In [1]:

```
#Add the required libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import statsmodels.formula.api as smf
from sklearn import linear_model
import scipy.stats as st
import numpy as np
import plotly.figure_factory as ff
```

In [2]:

```
# Task 1
data_a = pd.read_csv("election_train.csv")
data_b = pd.read_csv("demographics_train.csv")

data_a_tidy = pd.pivot_table(data_a,index = ['Year','State','County','Office'],columns
= 'Party',values = 'Votes',aggfunc='sum').reset_index()

data_a_tidy.head()
```

Out[2]:

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

In [3]:

```
# Task 2
data_a_tidy['County'] = data_a_tidy['County'].str.replace(' County','')
data_a_tidy['County'] = data_a_tidy['County'].str.lower()
data_b['County'] = data_b['County'].str.lower()

change_state = {'Arizona' : 'AZ', 'Connecticut' : 'CT', 'Delaware' : 'DE', 'Florida' :
'FL', 'Hawaii' : 'HI', 'Indiana' : 'IN', 'Maine' : 'ME', 'Maryland' : 'MD', 'Massachuse
tts' : 'MA', 'Michigan' : 'MI', 'Minnesota' : 'MN', 'Montana' : 'MT', 'Nebraska' : 'NE'
, 'Nevada' : 'NV', 'New Jersey' : 'NJ', 'New Mexico' : 'NM', 'New York' : 'NY', 'North
 Dakota' : 'ND', 'Ohio' : 'OH', 'Pennsylvania' : 'PA', 'Rhode Island' : 'RI', 'Tennesse
e' : 'TN', 'Texas' : 'TX', 'Utah' : 'UT', 'Vermont' : 'VT', 'Virginia' : 'VA', 'Washing
ton' : 'WA', 'West Virginia' : 'WV', 'Wisconsin' : 'WI', 'Wyoming' : 'WY'}
data_b['State'] = data_b['State'].map(change_state)

data_merge = pd.merge(data_a_tidy,data_b, how="inner", on=['County','State'], sort=True
)
data_merge.head()
```

Out[3]:

	Year	State	County	Office	Democratic	Republican	FIPS	Total Population	Citizen Voting-Age Population	F Hi or
0	2018	IN	adams	US Senator	3146.0	7511.0	18001	34813	0	93.
1	2018	ND	adams	US Senator	364.0	796.0	38001	2348	0	93.
2	2018	NE	adams	US Senator	3334.0	6487.0	31001	31536	0	87.
3	2018	OH	adams	US Senator	2635.0	6000.0	39001	28111	0	96.
4	2018	PA	adams	US Senator	14880.0	23419.0	42001	101759	78370	89.

5 rows × 21 columns

In [4]:

```
# Task 3
print(data_merge.info())
print(data_merge.Year.unique())
print(data_merge.Office.unique())
data_merge = data_merge.drop(columns=['Office','Year'])
```

```
<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
None
[2018]
['US Senator']
```

In [5]:

```
#Task 4
print(data_merge.isnull()) # Checking to see if the values are NULL
print(data_merge.isin([0]).sum()) # Printing out the missing values
print(data_merge.isin([0]).sum().sum()) # Printing out all the missing values
```

	State	County	Democratic	Republican	FIPS	Total Population	\
0	False	False	False	False	False	False	
1	False	False	False	False	False	False	
2	False	False	False	False	False	False	
3	False	False	False	False	False	False	
4	False	False	False	False	False	False	
5	False	False	False	False	False	False	
6	False	False	False	False	False	False	
7	False	False	False	False	False	False	
8	False	False	False	False	False	False	
9	False	False	False	False	False	False	
10	False	False	False	False	False	False	
11	False	False	False	False	False	False	
12	False	False	False	False	False	False	
13	False	False	False	False	False	False	
14	False	False	False	False	False	False	
15	False	False	False	False	False	False	
16	False	False	False	False	False	False	
17	False	False	False	False	False	False	
18	False	False	False	False	False	False	
19	False	False	False	False	False	False	
20	False	False	False	False	False	False	
21	False	False	False	False	False	False	
22	False	False	False	False	False	False	
23	False	False	False	False	False	False	
24	False	False	False	False	False	False	
25	False	False	False	False	False	False	
26	False	False	False	False	False	False	
27	False	False	False	False	False	False	
28	False	False	False	False	False	False	
29	False	False	False	False	False	False	
...	...	...	...	...	...	...	
1170	False	False	False	False	False	False	
1171	False	False	False	False	False	False	
1172	False	False	False	False	False	False	
1173	False	False	False	False	False	False	
1174	False	False	False	False	False	False	
1175	False	False	False	False	False	False	
1176	False	False	False	False	False	False	
1177	False	False	False	False	False	False	
1178	False	False	False	False	False	False	
1179	False	False	False	False	False	False	
1180	False	False	False	False	False	False	
1181	False	False	False	False	False	False	
1182	False	False	False	False	False	False	
1183	False	False	False	False	False	False	
1184	False	False	False	False	False	False	
1185	False	False	False	False	False	False	
1186	False	False	False	False	False	False	
1187	False	False	False	False	False	False	
1188	False	False	False	False	False	False	
1189	False	False	False	False	False	False	
1190	False	False	False	False	False	False	
1191	False	False	False	False	False	False	
1192	False	False	False	False	False	False	
1193	False	False	False	False	False	False	
1194	False	False	False	False	False	False	
1195	False	False	False	False	False	False	
1196	False	False	False	False	False	False	
1197	False	False	False	False	False	False	
1198	False	False	False	False	False	False	

```
1199  False   False       False       False  False             False

      Citizen Voting-Age Population  Percent White, not Hispanic or Latino
\
0                                False                                  False
1                                False                                  False
2                                False                                  False
3                                False                                  False
4                                False                                  False
5                                False                                  False
6                                False                                  False
7                                False                                  False
8                                False                                  False
9                                False                                  False
10                               False                                  False
11                               False                                  False
12                               False                                  False
13                               False                                  False
14                               False                                  False
15                               False                                  False
16                               False                                  False
17                               False                                  False
18                               False                                  False
19                               False                                  False
20                               False                                  False
21                               False                                  False
22                               False                                  False
23                               False                                  False
24                               False                                  False
25                               False                                  False
26                               False                                  False
27                               False                                  False
28                               False                                  False
29                               False                                  False
...                                ...                                    ...
1170                             False                                  False
1171                             False                                  False
1172                             False                                  False
1173                             False                                  False
1174                             False                                  False
1175                             False                                  False
1176                             False                                  False
1177                             False                                  False
1178                             False                                  False
1179                             False                                  False
1180                             False                                  False
1181                             False                                  False
1182                             False                                  False
1183                             False                                  False
1184                             False                                  False
1185                             False                                  False
1186                             False                                  False
1187                             False                                  False
1188                             False                                  False
1189                             False                                  False
1190                             False                                  False
1191                             False                                  False
1192                             False                                  False
1193                             False                                  False
1194                             False                                  False
1195                             False                                  False
```

```
1196                          False                                False
1197                          False                                False
1198                          False                                False
1199                          False                                False

      Percent Black, not Hispanic or Latino  Percent Hispanic or Latino  \
0                                     False                        False
1                                     False                        False
2                                     False                        False
3                                     False                        False
4                                     False                        False
5                                     False                        False
6                                     False                        False
7                                     False                        False
8                                     False                        False
9                                     False                        False
10                                    False                        False
11                                    False                        False
12                                    False                        False
13                                    False                        False
14                                    False                        False
15                                    False                        False
16                                    False                        False
17                                    False                        False
18                                    False                        False
19                                    False                        False
20                                    False                        False
21                                    False                        False
22                                    False                        False
23                                    False                        False
24                                    False                        False
25                                    False                        False
26                                    False                        False
27                                    False                        False
28                                    False                        False
29                                    False                        False
...                                     ...                          ...
1170                                  False                        False
1171                                  False                        False
1172                                  False                        False
1173                                  False                        False
1174                                  False                        False
1175                                  False                        False
1176                                  False                        False
1177                                  False                        False
1178                                  False                        False
1179                                  False                        False
1180                                  False                        False
1181                                  False                        False
1182                                  False                        False
1183                                  False                        False
1184                                  False                        False
1185                                  False                        False
1186                                  False                        False
1187                                  False                        False
1188                                  False                        False
1189                                  False                        False
1190                                  False                        False
1191                                  False                        False
1192                                  False                        False
1193                                  False                        False
```

```
1194                                False                       False
1195                                False                       False
1196                                False                       False
1197                                False                       False
1198                                False                       False
1199                                False                       False

      Percent Foreign Born  Percent Female  Percent Age 29 and Under  \
0                    False           False                      False
1                    False           False                      False
2                    False           False                      False
3                    False           False                      False
4                    False           False                      False
5                    False           False                      False
6                    False           False                      False
7                    False           False                      False
8                    False           False                      False
9                    False           False                      False
10                   False           False                      False
11                   False           False                      False
12                   False           False                      False
13                   False           False                      False
14                   False           False                      False
15                   False           False                      False
16                   False           False                      False
17                   False           False                      False
18                   False           False                      False
19                   False           False                      False
20                   False           False                      False
21                   False           False                      False
22                   False           False                      False
23                   False           False                      False
24                   False           False                      False
25                   False           False                      False
26                   False           False                      False
27                   False           False                      False
28                   False           False                      False
29                   False           False                      False
...                    ...             ...                        ...
1170                 False           False                      False
1171                 False           False                      False
1172                 False           False                      False
1173                 False           False                      False
1174                 False           False                      False
1175                 False           False                      False
1176                 False           False                      False
1177                 False           False                      False
1178                 False           False                      False
1179                 False           False                      False
1180                 False           False                      False
1181                 False           False                      False
1182                 False           False                      False
1183                 False           False                      False
1184                 False           False                      False
1185                 False           False                      False
1186                 False           False                      False
1187                 False           False                      False
1188                 False           False                      False
1189                 False           False                      False
1190                 False           False                      False
1191                 False           False                      False
```

```
1192                False          False                   False
1193                False          False                   False
1194                False          False                   False
1195                False          False                   False
1196                False          False                   False
1197                False          False                   False
1198                False          False                   False
1199                False          False                   False

      Percent Age 65 and Older  Median Household Income  Percent Unemploye
d  \
0                        False                   False              Fals
e
1                        False                   False              Fals
e
2                        False                   False              Fals
e
3                        False                   False              Fals
e
4                        False                   False              Fals
e
5                        False                   False              Fals
e
6                        False                   False              Fals
e
7                        False                   False              Fals
e
8                        False                   False              Fals
e
9                        False                   False              Fals
e
10                       False                   False              Fals
e
11                       False                   False              Fals
e
12                       False                   False              Fals
e
13                       False                   False              Fals
e
14                       False                   False              Fals
e
15                       False                   False              Fals
e
16                       False                   False              Fals
e
17                       False                   False              Fals
e
18                       False                   False              Fals
e
19                       False                   False              Fals
e
20                       False                   False              Fals
e
21                       False                   False              Fals
e
22                       False                   False              Fals
e
23                       False                   False              Fals
e
24                       False                   False              Fals
e
```

25	False	False	Fals e
26	False	False	Fals e
27	False	False	Fals e
28	False	False	Fals e
29	False	False	Fals e
...	...	...	...
1170	False	False	Fals e
1171	False	False	Fals e
1172	False	False	Fals e
1173	False	False	Fals e
1174	False	False	Fals e
1175	False	False	Fals e
1176	False	False	Fals e
1177	False	False	Fals e
1178	False	False	Fals e
1179	False	False	Fals e
1180	False	False	Fals e
1181	False	False	Fals e
1182	False	False	Fals e
1183	False	False	Fals e
1184	False	False	Fals e
1185	False	False	Fals e
1186	False	False	Fals e
1187	False	False	Fals e
1188	False	False	Fals e
1189	False	False	Fals e
1190	False	False	Fals e
1191	False	False	Fals e
1192	False	False	Fals e
1193	False	False	Fals e
1194	False	False	Fals

```
e
1195                    False                    False               Fals
e
1196                    False                    False               Fals
e
1197                    False                    False               Fals
e
1198                    False                    False               Fals
e
1199                    False                    False               Fals
e

      Percent Less than High School Degree  \
0                                    False
1                                    False
2                                    False
3                                    False
4                                    False
5                                    False
6                                    False
7                                    False
8                                    False
9                                    False
10                                   False
11                                   False
12                                   False
13                                   False
14                                   False
15                                   False
16                                   False
17                                   False
18                                   False
19                                   False
20                                   False
21                                   False
22                                   False
23                                   False
24                                   False
25                                   False
26                                   False
27                                   False
28                                   False
29                                   False
...                                    ...
1170                                 False
1171                                 False
1172                                 False
1173                                 False
1174                                 False
1175                                 False
1176                                 False
1177                                 False
1178                                 False
1179                                 False
1180                                 False
1181                                 False
1182                                 False
1183                                 False
1184                                 False
1185                                 False
1186                                 False
```

1187	False
1188	False
1189	False
1190	False
1191	False
1192	False
1193	False
1194	False
1195	False
1196	False
1197	False
1198	False
1199	False

	Percent Less than Bachelor's Degree	Percent Rural
0	False	False
1	False	False
2	False	False
3	False	False
4	False	False
5	False	False
6	False	False
7	False	False
8	False	False
9	False	False
10	False	False
11	False	False
12	False	False
13	False	False
14	False	False
15	False	False
16	False	False
17	False	False
18	False	False
19	False	False
20	False	False
21	False	False
22	False	False
23	False	False
24	False	False
25	False	False
26	False	False
27	False	False
28	False	False
29	False	False
...	...	...
1170	False	False
1171	False	False
1172	False	False
1173	False	False
1174	False	False
1175	False	False
1176	False	False
1177	False	False
1178	False	False
1179	False	False
1180	False	False
1181	False	False
1182	False	False
1183	False	False
1184	False	False

```
1185                                False          False
1186                                False          False
1187                                False          False
1188                                False          False
1189                                False          False
1190                                False          False
1191                                False          False
1192                                False          False
1193                                False          False
1194                                False          False
1195                                False          False
1196                                False          False
1197                                False          False
1198                                False          False
1199                                False          False

[1200 rows x 19 columns]
State                                   0
County                                  0
Democratic                              5
Republican                              5
FIPS                                    0
Total Population                        0
Citizen Voting-Age Population         680
Percent White, not Hispanic or Latino   0
Percent Black, not Hispanic or Latino  45
Percent Hispanic or Latino              5
Percent Foreign Born                    3
Percent Female                          0
Percent Age 29 and Under                0
Percent Age 65 and Older                0
Median Household Income                 0
Percent Unemployed                      3
Percent Less than High School Degree    0
Percent Less than Bachelor's Degree     0
Percent Rural                          19
dtype: int64
765
```

In [6]:

```
#Task 4-2

# Dropping the column due to a lot of missing values
data_merge = data_merge.drop(columns='Citizen Voting-Age Population')

# Replaces all the 0s with NaN
data_merge = data_merge.replace(0, np.nan)
data_merge = data_merge.interpolate(method='linear')

# Printing out to show that they are no more missing values
print(data_merge.isin([0]).sum())
```

```
State                                  0
County                                 0
Democratic                             0
Republican                             0
FIPS                                   0
Total 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 [7]:

```
#Task 5
def func(row) :
    if row['Democratic'] > row['Republican']:
        val = 1
    else:
        val = 0
    return val

data_merge['Party'] = data_merge.apply(func, axis = 1)
print(data_merge)
```

	State	County	Democratic	Republican	FIPS
0	IN	adams	3146.0	7511.0	18001
1	ND	adams	364.0	796.0	38001
2	NE	adams	3334.0	6487.0	31001
3	OH	adams	2635.0	6000.0	39001
4	PA	adams	14880.0	23419.0	42001
5	WA	adams	1365.0	2867.0	53001
6	WI	adams	4537.0	4854.0	55001
7	VT	addison	11964.0	4439.0	50001
8	MN	aitkin	4118.0	3808.0	27001
9	FL	alachua	74493.0	40599.0	12001
10	WY	albany	7576.0	6366.0	56001
11	VA	albemarle	35701.0	16371.0	51003
12	MI	alcona	1915.0	3541.0	26001
13	VA	alexandria city	53307.0	10734.0	51510
14	MI	alger	1911.0	2162.0	26003
15	NY	allegany	4844.0	8305.0	36003
16	VA	alleghany	1952.0	3433.0	51005
17	PA	allegheny	355907.0	176351.0	42003
18	IN	allen	55903.0	65927.0	18003
19	MI	alpena	5412.0	7380.0	26007
20	VA	amelia	1938.0	3823.0	51007
21	VA	amherst	4541.0	7779.0	51009
22	TX	anderson	3307.0	11335.0	48001
23	ME	androscoggin	22150.0	18931.0	23001
24	TX	angelina	7130.0	19166.0	48005
25	MD	anne arundel	122910.0	92401.0	24003
26	MN	anoka	87756.0	65707.0	27003
27	NE	antelope	396.0	2061.0	31003
28	MI	antrim	4953.0	7629.0	26009
29	AZ	apache	16298.0	7810.0	4001
...	...	...	...	...	...
1170	MN	wilkin	1251.0	1243.0	27167
1171	OH	williams	5272.0	7927.0	39171
1172	VA	williamsburg city	4530.0	1547.0	51830
1173	TN	williamson	42611.0	62039.0	47187
1174	TN	wilson	19447.0	32810.0	47189
1175	TX	wilson	4567.0	13025.0	48493
1176	VA	winchester city	5242.0	3657.0	51840
1177	CT	windham	20490.0	19032.0	9015
1178	VT	windham	14386.0	3673.0	50025
1179	VT	windsor	17354.0	6402.0	50027
1180	TX	winkler	321.0	1123.0	48495
1181	WI	winnebago	40185.0	35282.0	55139
1182	VA	wise	2860.0	7991.0	51195
1183	TX	wood	2635.0	13987.0	48499
1184	WI	wood	15992.0	16899.0	55141
1185	WV	wood	14189.0	13696.0	54107
1186	MD	worcester	9840.0	12886.0	24047
1187	MN	wright	26821.0	30572.0	27171
1188	NY	wyoming	3766.0	7553.0	36121
1189	PA	wyoming	3868.0	6582.0	42131
1190	WV	wyoming	2607.0	3096.0	54109
1191	VA	wythe	2879.0	7669.0	51197
1192	WA	yakima	29476.0	40958.0	53077
1193	AZ	yavapai	40160.0	65308.0	4025
1194	TX	yoakum	335.0	1558.0	48501
1195	ME	york	51387.0	32849.0	23031
1196	NE	york	1281.0	3659.0	31185
1197	PA	york	69272.0	95814.0	42133
1198	TX	young	821.0	5543.0	48503

```
1199    TX             zapata      1392.0       821.0  48505

      Total Population  Percent White, not Hispanic or Latino  \
0                34813                                93.740844
1                 2348                                93.100511
2                31536                                87.338280
3                28111                                96.673900
4               101759                                89.762085
5                19100                                35.670157
6                20294                                90.637627
7                36926                                93.107837
8                15722                                94.122885
9               256581                                62.460198
10               37836                                83.269902
11              104287                                77.424799
12               10461                                96.138037
13              151473                                52.036337
14                9396                                84.514687
15               47700                                94.582809
16               15919                                91.940449
17             1230360                                79.368234
18              365565                                74.923748
19               28929                                96.121539
20               12793                                72.203549
21               31999                                75.611738
22               57772                                60.029080
23              107376                                91.319289
24               87657                                61.672200
25              559737                                70.303375
26              341249                                83.336508
27                6421                                95.421274
28               23215                                95.179841
29               72346                                18.571863
...                ...                                      ...
1170              6479                                94.320111
1171             37270                                93.273410
1172             14988                                68.714972
1173            205645                                85.651973
1174            125616                                86.270857
1175             46444                                58.087159
1176             27349                                67.629529
1177            117078                                83.782606
1178             43609                                93.687083
1179             55894                                95.128279
1180              7723                                37.822090
1181            169487                                89.548461
1182             40074                                91.944902
1183             43198                                83.499236
1184             73621                                93.080779
1185             86262                                95.478890
1186             51441                                80.047044
1187            129922                                93.106633
1188             41239                                89.885788
1189             27975                                95.874888
1190             22537                                97.475263
1191             29171                                94.052312
1192            247681                                44.996588
1193            218586                                81.159361
1194              8316                                34.271284
1195            200536                                94.761539
1196             13842                                91.619708
```

1197	440604	84.725513
1198	18275	79.102599
1199	14335	5.866760

	Percent Black, not Hispanic or Latino	Percent Hispanic or Latino	\
0	0.709505	4.403527	
1	0.894378	0.851789	
2	0.821284	8.907281	
3	0.355733	0.522927	
4	1.352215	6.604821	
5	0.308901	61.884817	
6	3.129004	3.813935	
7	0.974923	2.071711	
8	0.368910	1.183056	
9	19.504562	8.994820	
10	1.345280	9.229305	
11	9.319474	5.592260	
12	0.248542	1.386101	
13	21.068441	16.741598	
14	7.322265	1.351639	
15	1.419287	1.551363	
16	5.207614	1.432251	
17	12.825027	1.890097	
18	11.330953	7.060304	
19	0.508141	1.296277	
20	23.528492	1.188150	
21	17.844308	2.184443	
22	21.197120	16.961504	
23	1.551557	1.733162	
24	14.582977	21.110693	
25	15.559093	7.039377	
26	5.091297	4.057741	
27	0.264756	3.145927	
28	0.323067	1.955632	
29	0.486551	5.947806	
...	...	...	
1170	0.725421	1.265627	
1171	1.078616	4.043467	
1172	14.738457	7.412597	
1173	4.229619	4.625933	
1174	6.652019	3.677875	
1175	1.332788	38.754199	
1176	9.514059	16.486892	
1177	1.866277	10.877364	
1178	0.807173	2.086725	
1179	0.713851	1.452750	
1180	2.563771	58.474686	
1181	1.996613	3.841003	
1182	5.140490	1.230224	
1183	5.037270	9.400898	
1184	0.709037	2.665000	
1185	1.123322	1.032900	
1186	13.603935	3.355300	
1187	1.139145	2.695463	
1188	5.160164	3.205703	
1189	0.868633	1.780161	
1190	0.678884	0.257355	
1191	3.383497	1.100408	
1192	0.740872	47.678667	
1193	0.518331	14.054880	
1194	0.679040	63.552189	

1195	0.839749	1.556828
1196	0.809132	4.681404
1197	5.254605	6.649735
1198	1.258550	17.668947
1199	0.013952	94.119288

	Percent Foreign Born	Percent Female	Percent Age 29 and Under
0	0.901962	50.389222	45.313532
1	1.916525	49.446337	32.282794
2	5.717275	49.242136	41.365424
3	0.640319	50.297037	37.319910
4	3.978027	50.636307	35.814031
5	23.801047	49.130890	51.513089
6	2.774219	46.018528	25.091160
7	4.387153	50.408926	36.359205
8	0.985880	49.014120	26.039944
9	10.144555	51.670623	48.551530
10	6.081510	47.756105	55.447193
11	9.917823	52.107166	39.463212
12	1.720677	49.555492	21.202562
13	27.777888	51.770943	34.734243
14	0.925926	45.008514	28.522776
15	2.301887	49.425577	41.396226
16	1.300333	51.077329	31.660280
17	5.620225	51.800367	36.392438
18	6.367404	51.219619	42.487109
19	1.161464	50.713817	31.684469
20	0.242320	48.080982	34.565778
21	1.712554	52.042251	34.785462
22	6.536038	38.757530	34.862909
23	2.986701	51.044926	36.995232
24	8.041571	51.213252	41.628164
25	8.270491	50.540700	38.635109
26	7.303758	49.959707	38.896524
27	2.071328	49.898770	34.605202
28	2.015938	50.273530	29.450786
29	1.719515	50.598513	45.854643
...	...	...	...
1170	0.478469	49.714462	34.526933
1171	1.333512	50.531258	36.576335
1172	9.067254	56.418468	58.853750
1173	6.826813	51.159522	39.357145
1174	4.134028	51.183766	37.068526
1175	4.484971	49.812678	38.215916
1176	11.645764	51.168233	41.602984
1177	5.269991	50.422795	37.389604
1178	3.107157	50.883992	32.426793
1179	3.594304	50.976849	30.704548
1180	13.712288	48.737537	47.028357
1181	3.318839	49.604394	39.787122
1182	1.170335	48.330588	37.363378
1183	3.433029	50.407426	31.811658
1184	2.326782	50.712433	34.645006
1185	1.001600	51.756277	34.726763
1186	4.593612	51.402578	30.036352
1187	2.880959	49.683656	41.620357
1188	1.804117	45.830403	35.194840
1189	1.276139	49.773012	35.145666
1190	0.257355	50.547988	33.877623
1191	0.901580	51.204964	32.508313
1192	18.119274	49.976381	46.587748

```
1193              6.456955       51.092476                28.717301
1194             25.637326       49.158249                46.284271
1195              3.068277       51.174851                33.230442
1196              1.719405       50.881376                38.289265
1197              3.599150       50.565133                36.839883
1198              5.389877       50.768810                36.990424
1199             29.180328       49.877921                50.931287

      Percent Age 65 and Older  Median Household Income  Percent Unemploye
d  \
0                    14.457243                    47572           5.46263
2
1                    22.274276                    53295           1.97784
8
2                    16.504947                    51201           5.38038
9
3                    16.395717                    34709           9.90799
7
4                    18.255879                    61927           5.52495
5
5                    10.376963                    47554           6.06894
9
6                    27.007983                    43554           9.48949
7
7                    16.974490                    61020           4.96146
4
8                    30.021626                    44524           6.98638
2
9                    12.430772                    44702           7.02015
4
10                    9.549107                    43043           4.57917
4
11                   16.454592                    70342           3.24457
5
12                   34.470892                    38160          10.29729
7
13                   10.017627                    89200           3.98333
8
14                   23.446147                    41270           7.61904
8
15                   16.792453                    44085           8.08327
7
16                   23.902255                    45538           4.56098
6
17                   17.403687                    54357           6.40696
1
18                   13.112032                    49574           7.04808
0
19                   21.625359                    39832           7.84705
3
20                   18.166185                    58269           4.03071
0
21                   18.866215                    47002           5.52695
9
22                   13.707332                    42146           4.34379
0
23                   15.848048                    48728           5.81262
6
24                   15.066680                    44185           8.31150
6
25                   13.405046                    91918           5.53196
```

7			
26	11.835346	73579	4.88194
9			
27	22.192805	46381	1.76361
4			
28	25.410295	48825	7.81578
4			
29	13.322091	32460	15.80743
3			
...	...	...	
...			
1170	18.598549	52963	1.97523
6			
1171	17.480547	45044	5.79099
9			
1172	15.158794	50091	9.07999
4			
1173	11.547570	100140	3.74988
2			
1174	14.487008	63426	5.37716
1			
1175	14.555163	68737	6.20907
7			
1176	15.134009	46466	5.12364
9			
1177	14.571482	60689	8.37552
4			
1178	19.172648	50917	6.04568
2			
1179	20.707053	54763	4.54163
4			
1180	11.070827	54261	4.13873
5			
1181	14.738003	53501	4.42067
8			
1182	15.768329	36352	10.37265
5			
1183	26.297514	47814	9.31908
0			
1184	18.608821	49926	5.66963
9			
1185	18.559737	43944	7.62545
8			
1186	25.475788	57227	7.71983
0			
1187	11.175936	75705	4.38956
7			
1188	15.512015	53612	7.13675
9			
1189	18.373548	53397	7.03230
9			
1190	17.846208	35469	12.07795
8			
1191	19.649652	42888	7.54525
0			
1192	12.748253	45700	8.34621
1			
1193	28.272625	46638	8.52598
6			
1194	11.423761	56655	9.13654
6			

```
1195                17.911996                    59132           5.11030
3
1196                18.841208                    55156           1.43207
9
1197                15.841663                    59853           6.24087
4
1198                19.387141                    46978           4.04804
3
1199                11.726543                    36976          12.63616
6

      Percent Less than High School Degree  \
0                                 15.757807
1                                  4.042806
2                                  9.088224
3                                 21.768031
4                                 12.207748
5                                 33.934880
6                                 12.695857
7                                  7.484449
8                                  8.652110
9                                  7.468059
10                                 4.210167
11                                 8.612515
12                                11.412535
13                                 8.620480
14                                10.614217
15                                10.952741
16                                15.537543
17                                 6.132273
18                                10.798136
19                                 9.607565
20                                18.747199
21                                17.252723
22                                19.827442
23                                10.560153
24                                20.102557
25                                 8.088706
26                                 6.524509
27                                 9.038290
28                                 9.311741
29                                21.758252
...                                    ...
1170                               7.664076
1171                              11.389654
1172                               5.315660
1173                               4.418839
1174                              10.211165
1175                              15.086262
1176                              16.514624
1177                              11.925239
1178                               8.547223
1179                               7.552595
1180                              31.932409
1181                               7.893714
1182                              24.758635
1183                              14.869281
1184                               7.658332
1185                              10.370080
1186                              10.366169
1187                               5.768514
```

1188	11.891069
1189	8.353634
1190	22.638690
1191	17.787306
1192	27.519935
1193	9.830672
1194	28.571429
1195	7.620950
1196	5.744431
1197	11.463427
1198	18.350979
1199	49.673777

	Percent Less than Bachelor's Degree	Percent Rural	Party
0	84.390221	53.674354	0
1	72.473246	100.000000	0
2	77.436983	22.484377	0
3	89.113368	89.026270	0
4	78.133265	53.675782	0
5	86.212165	40.159120	0
6	87.335109	100.000000	0
7	64.225999	78.430787	1
8	84.000985	100.000000	1
9	58.485040	21.193437	1
10	52.136794	11.939723	1
11	48.466284	45.022734	1
12	85.547240	98.921587	0
13	37.911181	83.983239	1
14	81.904762	69.044891	0
15	79.519853	78.739836	0
16	83.711604	52.393846	0
17	60.930419	2.486292	1
18	72.630294	11.879695	0
19	83.224586	51.827826	0
20	85.925594	100.000000	0
21	81.120249	63.663339	0
22	88.231639	67.063533	0
23	79.296169	43.366883	1
24	84.339812	43.082366	0
25	60.634981	5.304135	1
26	71.774386	13.500623	1
27	81.433660	100.000000	0
28	73.314054	100.000000	0
29	88.941063	74.061076	1
...	...	...	...
1170	78.670121	50.136861	1
1171	85.800768	63.628394	0
1172	43.290517	41.507289	1
1173	43.419470	19.386184	0
1174	71.139261	38.467274	0
1175	80.703618	85.861410	0
1176	68.728921	67.811223	1
1177	76.143316	49.761036	1
1178	64.734860	68.245232	1
1179	64.822447	75.588495	1
1180	89.688042	18.129395	0
1181	73.113141	13.431021	1
1182	85.575341	56.670366	0
1183	82.862361	74.199314	0
1184	79.617467	36.682765	0
1185	79.738786	26.774461	1

```
1186                            70.064222       35.526878       0
1187                            73.239335       32.507618       0
1188                            85.003383       64.082552       0
1189                            81.399716       83.455934       0
1190                            91.787560       88.884687       0
1191                            83.409509       75.330939       0
1192                            84.726624       23.519206       0
1193                            74.458362       33.197178       0
1194                            82.183908       37.327072       0
1195                            69.934920       56.843419       1
1196                            71.650858       44.141969       0
1197                            76.884656       24.720442       0
1198                            81.408406       33.568733       0
1199                            91.681305       23.534028       1

[1200 rows x 19 columns]
```

In [8]:

```
#Task 6
data_merge.groupby('Party', as_index = False)['Total Population'].mean()
democratic = data_merge[data_merge['Party'] == 1]
republican = data_merge[data_merge['Party'] == 0]

print(democratic.mean())
print(republican.mean())

#required measures
[statistic, pvalue] = st.ttest_ind(democratic['Total Population'],republican['Total Pop
ulation'], equal_var= False)

print([statistic,pvalue])
statistic, pvalue/2
```

```
Democratic                                 71193.172308
Republican                                 41322.861538
FIPS                                       37130.873846
Total Population                          300998.316923
Percent White, not Hispanic or Latino         69.683766
Percent Black, not Hispanic or Latino          9.269702
Percent Hispanic or Latino                    12.587391
Percent Foreign Born                           7.986330
Percent Female                                50.385433
Percent Age 29 and Under                      38.726959
Percent Age 65 and Older                      16.194826
Median Household Income                   53798.732308
Percent Unemployed                             6.908426
Percent Less than High School Degree          11.883760
Percent Less than Bachelor's Degree           71.968225
Percent Rural                                 39.118070
Party                                          1.000000
dtype: float64
Democratic                                  7969.218286
Republican                                 12685.142286
FIPS                                       38755.305143
Total Population                           53974.214857
Percent White, not Hispanic or Latino         82.597026
Percent Black, not Hispanic or Latino          4.454970
Percent Hispanic or Latino                     9.847969
Percent Foreign Born                           4.017041
Percent Female                                49.617156
Percent Age 29 and Under                      36.020984
Percent Age 65 and Older                      18.814997
Median Household Income                   48724.150857
Percent Unemployed                             6.425490
Percent Less than High School Degree          14.029195
Percent Less than Bachelor's Degree           81.103128
Percent Rural                                 63.431458
Party                                          0.000000
dtype: float64
[8.001207114045041, 2.0965719353509958e-14]
```

Out[8]:

```
(8.001207114045041, 1.0482859676754979e-14)
```

In [9]:

```
#Task 7
[statistic, pvalue] = st.ttest_ind(democratic['Median Household Income'],republican['Median Household Income'], equal_var= False)

print(democratic.mean())
print(republican.mean())

print([statistic,pvalue])
statistic, pvalue/2
```

```
Democratic                                 71193.172308
Republican                                 41322.861538
FIPS                                       37130.873846
Total Population                          300998.316923
Percent White, not Hispanic or Latino         69.683766
Percent Black, not Hispanic or Latino          9.269702
Percent Hispanic or Latino                    12.587391
Percent Foreign Born                           7.986330
Percent Female                                50.385433
Percent Age 29 and Under                      38.726959
Percent Age 65 and Older                      16.194826
Median Household Income                    53798.732308
Percent Unemployed                             6.908426
Percent Less than High School Degree          11.883760
Percent Less than Bachelor's Degree           71.968225
Percent Rural                                 39.118070
Party                                          1.000000
dtype: float64
Democratic                                  7969.218286
Republican                                 12685.142286
FIPS                                       38755.305143
Total Population                           53974.214857
Percent White, not Hispanic or Latino         82.597026
Percent Black, not Hispanic or Latino          4.454970
Percent Hispanic or Latino                     9.847969
Percent Foreign Born                           4.017041
Percent Female                                49.617156
Percent Age 29 and Under                      36.020984
Percent Age 65 and Older                      18.814997
Median Household Income                    48724.150857
Percent Unemployed                             6.425490
Percent Less than High School Degree          14.029195
Percent Less than Bachelor's Degree           81.103128
Percent Rural                                 63.431458
Party                                          0.000000
dtype: float64
[5.507012409466501, 6.173239891230373e-08]
```

Out[9]:

```
(5.507012409466501, 3.0866199456151866e-08)
```

In [10]:

```
#Task 8-1
data_merge['Party'].value_counts(normalize=True)
```

Out[10]:

```
0    0.729167
1    0.270833
Name: Party, dtype: float64
```

In [11]:

```
#Task 8-2
#Printing out the averages for each party
print(data_merge[data_merge['Party']==0].mean()) #Republican
print(" ")
print(data_merge[data_merge['Party']==1].mean()) #Democratic
```

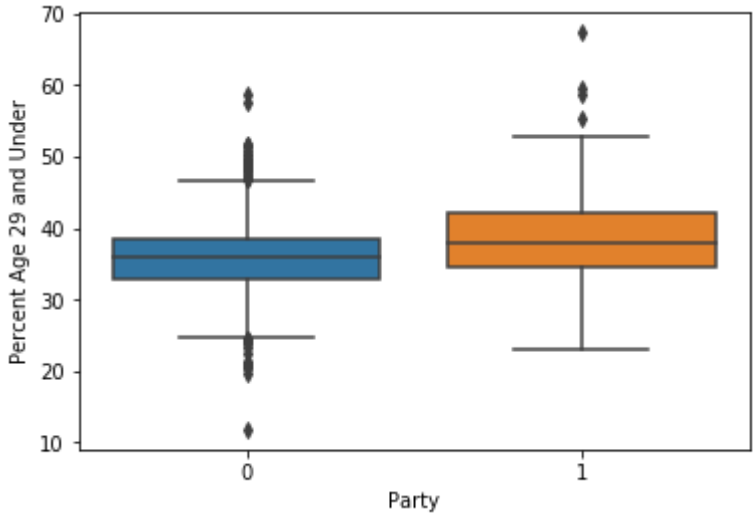
```
Democratic                                 7969.218286
Republican                                12685.142286
FIPS                                      38755.305143
Total Population                          53974.214857
Percent White, not Hispanic or Latino        82.597026
Percent Black, not Hispanic or Latino         4.454970
Percent Hispanic or Latino                    9.847969
Percent Foreign Born                          4.017041
Percent Female                               49.617156
Percent Age 29 and Under                     36.020984
Percent Age 65 and Older                     18.814997
Median Household Income                   48724.150857
Percent Unemployed                            6.425490
Percent Less than High School Degree         14.029195
Percent Less than Bachelor's Degree          81.103128
Percent Rural                                63.431458
Party                                         0.000000
dtype: float64

Democratic                                 71193.172308
Republican                                 41322.861538
FIPS                                       37130.873846
Total Population                          300998.316923
Percent White, not Hispanic or Latino         69.683766
Percent Black, not Hispanic or Latino          9.269702
Percent Hispanic or Latino                    12.587391
Percent Foreign Born                           7.986330
Percent Female                                50.385433
Percent Age 29 and Under                      38.726959
Percent Age 65 and Older                      16.194826
Median Household Income                   53798.732308
Percent Unemployed                             6.908426
Percent Less than High School Degree          11.883760
Percent Less than Bachelor's Degree           71.968225
Percent Rural                                 39.118070
Party                                          1.000000
dtype: float64
```

In [12]:

```
#Task 8-3
sns.boxplot(x='Party', y='Percent Age 29 and Under', data=data_merge)
```

Out[12]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1b24bffef60>
```

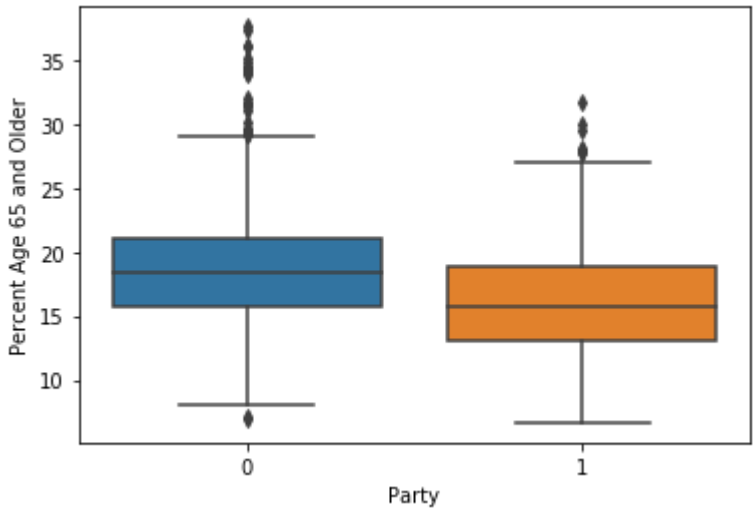


In [13]:

```
#Task 8-4
sns.boxplot(x='Party', y='Percent Age 65 and Older', data=data_merge)
```

Out[13]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1b24c0e6d68>
```

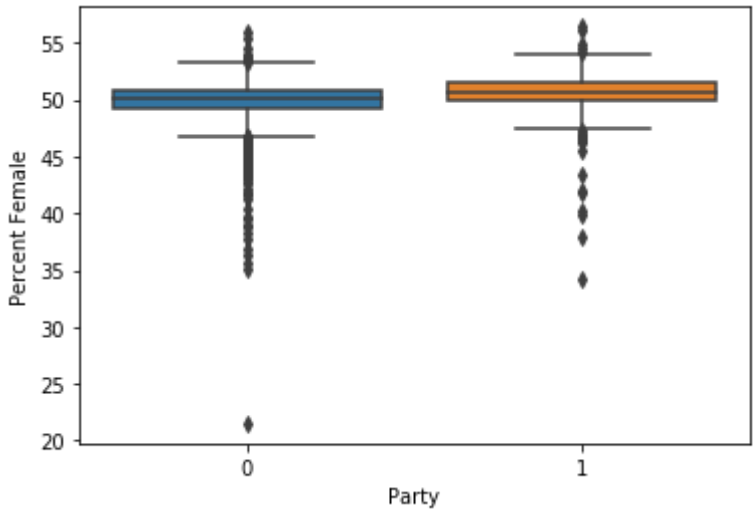


In [14]:

```
#Task 8-5
sns.boxplot(x='Party', y='Percent Female', data=data_merge)
```

Out[14]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1b24c170e10>
```

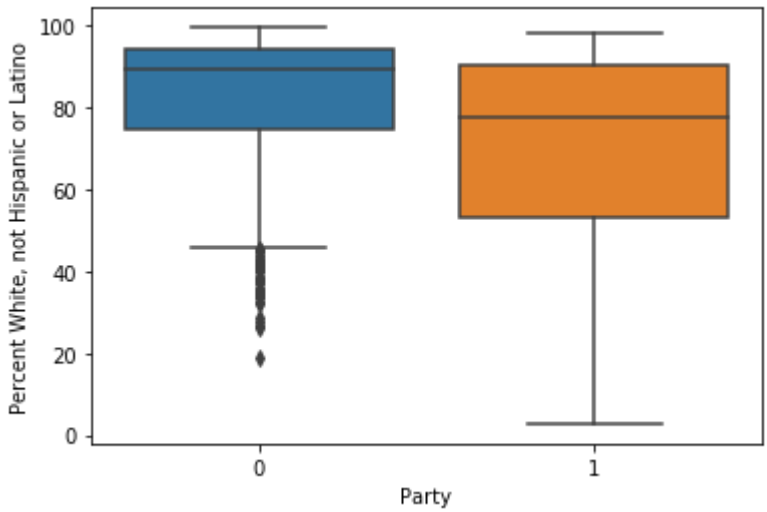


In [15]:

```
#Task 8-6
sns.boxplot(x='Party', y='Percent White, not Hispanic or Latino', data=data_merge)
```

Out[15]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1b24c1e1828>
```

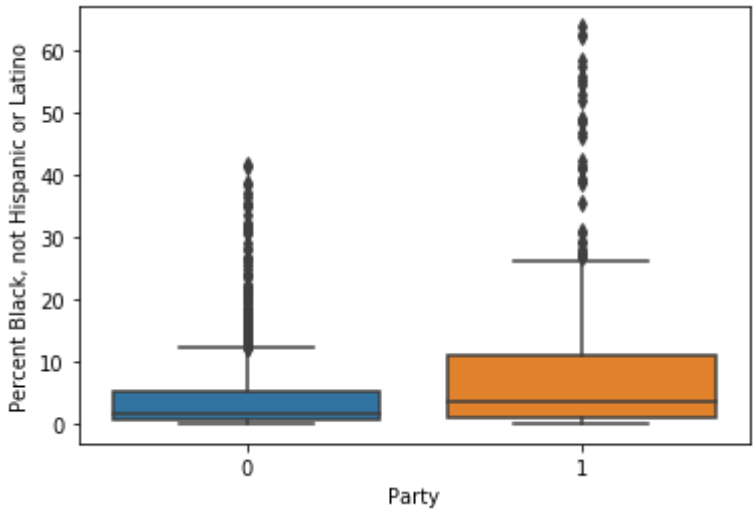


In [16]:

```
#Task 8-7
sns.boxplot(x='Party', y='Percent Black, not Hispanic or Latino', data=data_merge)
```

Out[16]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1b24c2649b0>
```

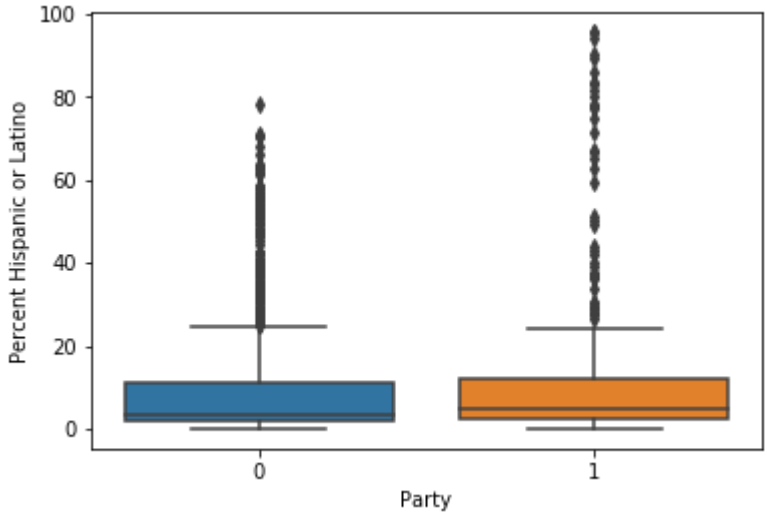


In [17]:

```
#Task 8-8
sns.boxplot(x='Party', y='Percent Hispanic or Latino', data=data_merge)
```

Out[17]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1b24c2be4e0>
```

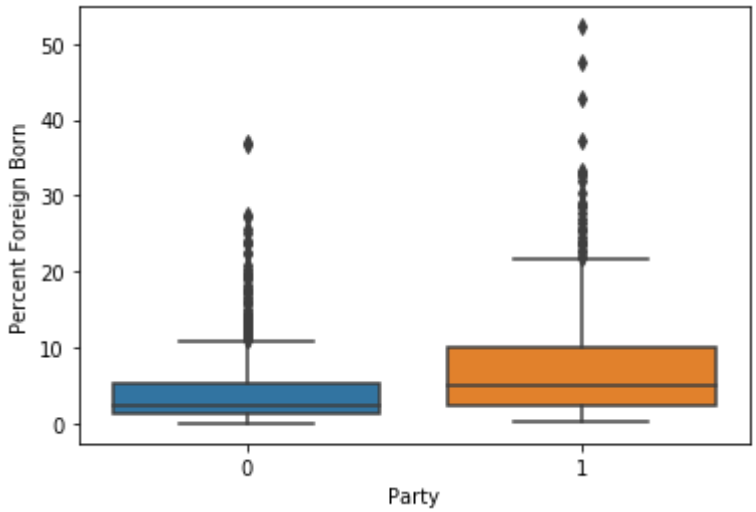


In [18]:

```
#Task 8-9
sns.boxplot(x='Party', y='Percent Foreign Born', data=data_merge)
```

Out[18]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1b24c351630>
```

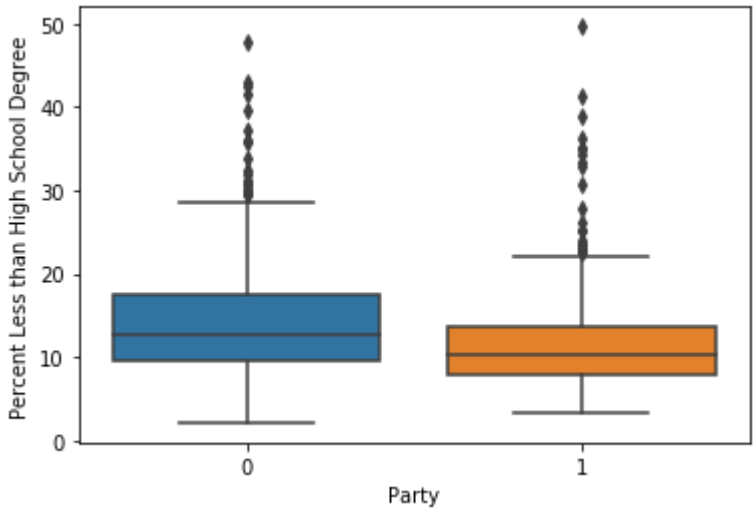


In [19]:

```
#Task 8-10
sns.boxplot(x='Party', y='Percent Less than High School Degree', data=data_merge)
```

Out[19]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1b24c3cd278>
```

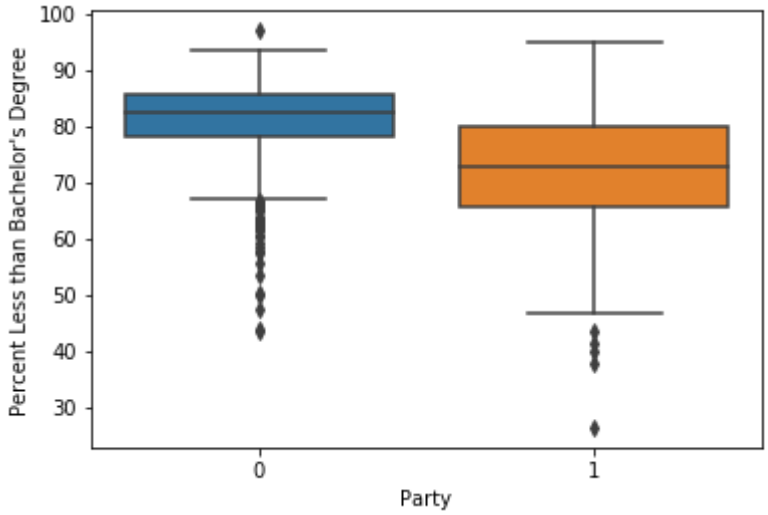


In [20]:

```
#Task 8-11
sns.boxplot(x='Party', y="Percent Less than Bachelor's Degree", data=data_merge)
```

Out[20]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1b24c435400>
```



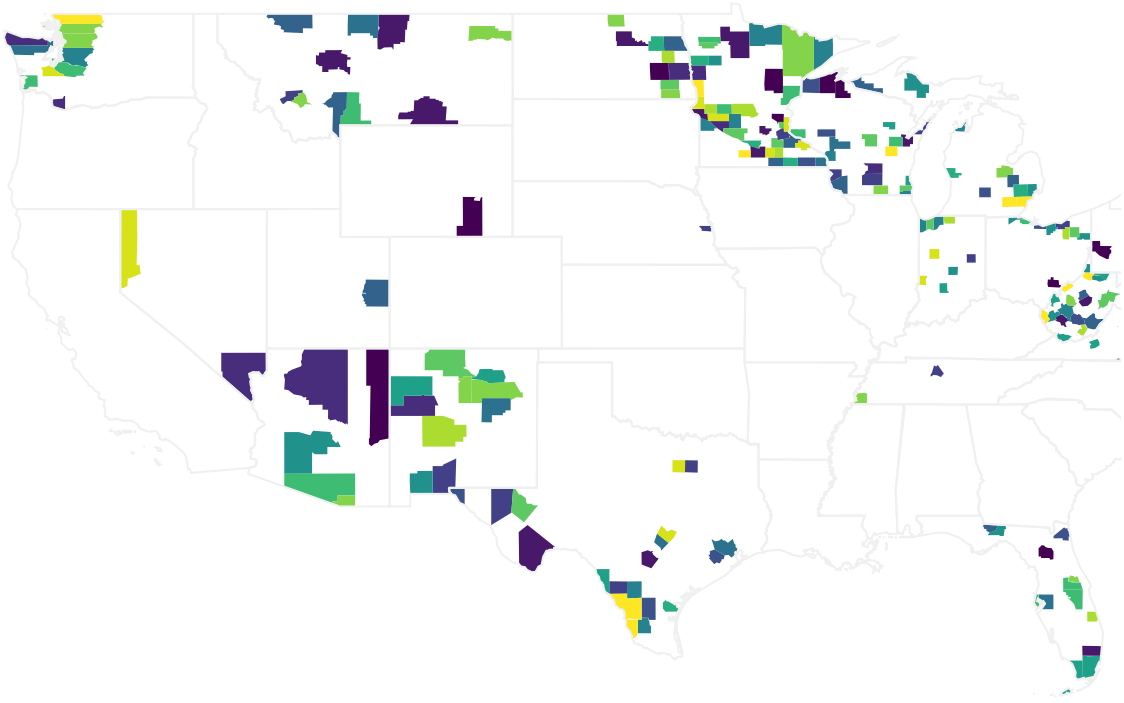
In [21]:

```
#Task 9
```

In [22]:

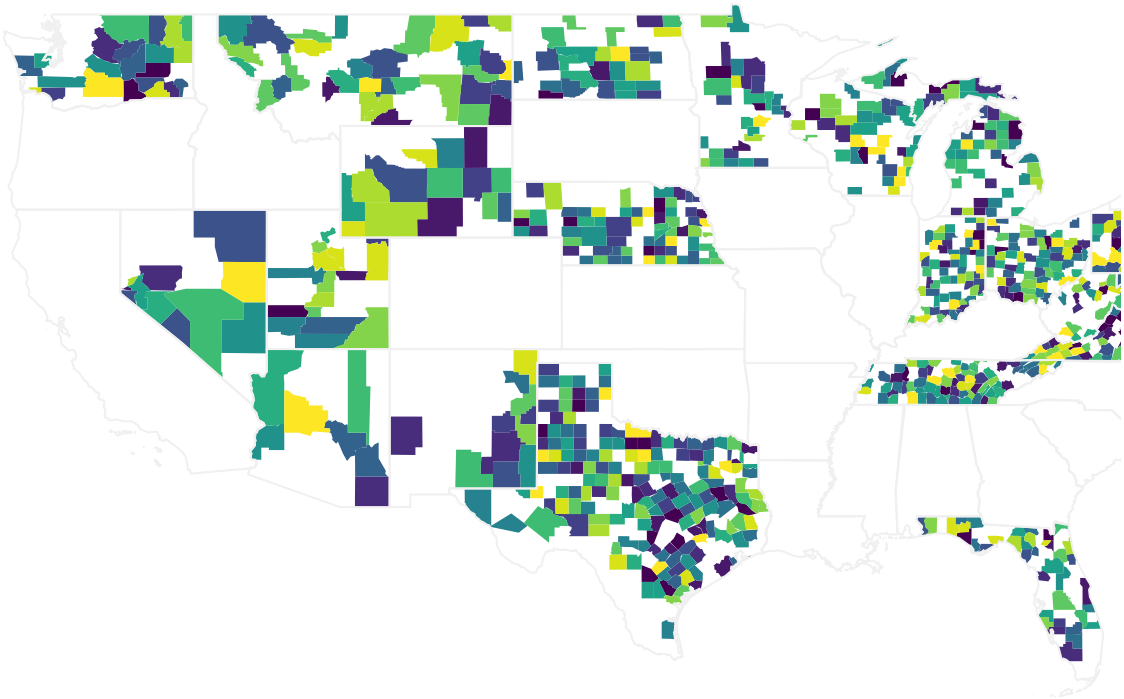
```
#Task 10-1 Democratic Counties
fips = democratic['FIPS']
values = range(len(fips))

fig = ff.create_choropleth(fips=fips, values=values)
fig.layout.template = None
fig.show()
```

```
C:\Users\Mehul\Anaconda3\lib\site-packages\pandas\core\frame.py:6692: Futu
reWarning:

Sorting because non-concatenation axis is not aligned. A future version
of pandas will change to not sort by default.

To accept the future behavior, pass 'sort=False'.

To retain the current behavior and silence the warning, pass 'sort=True'.
```



In [23]:

```
#Task 10-2-Republican Counties
fips = republican['FIPS']
values = range(len(fips))

fig = ff.create_choropleth(fips=fips, values=values)
fig.layout.template = None
fig.show()
```



In [25]:

```
#Task 10-3
fips = data_merge['FIPS'].tolist()
values = data_merge['Party'].tolist()

colorscale = ['#f54242', '#4287f5']

fig = ff.create_choropleth(fips=fips, values=values, scope=['usa','HI'],
                           title_text = 'Democratic vs Republican Counties',
                           colorscale=colorscale,
                           state_outline={'color': 'rgb(0,0,0)', 'width': 1},
                           county_outline={'color': 'rgb(0,0,0)', 'width': 0.5},
                           legend_title='Party')
fig.layout.template = None
fig.show()
```