data_clean_up

December 13, 2020

[3]:
```
import pandas as pd
import numpy as np
```

[82]:
```
parse_dates = ['date']

covid_us_county = pd.read_csv("covid_us_county.csv", parse_dates=parse_dates,␣
↪index_col=False)

covid_us_county = covid_us_county.drop('lat', 1)
covid_us_county = covid_us_county.drop('long', 1)

covid_us_county
```

[82]:

	state	state_code	county	date	cases	deaths
0	Alabama	AL	Autauga	2020-01-22	0	0
1	Alabama	AL	Autauga	2020-01-23	0	0
2	Alabama	AL	Autauga	2020-01-24	0	0
3	Alabama	AL	Autauga	2020-01-25	0	0
4	Alabama	AL	Autauga	2020-01-26	0	0
...	...	...	...	...	...	...
1048570	Wisconsin	WI	Buffalo	2020-05-13	5	1
1048571	Wisconsin	WI	Buffalo	2020-05-14	5	1
1048572	Wisconsin	WI	Buffalo	2020-05-15	5	1
1048573	Wisconsin	WI	Buffalo	2020-05-16	5	1
1048574	Wisconsin	WI	Buffalo	2020-05-17	5	1

[1048575 rows x 6 columns]

[83]:
```
covid_us_states = covid_us_county.groupby(['state','state_code','date'],␣
↪as_index=False).agg('sum').reset_index()
covid_us_states = covid_us_states.drop('index', 1)
covid_us_states.to_csv("covid_us_states.csv", index=False)
covid_us_states
```

[83]:

	state	state_code	date	cases	deaths
0	Alabama	AL	2020-01-22	0	0
1	Alabama	AL	2020-01-23	0	0
2	Alabama	AL	2020-01-24	0	0

```
3        Alabama          AL 2020-01-25      0       0
4        Alabama          AL 2020-01-26      0       0
...          ...         ...        ...    ...     ...
15822  Wisconsin          WI 2020-12-05  31005     234
15823  Wisconsin          WI 2020-12-06  31177     242
15824  Wisconsin          WI 2020-12-07  31206     243
15825  Wisconsin          WI 2020-12-08  31494     242
15826  Wisconsin          WI 2020-12-09  31796     241

[15827 rows x 5 columns]
```

[ ]:

[227]:
```
parse_dates2 = ['date', 'lastUpdateEt', 'dateModified',␣
 ↪'checkTimeEt','dateChecked']

us_states_covid19_daily = pd.read_csv("us_states_covid19_daily.csv",␣
 ↪parse_dates=parse_dates2, index_col=False)
us_states_covid19_daily.head()
```

[227]:
```
        date state  positive  probableCases   negative  pending  \
0 2020-12-06    AK   35720.0            NaN  1042056.0      NaN
1 2020-12-06    AL  269877.0        45962.0  1421126.0      NaN
2 2020-12-06    AR  170924.0        22753.0  1614979.0      NaN
3 2020-12-06    AS       0.0            NaN     2140.0      NaN
4 2020-12-06    AZ  364276.0        12590.0  2018813.0      NaN

  totalTestResultsSource  totalTestResults  hospitalizedCurrently  \
0       totalTestsViral         1077776.0                  164.0
1  totalTestsPeopleViral         1645041.0                 1927.0
2       totalTestsViral         1763150.0                 1076.0
3       totalTestsViral            2140.0                    NaN
4  totalTestsPeopleViral         2370499.0                 2977.0

   hospitalizedCumulative  ...   posNeg  deathIncrease  hospitalizedIncrease  \
0                   799.0  ...  1077776              0                     0
1                 26331.0  ...  1691003             12                     0
2                  9401.0  ...  1785903             40                    21
3                     NaN  ...     2140              0                     0
4                 28248.0  ...  2383089             25                   242

                                   hash  commercialScore  \
0  7b1d31e2756687bb9259b29195f1db6cdb321ea6                0
1  19454ed8fe28fc0a7948fc0771b2f3c846c1c92e                0
2  25fc83bffff5b32ba1a737be8e087fad9f4fde33                0
3  8c39eec317586b0c34fc2903e6a3891ecb00469e                0
4  7cf59da9e4bc31d905e179211313d08879880a85                0
```

```
  negativeRegularScore negativeScore positiveScore score  grade
0                    0             0             0     0    NaN
1                    0             0             0     0    NaN
2                    0             0             0     0    NaN
3                    0             0             0     0    NaN
4                    0             0             0     0    NaN

[5 rows x 55 columns]
```

[228]:
```
cols = ['date', 'state','positive','probableCases', 'negative',
        'totalTestResults','death','hospitalized', 'total']

us_states_covid19_daily_simple = us_states_covid19_daily[cols] #.fillna(0)
us_states_covid19_daily_simple
```

[228]:
```
            date state  positive  probableCases   negative  totalTestResults  \
0     2020-12-06    AK   35720.0            NaN  1042056.0         1077776.0
1     2020-12-06    AL  269877.0        45962.0  1421126.0         1645041.0
2     2020-12-06    AR  170924.0        22753.0  1614979.0         1763150.0
3     2020-12-06    AS       0.0            NaN     2140.0            2140.0
4     2020-12-06    AZ  364276.0        12590.0  2018813.0         2370499.0
...          ...   ...       ...            ...        ...               ...
15628 2020-01-24    WA       0.0            NaN        0.0               0.0
15629 2020-01-23    MA       NaN            NaN        NaN               2.0
15630 2020-01-23    WA       0.0            NaN        0.0               0.0
15631 2020-01-22    MA       NaN            NaN        NaN               1.0
15632 2020-01-22    WA       0.0            NaN        0.0               0.0

        death  hospitalized    total
0       143.0         799.0  1077776
1      3889.0       26331.0  1691003
2      2660.0        9401.0  1785903
3         0.0           NaN     2140
4      6950.0       28248.0  2383089
...       ...           ...      ...
15628     NaN           NaN        0
15629     NaN           NaN        0
15630     NaN           NaN        0
15631     NaN           NaN        0
15632     NaN           NaN        0

[15633 rows x 9 columns]
```

[229]:
```
covid_us_states
```

[229]:
```
           state state_code       date  cases  deaths
0        Alabama         AL 2020-01-22      0       0
1        Alabama         AL 2020-01-23      0       0
2        Alabama         AL 2020-01-24      0       0
```

```
3        Alabama          AL 2020-01-25      0       0
4        Alabama          AL 2020-01-26      0       0
...          ...         ...        ...    ...     ...
15822  Wisconsin          WI 2020-12-05  31005     234
15823  Wisconsin          WI 2020-12-06  31177     242
15824  Wisconsin          WI 2020-12-07  31206     243
15825  Wisconsin          WI 2020-12-08  31494     242
15826  Wisconsin          WI 2020-12-09  31796     241

[15827 rows x 5 columns]
```

[261]:
```
covid_tigeryi = pd.merge(covid_us_states, us_states_covid19_daily_simple,␣
↪how='inner',
                      left_on=['state_code','date'], right_on =␣
↪['state','date'])


covid_tigeryi = covid_tigeryi.rename({'state_x': 'state'}, axis=1)
covid_tigeryi = covid_tigeryi.drop('state_y', 1)
covid_tigeryi = covid_tigeryi.drop('deaths', 1)
covid_tigeryi
```

[261]:
```
           state state_code       date  cases  positive  probableCases  \
0        Alabama          AL 2020-03-07      0       0.0            NaN
1        Alabama          AL 2020-03-08      0       0.0            NaN
2        Alabama          AL 2020-03-09      0       0.0            NaN
3        Alabama          AL 2020-03-10      0       0.0            NaN
4        Alabama          AL 2020-03-11      3       0.0            NaN
...          ...         ...        ...    ...       ...            ...
13740  Wisconsin          WI 2020-12-02  30067  420930.0        25840.0
13741  Wisconsin          WI 2020-12-03  30435  426534.0        26826.0
13742  Wisconsin          WI 2020-12-04  30728  432207.0        27652.0
13743  Wisconsin          WI 2020-12-05  31005  437918.0        28532.0
13744  Wisconsin          WI 2020-12-06  31177  441067.0        28890.0

        negative  totalTestResults   death  hospitalized    total
0            NaN               0.0     NaN           NaN        0
1            NaN               0.0     NaN           NaN        0
2            NaN               0.0     NaN           NaN        0
3            0.0               0.0     NaN           NaN        0
4           10.0              10.0     NaN           NaN       10
...          ...               ...     ...           ...      ...
13740  2162616.0         4479199.0  3703.0       17569.0  2583704
13741  2169970.0         4517462.0  3773.0       17741.0  2596658
13742  2180320.0         4556566.0  3842.0       17943.0  2612622
13743  2188949.0         4595523.0  3934.0       18126.0  2626962
13744  2196722.0         4623227.0  3952.0       18216.0  2637884
```

```
[13745 rows x 11 columns]
```

[234]:
```
covid_tigeryi.loc[covid_tigeryi['state_code'] == 'NY']
```

[234]:
```
         state state_code       date   cases  positive  probableCases  \
8704  New York         NY 2020-03-02       0       0.0            NaN
8705  New York         NY 2020-03-03       1       1.0            NaN
8706  New York         NY 2020-03-04      10       1.0            NaN
8707  New York         NY 2020-03-05      21       3.0            NaN
8708  New York         NY 2020-03-06      24      25.0            NaN
...        ...        ...        ...     ...       ...            ...
8979  New York         NY 2020-12-02  664238  664238.0            NaN
8980  New York         NY 2020-12-03  674093  674093.0            NaN
8981  New York         NY 2020-12-04  685364  685364.0            NaN
8982  New York         NY 2020-12-05  696125  696125.0            NaN
8983  New York         NY 2020-12-06  705827  705827.0            NaN

        negative  totalTestResults    death  hospitalized     total
8704         NaN               0.0      NaN           NaN         0
8705         0.0               1.0      NaN           NaN         1
8706         9.0              10.0      NaN           NaN        34
8707        27.0              30.0      NaN           NaN        54
8708        97.0             122.0      NaN           NaN       358
...          ...               ...      ...           ...       ...
8979  19097486.0        19761724.0  26889.0       89995.0  19761724
8980  19291071.0        19965164.0  26955.0       89995.0  19965164
8981  19488097.0        20173461.0  27017.0       89995.0  20173461
8982  19692737.0        20388862.0  27089.0       89995.0  20388862
8983  19888867.0        20594694.0  27149.0       89995.0  20594694

[280 rows x 11 columns]
```

[235]:
```
covid_tigeryi.to_csv("covid_tigeryi.csv", index=False)
```

[168]:
```
# dtypes = {'dem_2020': 'float', 'rep_2020': 'float', 'other_2020':␣
 ↪'float','Total 2016 Votes': 'float', 'Total 2020 Votes': 'float'}
popular_vote_by_states = pd.read_csv("popular_vote_by_states.csv") # ,␣
 ↪dtype=dtypes
popular_vote_by_states = popular_vote_by_states.drop('X', 1)
popular_vote_by_states = popular_vote_by_states.drop('Y', 1)
popular_vote_by_states = popular_vote_by_states.drop('State_num', 1)
popular_vote_by_states = popular_vote_by_states.drop('Center_X', 1)
popular_vote_by_states = popular_vote_by_states.drop('Center_Y', 1)
popular_vote_by_states.head()
```

[168]:
```
        state called final  dem_2020  rep_2020  other_2020  dem_percent  \
0  U.S. Total      D    No  81281890  74222108     2884357        0.513
1     Arizona      D   Yes   1672143   1661686       53497        0.494
2     Florida      R   Yes   5297045   5668731      101680        0.479
```

```
3     Georgia      D   Yes   2474507   2461837       62138        0.495
4        Iowa      R   Yes    759061    897672       34138        0.449

   rep_percent  other_percent  dem_2020_margin  margin_shift  vote_change  \
0        0.469          0.018            0.045         0.024        0.159
1        0.491          0.016            0.003         0.039        0.316
2        0.512          0.009           -0.034        -0.022        0.175
3        0.493          0.012            0.003         0.054        0.221
4        0.531          0.020           -0.082         0.012        0.080

  stateid    EV   dem_2016   rep_2016  2016 Margin  Total 2016 Votes  \
0     NaN   NaN        NaN        NaN     0.020995         136639848
1      AZ  11.0  1161167.0  1252401.0    -0.035456           2573165
2      FL  29.0  4504975.0  4617886.0    -0.011986           9420039
3      GA  16.0  1877963.0  2089104.0    -0.051313           4092373
4      IA   6.0   653669.0   800983.0    -0.094068           1566031

   Total 2020 Votes
0         158388355
1           3387326
2          11067456
3           4998482
4           1690871
```

[272]:
```
popular_vote_by_states["dem_percent"] = popular_vote_by_states["dem_2020"] /␣
 ↪popular_vote_by_states["Total 2020 Votes"] * 100
popular_vote_by_states["rep_percent"] = popular_vote_by_states["rep_2020"] /␣
 ↪popular_vote_by_states["Total 2020 Votes"] * 100
popular_vote_by_states["other_percent"] = popular_vote_by_states["other_2020"] /
 ↪ popular_vote_by_states["Total 2020 Votes"] * 100
popular_vote_by_states["dem_2020_margin"] =␣
 ↪popular_vote_by_states["dem_percent"] - popular_vote_by_states["rep_percent"]
#popular_vote_by_states["vote_change"] = popular_vote_by_states["vote_change"]␣
 ↪* 100
#popular_vote_by_states["2016 Margin"] = popular_vote_by_states["2016 Margin"]␣
 ↪* 100
popular_vote_by_states["margin_shift"] =␣
 ↪popular_vote_by_states["dem_2020_margin"] - popular_vote_by_states["2016␣
 ↪Margin"]
popular_vote_by_states["vote_change"] = (popular_vote_by_states["Total 2020␣
 ↪Votes"] - popular_vote_by_states["Total 2016 Votes"]) /␣
 ↪popular_vote_by_states["Total 2016 Votes"] * 100

popular_vote_by_states.head()
```

[272]:
```
        state called final  dem_2020  rep_2020  other_2020  dem_percent  \
0  U.S. Total      D    No  81281890  74222108     2884357    51.318097
```

```
1     Arizona      D   Yes   1672143   1661686       53497    49.364691
2     Florida      R   Yes   5297045   5668731      101680    47.861451
3     Georgia      D   Yes   2474507   2461837       62138    49.505170
4        Iowa      R   Yes    759061    897672       34138    44.891716

   rep_percent  other_percent  dem_2020_margin  margin_shift  vote_change  \
0    46.860836       1.821066         4.457261      2.357805    15.916665
1    49.055981       1.579328         0.308710      3.854304    31.640451
2    51.219820       0.918730        -3.358369     -2.159743    17.488431
3    49.251693       1.243137         0.253477      5.384820    22.141408
4    53.089325       2.018959        -8.197609      1.209229     7.971745

  stateid    EV   dem_2016   rep_2016  2016 Margin  Total 2016 Votes  \
0     NaN   NaN        NaN        NaN     2.099456         136639848
1      AZ  11.0  1161167.0  1252401.0    -3.545595           2573165
2      FL  29.0  4504975.0  4617886.0    -1.198626           9420039
3      GA  16.0  1877963.0  2089104.0    -5.131343           4092373
4      IA   6.0   653669.0   800983.0    -9.406838           1566031

   Total 2020 Votes
0         158388355
1           3387326
2          11067456
3           4998482
4           1690871
```

[273]:
```
popular_vote_by_states.to_csv("election_tigeryi.csv", index=False)
```

[253]:
```
parse_dates3 = ['modeldate']
presidential_poll_averages_2020 = pd.read_csv("presidential_poll_averages_2020.
 ↪csv", parse_dates=parse_dates3)
presidential_poll_averages_2020 = presidential_poll_averages_2020.drop('cycle',␣
 ↪1)
presidential_poll_averages_2020 =  presidential_poll_averages_2020.
 ↪drop('pct_estimate', 1)
presidential_poll_averages_2020
```

[253]:
```
              state  modeldate       candidate_name  pct_trend_adjusted
0           Wyoming 2020-11-03  Joseph R. Biden Jr.            30.82599
1         Wisconsin 2020-11-03  Joseph R. Biden Jr.            52.09584
2     West Virginia 2020-11-03  Joseph R. Biden Jr.            33.51517
3        Washington 2020-11-03  Joseph R. Biden Jr.            59.39408
4          Virginia 2020-11-03  Joseph R. Biden Jr.            53.72101
...             ...        ...                  ...                 ...
29080   Connecticut 2020-02-27         Donald Trump            34.58325
29081      Colorado 2020-02-27         Donald Trump            44.07662
29082    California 2020-02-27         Donald Trump            34.69761
29083       Arizona 2020-02-27         Donald Trump            48.07208
```

```
29084        Alabama 2020-02-27         Donald Trump            59.14228

[29085 rows x 4 columns]
```

[254]:
```
presidential_poll_averages_2020 = presidential_poll_averages_2020.
 ↪rename({'candidate_name': 'index'}, axis=1)
presidential_poll_pct_trend_adjusted = presidential_poll_averages_2020.
 ↪pivot_table(index = ["state", "modeldate"],
                                          columns="index",
                                          values="pct_trend_adjusted").
 ↪reset_index()
presidential_poll_pct_trend_adjusted = presidential_poll_pct_trend_adjusted.
 ↪drop("Convention Bounce for Donald Trump", 1)
presidential_poll_pct_trend_adjusted = presidential_poll_pct_trend_adjusted.
 ↪drop("Convention Bounce for Joseph R. Biden Jr.", 1)
presidential_poll_pct_trend_adjusted = presidential_poll_pct_trend_adjusted.
 ↪rename({'modeldate': 'date'}, axis=1)
presidential_poll_pct_trend_adjusted = presidential_poll_pct_trend_adjusted.
 ↪rename({'Donald Trump': 'rep_poll_2020'}, axis=1)
presidential_poll_pct_trend_adjusted = presidential_poll_pct_trend_adjusted.
 ↪rename({'Joseph R. Biden Jr.': 'dem_poll_2020'}, axis=1)
```

[255]:
```
presidential_poll_pct_trend_adjusted["dem_lead"] =␣
 ↪presidential_poll_pct_trend_adjusted["dem_poll_2020"] -␣
 ↪presidential_poll_pct_trend_adjusted["rep_poll_2020"]
presidential_poll_pct_trend_adjusted
```

[255]:
```
index    state       date  rep_poll_2020  dem_poll_2020  dem_lead
0      Alabama 2020-02-27       59.14228       38.22156 -20.92072
1      Alabama 2020-02-28       59.17695       38.39190 -20.78505
2      Alabama 2020-02-29       59.23509       38.38301 -20.85208
3      Alabama 2020-03-01       59.23509       38.38301 -20.85208
4      Alabama 2020-03-02       59.23489       38.38627 -20.84862
...        ...        ...            ...            ...       ...
10731  Wyoming 2020-10-30       62.81937       30.50365 -32.31572
10732  Wyoming 2020-10-31       62.53460       30.55327 -31.98133
10733  Wyoming 2020-11-01       62.38977       30.87572 -31.51405
10734  Wyoming 2020-11-02       62.31048       30.82599 -31.48449
10735  Wyoming 2020-11-03       62.31048       30.82599 -31.48449

[10736 rows x 5 columns]
```

[258]:
```
presidential_poll_pct_trend_adjusted["dem_lead_rolling_7"] =␣
 ↪presidential_poll_pct_trend_adjusted["dem_lead"].rolling(7, min_periods = 0).
 ↪mean()

presidential_poll_pct_trend_adjusted
```

[258]:
```
index    state       date  rep_poll_2020  dem_poll_2020  dem_lead  \
0      Alabama 2020-02-27        59.14228       38.22156 -20.92072
1      Alabama 2020-02-28        59.17695       38.39190 -20.78505
2      Alabama 2020-02-29        59.23509       38.38301 -20.85208
3      Alabama 2020-03-01        59.23509       38.38301 -20.85208
4      Alabama 2020-03-02        59.23489       38.38627 -20.84862
...        ...        ...             ...            ...       ...
10731  Wyoming 2020-10-30        62.81937       30.50365 -32.31572
10732  Wyoming 2020-10-31        62.53460       30.55327 -31.98133
10733  Wyoming 2020-11-01        62.38977       30.87572 -31.51405
10734  Wyoming 2020-11-02        62.31048       30.82599 -31.48449
10735  Wyoming 2020-11-03        62.31048       30.82599 -31.48449

index  dem_lead_rolling_7
0               -20.920720
1               -20.852885
2               -20.852617
3               -20.852483
4               -20.851710
...                    ...
10731           -36.022699
10732           -35.313801
10733           -34.537570
10734           -33.784917
10735           -33.028180

[10736 rows x 6 columns]
```

[260]:
```
presidential_poll_pct_trend_adjusted.to_csv("poll_tigeryi.csv", index=False)
```

[275]:
```
covid_tigeryi["positive_rate"] = covid_tigeryi["positive"] /␣
↪covid_tigeryi["totalTestResults"] * 100
covid_tigeryi["positive_negative_ratio"] = covid_tigeryi["positive"] /␣
↪covid_tigeryi["negative"] * 100
covid_tigeryi
```

[275]:
```
           state state_code       date  cases  positive  probableCases  \
0        Alabama         AL 2020-03-07      0       0.0            NaN
1        Alabama         AL 2020-03-08      0       0.0            NaN
2        Alabama         AL 2020-03-09      0       0.0            NaN
3        Alabama         AL 2020-03-10      0       0.0            NaN
4        Alabama         AL 2020-03-11      3       0.0            NaN
...          ...        ...        ...    ...       ...            ...
13740  Wisconsin         WI 2020-12-02  30067  420930.0        25840.0
13741  Wisconsin         WI 2020-12-03  30435  426534.0        26826.0
13742  Wisconsin         WI 2020-12-04  30728  432207.0        27652.0
13743  Wisconsin         WI 2020-12-05  31005  437918.0        28532.0
13744  Wisconsin         WI 2020-12-06  31177  441067.0        28890.0
```

```
        negative  totalTestResults   death  hospitalized    total  \
0            NaN               0.0     NaN           NaN        0
1            NaN               0.0     NaN           NaN        0
2            NaN               0.0     NaN           NaN        0
3            0.0               0.0     NaN           NaN        0
4           10.0              10.0     NaN           NaN       10
...          ...               ...     ...           ...      ...
13740  2162616.0         4479199.0  3703.0       17569.0  2583704
13741  2169970.0         4517462.0  3773.0       17741.0  2596658
13742  2180320.0         4556566.0  3842.0       17943.0  2612622
13743  2188949.0         4595523.0  3934.0       18126.0  2626962
13744  2196722.0         4623227.0  3952.0       18216.0  2637884

       positive_rate  positive_negative_ratio
0                NaN                      NaN
1                NaN                      NaN
2                NaN                      NaN
3                NaN                      NaN
4           0.000000                 0.000000
...              ...                      ...
13740       9.397439                19.463927
13741       9.441895                19.656216
13742       9.485367                19.823099
13743       9.529231                20.005857
13744       9.540241                20.078417

[13745 rows x 13 columns]
```

[286]:
```
covid_tigeryi['positive_delta'] = covid_tigeryi['positive'] -␣
 ↪covid_tigeryi['positive'].shift(1)
covid_tigeryi['negative_delta'] = covid_tigeryi['negative'] -␣
 ↪covid_tigeryi['negative'].shift(1)
covid_tigeryi['totalTestResults_delta'] = covid_tigeryi['totalTestResults'] -␣
 ↪covid_tigeryi['totalTestResults'].shift(1)
covid_tigeryi['death_delta'] = covid_tigeryi['death'] - covid_tigeryi['death'].
 ↪shift(1)
covid_tigeryi["positive_rate_delta"] = covid_tigeryi["positive_delta"] /␣
 ↪covid_tigeryi["totalTestResults_delta"] * 100
covid_tigeryi["positive_negative_ratio_delta"] =␣
 ↪covid_tigeryi["positive_delta"] / covid_tigeryi["negative_delta"] * 100



covid_tigeryi.tail(10)
```

[286]:

	state	state_code	date	cases	positive	probableCases
13735	Wisconsin	WI	2020-11-27	28867	399526.0	23689.0
13736	Wisconsin	WI	2020-11-28	29155	404999.0	24129.0
13737	Wisconsin	WI	2020-11-29	29340	409054.0	24353.0
13738	Wisconsin	WI	2020-11-30	29411	411730.0	24495.0
13739	Wisconsin	WI	2020-12-01	29781	416365.0	25052.0
13740	Wisconsin	WI	2020-12-02	30067	420930.0	25840.0
13741	Wisconsin	WI	2020-12-03	30435	426534.0	26826.0
13742	Wisconsin	WI	2020-12-04	30728	432207.0	27652.0
13743	Wisconsin	WI	2020-12-05	31005	437918.0	28532.0
13744	Wisconsin	WI	2020-12-06	31177	441067.0	28890.0

	negative	totalTestResults	death	hospitalized	total
13735	2133700.0	4351880.0	3434.0	16715.0	2533395
13736	2138110.0	4373089.0	3464.0	16882.0	2543257
13737	2143790.0	4398150.0	3487.0	16999.0	2552975
13738	2149996.0	4426030.0	3494.0	17095.0	2561869
13739	2156410.0	4453101.0	3611.0	17372.0	2572906
13740	2162616.0	4479199.0	3703.0	17569.0	2583704
13741	2169970.0	4517462.0	3773.0	17741.0	2596658
13742	2180320.0	4556566.0	3842.0	17943.0	2612622
13743	2188949.0	4595523.0	3934.0	18126.0	2626962
13744	2196722.0	4623227.0	3952.0	18216.0	2637884

	positive_rate	positive_negative_ratio	positive_delta	negative_delta
13735	9.180538	18.724563	1422.0	8498.0
13736	9.261165	18.941916	5473.0	4410.0
13737	9.300592	19.080880	4055.0	5680.0
13738	9.302467	19.150268	2676.0	6206.0
13739	9.350001	19.308248	4635.0	6414.0
13740	9.397439	19.463927	4565.0	6206.0
13741	9.441895	19.656216	5604.0	7354.0
13742	9.485367	19.823099	5673.0	10350.0
13743	9.529231	20.005857	5711.0	8629.0
13744	9.540241	20.078417	3149.0	7773.0

	totalTestResults_delta	death_delta	positive_rate_delta
13735	33665.0	19.0	4.223971
13736	21209.0	30.0	25.805083
13737	25061.0	23.0	16.180520
13738	27880.0	7.0	9.598278
13739	27071.0	117.0	17.121643
13740	26098.0	92.0	17.491762
13741	38263.0	70.0	14.646003
13742	39104.0	69.0	14.507467
13743	38957.0	92.0	14.659753
13744	27704.0	18.0	11.366590

```
       positive_negative_ratio_delta
13735                      16.733349
13736                     124.104308
13737                      71.390845
13738                      43.119562
13739                      72.263798
13740                      73.557847
13741                      76.203427
13742                      54.811594
13743                      66.183799
13744                      40.512029
```

[284]:
```
covid_tigeryi.to_csv("covid_tigeryi.csv", index=False)
```

[241]:
```
covid_election_day_county = covid_us_county.loc[covid_us_county['date'] ==␣
↪'2020-11-03']
covid_election_day_county
```

[241]:
```
            state state_code    county       date  cases  deaths
286       Alabama         AL   Autauga 2020-11-03   2210      31
609       Alabama         AL   Baldwin 2020-11-03   7054      74
932       Alabama         AL   Barbour 2020-11-03   1077       9
1255      Alabama         AL      Bibb 2020-11-03    900      15
1578      Alabama         AL    Blount 2020-11-03   2204      25
...           ...        ...       ...        ...    ...     ...
1047129 Wisconsin         WI     Adams 2020-11-03    654       7
1047452 Wisconsin         WI   Ashland 2020-11-03    334       4
1047775 Wisconsin         WI    Barron 2020-11-03   1456      12
1048098 Wisconsin         WI  Bayfield 2020-11-03    326       3
1048421 Wisconsin         WI     Brown 2020-11-03  16937     120

[3246 rows x 6 columns]
```

[287]:
```
election_tigeryi = pd.read_csv("election_tigeryi.csv")
election_tigeryi.head()
```

[287]:
```
        state called final  dem_2020  rep_2020  other_2020  dem_percent  \
0  U.S. Total      D    No  81281890  74222108     2884357    51.318097
1     Arizona      D   Yes   1672143   1661686       53497    49.364691
2     Florida      R   Yes   5297045   5668731      101680    47.861451
3     Georgia      D   Yes   2474507   2461837       62138    49.505170
4        Iowa      R   Yes    759061    897672       34138    44.891716

   rep_percent  other_percent  dem_2020_margin  margin_shift  vote_change  \
0    46.860836       1.821066         4.457261      2.357805    15.916665
1    49.055981       1.579328         0.308710      3.854304    31.640451
2    51.219820       0.918730        -3.358369     -2.159743    17.488431
3    49.251693       1.243137         0.253477      5.384820    22.141408
```

```
4    53.089325       2.018959        -8.197609      1.209229     7.971745

  stateid    EV   dem_2016   rep_2016  2016 Margin  Total 2016 Votes  \
0     NaN   NaN        NaN        NaN     2.099456         136639848
1      AZ  11.0  1161167.0  1252401.0    -3.545595           2573165
2      FL  29.0  4504975.0  4617886.0    -1.198626           9420039
3      GA  16.0  1877963.0  2089104.0    -5.131343           4092373
4      IA   6.0   653669.0   800983.0    -9.406838           1566031

   Total 2020 Votes
0         158388355
1           3387326
2          11067456
3           4998482
4           1690871
```

[291]:
```
covid_election_day_county
```

[291]:
```
             state state_code    county       date  cases  deaths
286        Alabama         AL   Autauga 2020-11-03   2210      31
609        Alabama         AL   Baldwin 2020-11-03   7054      74
932        Alabama         AL   Barbour 2020-11-03   1077       9
1255       Alabama         AL      Bibb 2020-11-03    900      15
1578       Alabama         AL    Blount 2020-11-03   2204      25
...            ...        ...       ...        ...    ...     ...
1047129  Wisconsin         WI     Adams 2020-11-03    654       7
1047452  Wisconsin         WI   Ashland 2020-11-03    334       4
1047775  Wisconsin         WI    Barron 2020-11-03   1456      12
1048098  Wisconsin         WI  Bayfield 2020-11-03    326       3
1048421  Wisconsin         WI     Brown 2020-11-03  16937     120

[3246 rows x 6 columns]
```

[298]:
```
county_statistics = pd.read_csv("county_statistics.csv", index_col=None)
```

[299]:
```
county_statistics = county_statistics.drop('cases', 1)
county_statistics = county_statistics.drop('deaths', 1)
```

[306]:
```
county_statistics_merge = pd.merge(county_statistics,␣
↪covid_election_day_county, how='left',
                    left_on=['state','county'], right_on =␣
↪['state_code','county'])
```

[307]:
```
county_statistics_merge = county_statistics_merge.drop('state_code', 1)
county_statistics_merge = county_statistics_merge.rename({'state_x':␣
↪'state_code'}, axis=1)
county_statistics_merge = county_statistics_merge.rename({'state_y': 'state'},␣
↪axis=1)
county_statistics_merge = county_statistics_merge.drop('date', 1)
```

[313]:
```
county_statistics_merge = county_statistics_merge.
 ↪dropna(subset=['percentage20_Donald_Trump','percentage20_Joe_Biden'], axis=0)
county_statistics_merge = county_statistics_merge.
 ↪dropna(subset=['percentage16_Donald_Trump','percentage16_Hillary_Clinton'],␣
 ↪axis=0)
```

[322]:
```
county_statistics_merge["percentage16_Donald_Trump"] =␣
 ↪county_statistics_merge["votes16_Donald_Trump"] /␣
 ↪county_statistics_merge["total_votes16"] * 100
county_statistics_merge["percentage16_Hillary_Clinton"] =␣
 ↪county_statistics_merge["votes16_Hillary_Clinton"] /␣
 ↪county_statistics_merge["total_votes16"] * 100
county_statistics_merge["percentage20_Donald_Trump"] =␣
 ↪county_statistics_merge["votes20_Donald_Trump"] /␣
 ↪county_statistics_merge["total_votes20"] * 100
county_statistics_merge["percentage20_Joe_Biden"] =␣
 ↪county_statistics_merge["votes20_Joe_Biden"] /␣
 ↪county_statistics_merge["total_votes20"] * 100
```

[328]:
```
county_statistics_merge["dem_16_margin"] =␣
 ↪county_statistics_merge["percentage16_Hillary_Clinton"] -␣
 ↪county_statistics_merge["percentage16_Donald_Trump"]
county_statistics_merge["dem_20_margin"] =␣
 ↪county_statistics_merge["percentage20_Joe_Biden"] -␣
 ↪county_statistics_merge["percentage20_Donald_Trump"]
county_statistics_merge["dem_margin_shift"] =␣
 ↪county_statistics_merge["dem_20_margin"] -␣
 ↪county_statistics_merge["dem_16_margin"]
```

[330]:
```
county_statistics_merge.to_csv("demographic_tigeryi.csv", index=False)
```

[329]:
```
county_statistics_merge
```

[329]:

	county	state_code	percentage16_Donald_Trump	\
0	Abbeville	SC	62.868333	
1	Acadia	LA	77.262105	
2	Accomack	VA	54.471596	
3	Ada	ID	47.931611	
4	Adair	IA	65.336526	
...	...	...	...	
3106	Yuma	AZ	50.506637	
3107	Yuma	CO	80.510292	
3108	Zapata	TX	32.801532	
3109	Zavala	TX	20.412979	
3110	Ziebach	SD	48.041775	

	percentage16_Hillary_Clinton	total_votes16	votes16_Donald_Trump	\
0	34.613950	10724.0	6742.0	
1	20.587161	27386.0	21159.0	

```
2                       42.761028        15755.0                8582.0
3                       38.691733       195587.0               93748.0
4                       29.981378         3759.0                2456.0
...                           ...            ...                   ...
3106                    44.986383        40759.0               20586.0
3107                    15.030017         4664.0                3755.0
3108                    65.603063         3134.0                1028.0
3109                    77.669617         3390.0                 692.0
3110                    45.953003          766.0                 368.0

      votes16_Hillary_Clinton  percentage20_Donald_Trump  \
0                      3712.0                  66.074157
1                      5638.0                  79.493404
2                      6737.0                  54.150431
3                     75676.0                  50.387256
4                      1127.0                  69.734640
...                       ...                        ...
3106                  18336.0                  52.268841
3107                    701.0                  82.640737
3108                   2056.0                  52.547194
3109                   2633.0                  34.026033
3110                    352.0                  44.591611

      percentage20_Joe_Biden  total_votes20  ...  PublicWork  SelfEmployed  \
0                  32.984799        12433.0  ...        13.3           7.8
1                  19.148637        28425.0  ...        12.1           7.6
2                  44.739639        16938.0  ...        18.1           7.1
3                  46.470359       259389.0  ...        15.0           6.6
4                  28.615826         4183.0  ...        15.3          10.4
...                      ...            ...  ...         ...           ...
3106               46.125068        68427.0  ...        20.8           4.6
3107               15.619243         4885.0  ...        15.5          12.4
3108               47.064908         3867.0  ...        21.0          12.6
3109               65.403060         4379.0  ...        21.2           4.9
3110               53.090508          906.0  ...        48.2          18.4

      FamilyWork  Unemployment           state    cases  deaths  \
0            0.1           9.4  South Carolina    808.0    18.0
1            0.3           8.9       Louisiana   3291.0   103.0
2            0.2           5.4        Virginia   1229.0    19.0
3            0.1           4.3           Idaho  17828.0   184.0
4            0.5           3.0            Iowa    250.0     1.0
...          ...           ...             ...      ...     ...
3106         0.2          10.9         Arizona  14071.0   359.0
3107         0.5           2.4        Colorado    229.0     1.0
3108         0.4          12.4           Texas    361.0     9.0
3109         0.1          10.1           Texas    474.0    19.0
```

```
3110         1.3          27.0    South Dakota    111.0     2.0

      dem_16_margin  dem_20_margin  dem_margin_shift
0        -28.254383     -33.089359         -4.834976
1        -56.674943     -60.344767         -3.669824
2        -11.710568      -9.410792          2.299776
3         -9.239878      -3.916897          5.322981
4        -35.355148     -41.118814         -5.763667
...             ...            ...               ...
3106      -5.520253      -6.143774         -0.623520
3107     -65.480274     -67.021494         -1.541220
3108      32.801532      -5.482286        -38.283818
3109      57.256637      31.377027        -25.879610
3110      -2.088773       8.498896         10.587669

[3086 rows x 54 columns]
```

[335]:
```
usa_states_latitude_and_longitude = pd.
 ↪read_csv("usa_states_latitude_and_longitude.csv")
us_lat_long = usa_states_latitude_and_longitude[['usa_state','usa_state_code',␣
 ↪'usa_state_latitude', 'usa_state_longitude']].dropna()
```

[339]:
```
election_tiger = pd.merge(election_tigeryi, us_lat_long, how='left',
                      left_on=['state','stateid'], right_on =␣
 ↪['usa_state','usa_state_code'])
```

[341]:
```
election_tiger = election_tiger.rename({'stateid': 'state_code'}, axis=1)
election_tiger = election_tiger.drop('usa_state', 1)
election_tiger = election_tiger.drop('usa_state_code', 1)
```

[343]:
```
election_tiger.to_csv("election_tigeryi.csv", index=False)
```

[345]:
```
parse_dates = ['date']

poll_tigeryi = pd.read_csv("poll_tigeryi.csv", parse_dates=parse_dates)
```

[348]:
```
poll_tiger = pd.merge(poll_tigeryi, us_lat_long, how='left',
                      left_on=['state'], right_on = ['usa_state'])
```

[351]:
```
poll_tiger = poll_tiger.drop('usa_state', 1)
```

[352]:
```
poll_tiger["dem_lead_rolling_30"] = poll_tiger["dem_lead"].rolling(30,␣
 ↪min_periods = 0).mean()
```

[354]:
```
poll_tiger.to_csv("poll_tigeryi.csv", index=False)
```

[361]:
```
election_tiger2 = election_tiger[['state','called','dem_percent',␣
 ↪'rep_percent','dem_2020_margin']]

election_tiger2.head()
```

[361]:
```
         state called  dem_percent  rep_percent  dem_2020_margin
0  U.S. Total      D    51.318097    46.860836         4.457261
1     Arizona      D    49.364691    49.055981         0.308710
2     Florida      R    47.861451    51.219820        -3.358369
3     Georgia      D    49.505170    49.251693         0.253477
4        Iowa      R    44.891716    53.089325        -8.197609
```

[359]:
```
poll_tiger.head()
```

[359]:
```
     state       date  rep_poll_2020  dem_poll_2020  dem_lead  \
0  Alabama 2020-02-27        59.14228        38.22156 -20.92072
1  Alabama 2020-02-28        59.17695        38.39190 -20.78505
2  Alabama 2020-02-29        59.23509        38.38301 -20.85208
3  Alabama 2020-03-01        59.23509        38.38301 -20.85208
4  Alabama 2020-03-02        59.23489        38.38627 -20.84862

   dem_lead_rolling_7 usa_state_code  usa_state_latitude  usa_state_longitude  \
0          -20.920720             AL           32.318231           -86.902298
1          -20.852885             AL           32.318231           -86.902298
2          -20.852617             AL           32.318231           -86.902298
3          -20.852483             AL           32.318231           -86.902298
4          -20.851710             AL           32.318231           -86.902298

   dem_lead_rolling_30
0           -20.920720
1           -20.852885
2           -20.852617
3           -20.852483
4           -20.851710
```

[362]:
```
poll_tiger2 = pd.merge(poll_tiger, election_tiger2, how='left',
                       left_on=['state'], right_on = ['state'])

poll_tiger2.head()
```

[362]:
```
     state       date  rep_poll_2020  dem_poll_2020  dem_lead  \
0  Alabama 2020-02-27        59.14228        38.22156 -20.92072
1  Alabama 2020-02-28        59.17695        38.39190 -20.78505
2  Alabama 2020-02-29        59.23509        38.38301 -20.85208
3  Alabama 2020-03-01        59.23509        38.38301 -20.85208
4  Alabama 2020-03-02        59.23489        38.38627 -20.84862

   dem_lead_rolling_7 usa_state_code  usa_state_latitude  usa_state_longitude  \
0          -20.920720             AL           32.318231           -86.902298
1          -20.852885             AL           32.318231           -86.902298
2          -20.852617             AL           32.318231           -86.902298
3          -20.852483             AL           32.318231           -86.902298
4          -20.851710             AL           32.318231           -86.902298
```

```
   dem_lead_rolling_30 called  dem_percent  rep_percent  dem_2020_margin
0           -20.920720      R     36.56999    62.031643       -25.461653
1           -20.852885      R     36.56999    62.031643       -25.461653
2           -20.852617      R     36.56999    62.031643       -25.461653
3           -20.852483      R     36.56999    62.031643       -25.461653
4           -20.851710      R     36.56999    62.031643       -25.461653
```

[366]:
```
poll_tiger2['dem_result_poll_diff'] = poll_tiger2['dem_percent'] -␣
 ↪poll_tiger2['dem_poll_2020']
poll_tiger2['rep_result_poll_diff'] = poll_tiger2['rep_percent'] -␣
 ↪poll_tiger2['rep_poll_2020']
poll_tiger2['dem_lead_result_poll_diff'] = poll_tiger2['dem_2020_margin'] -␣
 ↪poll_tiger2['dem_lead']
poll_tiger2['dem_lead_rolling_7_diff'] = poll_tiger2['dem_2020_margin'] -␣
 ↪poll_tiger2['dem_lead_rolling_7']
poll_tiger2['dem_lead_rolling_30_diff'] = poll_tiger2['dem_2020_margin'] -␣
 ↪poll_tiger2['dem_lead_rolling_30']

poll_tiger2.head()
```

[366]:
```
      state       date  rep_poll_2020  dem_poll_2020  dem_lead  \
0  Alabama 2020-02-27        59.14228       38.22156 -20.92072
1  Alabama 2020-02-28        59.17695       38.39190 -20.78505
2  Alabama 2020-02-29        59.23509       38.38301 -20.85208
3  Alabama 2020-03-01        59.23509       38.38301 -20.85208
4  Alabama 2020-03-02        59.23489       38.38627 -20.84862

   dem_lead_rolling_7 usa_state_code  usa_state_latitude  usa_state_longitude  \
0          -20.920720             AL           32.318231           -86.902298
1          -20.852885             AL           32.318231           -86.902298
2          -20.852617             AL           32.318231           -86.902298
3          -20.852483             AL           32.318231           -86.902298
4          -20.851710             AL           32.318231           -86.902298

   dem_lead_rolling_30 called  dem_percent  rep_percent  dem_2020_margin  \
0           -20.920720      R     36.56999    62.031643       -25.461653
1           -20.852885      R     36.56999    62.031643       -25.461653
2           -20.852617      R     36.56999    62.031643       -25.461653
3           -20.852483      R     36.56999    62.031643       -25.461653
4           -20.851710      R     36.56999    62.031643       -25.461653

   dem_result_poll_diff  rep_result_poll_diff  dem_lead_result_poll_diff  \
0              -1.65157              2.889363                  -4.540933
1              -1.82191              2.854693                  -4.676603
2              -1.81302              2.796553                  -4.609573
3              -1.81302              2.796553                  -4.609573
4              -1.81628              2.796753                  -4.613033
```

```
   dem_lead_rolling_7_diff  dem_lead_rolling_30_diff
0                -4.540933                -4.540933
1                -4.608768                -4.608768
2                -4.609036                -4.609036
3                -4.609170                -4.609170
4                -4.609943                -4.609943
```

[369]:
```
poll_tiger2 = poll_tiger2.dropna(subset=['dem_percent','rep_percent'], axis=0)
```

[371]:
```
poll_tiger2.to_csv("poll_tigeryi.csv", index=False)
```

[372]:
```
county_statistics_merge = pd.read_csv("demographic_tigeryi.csv")

county_statistics_merge.head()
```

[372]:
```
       county state_code  percentage16_Donald_Trump  \
0  Abbeville         SC                  62.868333
1     Acadia         LA                  77.262105
2   Accomack         VA                  54.471596
3        Ada         ID                  47.931611
4      Adair         IA                  65.336526

   percentage16_Hillary_Clinton  total_votes16  votes16_Donald_Trump  \
0                     34.613950        10724.0                6742.0
1                     20.587161        27386.0               21159.0
2                     42.761028        15755.0                8582.0
3                     38.691733       195587.0               93748.0
4                     29.981378         3759.0                2456.0

   votes16_Hillary_Clinton  percentage20_Donald_Trump  percentage20_Joe_Biden  \
0                   3712.0                  66.074157               32.984799
1                   5638.0                  79.493404               19.148637
2                   6737.0                  54.150431               44.739639
3                  75676.0                  50.387256               46.470359
4                   1127.0                  69.734640               28.615826

   total_votes20  ...  PublicWork  SelfEmployed  FamilyWork  Unemployment  \
0        12433.0  ...        13.3           7.8         0.1           9.4
1        28425.0  ...        12.1           7.6         0.3           8.9
2        16938.0  ...        18.1           7.1         0.2           5.4
3       259389.0  ...        15.0           6.6         0.1           4.3
4         4183.0  ...        15.3          10.4         0.5           3.0

           state    cases  deaths  dem_16_margin  dem_20_margin  \
0  South Carolina    808.0    18.0     -28.254383     -33.089359
1       Louisiana   3291.0   103.0     -56.674943     -60.344767
2        Virginia   1229.0    19.0     -11.710568      -9.410792
3           Idaho  17828.0   184.0      -9.239878      -3.916897
4            Iowa    250.0     1.0     -35.355148     -41.118814
```

```
   dem_margin_shift
0         -4.834976
1         -3.669824
2          2.299776
3          5.322981
4         -5.763667

[5 rows x 54 columns]
```

[381]:
```
county_statistics_merge['cases'] = county_statistics_merge['cases'].fillna(0)
county_statistics_merge['deaths'] = county_statistics_merge['deaths'].fillna(0)
```

[382]:
```
county_statistics_merge['cases_rate'] = county_statistics_merge['cases'] /␣
 ↪county_statistics_merge['TotalPop'] * 100
county_statistics_merge['deaths_rate'] = county_statistics_merge['deaths'] /␣
 ↪county_statistics_merge['TotalPop'] * 100
county_statistics_merge['men_percent'] = county_statistics_merge['Men'] /␣
 ↪county_statistics_merge['TotalPop'] * 100
county_statistics_merge['women_percent'] = county_statistics_merge['Women'] /␣
 ↪county_statistics_merge['TotalPop'] * 100


county_statistics_merge.head()
```

[382]:
```
       county state_code  percentage16_Donald_Trump  \
0  Abbeville         SC                  62.868333
1     Acadia         LA                  77.262105
2   Accomack         VA                  54.471596
3        Ada         ID                  47.931611
4      Adair         IA                  65.336526

   percentage16_Hillary_Clinton  total_votes16  votes16_Donald_Trump  \
0                     34.613950        10724.0                6742.0
1                     20.587161        27386.0               21159.0
2                     42.761028        15755.0                8582.0
3                     38.691733       195587.0               93748.0
4                     29.981378         3759.0                2456.0

   votes16_Hillary_Clinton  percentage20_Donald_Trump  percentage20_Joe_Biden  \
0                   3712.0                  66.074157               32.984799
1                   5638.0                  79.493404               19.148637
2                   6737.0                  54.150431               44.739639
3                  75676.0                  50.387256               46.470359
4                   1127.0                  69.734640               28.615826

   total_votes20  ...           state    cases  deaths  dem_16_margin  \
0        12433.0  ...  South Carolina    808.0    18.0     -28.254383
```

```
1        28425.0  ...       Louisiana   3291.0   103.0     -56.674943
2        16938.0  ...        Virginia   1229.0    19.0     -11.710568
3       259389.0  ...           Idaho  17828.0   184.0      -9.239878
4         4183.0  ...            Iowa    250.0     1.0     -35.355148

   dem_20_margin  dem_margin_shift  cases_rate  deaths_rate  men_percent  \
0     -33.089359         -4.834976    3.259642     0.072616    48.588026
1     -60.344767         -3.669824    5.256601     0.164518    48.609580
2      -9.410792          2.299776    3.742387     0.057856    48.961632
3      -3.916897          5.322981    4.097289     0.042287    50.101237
4     -41.118814         -5.763667    3.476085     0.013904    49.388209

   women_percent
0      51.411974
1      51.390420
2      51.038368
3      49.898763
4      50.611791

[5 rows x 58 columns]
```

[383]:
```
county_statistics_merge.to_csv("demographic_tigeryi.csv", index=False)
```

[4]:
```
parse_dates = ['date']

covid_3 = pd.read_csv("covid_tigeryi.csv", parse_dates=parse_dates,␣
 ↪index_col=False)
```

[5]:
```
covid_3.head(10)
```

[5]:
```
     state state_code       date  cases  positive  probableCases  negative  \
0  Alabama         AL 2020-03-07      0       0.0            NaN       NaN
1  Alabama         AL 2020-03-08      0       0.0            NaN       NaN
2  Alabama         AL 2020-03-09      0       0.0            NaN       NaN
3  Alabama         AL 2020-03-10      0       0.0            NaN       0.0
4  Alabama         AL 2020-03-11      3       0.0            NaN      10.0
5  Alabama         AL 2020-03-12      4       0.0            NaN      10.0
6  Alabama         AL 2020-03-13      8       1.0            NaN      11.0
7  Alabama         AL 2020-03-14     15       6.0            NaN      22.0
8  Alabama         AL 2020-03-15     28      12.0            NaN      28.0
9  Alabama         AL 2020-03-16     36      28.0            NaN      28.0

   totalTestResults  death  hospitalized  total  positive_rate  \
0               0.0    NaN           NaN      0            NaN
1               0.0    NaN           NaN      0            NaN
2               0.0    NaN           NaN      0            NaN
3               0.0    NaN           NaN      0            NaN
4              10.0    NaN           NaN     10       0.000000
5              10.0    NaN           NaN     10       0.000000
```

```
6              12.0    NaN           NaN     12       8.333333
7              28.0    NaN           NaN     74      21.428571
8              40.0    0.0           NaN     86      30.000000
9              56.0    0.0           NaN     96      50.000000

   positive_negative_ratio  positive_delta  negative_delta  \
0                      NaN             NaN             NaN
1                      NaN             0.0             NaN
2                      NaN             0.0             NaN
3                      NaN             0.0             NaN
4                 0.000000             0.0            10.0
5                 0.000000             0.0             0.0
6                 9.090909             1.0             1.0
7                27.272727             5.0            11.0
8                42.857143             6.0             6.0
9               100.000000            16.0             0.0

   totalTestResults_delta  death_delta  positive_rate_delta  \
0                     NaN          NaN                  NaN
1                     0.0          NaN                  NaN
2                     0.0          NaN                  NaN
3                     0.0          NaN                  NaN
4                    10.0          NaN               0.0000
5                     0.0          NaN                  NaN
6                     2.0          NaN               0.5000
7                    16.0          NaN               0.3125
8                    12.0          NaN               0.5000
9                    16.0          0.0               1.0000

   positive_negative_ratio_delta
0                            NaN
1                            NaN
2                            NaN
3                            NaN
4                       0.000000
5                            NaN
6                       1.000000
7                       0.454545
8                       1.000000
9                            NaN
```

[6]:
```
#pd.set_option('use_inf_as_na', True)

covid_3['positive_rate_delta'] = covid_3['positive_delta'] /␣
 ↪covid_3['totalTestResults_delta'] * 100
covid_3['positive_negative_ratio_delta'] = covid_3['positive_delta'] /␣
 ↪covid_3['negative_delta'] * 100
```

```
covid_3 = covid_3.replace(np.inf, np.nan)
#covid_3.replace(np.inf, np.nan)
```

[7]:
```
covid_3.head()
```

[7]:
```
     state state_code       date  cases  positive  probableCases  negative  \
0  Alabama         AL 2020-03-07      0       0.0            NaN       NaN
1  Alabama         AL 2020-03-08      0       0.0            NaN       NaN
2  Alabama         AL 2020-03-09      0       0.0            NaN       NaN
3  Alabama         AL 2020-03-10      0       0.0            NaN       0.0
4  Alabama         AL 2020-03-11      3       0.0            NaN      10.0

   totalTestResults  death  hospitalized  total  positive_rate  \
0               0.0    NaN           NaN      0            NaN
1               0.0    NaN           NaN      0            NaN
2               0.0    NaN           NaN      0            NaN
3               0.0    NaN           NaN      0            NaN
4              10.0    NaN           NaN     10            0.0

   positive_negative_ratio  positive_delta  negative_delta  \
0                      NaN             NaN             NaN
1                      NaN             0.0             NaN
2                      NaN             0.0             NaN
3                      NaN             0.0             NaN
4                      0.0             0.0            10.0

   totalTestResults_delta  death_delta  positive_rate_delta  \
0                     NaN          NaN                  NaN
1                     0.0          NaN                  NaN
2                     0.0          NaN                  NaN
3                     0.0          NaN                  NaN
4                    10.0          NaN                  0.0

   positive_negative_ratio_delta
0                            NaN
1                            NaN
2                            NaN
3                            NaN
4                            0.0
```

[8]:
```
covid_3.to_csv("covid_tigeryi.csv", index=False)
```

[ ]:

[ ]:

[ ]:

[ ]: