Question 3

In [70]:
```
import plotly.express as px
import plotly.figure_factory as ff
import pandas as pd
from urllib.request import urlopen
import json

# Load a GeoJSON file that contains US counties' shapes and locations.
# In this GeoJSON file, the location property is "id", which stores th
with urlopen('https://raw.githubusercontent.com/plotly/datasets/master
    counties = json.load(response)

scope = ['Georgia']
df = pd.read_csv('/Users/shilp/Downloads/Project3_2/election-context-2
```

In [71]:
```
df.head(5)
```

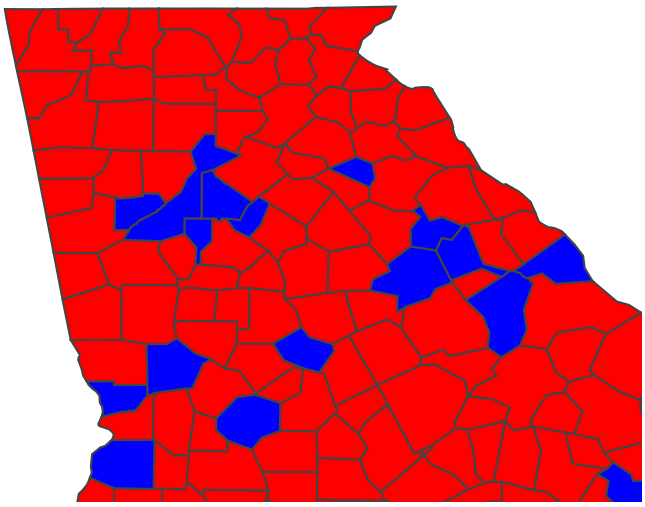
Out[71]:

	state	county	fips	trump16	clinton16	otherpres16	romney12	obama12	otherpres12
0	Alabama	Autauga	1001	18172	5936	865	17379	6363	190
1	Alabama	Baldwin	1003	72883	18458	3874	66016	18424	898
2	Alabama	Barbour	1005	5454	4871	144	5550	5912	47
3	Alabama	Bibb	1007	6738	1874	207	6132	2202	86
4	Alabama	Blount	1009	22859	2156	573	20757	2970	279

5 rows × 39 columns

In [72]:
```
data = pd.read_csv('/Users/shilp/Downloads/Project3_2/election-context
with urlopen('https://raw.githubusercontent.com/plotly/datasets/master
    counties = json.load(response)
data['winning'] = data[['demsen16', 'repsen16', 'othersen16']].idxmax(
fig = px.choropleth(data[data['state'] == 'Georgia'], geojson=counties
fig.update_geos(fitbounds="locations", visible=False)
fig.update_layout(title='Senate representatives')
fig.show()
```

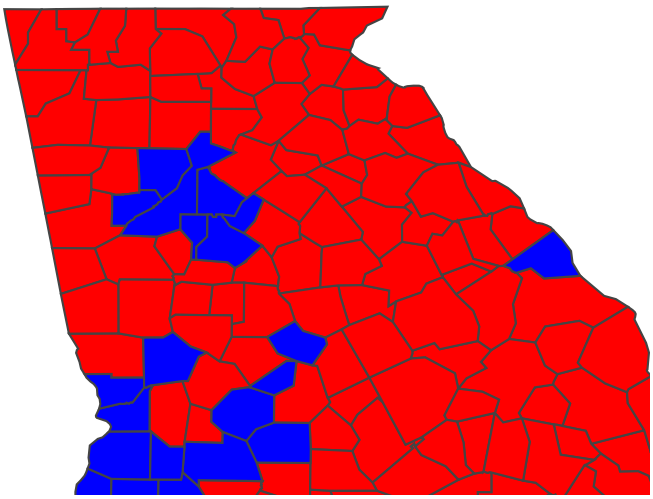
Senate representatives



Question 3 II

In [73]:
```
data['winning'] = data[['demhouse16', 'rephouse16', 'otherhouse16']].i
fig = px.choropleth(data[data['state'] == 'Georgia'], geojson=counties
fig.update_geos(fitbounds="locations", visible=False)
fig.update_layout(title='House representatives')
fig.show()
```

House representatives



Question 4

In [60]:
```
import plotly.figure_factory as ff

import numpy as np
import pandas as pd
dx = pd.read_csv('/Users/shilp/Downloads/Project3_2/1976-2018-senate.c
df = dx[dx['year']==2018]
df.head(5)
```

Out[60]:

	year	state	state_po	state_fips	state_cen	state_ic	office	district	stage	speci
3269	2018	Arizona	AZ	4	86	61	US Senate	statewide	gen	Fals
3270	2018	Arizona	AZ	4	86	61	US Senate	statewide	gen	Fals
3271	2018	Arizona	AZ	4	86	61	US Senate	statewide	gen	Fals
3272	2018	Arizona	AZ	4	86	61	US Senate	statewide	gen	Fals
3273	2018	California	CA	6	93	71	US Senate	statewide	gen	Fals

In [61]:
```
df["party"] = df["party"].astype('category')
df.dtypes
```

Out[61]:
```
year                int64
state              object
state_po           object
state_fips          int64
state_cen           int64
state_ic            int64
office             object
district           object
stage              object
special              bool
candidate          object
party            category
writein              bool
mode               object
candidatevotes      int64
totalvotes          int64
unofficial           bool
version           float64
dtype: object
```

In [62]:
```
df["type_party"] = df["party"].cat.codes
df.head()
```

Out[62]:

	year	state	state_po	state_fips	state_cen	state_ic	office	district	stage	speci
3269	2018	Arizona	AZ	4	86	61	US Senate	statewide	gen	Fals
3270	2018	Arizona	AZ	4	86	61	US Senate	statewide	gen	Fals
3271	2018	Arizona	AZ	4	86	61	US Senate	statewide	gen	Fals
3272	2018	Arizona	AZ	4	86	61	US Senate	statewide	gen	Fals
3273	2018	California	CA	6	93	71	US Senate	statewide	gen	Fals

In [63]:
```
df.dtypes
```

Out[63]:
```
year                int64
state              object
state_po           object
state_fips          int64
state_cen           int64
state_ic            int64
office             object
district           object
stage              object
special              bool
candidate          object
party            category
writein              bool
mode               object
candidatevotes      int64
totalvotes          int64
unofficial           bool
version           float64
type_party           int8
dtype: object
```

In [64]:
```
df["percentage"] = (df["candidatevotes"]/df["totalvotes"]) *100
print(df)
```

```
      year       state state_po  state_fips  state_cen  state_ic
office  \
3269  2018     Arizona       AZ           4         86        61  US
Senate
```

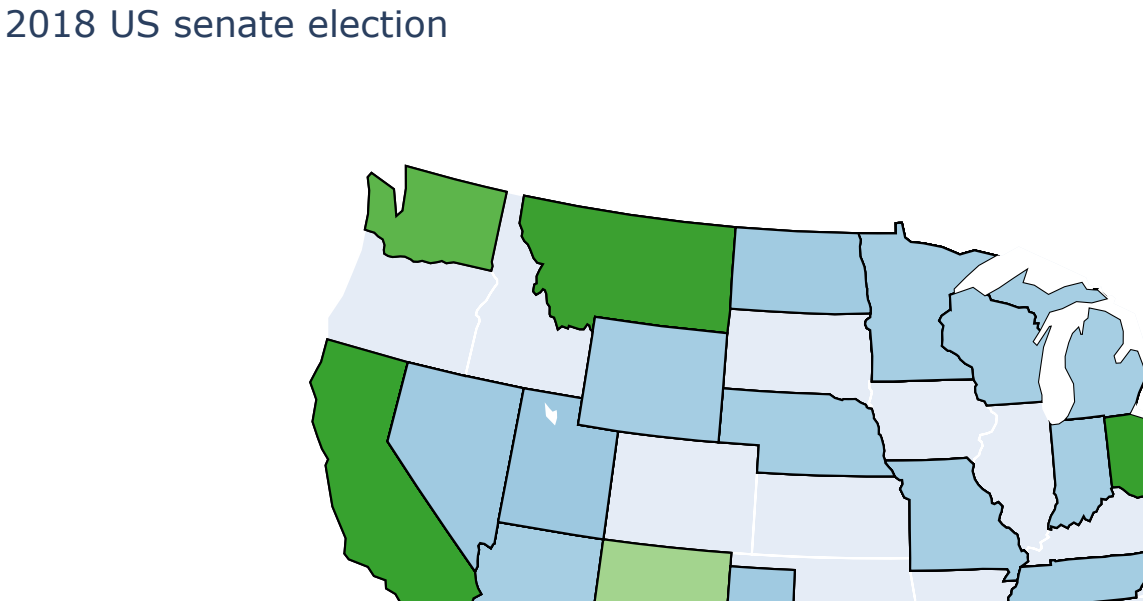
```
3270  2018     Arizona       AZ           4         86        61  US
Senate
3271  2018     Arizona       AZ           4         86        61  US
Senate
3272  2018     Arizona       AZ           4         86        61  US
Senate
3273  2018  California       CA           6         93        71  US
Senate
...    ...         ...      ...         ...       ...       ...
...
3416  2018   Wisconsin       WI          55         35        25  US
Senate
3417  2018     Wyoming       WY          56         83        68  US
Senate
3418  2018     Wyoming       WY          56         83        68  US
Senate
3419  2018     Wyoming       WY          56         83        68  US
Senate
3420  2018     Wyoming       WY          56         83        68  US
Senate

       district stage  special         candidate        party  writei
n   mode  \
3269  statewide   gen    False    Martha McSally   republican    Fals
e  total
3270  statewide   gen    False    Kyrsten Sinema     democrat    Fals
e  total
3271  statewide   gen    False      Angela Green        green    Fals
e  total
3272  statewide   gen    False               NaN          NaN     Tru
e  total
3273  statewide   gen    False  Dianne Feinstein     democrat    Fals
e  total
...         ...   ...      ...               ...          ...      ..
.    ...
3416  statewide   gen    False               NaN          NaN     Tru
e  total
3417  statewide   gen    False     John Barrasso   republican    Fals
e  total
3418  statewide   gen    False      Gary Trauner     democrat    Fals
e  total
3419  statewide   gen    False    Joseph Porambo  libertarian    Fals
e  total
3420  statewide   gen    False               NaN          NaN     Tru
e  total

      candidatevotes  totalvotes  unofficial     version  type_party
\
3269         1135200     2384308       False  20190110.0          19
3270         1191100     2384308       False  20190110.0           2
```

```
3271          57442     2384308       False  20190110.0          7
3272            566     2384308       False  20190110.0         -1
3273        6019422    11113364       False  20190110.0          2
...             ...         ...         ...         ...        ...
3416             42     2657841       False  20190110.0         -1
3417         136210      203420       False  20190110.0         19
3418          61227      203420       False  20190110.0          2
3419           5658      203420       False  20190110.0         12
3420            325      203420       False  20190110.0         -1

      percentage
3269   47.611299
3270   49.955794
3271    2.409169
3272    0.023739
3273   54.163816
...          ...
3416    0.001580
3417   66.959984
3418   30.098810
3419    2.781437
3420    0.159768

[152 rows x 20 columns]
```

In [68]:
```
import plotly.graph_objects as go
```

In [69]:
```
for col in df.columns:
    df[col] = df[col].astype(str)

df['text'] = df['state_po'] + '<br>' + \
    'winning candidate's name ' + df['candidate'] + '<br>' + \
    'party affiliation ' + df['party']

fig = go.Figure(data=go.Choropleth(
    locations=df['state_po'],

    z=df['percentage'].astype(float),

    locationmode='USA-states',
    colorscale=[[0, "rgb(166,206,227)"],
                [0.25, "rgb(31,120,180)"],
                [0.45, "rgb(178,223,138)"],
                [0.65, "rgb(51,160,44)"],
                [0.85, "rgb(251,154,153)"],
                [1, "rgb(227,26,28)"]],

    autocolorscale=False,
    text=df['text'], # hover text
    marker_line_color='black', # line markers between states
```

```
    marker_line_color= 'black' , # line markers between states
    colorbar_title="vote percentage",
    #color = df['percentage'],
))

fig.update_layout(
    title_text='2018 US senate election',
    geo = dict(
        scope='usa',
        projection=go.layout.geo.Projection(type = 'albers usa'),
        showlakes=True, # lakes
        lakecolor='rgb(255, 255, 255)'),

)

fig.show()
```



In [ ]:
```
Question 5
```

In [35]:
```
import plotly.graph_objects as go

import pandas as pd

df = pd.read_csv('/Users/shilp/Downloads/Project3_2/1962_2006_walmart_
df.head(5)
df_filtered = df.query('YEAR>=2000')
df= df_filtered
df['text'] = df['type_store'] + '' + df['STRCITY'] + ', ' + df['STRSTA
df.head(5)
```

Out[35]:

	storenum	OPENDATE	date_super	conversion	st	county	STREETADDR	STRCITY	S
2352	2815	5/17/00	NaN	NaN	17	31	1460 GOLF ROAD	Rolling Meadows	
2353	2988	10/11/00	10/11/00	1.0	13	295	2625 NO. HWY 27	LaFayette	
2354	2609	3/15/00	NaN	NaN	6	37	2770 CARSON STREET	Lakewood	
2355	2548	7/19/00	7/19/00	0.0	21	195	28402 U.S. HWY 119 NO	South Williamson	
2356	2365	1/26/00	1/26/00	1.0	42	89	500 ROUTE 940	Mount Pocono	

In [36]:
```
df["type_store"] = df["type_store"].astype('category')
df.dtypes
```

Out[36]:
```
storenum         int64
OPENDATE        object
date_super      object
conversion     float64
st               int64
county           int64
STREETADDR      object
STRCITY         object
STRSTATE        object
ZIPCODE          int64
type_store    category
LAT            float64
LON            float64
MONTH            int64
DAY              int64
YEAR             int64
text            object
dtype: object
```

In [37]:
```
df["type_store_cat"] = df["type_store"].cat.codes
df.head()
```

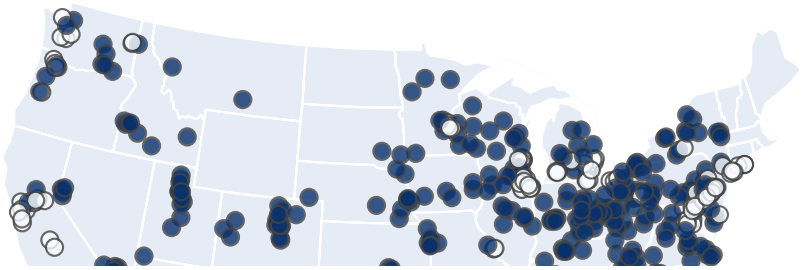
Out[37]:

	storenum	OPENDATE	date_super	conversion	st	county	STREETADDR	STRCITY	S
2352	2815	5/17/00	NaN	NaN	17	31	1460 GOLF ROAD	Rolling Meadows	
2353	2988	10/11/00	10/11/00	1.0	13	295	2625 NO. HWY 27	LaFayette	
2354	2609	3/15/00	NaN	NaN	6	37	2770 CARSON STREET	Lakewood	
2355	2548	7/19/00	7/19/00	0.0	21	195	28402 U.S. HWY 119 NO	South Williamson	
2356	2365	1/26/00	1/26/00	1.0	42	89	500 ROUTE 940	Mount Pocono	

In [38]:
```
fig = go.Figure(data=go.Scattergeo(
        lon = df['LON'],
        lat = df['LAT'],
        text = df['text'],
        mode = 'markers',
     #   marker_color = df['type_store_cat'],
    #marker=dict(size=10, color='green'),
    #                    legendgroup='Buy', showlegend=True, name='Buy'
    marker = dict(
            size = 8,
            opacity = 0.8,
            reversescale = True,
            autocolorscale = False,
            symbol = 'circle',
            line = dict(
                width=1,
                color='rgba(102, 102, 102)'
            ),
            colorscale = 'Blues',
           cmin = 0,
            color = df['type_store_cat'],
        #color='green'
            cmax = df['type_store_cat'].max(),
           colorbar_title="Store Type 1= Wal-Mart 0= Supercenter"
        )))

fig.update_layout(
        title = 'Walmart store opened since 2000',
        geo_scope='usa',
```

```
    #legend_title='<b> Store Type </b>',
    #showlegend=True,

     )
fig.show()
#1= Wal-Mart
#0= Supercenter
```

Walmart store opened since 2000



In [ ]: