1411.0
1176       5943.0
1177      12620.0
1178       4017.0
1179       1961.0
1180       1899.0
1181       1868.0
1182       2642.0
1183       1502.0
1184       1603.0
1185       4010.0
1186       5954.0
1187       2135.0
1188      13696.0
1189       3096.0
1190       6366.0
1191      11020.0
1192       3673.0
1193       3959.0
1194       9262.0
1195       3658.0
1196       3085.0
1197       5846.0
1198      16359.0
1199        980.0
1200       2850.0
1201       2653.0
1202       8577.0
1203       4713.0
1204       2423.0

[1205 rows x 6 columns]
```

In [6]:
```
#2ii-Rename column names in election - County
data_election['County'] = data_election['County'].str.replace('County', '')
data_election['State'] = data_election['State'].str.strip().str.lower()
data_election['County'] = data_election['County'].str.strip().str.lower()
print(data_election)

data_demographics['State'] = data_demographics['State'].str.strip().str.lower()
data_demographics['County'] = data_demographics['County'].str.strip().str.lower()
print(data_demographics)
```

	Party	Year	State	County	Office	Democratic	Republican
0		2018	arizona	apache	US Senator	16298.0	7810.0
1		2018	arizona	cochise	US Senator	17383.0	26929.0
2		2018	arizona	coconino	US Senator	34240.0	19249.0
3		2018	arizona	gila	US Senator	7643.0	12180.0
4		2018	arizona	graham	US Senator	3368.0	6870.0
5		2018	arizona	la paz	US Senator	1609.0	3265.0
6		2018	arizona	maricopa	US Senator	732671.0	672505.0
7		2018	arizona	mohave	US Senator	19214.0	50209.0
8		2018	arizona	navajo	US Senator	16624.0	18767.0
9		2018	arizona	pima	US Senator	221242.0	160550.0
10		2018	arizona	santa cruz	US Senator	9241.0	3828.0
11		2018	arizona	yavapai	US Senator	40160.0	65308.0
12		2018	connecticut	fairfield	US Senator	210899.0	131321.0
13		2018	connecticut	hartford	US Senator	203591.0	123864.0
14		2018	connecticut	middlesex	US Senator	42383.0	32836.0
15		2018	connecticut	new haven	US Senator	179714.0	126004.0
16		2018	connecticut	tolland	US Senator	34732.0	28046.0
17		2018	connecticut	windham	US Senator	20490.0	19032.0
18		2018	delaware	sussex	US Senator	40675.0	50391.0
19		2018	florida	alachua	US Senator	74493.0	40599.0
20		2018	florida	baker	US Senator	1945.0	8579.0
21		2018	florida	bay	US Senator	16723.0	46681.0
22		2018	florida	bradford	US Senator	2879.0	7576.0
23		2018	florida	brevard	US Senator	121112.0	160305.0
24		2018	florida	broward	US Senator	472239.0	211397.0
25		2018	florida	charlotte	US Senator	33525.0	52916.0
26		2018	florida	citrus	US Senator	22660.0	48008.0
27		2018	florida	collier	US Senator	54390.0	101266.0
28		2018	florida	desoto	US Senator	3328.0	5503.0
29		2018	florida	dixie	US Senator	1322.0	4442.0
...	...	...	...	...	...	...	...
1175		2018	west virginia	pocahontas	US Senator	1269.0	1411.0
1176		2018	west virginia	preston	US Senator	3686.0	5943.0
1177		2018	west virginia	raleigh	US Senator	10581.0	12620.0
1178		2018	west virginia	randolph	US Senator	4472.0	4017.0
1179		2018	west virginia	ritchie	US Senator	1082.0	1961.0
1180		2018	west virginia	roane	US Senator	2165.0	1899.0
1181		2018	west virginia	summers	US Senator	2069.0	1868.0
1182		2018	west virginia	taylor	US Senator	2376.0	2642.0
1183		2018	west virginia	tucker	US Senator	1469.0	1502.0
1184		2018	west virginia	tyler	US Senator	1065.0	1603.0
1185		2018	west virginia	upshur	US Senator	3102.0	4010.0
1186		2018	west virginia	wayne	US Senator	6395.0	5954.0
1187		2018	west virginia	wetzel	US Senator	2518.0	2135.0
1188		2018	west virginia	wood	US Senator	14189.0	13696.0
1189		2018	west virginia	wyoming	US Senator	2607.0	3096.0
1190		2018	wyoming	albany	US Senator	7576.0	6366.0
1191		2018	wyoming	campbell	US Senator	1628.0	11020.0
1192		2018	wyoming	carbon	US Senator	1359.0	3673.0
1193		2018	wyoming	converse	US Senator	834.0	3959.0
1194		2018	wyoming	fremont	US Senator	4734.0	9262.0
1195		2018	wyoming	goshen	US Senator	1020.0	3658.0
1196		2018	wyoming	johnson	US Senator	722.0	3085.0
1197		2018	wyoming	lincoln	US Senator	1152.0	5846.0
1198		2018	wyoming	natrona	US Senator	7285.0	16359.0
1199		2018	wyoming	niobrara	US Senator	144.0	980.0
1200		2018	wyoming	platte	US Senator	801.0	2850.0
1201		2018	wyoming	sublette	US Senator	668.0	2653.0
1202		2018	wyoming	sweetwater	US Senator	3943.0	8577.0
1203		2018	wyoming	uinta	US Senator	1371.0	4713.0
1204		2018	wyoming	washakie	US Senator	588.0	2423.0

[1205 rows x 6 columns]

	State	County	FIPS	Total Population	\
0	wisconsin	la crosse	55063	117538	
1	virginia	alleghany	51005	15919	
2	indiana	fountain	18045	16741	
3	ohio	geauga	39055	94020	
4	wisconsin	jackson	55053	20566	
5	texas	baylor	48023	3639	
6	nebraska	madison	31119	35125	
7	hawaii	hawaii	15001	193680	
8	tennessee	henry	47079	32291	
9	michigan	oceana	26127	26152	
10	nebraska	pierce	31139	7179	
11	texas	jack	48237	8866	
12	florida	walton	12131	61528	
13	virginia	washington	51191	54562	
14	florida	escambia	12033	309574	
15	texas	wheeler	48483	5642	
16	arizona	yavapai	4025	218586	
17	nebraska	loup	31115	542	
18	michigan	antrim	26009	23215	
19	minnesota	wabasha	27157	21327	
20	nevada	elko	32007	52029	
21	wisconsin	dodge	55027	88404	
22	minnesota	lake of the woods	27077	3901	

23	west virginia	tucker	54093	6922
24	virginia	lexington city	51678	7036
25	florida	gilchrist	12041	17033
26	tennessee	claiborne	47025	31701
27	connecticut	middlesex	9007	164438
28	utah	sevier	49041	20913
29	massachusetts	hampshire	25015	161035
...	...	...	...	...
1186	west virginia	wood	54107	86262
1187	florida	orange	12095	1256055
1188	north dakota	walsh	38099	10995
1189	indiana	white	18181	24265
1190	montana	carter	30011	1295
1191	indiana	pulaski	18131	12910
1192	washington	lewis	53041	75724
1193	pennsylvania	greene	42059	37669
1194	texas	yoakum	48501	8316
1195	montana	pondera	30073	6166
1196	north dakota	dickey	38021	5160
1197	texas	refugio	48391	7315
1198	virginia	smyth	51173	31513
1199	virginia	charlotte	51037	12232
1200	michigan	grand traverse	26055	90715
1201	indiana	rush	18139	16873
1202	montana	deer lodge	30023	9176
1203	ohio	washington	39167	61154
1204	wisconsin	portage	55097	70551
1205	new york	monroe	36055	749236
1206	florida	broward	12011	1863780
1207	washington	wahkiakum	53069	4051
1208	new york	tioga	36107	49649
1209	montana	carbon	30009	10340
1210	texas	johnson	48251	157544
1211	montana	lincoln	30053	19268
1212	ohio	tuscarawas	39157	92579
1213	michigan	newaygo	26123	47957
1214	tennessee	lauderdale	47097	27261
1215	texas	sabine	48403	10367

	Citizen Voting-Age Population	Percent White, not Hispanic or Latino
0	0	90.537528
1	12705	91.940449
2	12750	95.705155
3	0	95.837056
4	15835	86.662453
5	0	86.644683
6	24885	81.249822
7	0	30.401694
8	25285	87.662197
9	18930	82.486999
10	5385	96.893718
11	6535	78.411911
12	47490	84.447731
13	0	95.484037
14	0	65.219624
15	3785	68.397731
16	0	81.159361
17	435	97.970480
18	0	95.179841
19	16385	94.926619
20	34740	67.354744
21	0	90.629383
22	0	93.975904
23	5690	97.688529
24	6335	80.542922
25	0	87.295250
26	25350	95.495410
27	0	84.738929
28	0	92.244059
29	0	84.610799
...	...	...
1186	67640	95.478890
1187	854605	42.656970
1188	8280	85.657117
1189	0	90.245209
1190	0	96.602317
1191	9870	94.771495
1192	56900	84.489726
1193	0	92.598689
1194	0	34.271284
1195	0	80.960104
1196	0	93.720930
1197	0	42.761449
1198	24960	94.376924
1199	9540	67.282538
1200	70710	92.673758
1201	0	96.325490
1202	0	90.747602

1203	0	95.364163
1204	0	91.815850
1205	0	71.329327
1206	0	39.245351
1207	0	89.681560
1208	38525	95.454088
1209	0	94.787234
1210	0	74.343041
1211	15640	93.351671
1212	70485	95.155489
1213	0	90.716684
1214	0	60.456330
1215	0	86.341275

	Percent Black, not Hispanic or Latino	Percent Hispanic or Latino	\
0	1.214075	1.724549	
1	5.207614	1.432251	
2	0.400215	2.359477	
3	1.256116	1.294405	
4	1.983857	3.082758	
5	1.841165	8.353943	
6	1.155872	14.217794	
7	0.547811	12.405514	
8	8.599919	2.201852	
9	1.131845	14.419547	
10	0.222872	1.587965	
11	4.376269	15.880893	
12	4.950592	5.888376	
13	1.268282	1.414904	
14	21.532816	5.389988	
15	2.658632	26.462247	
16	0.518331	14.054880	
17	0.000000	0.000000	
18	0.323067	1.955632	
19	0.150045	2.874291	
20	1.060947	24.067347	
21	2.085878	4.485091	
22	0.102538	1.230454	
23	0.173360	0.650101	
24	8.996589	4.178511	
25	5.988376	5.336699	
26	1.050440	1.113530	
27	4.664980	5.641032	
28	0.344283	4.882131	
29	2.683268	5.253516	
...	...	...	
1186	1.123322	1.032900	
1187	19.981211	29.338126	
1188	0.300136	10.759436	
1189	0.461570	7.554090	
1190	0.000000	1.389961	
1191	0.209140	2.788536	
1192	0.662934	9.629708	
1193	4.149300	1.422921	
1194	0.000000	63.552189	
1195	0.389231	1.881284	
1196	1.472868	3.333333	
1197	4.743677	49.323308	
1198	2.068987	1.948402	
1199	30.657292	0.948332	
1200	1.234636	2.621397	
1201	1.149766	1.428317	
1202	0.348736	3.073235	
1203	1.100500	1.013834	
1204	0.687446	2.951057	
1205	14.497969	8.120272	
1206	27.214639	27.564841	
1207	0.123426	4.640829	
1208	0.713005	1.695905	
1209	0.009671	2.379110	
1210	2.487559	19.762733	
1211	0.057089	2.678015	
1212	0.804718	2.349345	
1213	1.317847	5.728048	
1214	34.789626	2.380690	
1215	7.080158	3.839105	

	Percent Foreign Born	Percent Female	Percent Age 29 and Under	\
0	2.976059	51.171536	43.241335	
1	1.300333	51.077329	31.660280	
2	1.547100	49.770026	35.899887	
3	2.578175	50.678579	36.281642	
4	1.376058	46.649810	36.292911	
5	2.473207	51.662545	30.090684	
6	6.784342	50.448399	41.432028	
7	11.003717	50.143019	36.008881	
8	1.560806	51.441578	33.238364	
9	5.578923	49.395840	36.643469	

10	0.780053	49.658727	36.634629
11	5.549289	43.187458	38.732236
12	5.759979	49.349889	33.165388
13	1.611011	50.500348	32.005792
14	4.747492	50.299444	41.364585
15	9.517901	49.202410	39.188231
16	6.456955	51.092476	28.717301
17	0.000000	52.398524	30.996310
18	2.015938	50.273530	29.450786
19	1.355090	50.171145	34.594645
20	8.602895	47.948260	44.928790
21	1.884530	47.788562	34.642098
22	1.563702	47.962061	28.479877
23	0.592314	50.245594	29.731292
24	3.936896	43.447982	67.367823
25	2.906123	47.390360	38.307990
26	1.164001	51.200278	35.673954
27	7.652732	51.098894	33.241708
28	1.989193	48.893033	43.877014
29	7.846120	53.230664	44.243177
...	...	...	...
1186	1.001600	51.756277	34.726763
1187	20.157318	50.827551	43.005123
1188	3.919964	48.722146	34.588449
1189	4.314857	50.294663	36.130229
1190	0.077220	51.737452	30.810811
1191	0.805577	49.310612	36.018590
1192	4.512440	50.158470	35.731076
1193	0.809684	48.084632	35.267727
1194	25.637326	49.158249	46.284271
1195	2.189426	49.724295	37.528381
1196	1.763566	49.806202	36.531008
1197	2.734108	50.799727	36.896787
1198	1.256624	50.423635	33.316409
1199	0.662198	49.959124	35.080118
1200	2.421871	50.830623	35.075787
1201	1.060866	50.974930	36.745096
1202	1.874455	46.785092	29.980384
1203	1.430814	50.596854	34.483435
1204	3.119729	49.846211	42.641493
1205	8.460752	51.694126	39.551223
1206	32.715718	51.376504	36.701971
1207	3.505307	51.246606	26.783510
1208	2.229652	50.554895	34.465951
1209	1.276596	49.990329	28.143133
1210	6.138603	50.168842	41.296400
1211	2.345858	49.974050	27.979033
1212	1.650482	50.823621	36.475875
1213	1.978856	49.656984	36.620306
1214	1.757089	47.734859	39.528997
1215	0.752387	50.506415	31.243368

	Percent Age 65 and Older	Median Household Income	Percent Unemployed	\
0	14.702479	51477	4.796952	
1	23.902255	45538	4.560986	
2	18.941521	45924	7.978789	
3	18.028079	74165	4.036902	
4	17.587280	49608	5.569698	
5	24.402308	34382	1.377410	
6	15.404982	48673	3.094085	
7	17.580545	53936	7.437632	
8	21.476572	38378	7.912277	
9	19.088406	41952	8.275556	
10	18.540187	53125	3.981142	
11	15.677871	50390	4.680468	
12	18.783318	46910	7.295474	
13	20.301675	43835	6.425857	
14	15.774581	46117	8.297930	
15	17.795108	51082	5.424099	
16	28.272625	46638	8.525986	
17	24.538745	56750	0.335570	
18	25.410295	48825	7.815784	
19	18.807146	58865	3.382303	
20	9.229468	74672	4.896851	
21	16.304692	54111	4.849356	
22	21.814919	45732	3.888099	
23	22.594626	43529	8.190184	
24	14.198408	34464	3.926590	
25	19.567898	40881	9.732220	
26	18.390587	33428	8.446716	
27	17.703937	79837	5.257635	
28	15.583608	48872	5.340291	
29	14.755177	62608	7.146733	
...	...	...	...	
1186	18.559737	43944	7.625458	
1187	10.741090	49391	8.072223	
1188	20.627558	51181	3.314246	
1189	18.569957	51547	5.421587	

1190	24.942085	46985	2.656250
1191	18.086754	44884	7.046070
1192	19.609371	44526	9.684249
1193	17.104250	49116	6.993228
1194	11.423761	56655	9.136546
1195	19.656179	42193	8.497409
1196	21.705426	55882	2.549773
1197	21.052632	50145	7.469005
1198	19.994923	38906	7.343413
1199	20.372793	33837	6.997030
1200	17.003803	55597	5.848530
1201	17.412434	46380	5.658982
1202	20.891456	39212	3.727181
1203	19.151650	44763	4.689885
1204	14.644725	52411	6.178216
1205	15.491914	53568	7.306146
1206	15.371879	52954	8.710312
1207	31.325599	48116	6.971514
1208	17.925839	58115	6.720605
1209	22.746615	52869	3.743112
1210	13.209643	59095	6.337092
1211	24.828732	35461	11.570571
1212	17.836658	46992	5.734798
1213	17.728382	44049	8.885564
1214	13.891640	32353	11.427738
1215	28.995852	32500	8.259134

	Percent Less than High School Degree \
0	5.474767
1	15.537543
2	12.032155
3	8.928599
4	11.792912
5	13.066667
6	11.678672
7	8.362326
8	15.710841
9	14.335429
10	5.227552
11	21.130221
12	15.083974
13	15.913833
14	10.053723
15	19.678609
16	9.830672
17	2.736318
18	9.311741
19	6.811536
20	16.622189
21	10.956491
22	10.309278
23	12.702390
24	14.962495
25	17.294025
26	22.759203
27	6.153444
28	10.556637
29	5.741345
...	...
1186	10.370080
1187	12.319191
1188	13.114754
1189	10.394650
1190	7.991588
1191	12.931034
1192	12.929994
1193	13.277047
1194	28.571429
1195	10.801144
1196	12.319456
1197	20.225171
1198	17.449694
1199	20.028275
1200	5.543796
1201	15.398535
1202	10.536314
1203	9.795325
1204	6.784049
1205	9.616412
1206	11.688725
1207	6.868490
1208	9.753866
1209	6.142694
1210	16.183214
1211	11.891078
1212	14.068037
1213	14.194114
1214	24.485474

```
1215                            14.643237

      Percent Less than Bachelor's Degree  Percent Rural
0                               67.529757      16.827753
1                               83.711604      52.393846
2                               85.538940      65.951276
3                               62.730824      63.968990
4                               86.129256      72.238251
5                               73.409524     100.000000
6                               80.345221      27.904576
7                               72.408979      37.990804
8                               85.305261      66.848129
9                               82.947912      89.890854
10                              78.618286     100.000000
11                              89.221949      53.217603
12                              73.691263      63.125193
13                              75.223085      71.676143
14                              75.116986       8.302225
15                              85.879874     100.000000
16                              74.458362      33.197178
17                              78.109453     100.000000
18                              73.314054     100.000000
19                              78.761297      64.541428
20                              81.880924      37.850793
21                              83.745359      48.641828
22                              83.024055     100.000000
23                              85.794140     100.000000
24                              55.112515       0.000000
25                              87.654646      83.918767
26                              86.813089      71.554962
27                              59.071541      24.540066
28                              83.495524      64.714931
29                              55.277464      27.443699
...                                   ...           ...
1186                            79.738786      26.774461
1187                            68.077150       2.042138
1188                            83.281291      62.055940
1189                            84.315482      67.966563
1190                            82.860147     100.000000
1191                            89.330497      80.943143
1192                            84.490042      60.654695
1193                            82.669606      66.832963
1194                            82.183908      37.327072
1195                            77.324750      58.101739
1196                            72.047579     100.000000
1197                            89.203860      60.679940
1198                            85.209810      75.285643
1199                            86.958058     100.000000
1200                            68.328672      48.023820
1201                            85.971564      61.160304
1202                            84.100541      33.641643
1203                            81.590788      56.562207
1204                            69.694941      36.031649
1205                            63.335285       6.449975
1206                            68.984269       0.016933
1207                            84.928385     100.000000
1208                            74.963552      65.658680
1209                            70.838103     100.000000
1210                            82.255463      37.874170
1211                            81.041325      79.793773
1212                            85.020218      41.580437
1213                            85.660708      83.840281
1214                            91.425468      58.662592
1215                            87.060703     100.000000

[1216 rows x 17 columns]
```

In [7]:
```
#2iii-Merge elections and demographics
data_whole = pd.merge(data_election, data_demographics, how = 'inner',
                      on = ['State', 'County'], sort = True)
print(data_whole)
data_whole.to_excel("output.xlsx")
```

	Year	State	County	Office	Democratic	Republican
0	2018	arizona	apache	US Senator	16298.0	7810.0
1	2018	arizona	cochise	US Senator	17383.0	26929.0
2	2018	arizona	coconino	US Senator	34240.0	19249.0
3	2018	arizona	gila	US Senator	7643.0	12180.0
4	2018	arizona	graham	US Senator	3368.0	6870.0
5	2018	arizona	la paz	US Senator	1609.0	3265.0
6	2018	arizona	maricopa	US Senator	732671.0	672505.0
7	2018	arizona	mohave	US Senator	19214.0	50209.0
8	2018	arizona	navajo	US Senator	16624.0	18767.0
9	2018	arizona	pima	US Senator	221242.0	160550.0
10	2018	arizona	santa cruz	US Senator	9241.0	3828.0
11	2018	arizona	yavapai	US Senator	40160.0	65308.0
12	2018	connecticut	fairfield	US Senator	210899.0	131321.0
13	2018	connecticut	hartford	US Senator	203591.0	123864.0
14	2018	connecticut	middlesex	US Senator	42383.0	32836.0
15	2018	connecticut	new haven	US Senator	179714.0	126004.0
16	2018	connecticut	tolland	US Senator	34732.0	28046.0
17	2018	connecticut	windham	US Senator	20490.0	19032.0
18	2018	delaware	sussex	US Senator	40675.0	50391.0
19	2018	florida	alachua	US Senator	74493.0	40599.0
20	2018	florida	baker	US Senator	1945.0	8579.0
21	2018	florida	bay	US Senator	16723.0	46681.0
22	2018	florida	bradford	US Senator	2879.0	7576.0
23	2018	florida	brevard	US Senator	121112.0	160305.0
24	2018	florida	broward	US Senator	472239.0	211397.0
25	2018	florida	charlotte	US Senator	33525.0	52916.0
26	2018	florida	citrus	US Senator	22660.0	48008.0
27	2018	florida	collier	US Senator	54390.0	101266.0
28	2018	florida	desoto	US Senator	3328.0	5503.0
29	2018	florida	dixie	US Senator	1322.0	4442.0
...	...	...	...	...	...	...
1170	2018	wisconsin	portage	US Senator	20170.0	14510.0
1171	2018	wisconsin	racine	US Senator	45397.0	41213.0
1172	2018	wisconsin	rock	US Senator	42616.0	25322.0
1173	2018	wisconsin	rusk	US Senator	2654.0	3381.0
1174	2018	wisconsin	sawyer	US Senator	3938.0	4159.0
1175	2018	wisconsin	sheboygan	US Senator	24183.0	28667.0
1176	2018	wisconsin	st. croix	US Senator	18765.0	21069.0
1177	2018	wisconsin	taylor	US Senator	3028.0	5005.0
1178	2018	wisconsin	walworth	US Senator	20299.0	24844.0
1179	2018	wisconsin	washington	US Senator	23072.0	47102.0
1180	2018	wisconsin	waukesha	US Senator	84147.0	136190.0
1181	2018	wisconsin	waupaca	US Senator	9509.0	12792.0
1182	2018	wisconsin	waushara	US Senator	4434.0	6106.0
1183	2018	wisconsin	winnebago	US Senator	40185.0	35282.0
1184	2018	wisconsin	wood	US Senator	15992.0	16899.0
1185	2018	wyoming	albany	US Senator	7576.0	6366.0
1186	2018	wyoming	campbell	US Senator	1628.0	11020.0
1187	2018	wyoming	carbon	US Senator	1359.0	3673.0
1188	2018	wyoming	converse	US Senator	834.0	3959.0
1189	2018	wyoming	fremont	US Senator	4734.0	9262.0
1190	2018	wyoming	goshen	US Senator	1020.0	3658.0
1191	2018	wyoming	johnson	US Senator	722.0	3085.0
1192	2018	wyoming	lincoln	US Senator	1152.0	5846.0
1193	2018	wyoming	natrona	US Senator	7285.0	16359.0
1194	2018	wyoming	niobrara	US Senator	144.0	980.0
1195	2018	wyoming	platte	US Senator	801.0	2850.0
1196	2018	wyoming	sublette	US Senator	668.0	2653.0
1197	2018	wyoming	sweetwater	US Senator	3943.0	8577.0
1198	2018	wyoming	uinta	US Senator	1371.0	4713.0
1199	2018	wyoming	washakie	US Senator	588.0	2423.0

	FIPS	Total Population	Citizen Voting-Age Population
0	4001	72346	0
1	4003	128177	92915
2	4005	138064	104265
3	4007	53179	0
4	4009	37529	0
5	4012	20304	15245
6	4013	4088549	2723565
7	4015	203629	0
8	4017	108209	76280
9	4019	1003338	0
10	4023	46547	27155
11	4025	218586	0
12	9001	941618	0
13	9003	895699	644940
14	9007	164438	0
15	9009	860874	631715
16	9013	151689	0
17	9015	117078	0
18	10005	211224	0
19	12001	256581	197720
20	12003	27312	20415
21	12005	178361	135795
22	12007	26919	0
23	12009	560683	438510

```
24    12011           1863780                          0
25    12015            169642                     141230
26    12017            140453                          0
27    12021            348236                          0
28    12027             35134                          0
29    12029             16084                      12890
...     ...               ...                        ...
1170  55097             70551                          0
1171  55101            194873                          0
1172  55105            160986                          0
1173  55107             14272                          0
1174  55113             16438                      13055
1175  55117            115269                      86115
1176  55109             86726                          0
1177  55119             20521                      15445
1178  55127            103021                          0
1179  55131            133422                     101285
1180  55133            395377                     298750
1181  55135             51974                          0
1182  55137             24238                          0
1183  55139            169487                          0
1184  55141             73621                          0
1185  56001             37836                      30070
1186  56005             48473                          0
1187  56007             15696                      11335
1188  56009             14223                          0
1189  56013             40683                      30170
1190  56015             13546                          0
1191  56019              8572                       6590
1192  56023             18543                          0
1193  56025             80871                      60415
1194  56027              2498                       1995
1195  56031              8740                       6830
1196  56035             10032                          0
1197  56037             44812                      30565
1198  56041             20893                      14355
1199  56043              8351                          0

      Percent White, not Hispanic or Latino  ...  Percent Hispanic or Latino  \
0                                 18.571863  ...                    5.947806
1                                 56.299492  ...                   34.403208
2                                 54.619597  ...                   13.711033
3                                 63.222325  ...                   18.548675
4                                 51.461536  ...                   32.097844
5                                 58.884949  ...                   26.182033
6                                 56.918114  ...                   30.286833
7                                 78.252606  ...                   15.708470
8                                 41.927196  ...                   11.049913
9                                 53.271579  ...                   36.105978
10                                15.274883  ...                   83.219112
11                                81.159361  ...                   14.054880
12                                63.509512  ...                   18.636007
13                                63.252052  ...                   16.897976
14                                84.738929  ...                    5.641032
15                                64.782767  ...                   16.790959
16                                85.892187  ...                    5.029369
17                                83.782606  ...                   10.877364
18                                74.829091  ...                    9.195925
19                                62.460198  ...                    8.994820
20                                82.088459  ...                    2.317663
21                                77.687387  ...                    5.728270
22                                74.787325  ...                    3.796575
23                                75.798803  ...                    9.301691
24                                39.245351  ...                   27.564841
25                                84.643543  ...                    6.632202
26                                88.871010  ...                    5.121998
27                                64.222539  ...                   26.634524
28                                55.040701  ...                   30.688222
29                                85.041035  ...                    3.780154
...                                     ...  ...                         ...
1170                              91.815850  ...                    2.951057
1171                              73.058864  ...                   12.452212
1172                              83.642056  ...                    8.207546
1173                              95.557735  ...                    1.604540
1174                              77.302592  ...                    2.269132
1175                              85.579818  ...                    5.975588
1176                              94.420358  ...                    2.180430
1177                              96.549876  ...                    1.744554
1178                              85.711651  ...                   10.894866
1179                              93.426871  ...                    2.972523
1180                              89.471568  ...                    4.495709
1181                              94.860892  ...                    2.880286
1182                              89.817642  ...                    6.213384
1183                              89.548461  ...                    3.841003
1184                              93.080779  ...                    2.665000
1185                              83.269902  ...                    9.229305
1186                              87.774637  ...                    8.357230
1187                              77.924312  ...                   17.647808
```

```
1188                          88.849047  ...                   7.691767
1189                          70.198855  ...                   6.779244
1190                          86.409272  ...                  10.519711
1191                          91.565562  ...                   2.134858
1192                          92.600982  ...                   4.416761
1193                          87.026252  ...                   8.198242
1194                          89.511609  ...                   4.243395
1195                          89.359268  ...                   7.814645
1196                          91.646730  ...                   7.814992
1197                          79.815674  ...                  15.859591
1198                          87.718375  ...                   8.959939
1199                          82.397318  ...                  13.962400

      Percent Foreign Born  Percent Female  Percent Age 29 and Under  \
0                 1.719515       50.598513                 45.854643
1                11.458374       49.069646                 37.902276
2                 4.825298       50.581614                 48.946141
3                 4.249798       50.296170                 32.238290
4                 4.385942       46.313518                 46.393456
5                11.372143       48.946020                 28.073286
6                14.729333       50.549278                 41.886620
7                 6.969047       49.676618                 30.485835
8                 2.914730       49.846131                 43.243168
9                12.903428       50.807405                 40.087388
10               32.644424       52.125808                 43.300320
11                6.456955       51.092476                 28.717301
12               21.154120       51.308280                 38.019876
13               15.011070       51.492075                 37.389458
14                7.652732       51.098894                 33.241708
15               12.073428       51.759956                 37.982562
16                6.769113       49.967367                 42.301024
17                5.269991       50.422795                 37.389604
18                6.671117       51.495569                 31.746866
19               10.144555       51.670623                 48.551530
20                1.475542       47.715290                 41.344464
21                5.479898       50.280050                 37.739192
22                2.306921       44.741632                 35.343066
23                8.632685       51.140841                 31.972612
24               32.715718       51.376504                 36.701971
25               10.185567       51.260891                 22.388913
26                5.413911       51.667818                 24.581889
27               23.816894       50.881299                 29.998909
28               17.504412       43.345477                 37.513520
29                2.785377       45.417807                 31.236011
...                    ...             ...                       ...
1170              3.119729       49.846211                 42.641493
1171              5.064837       50.538043                 38.058120
1172              4.583007       50.796343                 38.521362
1173              1.891816       49.691704                 31.747478
1174              1.502616       49.306485                 30.800584
1175              5.754366       49.745378                 36.383590
1176              2.453705       50.064571                 38.543228
1177              1.461917       48.842649                 35.100629
1178              6.097786       50.039312                 39.967579
1179              2.717693       50.384494                 35.089416
1180              4.902410       50.893198                 34.964350
1181              1.816293       49.265017                 33.189672
1182              2.784883       47.520422                 30.006601
1183              3.318839       49.604394                 39.787122
1184              2.326782       50.712433                 34.645006
1185              6.081510       47.756105                 55.447193
1186              3.645328       48.008170                 45.326264
1187              5.198777       45.839704                 39.181957
1188              2.706883       49.933207                 38.515081
1189              1.339626       49.907824                 39.751247
1190              2.724051       47.091392                 35.914661
1191              1.656556       46.966869                 32.571162
1192              2.151755       48.773122                 38.715418
1193              2.729038       49.421919                 40.688257
1194              0.280224       54.443555                 36.269015
1195              2.780320       47.711670                 32.700229
1196              2.053429       46.949761                 36.393541
1197              5.509685       47.824244                 44.153352
1198              3.986981       49.327526                 43.205858
1199              3.783978       51.359119                 34.774279

      Percent Age 65 and Older  Median Household Income  Percent Unemployed  \
0                    13.322091                    32460           15.807433
1                    19.756275                    45383            8.567108
2                    10.873943                    51106            8.238305
3                    26.397638                    40593           12.129932
4                    12.315809                    47422           14.424104
5                    36.056935                    36321           10.599013
6                    13.837843                    55676            6.808454
7                    26.858650                    39856           11.680953
8                    15.745456                    36868           18.525791
9                    17.801778                    46764            9.214114
10                   15.895761                    38941            9.749896
```

11	28.272625	46638	8.525986
12	14.383115	86670	8.211898
13	15.644876	68027	8.237243
14	17.703937	79837	5.257635
15	15.604490	62715	8.545483
16	13.868507	80129	6.267325
17	14.571482	60689	8.375524
18	24.214578	54218	7.108621
19	12.430772	44702	7.020154
20	12.910076	53327	6.832522
21	16.103857	48577	7.475668
22	17.452357	43373	10.563312
23	22.491140	49914	9.278023
24	15.371879	52954	8.710312
25	37.622759	44865	10.088979
26	35.105694	39054	11.648012
27	29.589704	59783	6.408868
28	19.408550	35513	8.291825
29	21.729669	34634	5.306200
...	...	...	...
1170	14.644725	52411	6.178216
1171	14.722409	56359	7.478894
1172	15.032984	50968	7.390601
1173	22.323430	39904	4.930617
1174	23.317922	41869	7.665187
1175	16.257623	54392	5.054640
1176	12.209718	73743	4.081465
1177	18.088787	47306	4.478597
1178	15.303676	55575	6.442507
1179	15.645096	70325	4.019701
1180	16.492866	78268	3.665774
1181	19.523223	52441	4.546458
1182	22.353329	46581	6.336312
1183	14.738003	53501	4.420678
1184	18.608821	49926	5.669639
1185	9.549107	43043	4.579174
1186	6.954387	80822	4.669432
1187	13.774210	56972	4.141937
1188	13.668003	66737	5.282284
1189	16.409803	53559	7.344324
1190	20.389783	44883	6.918819
1191	20.496967	54594	4.512276
1192	14.382786	64579	5.618095
1193	12.825364	56983	4.861868
1194	18.254604	40640	0.457875
1195	22.013730	41051	3.901047
1196	13.337321	76004	2.786971
1197	9.417120	68233	5.072255
1198	10.678218	53323	6.390755
1199	19.650341	46212	7.441860

	Percent Less than High School Degree \
0	21.758252
1	13.409171
2	11.085381
3	15.729958
4	14.580797
5	24.842215
6	13.051927
7	16.145850
8	18.494087
9	12.252238
10	25.206726
11	9.830672
12	10.103521
13	10.736314
14	6.153444
15	10.503197
16	6.214203
17	11.925239
18	13.792481
19	7.468059
20	17.871127
21	11.320101
22	23.212996
23	8.823708
24	11.688725
25	10.578890
26	13.288908
27	14.321540
28	29.475750
29	22.236087
...	...
1170	6.784049
1171	11.170104
1172	10.591406
1173	12.513379
1174	8.901562

1175	8.367242
1176	4.242959
1177	12.358524
1178	9.074810
1179	6.072757
1180	4.116150
1181	9.145423
1182	13.760962
1183	7.893714
1184	7.658332
1185	4.210167
1186	8.330054
1187	9.579879
1188	9.758393
1189	8.537172
1190	8.390574
1191	5.105750
1192	6.949996
1193	8.332255
1194	11.054422
1195	9.675889
1196	4.658830
1197	9.314606
1198	10.361224
1199	12.577108

	Percent Less than Bachelor's Degree	Percent Rural
0	88.941063	74.061076
1	76.837055	36.301067
2	65.791439	31.466066
3	82.262624	41.062000
4	86.675944	46.437399
5	89.563407	56.327786
6	69.031137	2.363800
7	88.121178	22.963644
8	85.507970	54.138242
9	69.199391	7.523491
10	77.506775	26.883172
11	74.458362	33.197178
12	53.637538	4.584061
13	63.133913	5.408640
14	59.071541	24.540066
15	65.711800	3.639981
16	61.140250	38.207884
17	76.143316	49.761036
18	76.115066	41.303609
19	58.485040	21.193437
20	87.235709	59.491057
21	77.466127	11.999858
22	88.220732	75.529453
23	72.271467	5.074387
24	68.984269	0.016933
25	78.396083	8.884347
26	82.154355	34.518820
27	66.118970	8.485942
28	89.960743	46.242327
29	93.602862	77.024723
...	...	...
1170	69.694941	36.031649
1171	76.027600	12.277901
1172	79.234347	20.416513
1173	85.209692	76.624873
1174	77.865939	84.200036
1175	76.378160	28.284866
1176	66.600965	53.168534
1177	86.537786	80.443714
1178	72.181794	34.197089
1179	70.577928	30.784687
1180	58.376564	9.860448
1181	82.335700	64.939897
1182	84.915338	89.496244
1183	73.113141	13.431021
1184	79.617467	36.682765
1185	52.136794	11.939723
1186	80.935983	29.072464
1187	79.681613	41.630469
1188	84.468152	55.360370
1189	77.648670	51.424370
1190	77.036880	45.995924
1191	74.537343	49.025557
1192	79.631291	82.729482
1193	77.786814	14.449304
1194	81.009070	100.000000
1195	80.300395	58.647744
1196	75.645069	100.000000
1197	78.628507	10.916313
1198	81.793082	43.095937
1199	78.923920	35.954529

```
[1200 rows x 21 columns]
```

In [8]:
```
#3 -  no.of variables, type of variables, irrelevant & redundant attributes
#3i - no.of variables
data_whole.shape
#Answer: 1200 rows & 21 columns
```

Out[8]:
```
(1200, 21)
```

In [9]:
```
#3ii - type of variables
data_whole.info()
#Answer:float64, int64, object
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1200 entries, 0 to 1199
Data columns (total 21 columns):
Year                                1200 non-null int64
State                               1200 non-null object
County                              1200 non-null object
Office                              1200 non-null object
Democratic                          1200 non-null float64
Republican                          1200 non-null float64
FIPS                                1200 non-null int64
Total Population                    1200 non-null int64
Citizen Voting-Age Population       1200 non-null int64
Percent White, not Hispanic or Latino    1200 non-null float64
Percent Black, not Hispanic or Latino    1200 non-null float64
Percent Hispanic or Latino          1200 non-null float64
Percent Foreign Born                1200 non-null float64
Percent Female                      1200 non-null float64
Percent Age 29 and Under            1200 non-null float64
Percent Age 65 and Older            1200 non-null float64
Median Household Income             1200 non-null int64
Percent Unemployed                  1200 non-null float64
Percent Less than High School Degree     1200 non-null float64
Percent Less than Bachelor's Degree      1200 non-null float64
Percent Rural                       1200 non-null float64
dtypes: float64(13), int64(5), object(3)
memory usage: 206.2+ KB
```

In [10]:
```
#3iii - Irrelevant & Redundant attributes
data_whole = data_whole.drop(['Year', 'Office'], axis=1)
print(data_whole)
'''
Answer:
Year and office attributes have same values. So,it doesn't give much information about the data.
So we drop columns year & office (as same row values)
'''
```

	State	County	Democratic	Republican	FIPS	\
0	arizona	apache	16298.0	7810.0	4001	
1	arizona	cochise	17383.0	26929.0	4003	
2	arizona	coconino	34240.0	19249.0	4005	
3	arizona	gila	7643.0	12180.0	4007	
4	arizona	graham	3368.0	6870.0	4009	
5	arizona	la paz	1609.0	3265.0	4012	
6	arizona	maricopa	732671.0	672505.0	4013	
7	arizona	mohave	19214.0	50209.0	4015	
8	arizona	navajo	16624.0	18767.0	4017	
9	arizona	pima	221242.0	160550.0	4019	
10	arizona	santa cruz	9241.0	3828.0	4023	
11	arizona	yavapai	40160.0	65308.0	4025	
12	connecticut	fairfield	210899.0	131321.0	9001	
13	connecticut	hartford	203591.0	123864.0	9003	
14	connecticut	middlesex	42383.0	32836.0	9007	
15	connecticut	new haven	179714.0	126004.0	9009	
16	connecticut	tolland	34732.0	28046.0	9013	
17	connecticut	windham	20490.0	19032.0	9015	
18	delaware	sussex	40675.0	50391.0	10005	
19	florida	alachua	74493.0	40599.0	12001	
20	florida	baker	1945.0	8579.0	12003	
21	florida	bay	16723.0	46681.0	12005	
22	florida	bradford	2879.0	7576.0	12007	
23	florida	brevard	121112.0	160305.0	12009	
24	florida	broward	472239.0	211397.0	12011	
25	florida	charlotte	33525.0	52916.0	12015	
26	florida	citrus	22660.0	48008.0	12017	
27	florida	collier	54390.0	101266.0	12021	
28	florida	desoto	3328.0	5503.0	12027	
29	florida	dixie	1322.0	4442.0	12029	
...	...	...	...	...	...	
1170	wisconsin	portage	20170.0	14510.0	55097	
1171	wisconsin	racine	45397.0	41213.0	55101	
1172	wisconsin	rock	42616.0	25322.0	55105	
1173	wisconsin	rusk	2654.0	3381.0	55107	
1174	wisconsin	sawyer	3938.0	4159.0	55113	
1175	wisconsin	sheboygan	24183.0	28667.0	55117	
1176	wisconsin	st. croix	18765.0	21069.0	55109	
1177	wisconsin	taylor	3028.0	5005.0	55119	
1178	wisconsin	walworth	20299.0	24844.0	55127	
1179	wisconsin	washington	23072.0	47102.0	55131	
1180	wisconsin	waukesha	84147.0	136190.0	55133	
1181	wisconsin	waupaca	9509.0	12792.0	55135	
1182	wisconsin	waushara	4434.0	6106.0	55137	
1183	wisconsin	winnebago	40185.0	35282.0	55139	
1184	wisconsin	wood	15992.0	16899.0	55141	
1185	wyoming	albany	7576.0	6366.0	56001	
1186	wyoming	campbell	1628.0	11020.0	56005	
1187	wyoming	carbon	1359.0	3673.0	56007	
1188	wyoming	converse	834.0	3959.0	56009	
1189	wyoming	fremont	4734.0	9262.0	56013	
1190	wyoming	goshen	1020.0	3658.0	56015	
1191	wyoming	johnson	722.0	3085.0	56019	
1192	wyoming	lincoln	1152.0	5846.0	56023	
1193	wyoming	natrona	7285.0	16359.0	56025	
1194	wyoming	niobrara	144.0	980.0	56027	
1195	wyoming	platte	801.0	2850.0	56031	
1196	wyoming	sublette	668.0	2653.0	56035	
1197	wyoming	sweetwater	3943.0	8577.0	56037	
1198	wyoming	uinta	1371.0	4713.0	56041	
1199	wyoming	washakie	588.0	2423.0	56043	

	Total Population	Citizen Voting-Age Population	\
0	72346	0	
1	128177	92915	
2	138064	104265	
3	53179	0	
4	37529	0	
5	20304	15245	
6	4088549	2723565	
7	203629	0	
8	108209	76280	
9	1003338	0	
10	46547	27155	
11	218586	0	
12	941618	0	
13	895699	644940	
14	164438	0	
15	860874	631715	
16	151689	0	
17	117078	0	
18	211224	0	
19	256581	197720	
20	27312	20415	
21	178361	135795	
22	26919	0	
23	560683	438510	

24	1863780	0
25	169642	141230
26	140453	0
27	348236	0
28	35134	0
29	16084	12890
...	...	...
1170	70551	0
1171	194873	0
1172	160986	0
1173	14272	0
1174	16438	13055
1175	115269	86115
1176	86726	0
1177	20521	15445
1178	103021	0
1179	133422	101285
1180	395377	298750
1181	51974	0
1182	24238	0
1183	169487	0
1184	73621	0
1185	37836	30070
1186	48473	0
1187	15696	11335
1188	14223	0
1189	40683	30170
1190	13546	0
1191	8572	6590
1192	18543	0
1193	80871	60415
1194	2498	1995
1195	8740	6830
1196	10032	0
1197	44812	30565
1198	20893	14355
1199	8351	0

	Percent White, not Hispanic or Latino	\
0	18.571863	
1	56.299492	
2	54.619597	
3	63.222325	
4	51.461536	
5	58.884949	
6	56.918114	
7	78.252606	
8	41.927196	
9	53.271579	
10	15.274883	
11	81.159361	
12	63.509512	
13	63.252052	
14	84.738929	
15	64.782767	
16	85.892187	
17	83.782606	
18	74.829091	
19	62.460198	
20	82.088459	
21	77.687387	
22	74.787325	
23	75.798803	
24	39.245351	
25	84.643543	
26	88.871010	
27	64.222539	
28	55.040701	
29	85.041035	
...	...	
1170	91.815850	
1171	73.058864	
1172	83.642056	
1173	95.557735	
1174	77.302592	
1175	85.579818	
1176	94.420358	
1177	96.549876	
1178	85.711651	
1179	93.426871	
1180	89.471568	
1181	94.860892	
1182	89.817642	
1183	89.548461	
1184	93.080779	
1185	83.269902	
1186	87.774637	
1187	77.924312	

```
1188                            88.849047
1189                            70.198855
1190                            86.409272
1191                            91.565562
1192                            92.600982
1193                            87.026252
1194                            89.511609
1195                            89.359268
1196                            91.646730
1197                            79.815674
1198                            87.718375
1199                            82.397318

      Percent Black, not Hispanic or Latino  Percent Hispanic or Latino  \
0                                  0.486551                   5.947806
1                                  3.714395                  34.403208
2                                  1.342855                  13.711033
3                                  0.552850                  18.548675
4                                  1.811932                  32.097844
5                                  0.379236                  26.182033
6                                  5.013612                  30.286833
7                                  0.951731                  15.708470
8                                  0.672772                  11.049913
9                                  3.199719                  36.105978
10                                 0.199798                  83.219112
11                                 0.518331                  14.054880
12                                10.456151                  18.636007
13                                12.657935                  16.897976
14                                 4.664980                   5.641032
15                                12.177276                  16.790959
16                                 2.811674                   5.029369
17                                 1.866277                  10.877364
18                                12.284589                   9.195925
19                                19.504562                   8.994820
20                                13.876684                   2.317663
21                                10.527526                   5.728270
22                                19.250344                   3.796575
23                                 9.838001                   9.301691
24                                27.214639                  27.564841
25                                 5.535775                   6.632202
26                                 2.817313                   5.121998
27                                 6.775003                  26.634524
28                                12.904878                  30.688222
29                                 7.504352                   3.780154
...                                     ...                        ...
1170                               0.687446                   2.951057
1171                              10.764960                  12.452212
1172                               3.404644                   8.207546
1173                               0.987948                   1.604540
1174                               0.663098                   2.269132
1175                               1.474811                   5.975588
1176                               0.785232                   2.180430
1177                               0.375225                   1.744554
1178                               0.990089                  10.894866
1179                               1.077034                   2.972523
1180                               1.402712                   4.495709
1181                               0.352099                   2.880286
1182                               1.547157                   6.213384
1183                               1.996613                   3.841003
1184                               0.709037                   2.665000
1185                               1.345280                   9.229305
1186                               0.742681                   8.357230
1187                               1.242355                  17.647808
1188                               0.007031                   7.691767
1189                               0.516186                   6.779244
1190                               0.147645                  10.519711
1191                               0.443304                   2.134858
1192                               0.210322                   4.416761
1193                               1.158635                   8.198242
1194                               0.120096                   4.243395
1195                               0.057208                   7.814645
1196                               0.000000                   7.814992
1197                               0.865840                  15.859591
1198                               0.186665                   8.959939
1199                               0.790325                  13.962400

      Percent Foreign Born  Percent Female  Percent Age 29 and Under  \
0                 1.719515       50.598513                 45.854643
1                11.458374       49.069646                 37.902276
2                 4.825298       50.581614                 48.946141
3                 4.249798       50.296170                 32.238290
4                 4.385942       46.313518                 46.393456
5                11.372143       48.946020                 28.073286
6                14.729333       50.549278                 41.886620
7                 6.969047       49.676618                 30.485835
8                 2.914730       49.846131                 43.243168
9                12.903428       50.807405                 40.087388
10               32.644424       52.125808                 43.300320
```

11	6.456955	51.092476	28.717301
12	21.154120	51.308280	38.019876
13	15.011070	51.492075	37.389458
14	7.652732	51.098894	33.241708
15	12.073428	51.759956	37.982562
16	6.769113	49.967367	42.301024
17	5.269991	50.422795	37.389604
18	6.671117	51.495569	31.746866
19	10.144555	51.670623	48.551530
20	1.475542	47.715290	41.344464
21	5.479898	50.280050	37.739192
22	2.306921	44.741632	35.343066
23	8.632685	51.140841	31.972612
24	32.715718	51.376504	36.701971
25	10.185567	51.260891	22.388913
26	5.413911	51.667818	24.581889
27	23.816894	50.881299	29.998909
28	17.504412	43.345477	37.513520
29	2.785377	45.417807	31.236011
...	...	...	...
1170	3.119729	49.846211	42.641493
1171	5.064837	50.538043	38.058120
1172	4.583007	50.796343	38.521362
1173	1.891816	49.691704	31.747478
1174	1.502616	49.306485	30.800584
1175	5.754366	49.745378	36.383590
1176	2.453705	50.064571	38.543228
1177	1.461917	48.842649	35.100629
1178	6.097786	50.039312	39.967579
1179	2.717693	50.384494	35.089416
1180	4.902410	50.893198	34.964350
1181	1.816293	49.265017	33.189672
1182	2.784883	47.520422	30.006601
1183	3.318839	49.604394	39.787122
1184	2.326782	50.712433	34.645006
1185	6.081510	47.756105	55.447193
1186	3.645328	48.008170	45.326264
1187	5.198777	45.839704	39.181957
1188	2.706883	49.933207	38.515081
1189	1.339626	49.907824	39.751247
1190	2.724051	47.091392	35.914661
1191	1.656556	46.966869	32.571162
1192	2.151755	48.773122	38.715418
1193	2.729038	49.421919	40.688257
1194	0.280224	54.443555	36.269015
1195	2.780320	47.711670	32.700229
1196	2.053429	46.949761	36.393541
1197	5.509685	47.824244	44.153352
1198	3.986981	49.327526	43.205858
1199	3.783978	51.359119	34.774279

	Percent Age 65 and Older	Median Household Income	Percent Unemployed	\
0	13.322091	32460	15.807433	
1	19.756275	45383	8.567108	
2	10.873943	51106	8.238305	
3	26.397638	40593	12.129932	
4	12.315809	47422	14.424104	
5	36.056935	36321	10.599013	
6	13.837843	55676	6.808454	
7	26.858650	39856	11.680953	
8	15.745456	36868	18.525791	
9	17.801778	46764	9.214114	
10	15.895761	38941	9.749896	
11	28.272625	46638	8.525986	
12	14.383115	86670	8.211898	
13	15.644876	68027	8.237243	
14	17.703937	79837	5.257635	
15	15.604490	62715	8.545483	
16	13.868507	80129	6.267325	
17	14.571482	60689	8.375524	
18	24.214578	54218	7.108621	
19	12.430772	44702	7.020154	
20	12.910076	53327	6.832522	
21	16.103857	48577	7.475668	
22	17.452357	43373	10.563312	
23	22.491140	49914	9.278023	
24	15.371879	52954	8.710312	
25	37.622759	44865	10.088979	
26	35.105694	39054	11.648012	
27	29.589704	59783	6.408868	
28	19.408550	35513	8.291825	
29	21.729669	34634	5.306200	
...	...	...	...	
1170	14.644725	52411	6.178216	
1171	14.722409	56359	7.478894	
1172	15.032984	50968	7.390601	
1173	22.323430	39904	4.930617	
1174	23.317922	41869	7.665187	

1175	16.257623	54392	5.054640
1176	12.209718	73743	4.081465
1177	18.088787	47306	4.478597
1178	15.303676	55575	6.442507
1179	15.645096	70325	4.019701
1180	16.492866	78268	3.665774
1181	19.523223	52441	4.546458
1182	22.353329	46581	6.336312
1183	14.738003	53501	4.420678
1184	18.608821	49926	5.669639
1185	9.549107	43043	4.579174
1186	6.954387	80822	4.669432
1187	13.774210	56972	4.141937
1188	13.668003	66737	5.282284
1189	16.409803	53559	7.344324
1190	20.389783	44883	6.918819
1191	20.496967	54594	4.512276
1192	14.382786	64579	5.618095
1193	12.825364	56983	4.861868
1194	18.254604	40640	0.457875
1195	22.013730	41051	3.901047
1196	13.337321	76004	2.786971
1197	9.417120	68233	5.072255
1198	10.678218	53323	6.390755
1199	19.650341	46212	7.441860

	Percent Less than High School Degree
0	21.758252
1	13.409171
2	11.085381
3	15.729958
4	14.580797
5	24.842215
6	13.051927
7	16.145850
8	18.494087
9	12.252238
10	25.206726
11	9.830672
12	10.103521
13	10.736314
14	6.153444
15	10.503197
16	6.214203
17	11.925239
18	13.792481
19	7.468059
20	17.871127
21	11.320101
22	23.212996
23	8.823708
24	11.688725
25	10.578890
26	13.288908
27	14.321540
28	29.475750
29	22.236087
...	...
1170	6.784049
1171	11.170104
1172	10.591406
1173	12.513379
1174	8.901562
1175	8.367242
1176	4.242959
1177	12.358524
1178	9.074810
1179	6.072757
1180	4.116150
1181	9.145423
1182	13.760962
1183	7.893714
1184	7.658332
1185	4.210167
1186	8.330054
1187	9.579879
1188	9.758393
1189	8.537172
1190	8.390574
1191	5.105750
1192	6.949996
1193	8.332255
1194	11.054422
1195	9.675889
1196	4.658830
1197	9.314606
1198	10.361224
1199	12.577108

	Percent Less than Bachelor's Degree	Percent Rural
0	88.941063	74.061076
1	76.837055	36.301067
2	65.791439	31.466066
3	82.262624	41.062000
4	86.675944	46.437399
5	89.563407	56.327786
6	69.031137	2.363800
7	88.121178	22.963644
8	85.507970	54.138242
9	69.199391	7.523491
10	77.506775	26.883172
11	74.458362	33.197178
12	53.637538	4.584061
13	63.133913	5.408640
14	59.071541	24.540066
15	65.711800	3.639981
16	61.140250	38.207884
17	76.143316	49.761036
18	76.115066	41.303609
19	58.485040	21.193437
20	87.235709	59.491057
21	77.466127	11.999858
22	88.220732	75.529453
23	72.271467	5.074387
24	68.984269	0.016933
25	78.396083	8.884347
26	82.154355	34.518820
27	66.118970	8.485942
28	89.960743	46.242327
29	93.602862	77.024723
...	...	...
1170	69.694941	36.031649
1171	76.027600	12.277901
1172	79.234347	20.416513
1173	85.209692	76.624873
1174	77.865939	84.200036
1175	76.378160	28.284866
1176	66.600965	53.168534
1177	86.537786	80.443714
1178	72.181794	34.197089
1179	70.577928	30.784687
1180	58.376564	9.860448
1181	82.335700	64.939897
1182	84.915338	89.496244
1183	73.113141	13.431021
1184	79.617467	36.682765
1185	52.136794	11.939723
1186	80.935983	29.072464
1187	79.681613	41.630469
1188	84.468152	55.360370
1189	77.648670	51.424370
1190	77.036880	45.995924
1191	74.537343	49.025557
1192	79.631291	82.729482
1193	77.786814	14.449304
1194	81.009070	100.000000
1195	80.300395	58.647744
1196	75.645069	100.000000
1197	78.628507	10.916313
1198	81.793082	43.095937
1199	78.923920	35.954529

[1200 rows x 19 columns]

In [11]:
```
#4 - Missing values
# Count missing values
num_rows = data_whole.shape[0]
print(num_rows - data_whole.count())
'''
Answer: No missing values but includes 0's
'''
```

```
State                                   0
County                                  0
Democratic                              0
Republican                              0
FIPS                                    0
Total Population                        0
Citizen Voting-Age Population           0
Percent White, not Hispanic or Latino   0
Percent Black, not Hispanic or Latino   0
Percent Hispanic or Latino              0
Percent Foreign Born                    0
Percent Female                          0
Percent Age 29 and Under                0
Percent Age 65 and Older                0
Median Household Income                 0
Percent Unemployed                      0
Percent Less than High School Degree    0
Percent Less than Bachelor's Degree     0
Percent Rural                           0
dtype: int64
```

In [12]:
```
#4i - Citizen Voting-Age Population
d = data_whole.loc[data_whole['Citizen Voting-Age Population'] < data_whole['Democratic'] + data_whole['Republican']]
#d = data_whole.loc[data_whole['Citizen Voting-Age Population'] == 0]
d = d[['Citizen Voting-Age Population','Democratic','Republican']]
print(d)
data_whole = data_whole.drop(['Citizen Voting-Age Population'], axis=1)
data_whole.to_excel("output.xlsx")
'''
Answer:
Citizen voting population has many zeros.
When we consdier citizen voting age population, when it is less than democratic
plus republican votes, it has 677 rows with zero values.
So, out of 1200 rows 677 is more than 50% of the data.
Hence, we drop the column 'Citizen Voting-Age Population'
'''
```

```
      Citizen Voting-Age Population  Democratic  Republican
0                                 0     16298.0      7810.0
3                                 0      7643.0     12180.0
4                                 0      3368.0      6870.0
7                                 0     19214.0     50209.0
9                                 0    221242.0    160550.0
11                                0     40160.0     65308.0
12                                0    210899.0    131321.0
14                                0     42383.0     32836.0
16                                0     34732.0     28046.0
17                                0     20490.0     19032.0
18                                0     40675.0     50391.0
22                                0      2879.0      7576.0
24                                0    472239.0    211397.0
26                                0     22660.0     48008.0
27                                0     54390.0    101266.0
28                                0      3328.0      5503.0
31                                0     52891.0     75947.0
32                                0     21419.0     31467.0
33                                0      2011.0      3404.0
34                                0     13945.0      6051.0
35                                0      1633.0      5703.0
36                                0      1156.0      2666.0
38                                0      1671.0      2856.0
39                                0      3542.0      5304.0
42                                0     29195.0     44798.0
43                                0      5182.0     10791.0
44                                0       623.0      2195.0
45                                0     61402.0     93537.0
48                                0     68877.0     94390.0
49                                0     59025.0     95592.0
...                             ...         ...         ...
1139                              0      3852.0      2703.0
1141                              0     16007.0     21482.0
1146                              0      1879.0      2062.0
1148                              0      3222.0      4893.0
1149                              0      1459.0      1605.0
1154                              0      4287.0      5158.0
1155                              0     35731.0     21160.0
1159                              0     16541.0     18795.0
1160                              0     29084.0     32661.0
1161                              0      7520.0      9731.0
1165                              0      9338.0     10112.0
1166                              0     41618.0     41354.0
1167                              0     21464.0     29480.0
1168                              0      1529.0      1578.0
1170                              0     20170.0     14510.0
1171                              0     45397.0     41213.0
1172                              0     42616.0     25322.0
1173                              0      2654.0      3381.0
1176                              0     18765.0     21069.0
1178                              0     20299.0     24844.0
1181                              0      9509.0     12792.0
1182                              0      4434.0      6106.0
1183                              0     40185.0     35282.0
1184                              0     15992.0     16899.0
1186                              0      1628.0     11020.0
1188                              0       834.0      3959.0
1190                              0      1020.0      3658.0
1192                              0      1152.0      5846.0
1196                              0       668.0      2653.0
1199                              0       588.0      2423.0

[677 rows x 3 columns]
```

In [13]:
```
#5 - Party variable (Democratic - 1; Republican - 0)
# Calculate Party
data_whole['Party'] = np.where(data_whole['Democratic']>data_whole['Republican'], '1', '0')
data_whole.to_excel("output.xlsx")
print(data_whole)
```

	State	County	Democratic	Republican	FIPS	\
0	arizona	apache	16298.0	7810.0	4001	
1	arizona	cochise	17383.0	26929.0	4003	
2	arizona	coconino	34240.0	19249.0	4005	
3	arizona	gila	7643.0	12180.0	4007	
4	arizona	graham	3368.0	6870.0	4009	
5	arizona	la paz	1609.0	3265.0	4012	
6	arizona	maricopa	732671.0	672505.0	4013	
7	arizona	mohave	19214.0	50209.0	4015	
8	arizona	navajo	16624.0	18767.0	4017	
9	arizona	pima	221242.0	160550.0	4019	
10	arizona	santa cruz	9241.0	3828.0	4023	
11	arizona	yavapai	40160.0	65308.0	4025	
12	connecticut	fairfield	210899.0	131321.0	9001	
13	connecticut	hartford	203591.0	123864.0	9003	
14	connecticut	middlesex	42383.0	32836.0	9007	
15	connecticut	new haven	179714.0	126004.0	9009	
16	connecticut	tolland	34732.0	28046.0	9013	
17	connecticut	windham	20490.0	19032.0	9015	
18	delaware	sussex	40675.0	50391.0	10005	
19	florida	alachua	74493.0	40599.0	12001	
20	florida	baker	1945.0	8579.0	12003	
21	florida	bay	16723.0	46681.0	12005	
22	florida	bradford	2879.0	7576.0	12007	
23	florida	brevard	121112.0	160305.0	12009	
24	florida	broward	472239.0	211397.0	12011	
25	florida	charlotte	33525.0	52916.0	12015	
26	florida	citrus	22660.0	48008.0	12017	
27	florida	collier	54390.0	101266.0	12021	
28	florida	desoto	3328.0	5503.0	12027	
29	florida	dixie	1322.0	4442.0	12029	
...	...	...	...	...	...	
1170	wisconsin	portage	20170.0	14510.0	55097	
1171	wisconsin	racine	45397.0	41213.0	55101	
1172	wisconsin	rock	42616.0	25322.0	55105	
1173	wisconsin	rusk	2654.0	3381.0	55107	
1174	wisconsin	sawyer	3938.0	4159.0	55113	
1175	wisconsin	sheboygan	24183.0	28667.0	55117	
1176	wisconsin	st. croix	18765.0	21069.0	55109	
1177	wisconsin	taylor	3028.0	5005.0	55119	
1178	wisconsin	walworth	20299.0	24844.0	55127	
1179	wisconsin	washington	23072.0	47102.0	55131	
1180	wisconsin	waukesha	84147.0	136190.0	55133	
1181	wisconsin	waupaca	9509.0	12792.0	55135	
1182	wisconsin	waushara	4434.0	6106.0	55137	
1183	wisconsin	winnebago	40185.0	35282.0	55139	
1184	wisconsin	wood	15992.0	16899.0	55141	
1185	wyoming	albany	7576.0	6366.0	56001	
1186	wyoming	campbell	1628.0	11020.0	56005	
1187	wyoming	carbon	1359.0	3673.0	56007	
1188	wyoming	converse	834.0	3959.0	56009	
1189	wyoming	fremont	4734.0	9262.0	56013	
1190	wyoming	goshen	1020.0	3658.0	56015	
1191	wyoming	johnson	722.0	3085.0	56019	
1192	wyoming	lincoln	1152.0	5846.0	56023	
1193	wyoming	natrona	7285.0	16359.0	56025	
1194	wyoming	niobrara	144.0	980.0	56027	
1195	wyoming	platte	801.0	2850.0	56031	
1196	wyoming	sublette	668.0	2653.0	56035	
1197	wyoming	sweetwater	3943.0	8577.0	56037	
1198	wyoming	uinta	1371.0	4713.0	56041	
1199	wyoming	washakie	588.0	2423.0	56043	

	Total Population	Percent White, not Hispanic or Latino	\
0	72346	18.571863	
1	128177	56.299492	
2	138064	54.619597	
3	53179	63.222325	
4	37529	51.461536	
5	20304	58.884949	
6	4088549	56.918114	
7	203629	78.252606	
8	108209	41.927196	
9	1003338	53.271579	
10	46547	15.274883	
11	218586	81.159361	
12	941618	63.509512	
13	895699	63.252052	
14	164438	84.738929	
15	860874	64.782767	
16	151689	85.892187	
17	117078	83.782606	
18	211224	74.829091	
19	256581	62.460198	
20	27312	82.088459	
21	178361	77.687387	
22	26919	74.787325	
23	560683	75.798803	

```
24            1863780                          39.245351
25             169642                          84.643543
26             140453                          88.871010
27             348236                          64.222539
28              35134                          55.040701
29              16084                          85.041035
...               ...                                ...
1170            70551                          91.815850
1171           194873                          73.058864
1172           160986                          83.642056
1173            14272                          95.557735
1174            16438                          77.302592
1175           115269                          85.579818
1176            86726                          94.420358
1177            20521                          96.549876
1178           103021                          85.711651
1179           133422                          93.426871
1180           395377                          89.471568
1181            51974                          94.860892
1182            24238                          89.817642
1183           169487                          89.548461
1184            73621                          93.080779
1185            37836                          83.269902
1186            48473                          87.774637
1187            15696                          77.924312
1188            14223                          88.849047
1189            40683                          70.198855
1190            13546                          86.409272
1191             8572                          91.565562
1192            18543                          92.600982
1193            80871                          87.026252
1194             2498                          89.511609
1195             8740                          89.359268
1196            10032                          91.646730
1197            44812                          79.815674
1198            20893                          87.718375
1199             8351                          82.397318

      Percent Black, not Hispanic or Latino  Percent Hispanic or Latino  \
0                                  0.486551                   5.947806
1                                  3.714395                  34.403208
2                                  1.342855                  13.711033
3                                  0.552850                  18.548675
4                                  1.811932                  32.097844
5                                  0.379236                  26.182033
6                                  5.013612                  30.286833
7                                  0.951731                  15.708470
8                                  0.672772                  11.049913
9                                  3.199719                  36.105978
10                                 0.199798                  83.219112
11                                 0.518331                  14.054880
12                                10.456151                  18.636007
13                                12.657935                  16.897976
14                                 4.664980                   5.641032
15                                12.177276                  16.790959
16                                 2.811674                   5.029369
17                                 1.866277                  10.877364
18                                12.284589                   9.195925
19                                19.504562                   8.994820
20                                13.876684                   2.317663
21                                10.527526                   5.728270
22                                19.250344                   3.796575
23                                 9.838001                   9.301691
24                                27.214639                  27.564841
25                                 5.535775                   6.632202
26                                 2.817313                   5.121998
27                                 6.775003                  26.634524
28                                12.904878                  30.688222
29                                 7.504352                   3.780154
...                                     ...                        ...
1170                               0.687446                   2.951057
1171                              10.764960                  12.452212
1172                               3.404644                   8.207546
1173                               0.987948                   1.604540
1174                               0.663098                   2.269132
1175                               1.474811                   5.975588
1176                               0.785232                   2.180430
1177                               0.375225                   1.744554
1178                               0.990089                  10.894866
1179                               1.077034                   2.972523
1180                               1.402712                   4.495709
1181                               0.352099                   2.880286
1182                               1.547157                   6.213384
1183                               1.996613                   3.841003
1184                               0.709037                   2.665000
1185                               1.345280                   9.229305
1186                               0.742681                   8.357230
1187                               1.242355                  17.647808
```

1188	0.007031	7.691767
1189	0.516186	6.779244
1190	0.147645	10.519711
1191	0.443304	2.134858
1192	0.210322	4.416761
1193	1.158635	8.198242
1194	0.120096	4.243395
1195	0.057208	7.814645
1196	0.000000	7.814992
1197	0.865840	15.859591
1198	0.186665	8.959939
1199	0.790325	13.962400

	Percent Foreign Born	Percent Female	Percent Age 29 and Under	\
0	1.719515	50.598513	45.854643	
1	11.458374	49.069646	37.902276	
2	4.825298	50.581614	48.946141	
3	4.249798	50.296170	32.238290	
4	4.385942	46.313518	46.393456	
5	11.372143	48.946020	28.073286	
6	14.729333	50.549278	41.886620	
7	6.969047	49.676618	30.485835	
8	2.914730	49.846131	43.243168	
9	12.903428	50.807405	40.087388	
10	32.644424	52.125808	43.300320	
11	6.456955	51.092476	28.717301	
12	21.154120	51.308280	38.019876	
13	15.011070	51.492075	37.389458	
14	7.652732	51.098894	33.241708	
15	12.073428	51.759956	37.982562	
16	6.769113	49.967367	42.301024	
17	5.269991	50.422795	37.389604	
18	6.671117	51.495569	31.746866	
19	10.144555	51.670623	48.551530	
20	1.475542	47.715290	41.344464	
21	5.479898	50.280050	37.739192	
22	2.306921	44.741632	35.343066	
23	8.632685	51.140841	31.972612	
24	32.715718	51.376504	36.701971	
25	10.185567	51.260891	22.388913	
26	5.413911	51.667818	24.581889	
27	23.816894	50.881299	29.998909	
28	17.504412	43.345477	37.513520	
29	2.785377	45.417807	31.236011	
...	...	...	...	
1170	3.119729	49.846211	42.641493	
1171	5.064837	50.538043	38.058120	
1172	4.583007	50.796343	38.521362	
1173	1.891816	49.691704	31.747478	
1174	1.502616	49.306485	30.800584	
1175	5.754366	49.745378	36.383590	
1176	2.453705	50.064571	38.543228	
1177	1.461917	48.842649	35.100629	
1178	6.097786	50.039312	39.967579	
1179	2.717693	50.384494	35.089416	
1180	4.902410	50.893198	34.964350	
1181	1.816293	49.265017	33.189672	
1182	2.784883	47.520422	30.006601	
1183	3.318839	49.604394	39.787122	
1184	2.326782	50.712433	34.645006	
1185	6.081510	47.756105	55.447193	
1186	3.645328	48.008170	45.326264	
1187	5.198777	45.839704	39.181957	
1188	2.706883	49.933207	38.515081	
1189	1.339626	49.907824	39.751247	
1190	2.724051	47.091392	35.914661	
1191	1.656556	46.966869	32.571162	
1192	2.151755	48.773122	38.715418	
1193	2.729038	49.421919	40.688257	
1194	0.280224	54.443555	36.269015	
1195	2.780320	47.711670	32.700229	
1196	2.053429	46.949761	36.393541	
1197	5.509685	47.824244	44.153352	
1198	3.986981	49.327526	43.205858	
1199	3.783978	51.359119	34.774279	

	Percent Age 65 and Older	Median Household Income	Percent Unemployed	\
0	13.322091	32460	15.807433	
1	19.756275	45383	8.567108	
2	10.873943	51106	8.238305	
3	26.397638	40593	12.129932	
4	12.315809	47422	14.424104	
5	36.056935	36321	10.599013	
6	13.837843	55676	6.808454	
7	26.858650	39856	11.680953	
8	15.745456	36868	18.525791	
9	17.801778	46764	9.214114	
10	15.895761	38941	9.749896	

11	28.272625	46638	8.525986
12	14.383115	86670	8.211898
13	15.644876	68027	8.237243
14	17.703937	79837	5.257635
15	15.604490	62715	8.545483
16	13.868507	80129	6.267325
17	14.571482	60689	8.375524
18	24.214578	54218	7.108621
19	12.430772	44702	7.020154
20	12.910076	53327	6.832522
21	16.103857	48577	7.475668
22	17.452357	43373	10.563312
23	22.491140	49914	9.278023
24	15.371879	52954	8.710312
25	37.622759	44865	10.088979
26	35.105694	39054	11.648012
27	29.589704	59783	6.408868
28	19.408550	35513	8.291825
29	21.729669	34634	5.306200
...	...	...	...
1170	14.644725	52411	6.178216
1171	14.722409	56359	7.478894
1172	15.032984	50968	7.390601
1173	22.323430	39904	4.930617
1174	23.317922	41869	7.665187
1175	16.257623	54392	5.054640
1176	12.209718	73743	4.081465
1177	18.088787	47306	4.478597
1178	15.303676	55575	6.442507
1179	15.645096	70325	4.019701
1180	16.492866	78268	3.665774
1181	19.523223	52441	4.546458
1182	22.353329	46581	6.336312
1183	14.738003	53501	4.420678
1184	18.608821	49926	5.669639
1185	9.549107	43043	4.579174
1186	6.954387	80822	4.669432
1187	13.774210	56972	4.141937
1188	13.668003	66737	5.282284
1189	16.409803	53559	7.344324
1190	20.389783	44883	6.918819
1191	20.496967	54594	4.512276
1192	14.382786	64579	5.618095
1193	12.825364	56983	4.861868
1194	18.254604	40640	0.457875
1195	22.013730	41051	3.901047
1196	13.337321	76004	2.786971
1197	9.417120	68233	5.072255
1198	10.678218	53323	6.390755
1199	19.650341	46212	7.441860

	Percent Less than High School Degree
0	21.758252
1	13.409171
2	11.085381
3	15.729958
4	14.580797
5	24.842215
6	13.051927
7	16.145850
8	18.494087
9	12.252238
10	25.206726
11	9.830672
12	10.103521
13	10.736314
14	6.153444
15	10.503197
16	6.214203
17	11.925239
18	13.792481
19	7.468059
20	17.871127
21	11.320101
22	23.212996
23	8.823708
24	11.688725
25	10.578890
26	13.288908
27	14.321540
28	29.475750
29	22.236087
...	...
1170	6.784049
1171	11.170104
1172	10.591406
1173	12.513379
1174	8.901562

```
1175                            8.367242
1176                            4.242959
1177                           12.358524
1178                            9.074810
1179                            6.072757
1180                            4.116150
1181                            9.145423
1182                           13.760962
1183                            7.893714
1184                            7.658332
1185                            4.210167
1186                            8.330054
1187                            9.579879
1188                            9.758393
1189                            8.537172
1190                            8.390574
1191                            5.105750
1192                            6.949996
1193                            8.332255
1194                           11.054422
1195                            9.675889
1196                            4.658830
1197                            9.314606
1198                           10.361224
1199                           12.577108
```

	Percent Less than Bachelor's Degree	Percent Rural	Party
0	88.941063	74.061076	1
1	76.837055	36.301067	0
2	65.791439	31.466066	1
3	82.262624	41.062000	0
4	86.675944	46.437399	0
5	89.563407	56.327786	0
6	69.031137	2.363800	1
7	88.121178	22.963644	0
8	85.507970	54.138242	0
9	69.199391	7.523491	1
10	77.506775	26.883172	1
11	74.458362	33.197178	0
12	53.637538	4.584061	1
13	63.133913	5.408640	1
14	59.071541	24.540066	1
15	65.711800	3.639981	1
16	61.140250	38.207884	1
17	76.143316	49.761036	1
18	76.115066	41.303609	0
19	58.485040	21.193437	1
20	87.235709	59.491057	0
21	77.466127	11.999858	0
22	88.220732	75.529453	0
23	72.271467	5.074387	0
24	68.984269	0.016933	1
25	78.396083	8.884347	0
26	82.154355	34.518820	0
27	66.118970	8.485942	0
28	89.960743	46.242327	0
29	93.602862	77.024723	0
...	...	...	...
1170	69.694941	36.031649	1
1171	76.027600	12.277901	1
1172	79.234347	20.416513	1
1173	85.209692	76.624873	0
1174	77.865939	84.200036	0
1175	76.378160	28.284866	0
1176	66.600965	53.168534	0
1177	86.537786	80.443714	0
1178	72.181794	34.197089	0
1179	70.577928	30.784687	0
1180	58.376564	9.860448	0
1181	82.335700	64.939897	0
1182	84.915338	89.496244	0
1183	73.113141	13.431021	1
1184	79.617467	36.682765	0
1185	52.136794	11.939723	1
1186	80.935983	29.072464	0
1187	79.681613	41.630469	0
1188	84.468152	55.360370	0
1189	77.648670	51.424370	0
1190	77.036880	45.995924	0
1191	74.537343	49.025557	0
1192	79.631291	82.729482	0
1193	77.786814	14.449304	0
1194	81.009070	100.000000	0
1195	80.300395	58.647744	0
1196	75.645069	100.000000	0
1197	78.628507	10.916313	0
1198	81.793082	43.095937	0
1199	78.923920	35.954529	0

```
[1200 rows x 19 columns]
```

In [14]:
```
#6 - Hypothesis Testing - Total Population - Democratic & Republican counties
#6i - Mean of Democratic counties
d_democratic = data_whole.loc[data_whole['Democratic']>data_whole['Republican']]
d_democratic_tp = d_democratic[['Total Population', 'Party']]
sample_mean = d_democratic_tp['Total Population'].mean()
print("Mean population for Democratic Counties:", sample_mean)
```

```
Mean population for Democratic Counties: 300998.3169230769
```

In [15]:
```
#6ii - Mean of Republican counties
d_republican = data_whole.loc[data_whole['Democratic']<data_whole['Republican']]
d_republican_tp = d_republican[['Total Population', 'Party']]
sample_mean = d_republican_tp['Total Population'].mean()
print("Mean population for Republican Counties:", sample_mean)
```

```
Mean population for Republican Counties: 53864.6724137931
```

In [16]:
```
'''
Mean population for Democratic Counties: 300998.3169230769
Mean population for Republican Counties: 53864.6724137931

Therefore clearly, mean population for democratic counties is greater than mean population for republican counties.
'''
```

In [17]:
```
#6iii - Compute sample standard deviations of Democratic & Republican
sample_std_democratic = d_democratic_tp['Total Population'].std()
sample_std_republican = d_republican_tp['Total Population'].std()
print("Sample standard deviation of democratic counties population:", sample_std_democratic)
print("Sample standard deviation of republican counties population:", sample_std_republican)
```

```
Sample standard deviation of democratic counties population: 553600.0257123302
Sample standard deviation of republican counties population: 94192.57279397613
```

In [18]:
```
#6iv - Compute p-value using scipy.stats.ttest_ind
import scipy.stats as st
[statistic, pvalue] = st.ttest_ind(d_democratic_tp['Total Population'], d_republican_tp['Total Population'], equal_var =
False)
print("t-test statistic:", statistic)
print("pvalue:", pvalue)
```

```
t-test statistic: 8.004638577960957
pvalue: 2.0478717602973023e-14
```

In [19]:
```
'''
6th Answer Conclusion: Total Population
Null Hypothesis: μ(Democratic population) = μ(Republican population)
Alternate Hypothesis: μ(Democratic population) != μ(Republican population)
t-test statistic: 8.004638577960957
pvalue: 2.0478717602973023e-14
Since pvalue < α (Statistically signficant)
Therefore, we reject null hypothesis.
There is sufficient evidence to conclude that the mean population of
democratic counties is different from republican counties.
'''
```

In [20]:
```
#7 - Hypothesis Testing - Median Household income - Democratic & Republican counties
#7i - Mean of Democratic counties
d_democratic_mhhi = d_democratic[['Median Household Income', 'Party']]
sample_mean = d_democratic_mhhi['Median Household Income'].mean()
print("Mean Median Household Income for Democratic Counties:", sample_mean)
```

```
Mean Median Household Income for Democratic Counties: 53798.732307692306
```

In [21]:
```
#7ii - Mean of Republican counties
d_republican_mhhi = d_republican[['Median Household Income', 'Party']]
sample_mean = d_republican_mhhi['Median Household Income'].mean()
print("Mean Median Household Income for Republican Counties:", sample_mean)
```

```
Mean Median Household Income for Republican Counties: 48746.81954022989
```

In [22]:
```
'''
Mean Median Household Income for Democratic Counties: 53798.732307692306
Mean Median Household Income for Republican Counties: 48746.81954022989

Therefore clearly, mean median household income for democratic counties is
greater than mean population for republican counties.
'''
```

In [23]:
```
#7iii - Compute sample standard deviations of Democratic & Republican
sample_std_democratic = d_democratic_mhhi['Median Household Income'].std()
sample_std_republican = d_republican_mhhi['Median Household Income'].std()
print("Sample standard deviation of democratic counties Median Household Income:", sample_std_democratic)
print("Sample standard deviation of republican counties Median Household Income:", sample_std_republican)
```

```
Sample standard deviation of democratic counties Median Household Income: 15289.130077404621
Sample standard deviation of republican counties Median Household Income: 10670.729411820219
```

In [24]:
```
#7iv - Compute p-value using scipy.stats.ttest_ind
import scipy.stats as st
[statistic, pvalue] = st.ttest_ind(d_democratic_mhhi['Median Household Income'], d_republican_mhhi['Median Household Income'], equal_var = False)
print("t-test statistic:", statistic)
print("pvalue:", pvalue)
```

```
t-test statistic: 5.479141589767387
pvalue: 7.149437363182572e-08
```

In [25]:
```
'''
7th Answer Conclusion: Total Population
Null Hypothesis: μ(Democratic median household income) = μ(Republican median household income)
Alternate Hypothesis: μ(Democratic median household income) != μ(Republican median household income)
t-test statistic: 5.479141589767387
pvalue: 7.149437363182598e-08
Since pvalue < α (Statistically signficant)
Therefore, we reject null hypothesis.
There is sufficient evidence to conclude that the mean median household income of
democratic counties is different from republican counties.
'''
```

In [26]:
```
#8i - Descriptive Statistics

#-----------------------------------------------------------------Part1(Statistic Descriptions)-----------------------------
-----------------------------------------

#Create a dataframe that contains variables we are interested in age, gender, race and ethinitcy, and education
data_stat = data_whole.drop(columns= ['Democratic', 'Republican', 'FIPS', 'Total Population', 'Median Household Income',
'Percent Unemployed', 'Percent Rural' ])

#Get percentage of males
percent_male = []
for i in range(len(data_stat['Percent Female'])):
    percent_male.append(100 - data_stat.iloc[i,7])
data_stat['Percent Male'] = percent_male

#Reorder the columns so male and female percentages are close
data_stat = data_stat[['Percent White, not Hispanic or Latino','Percent Black, not Hispanic or Latino','Percent Hispanic
or Latino','Percent Foreign Born', 'Percent Female', 'Percent Male','Percent Age 29 and Under','Percent Age 65 and Olde
r','Percent Less than High School Degree','Percent Less than Bachelor\'s Degree', 'Party']]

#Get the information for each county
democratic_counties = data_stat.loc[data_whole['Democratic']>data_whole['Republican']]
republican_counties = data_stat.loc[data_whole['Democratic']<data_whole['Republican']]

#8i - Compute the statistic description for democratic county
democratic_counties.describe().transpose()
```

Out[26]:

	count	mean	std	min	25%	50%	75%	max
Percent White, not Hispanic or Latino	325.0	69.683766	24.981502	2.776702	53.271579	77.786090	90.300749	98.063495
Percent Black, not Hispanic or Latino	325.0	9.242649	13.351340	0.000000	0.839103	3.485992	11.058843	63.953279
Percent Hispanic or Latino	325.0	12.587391	19.575030	0.193349	2.531017	5.039747	11.857116	95.479801
Percent Foreign Born	325.0	7.986330	8.330740	0.179769	2.470508	5.105490	10.144555	52.229868
Percent Female	325.0	50.385433	2.149359	34.245291	49.854280	50.653830	51.492075	56.418468
Percent Male	325.0	61.273041	6.252786	32.632177	57.838838	61.925849	65.511556	76.843548
Percent Age 29 and Under	325.0	38.726959	6.252786	23.156452	34.488444	38.074151	42.161162	67.367823
Percent Age 65 and Older	325.0	16.194826	4.282422	6.653188	13.106233	15.698087	18.806426	31.642106
Percent Less than High School Degree	325.0	11.883760	6.505613	3.215803	7.893714	10.370080	13.637059	49.673777
Percent Less than Bachelor's Degree	325.0	71.968225	11.192404	26.335440	65.711800	72.736143	79.903653	94.849957

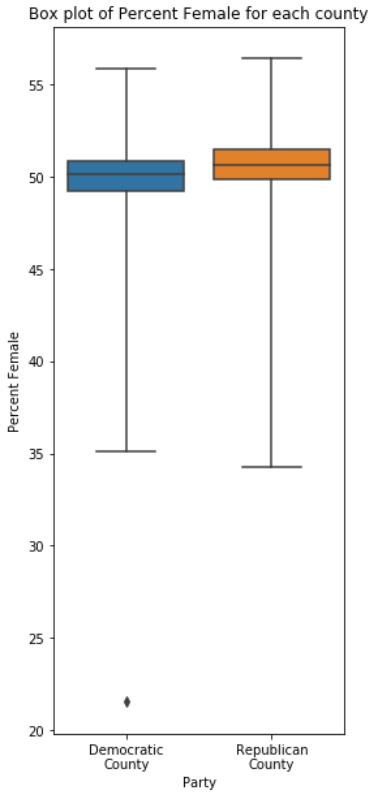
In [27]:
```
#8ii - Compute the statistic description for republican county
republican_counties.describe().transpose()
```

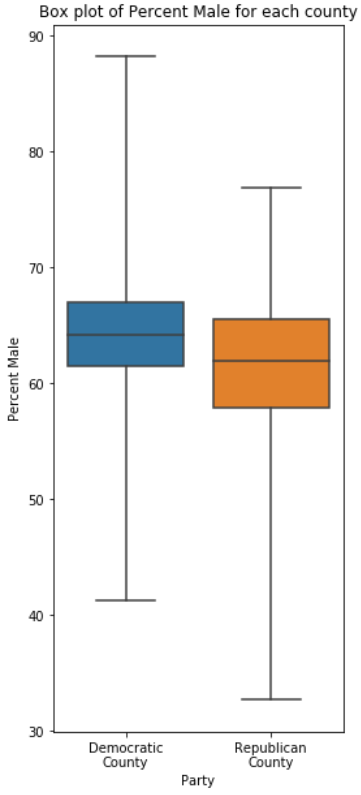
Out[27]:

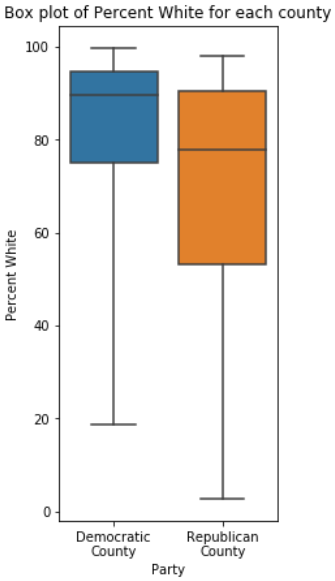
	count	mean	std	min	25%	50%	75%	max
Percent White, not Hispanic or Latino	870.0	82.656646	16.056122	18.758977	75.016397	89.434849	94.466596	99.627329
Percent Black, not Hispanic or Latino	870.0	4.189241	6.721695	0.000000	0.460419	1.318311	4.753831	41.563041
Percent Hispanic or Latino	870.0	9.733094	14.049576	0.000000	1.704539	3.427435	10.709696	78.397012
Percent Foreign Born	870.0	3.990096	4.507786	0.000000	1.320101	2.326317	5.149429	37.058317
Percent Female	870.0	49.630898	2.429013	21.513413	49.222905	50.176792	50.829770	55.885023
Percent Male	870.0	63.994281	5.181522	41.250884	61.460213	64.153468	67.016348	88.157895
Percent Age 29 and Under	870.0	36.005719	5.181522	11.842105	32.983652	35.846532	38.539787	58.749116
Percent Age 65 and Older	870.0	18.828267	4.733155	6.954387	15.784982	18.377896	21.112847	37.622759
Percent Less than High School Degree	870.0	14.009112	6.303126	2.134454	9.662491	12.572435	17.447168	47.812773
Percent Less than Bachelor's Degree	870.0	81.095427	6.815537	43.419470	78.108424	82.406700	85.546272	97.014925

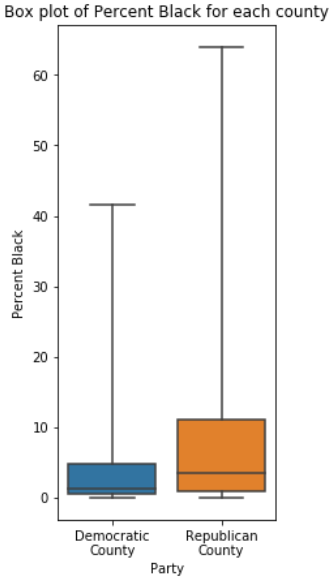
In [28]:
```
#8ii - Plotting
#-------------------------------------------------------------Part2(Plot)-------------------------------------------------
-----------------------------------

plt.figure(figsize=(4, 10))
ax = sns.boxplot(x = 'Party', y = 'Percent Female', data = data_stat, whis=10)
ax.set(title = 'Box plot of Percent Female for each county', xlabel = 'Party', ylabel = 'Percent Female')
ax.set_xticklabels(['Democratic\nCounty', 'Republican\nCounty'])

plt.figure(figsize=(4, 10))
ax = sns.boxplot(x = 'Party', y = 'Percent Male', data = data_stat, whis=10)
ax.set(title = 'Box plot of Percent Male for each county', xlabel = 'Party', ylabel = 'Percent Male')
ax.set_xticklabels(['Democratic\nCounty', 'Republican\nCounty'])

plt.figure(figsize=(3, 7))
ax = sns.boxplot(x = 'Party', y = 'Percent White, not Hispanic or Latino', data = data_stat, whis=10)
ax.set(title = 'Box plot of Percent White for each county', xlabel = 'Party', ylabel = 'Percent White')
ax.set_xticklabels(['Democratic\nCounty', 'Republican\nCounty'])

plt.figure(figsize=(3, 7))
ax = sns.boxplot(x = 'Party', y = 'Percent Black, not Hispanic or Latino', data = data_stat, whis=10)
ax.set(title = 'Box plot of Percent Black for each county', xlabel = 'Party', ylabel = 'Percent Black')
ax.set_xticklabels(['Democratic\nCounty', 'Republican\nCounty'])

plt.figure(figsize=(3, 7))
ax = sns.boxplot(x = 'Party', y = 'Percent Hispanic or Latino', data = data_stat, whis=10)
ax.set(title = 'Box plot of Percent Hispanic\nor Latino for each county', xlabel = 'Party', ylabel = 'Percent Hispanic o
r Latino')
ax.set_xticklabels(['Democratic\nCounty', 'Republican\nCounty'])

plt.figure(figsize=(3, 7))
ax = sns.boxplot(x = 'Party', y = 'Percent Foreign Born', data = data_stat, whis=10)
ax.set(title = 'Box plot of Percent Foreign for each county', xlabel = 'Party', ylabel = 'Percent Foreign')
ax.set_xticklabels(['Democratic\nCounty', 'Republican\nCounty'])

plt.figure(figsize=(3, 7))
ax = sns.boxplot(x = 'Party', y = 'Percent Age 29 and Under', data = data_stat, whis=10)
ax.set(title = 'Box plot of Percent Age 29\nand Under for each county', xlabel = 'Party', ylabel = 'Percent 29 and Unde
r')
ax.set_xticklabels(['Democratic\nCounty', 'Republican\nCounty'])

plt.figure(figsize=(3, 7))
ax = sns.boxplot(x = 'Party', y = 'Percent Age 65 and Older', data = data_stat, whis=10)
ax.set(title = 'Box plot of Percent Age 65\nand Older for each county', xlabel = 'Party', ylabel = 'Percent 65 and Olde
r')
ax.set_xticklabels(['Democratic\nCounty', 'Republican\nCounty'])

plt.figure(figsize=(3, 7))
ax = sns.boxplot(x = 'Party', y = 'Percent Less than High School Degree', data = data_stat, whis=10)
ax.set(title = 'Box plot of Percent Less than \nHigh School Degree for each county', xlabel = 'Party', ylabel = 'Percent
Less Than Highschool')
ax.set_xticklabels(['Democratic\nCounty', 'Republican\nCounty'])

plt.figure(figsize=(3, 7))
ax = sns.boxplot(x = 'Party', y = 'Percent Less than Bachelor\'s Degree', data = data_stat, whis=10)
ax.set(title = 'Box plot of Percent Less than \nBachelor Degree for each county', xlabel = 'Party', ylabel = 'Percent Le
ss Than Bachelor')
ax.set_xticklabels(['Democratic\nCounty', 'Republican\nCounty'])
```

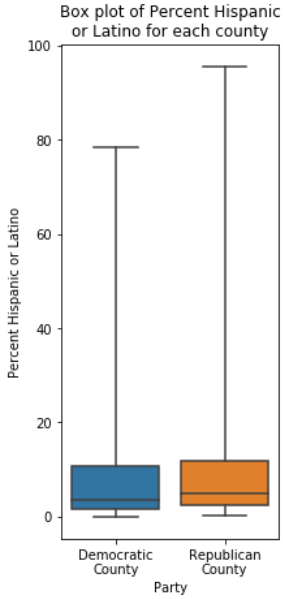
Out[28]: [Text(0, 0, 'Democratic\nCounty'), Text(0, 0, 'Republican\nCounty')]

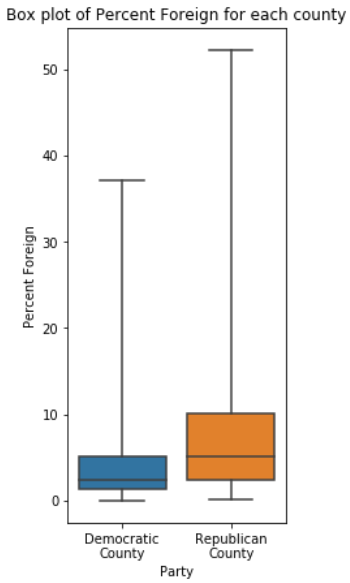


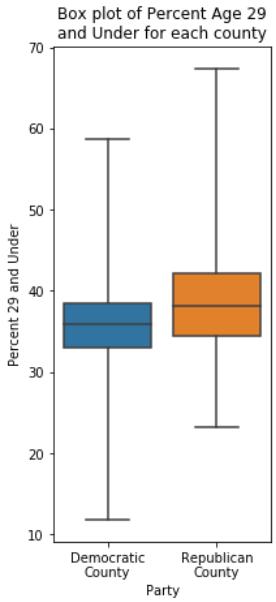


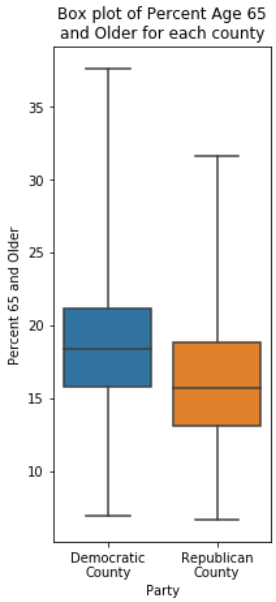

Box plot of Percent White for each county
Percent White
100
80
60
40
20
0
Democratic County
Republican County
Party

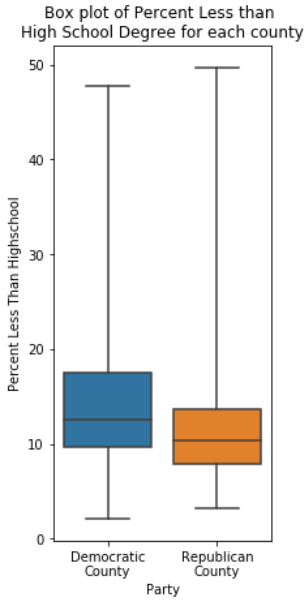

Box plot of Percent Black for each county
Percent Black
60
50
40
30
20
10
0
Democratic County
Republican County
Party

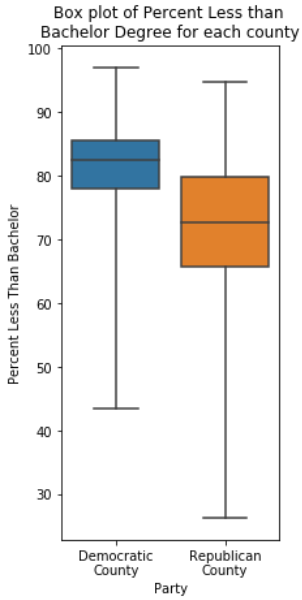

Box plot of Percent Hispanic or Latino for each county
Percent Hispanic or Latino
100
80
60
40
20
0
Democratic County
Republican County
Party


Box plot of Percent Foreign for each county
Percent Foreign
0
10
20
30
40
50
Democratic County
Republican County
Party


Box plot of Percent Age 29 and Under for each county
Percent 29 and Under
10
20
30
40
50
60
70
Democratic County
Republican County
Party


Box plot of Percent Age 65 and Older for each county
Percent 65 and Older
10
15
20
25
30
35
Democratic County
Republican County
Party





In [29]:
```
'''
#9 Answer:
The variables that are important to determine whether a county is marked Democratic or Republican are
Race and Ethnicity, Education and Age.
The percentage of White make up about 70 percent of Republicans and 83 percent of Democrats.
Education also plays a very large part, 72 percent of Democrats have a degree less than a bachelor's
degree and 81 percent of Republicans have a degree less than a bachelor's degree.
Finally people age 29 and younger makeup 39 percent of Republicans and 36 percent of Democrats.
'''
```

In [31]:
```
#10 - Create a map of Democratic & Republic counties with FIPS codes based on the dataset
import plotly.figure_factory as ff
from plotly.offline import init_notebook_mode, iplot # Needed to render the figure when exporting to HTML
init_notebook_mode(connected=True)

fips = data_whole['FIPS'].tolist()
party_values = data_whole['Party'].map({'0': 'Republican', '1': 'Democratic'})
colorscale = ["#1689E0", "#D13D3F"]
figure = ff.create_choropleth(fips=fips,
                              values=party_values,
                              colorscale=colorscale,
                              county_outline={'color': '#000000', 'width': 0.5},
                              show_hover=False,
                              title='Political Party by Counties',
                              legend_title='Political Party')
figure.layout.template = None
iplot(figure, validate=False) # Displaying figure even when exported to HTML
```