phrgina﷽﷽﷽﷽﷽ilar patterns amongG4063 Assignment 3

Xinran Wang

xw2418

In the tweets origins plots, the tweets are presented as dots and darker color of the clustering means more tweets within this area. I use blue for candidates from Democratic Party and red for candidates from the Republican Party. The plots display similar patterns among all five candidates. Most of the tweets come from inside the United States. Especially in the east side, California and Washington, these areas have high dot density and the colors are significantly darker. For other areas in the world, there are a number of tweets in the northern hemisphere around the zero longitude. Not a lot of tweets appear to be in the eastern and southern hemispheres. One reason might be that Twitter is not very accessible in these areas. The range of distribution is also similar for all candidates: the latitude goes from 50 to -50.

In the tweet distribution for Clinton, the clustering around longitude 0 and latitude 50 is moderately dense for tweets outside of US. When I locate this area on the map, no specific area corresponds to this position. I guess these tweets are due to random location selected by the users.

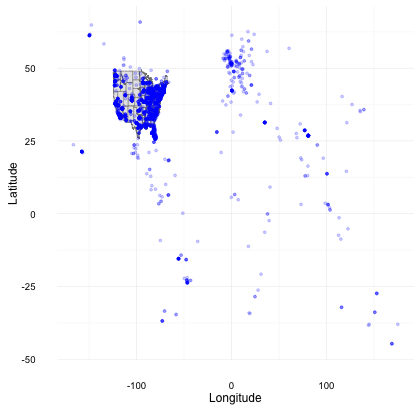


图 1The Tweets Origins Plot for Clinton

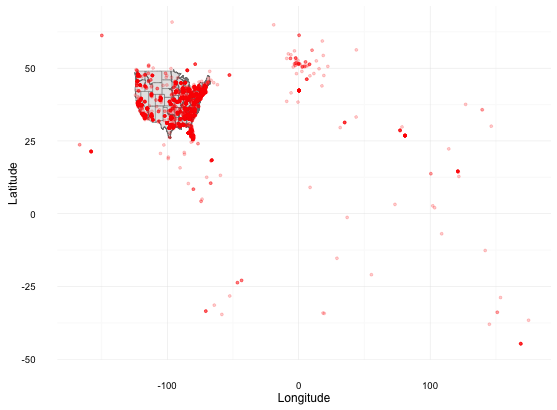


图 2The Tweets Origins Plot for Cruz

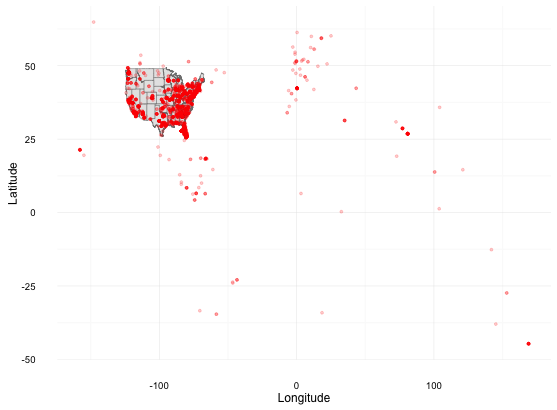


图 3The Tweets Origins Plot for Rubio



图 4The Tweets Origins Plot for Sanders

Trump has significantly more tweets compared to other candidates, which can be easily tell from the density of the tweets in the plot. The actual data also shows that Trump has over 60,000 parsed tweets compared to other candidates 10,000-20,000. Similarly to Clinton, there is a cluster around latitude 50 longitude 0 that could be the result of random location. Other than this area and the U.S, there are also a number of tweets in Europe and southern hemisphere, whereas other candidates only have a few dots sparsely distributed considering the global range. It is without doubt that Trump is the most debatable and widely discussed candidate.

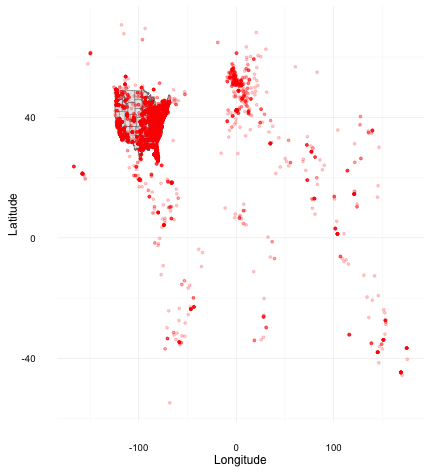


图 5The Tweets Origins Plot for Trump

While the tweets origins maps allow me to compare the overall number and distribution of tweets between candidates, the color-coded map of US shows the differences between states and reveals more specific information about each candidate. The maps are created based on the number of geolocated tweets. The counts are converted to proportions and darker color represents higher frequency.

For Clinton, most tweets come from New York. California and Florida rank the second and third in tweets frequency. Clinton’s home state is New York and she indeed receives more attention in tweets from New York. Except for New York, California, Florida and Texas, other states have much lower frequencies. Generally the colors of the states are darker within the east side.

> Frequency Count of Tweets for Clinton

            alabama              arizona             arkansas           california             colorado          connecticut

                 32                   64                   18                  382                   66                   23

           delaware district of columbia              florida              georgia                idaho             illinois

                  3                   38                  196                  116                   13                  130

            indiana                 iowa               kansas             kentucky            louisiana                maine

                 62                   23                   22                   37                   39                   22

           maryland        massachusetts             michigan            minnesota          mississippi             missouri

                 52                   61                   77                   63                   20                   47

            montana             nebraska               nevada        new hampshire           new jersey           new mexico

                  8                    6                   92                   23                   74                   12

           new york       north carolina         north dakota                 ohio             oklahoma               oregon

                426                   99                    2                  110                   24                   54

       pennsylvania         rhode island       south carolina         south dakota            tennessee                texas

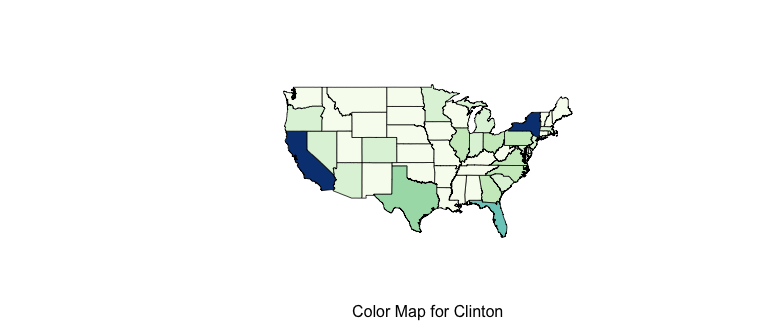
                114                   18                   68                    1                   36                  187

               utah              vermont             virginia           washington        west virginia            wisconsin

                 19                   14                  102                   37                    9                   22

            wyoming

                  3



Cruz’s home state is Texas where he receives the most tweets discussions. California, New York and Idaho are also high in tweets frequencies. The plot assigns the darkest color to Texas which makes it stands out from all other states. This means that Texas has significantly more tweets. Similar to Clinton’s plot, the states along east coast and west coast are more active in Twitter discussion. For the states in the middle area, Kansas is more active than Colorado and the situation is the opposite for Clinton. The fact that they come from different political party may results in the difference.

> poly.counts(tw\_points\_cruz, usa)

            alabama              arizona             arkansas           california             colorado          connecticut

                 76                   95                   29                  310                   42                   34

           delaware district of columbia              florida              georgia                idaho             illinois

                  0                   18                  184                  224                   17                   87

            indiana                 iowa               kansas             kentucky            louisiana                maine

                 63                   49                   55                   60                   34                   19

           maryland        massachusetts             michigan            minnesota          mississippi             missouri

                 39                   66                   63                   44                   23                   63

            montana             nebraska               nevada        new hampshire           new jersey           new mexico

                 16                    5                   66                   23                   95                    6

           new york       north carolina         north dakota                 ohio             oklahoma               oregon

                236                  134                    4                  117                   23                   63

       pennsylvania         rhode island       south carolina         south dakota            tennessee                texas

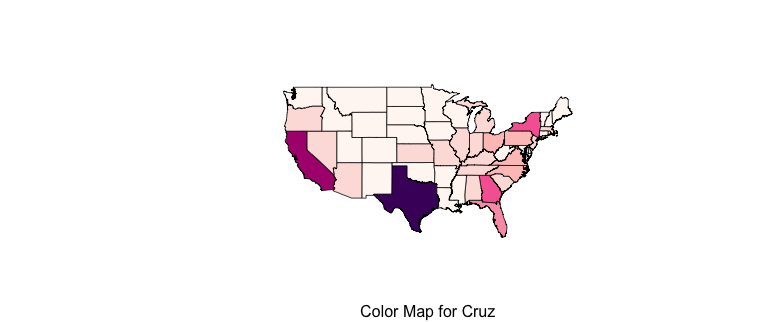
                141                   11                   78                    2                   86                  452

               utah              vermont             virginia           washington        west virginia            wisconsin

                 22                    4                  126                   37                   14                   47

            wyoming

                  3



Rubio receives the most Twitter discussion from Florida which is also his home state. California comes the second and Texas comes the third. New York and Georgia are also relatively active. The pattern within the middle area still holds: Rubio, as a candidate from the Republican Party, receives more tweets from Kansas than Colorado.

> > poly.counts(tw\_points\_rubio, usa)

            alabama              arizona             arkansas           california             colorado          connecticut

                 52                   68                   22                  285                   31                   19

           delaware district of columbia              florida              georgia                idaho             illinois

                  0                   31                  314                  156                    7                   81

            indiana                 iowa               kansas             kentucky            louisiana                maine

                 56                   10                   37                   39                   23                   14

           maryland        massachusetts             michigan            minnesota          mississippi             missouri

                 39                   62                   30                   50                   20                   56

            montana             nebraska               nevada        new hampshire           new jersey           new mexico

                  5                    8                   47                    9                   73                    1

           new york       north carolina         north dakota                 ohio             oklahoma               oregon

                165                  100                    3                   56                   28                   33

       pennsylvania         rhode island       south carolina         south dakota            tennessee                texas

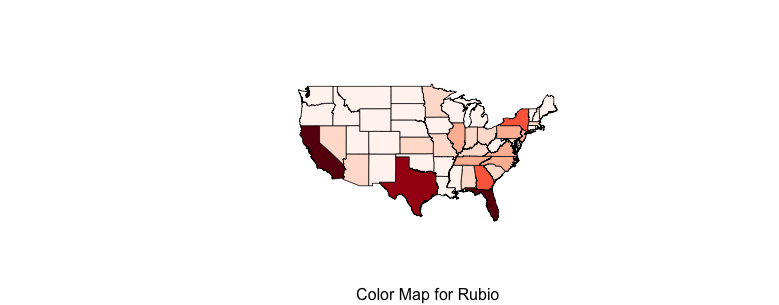
                103                    5                   67                    0                   79                  261

               utah              vermont             virginia           washington        west virginia            wisconsin

                  3                    4                   76                   20                   10                   14

            wyoming

                  4



Sanders’s home state is Vermont and he has the most tweets in New York. California, Texas, Florida and Illinois take the following spaces in frequency rank. Considering the case with Clinton, Twitter users from New York are most active in discussing the Democratic Party candidates. California is active in every case, which implies that people from California in general uses the social network more frequently.

> poly.counts(tw\_points\_sanders, usa)

            alabama              arizona             arkansas           california             colorado          connecticut

                 25                   90                   22                  454                   43                   23

           delaware district of columbia              florida              georgia                idaho             illinois

                  3                   29                  159                   92                    5                  162

            indiana                 iowa               kansas             kentucky            louisiana                maine

                 75                   32                   41                   36                   40                   23

           maryland        massachusetts             michigan            minnesota          mississippi             missouri

                 39                   90                   97                   36                   10                   64

            montana             nebraska               nevada        new hampshire           new jersey           new mexico

                  4                   13                   83                   32                   75                   10

           new york       north carolina         north dakota                 ohio             oklahoma               oregon

                407                   77                    3                  116                   16                   69

       pennsylvania         rhode island       south carolina         south dakota            tennessee                texas

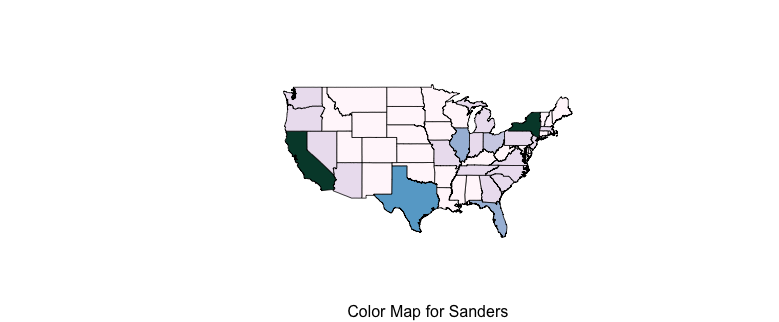
                 92                   23                   55                    1                   53                  243

               utah              vermont             virginia           washington        west virginia            wisconsin

                 19                   45                   98                   75                    8                   41

            wyoming

                  1



Since Trump has about four times of tweets of other candidates, the frequency of tweets is higher in every state and the pattern should be easier to detect. The home state for Trump is New York but he receives most tweets from California. States coming next are Texas, New York and Florida. For states in the middle areas, the colors are almost identical. The plots of all five candidates confirms that east and west coast are more active in Twitter using or cares about the election more.

> poly.counts(tw\_points\_trump, usa)

            alabama              arizona             arkansas           california             colorado          connecticut

                173                  314                   90                 1128                  120                  118

           delaware district of columbia              florida              georgia                idaho             illinois

                  3                   78                  866                  523                   41                  364

            indiana                 iowa               kansas             kentucky            louisiana                maine

                209                  103                  110                  194                  147                   53

           maryland        massachusetts             michigan            minnesota          mississippi             missouri

                147                  237                  187                  130                   75                  159

            montana             nebraska               nevada        new hampshire           new jersey           new mexico

                 27                   46                  204                   70                  360                   28

           new york       north carolina         north dakota                 ohio             oklahoma               oregon

                969                  364                   14                  464                  100                  148

       pennsylvania         rhode island       south carolina         south dakota            tennessee                texas

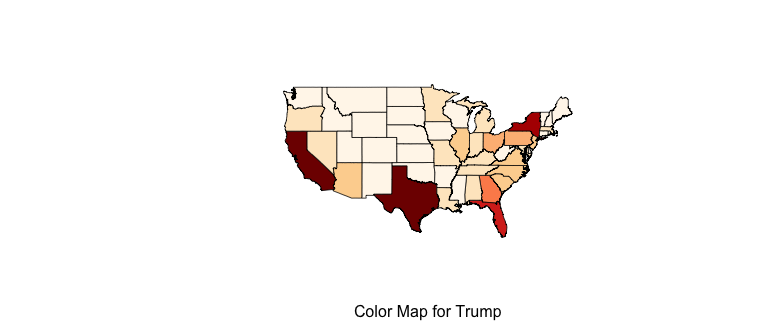
                446                   37                  270                    8                  231                 1096

               utah              vermont             virginia           washington        west virginia            wisconsin

                 44                   27                  323                  123                   57                  121

            wyoming

                  6



On the date when the tweets are collected, the states holding primaries are New Hampshire, South Carolina, Nevada, Alabama, Arkansas, Georgia, Massachusetts, Oklahoma, Tennessee, Texas, Vermont, Virginia, Louisiana, Michigan, Mississippi, Florida, Illinois, Missouri, North Carolina and Ohio. It is not necessarily true that states holding primaries appear to have more tweets. For example, New York and California are not holding primaries yet still rank top in the tweets frequencies for all candidates. The difference is not significant between states with primaries and without. In fact, I think the number of tweets depends on the social and cultural situation and political characteristic of each state. If the state is the candidate’s home state or is very active in politics, then the number of tweets is larger. Such is the case New York and California. these two cities are the major cities in the U.S and are socially and culturally active. This fact may imply that Twitter has not yet becomes the most influential tool in political election as the number of tweets is not significantly affected by the schedule of the campaign. Candidates may develop methods to promote the use of social network platform and make use of Twitter more effectively.