

Twitter Conversations about SARS-Cov-2: What is Going on Out There in Illinois?

Adee Athiyaman¹

Research Brief, Short Paper

ISSN 2687-8844

Vol. 2, No. 5 (2020, April 7)



Abstract

We study Twitter conversations among Illinoisans about Coronavirus through the systematic analysis of tweets collected during March 21, 2020 through April 5, 2020, the first two weeks of stay-at-home order issued by Governor J. B. Pritzker. The subjective information in the tweet is processed to identify the strength of the positive and negative tone of the tweet. Results reveal that rural residents are more relaxed about the Coronavirus pandemic than their urban counterparts.

1.0. Data and Methodology for Data Analysis

Tweets that had the keywords listed in Table 1 were extracted using the standard API provided by Twitter. To get tweets by counties, geographical codes for the county's courthouse were specified and all tweets originating within a ten-mile radius of the courthouse were gathered². Metro residents tweeted almost four times more about the virus than the nonmetro residents³. In all, 69,638 tweets were assembled, of which 68,235 tweets (98% of the total) were unique.

Table 1: Profile of Tweets on Coronavirus: March 21, 2020 through April 5, 2020

Keywords	Number of Hits	Percentage Non-Metro	Percentage Metro
Covid	9,212	23	77
Covid19	16,022	22	78
Covid-19	5,261	22	78
Coronavirus	28,476	23	77
#Covid	2,296	19	81
#Covid19	1,628	21	79
Others including #Coronavirus	6743	16	84
Total	69, 638	22	78

¹ Professor, Illinois Institute for Rural Affairs, Western Illinois University.

² Geo coordinates for the counties were obtained from www.lat-long.com.

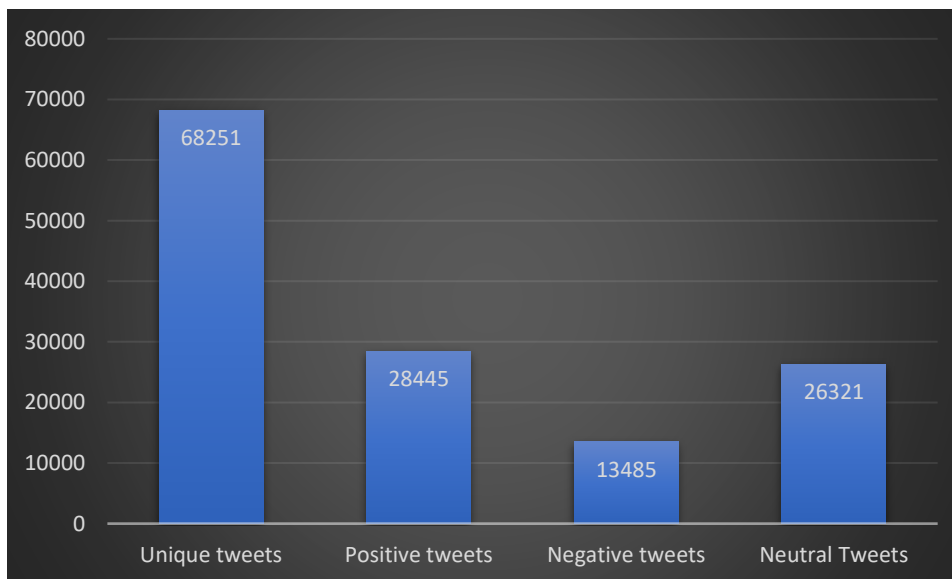
³ The metro / nonmetro classifications of the counties were based on rural-urban continuum codes; see <https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx>.

The methodology for sentiment analysis is basically the process of determining the emotion of the “tweeter”, positive, neutral, or negative. A python algorithm was used to analyze the polarity of the tweets. Polarity scores range from -1 (a negative tweet) to +1 (a positive tweet); tweets were summed to get county, metro and nonmetro scores⁴.

2.0. Insights from Data Analysis

The Centre for Disease Control states that COVID-19 is a new disease, and older adults and people of any age who have serious underlying medical conditions might be at higher risk for severe illness from COVID-19⁵. In spite of this statement, majority of the tweets about the illness were either neutral or positive (Figure 1). This is a good situation for Illinois, neutral to positive emotions among the general public signal that they are relaxed or relieved that the virus situation is not out of control and/or it can be managed; examples of neutral, positive, and negative tweets are given in the “Note” to Figure 1.

Figure 1: Number of Unique, Positive, Negative, and Neutral Tweets



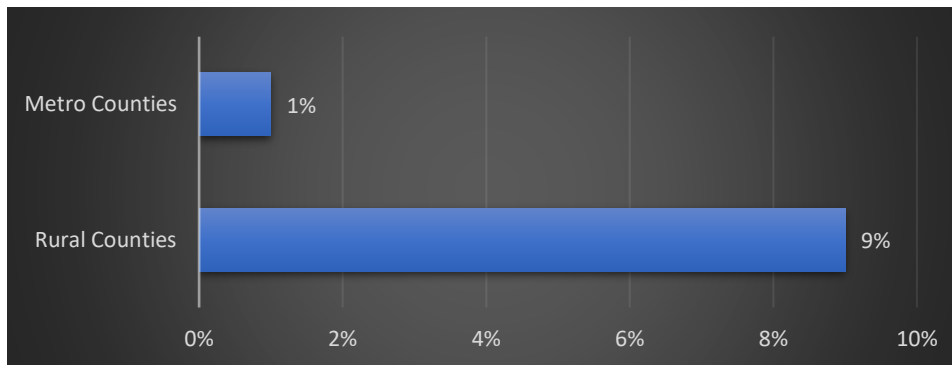
Note: Neutral tweet: China should pay the damages caused by this Covid-19 pandemic;
Positive tweet: Donations pour in for the Knox County first responders during COVID-19 Pandemic;
Negative tweet: Yeah coronavirus is bad but have you thought about the national debt?

⁴ For more information on sentiment analysis, see Athiyaman (2020). Social Media Data for Market Analysis: The Case of Higher Education Institutions. Available at: http://research.iira.org/wp-content/uploads/2019/12/HEpaper_2019.pdf

⁵⁵ See <https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/groups-at-higher-risk.html>

At the county level, nonmetro residents have a more positive outlook about the Coronavirus situation than the metro residents (Figure 2). This could be because of the lower number of Coronavirus cases and associated deaths in the rural regions. For instance, the median number of days the virus has been active in the rural regions as at April 8, 2020 is ten, one-half of the median number of days the virus has been active in the metro regions. Similarly, the median number of positive Coronavirus cases in the metros are four to five times more than the median number of infected persons in the rural regions. As far as the number of deaths in the rural areas, Christian county had the most, four deaths in 19 days. In contrast, in the metros, Cook County registered 312 deaths in the 75 days the virus has been active in the county. Table 2 shows the three worst affected counties in both the metro and the nonmetro regions of Illinois. Appendix 1 provides information for all of the affected metro and nonmetro counties.

Figure 2: Illinoisans’ Emotions Associated with Covid-19: Polarity Scores



Note: Scores can range from -1 to +1; positive polarity scores indicate positive feelings.

Table 2: Covid-19 Spread as at April 8, 2020: Three Worst Affected Metro and Nonmetro Counties and the Compound, Daily Growth Rates of Cases and Deaths

Rural County	First Case	Number of Cases as at 4/8/2020	Number of Deaths	CGR, Cases	CGR, Deaths
Christian	3/20/2020	24	4	16%	7%
Randolph	3/30/2020	31	0	34%	NA
Whiteside	3/15/2020	21	1	12%	~0%
Metro County	First Case	Number of Cases as at 4/8/2020	Number of Deaths	CGR, Cases	CGR, Deaths
Cook	1/24/2020	10,520	312	12%	8%
Lake	3/11/2020	1,041	23	24%	11%
Will	3/16/2020	914	35	28%	15%

3.0. Discussion

We believe that this is one of the first research analyzing the sentiment expressed by Illinoisans on social media posts about Covid-19. Rural Illinoisans are more positive in their thinking and feelings about the virus than their urban counterparts. This is mainly due to the lower number of infected cases in the rural areas. A careful analysis of the infection data shows that the rate of spread in nonmetro counties is about two-thirds of the rate of spread in metro areas; the virus spreads at a daily, average compound growth rate of 18% in urban geographies and 12% in rural areas (Appendix 1); how and when this will end remain uncertain.

Health indicators suggest that the problem could become severe in the rural regions; rural areas have very few resources available to stop community spread, other than closing businesses and schools, and curtailing large gatherings. Public officials are advising people to contact their doctor if they are sick, but there are very few physicians in rural Illinois; data from the Illinois Department of Financial and Professional Regulation show that a typical rural county has 21 physicians, compared to 244 physicians on average for an urban county.

The case-fatality rate (CFR), number of deaths given infected cases, is also higher in the rural counties (Appendix 2). For example, in 11 days Montgomery County registered a CFR of 25%. In contrast, in Cook County the virus has been active for 11 weeks, yet the CFR for the county is a low 3%.

To return to work, education, and other activities as soon as possible, Illinois needs information not only about the strains the virus cause its healthcare system, but also about the moods of its citizens and their preparedness to combat the illness. For the latter, research such as this is needed.

Appendix 1: Covid-19 Status of the Illinois Counties as at April 8, 2020

(i) Rural Counties

County	First Case Reported on...	No of Days	Number of Cases	Number of Deaths	CGR, Cases	CGR, Deaths
Adams	2020-03-20	20	9	0	11%	NA
Bureau	2020-03-27	13	4	0	11%	NA
Carroll	2020-03-28	12	5	1	13%	0
Christian	2020-03-20	20	24	4	16%	7%
Clark	2020-03-30	10	4	0	14%	NA
Coles	2020-04-07	2	2	0	35%	NA
Crawford	2020-03-30	10	5	0	16%	NA
Cumberland	2020-03-14	26	1	0	0%	NA
Douglas	2020-03-25	15	10	0	15%	NA
Effingham	2020-04-03	6	2	0	12%	NA
Fayette	2020-03-28	12	4	0	12%	NA
Franklin	2020-03-26	14	2	0	5%	NA
Gallatin	2020-04-05	4	1	0	0%	NA
Hancock	2020-04-08	1	1	0	0%	NA
Iroquois	2020-03-27	13	7	0	15%	NA
Jasper	2020-04-04	5	4	0	28%	NA
Jefferson	2020-04-06	3	3	0	37%	NA
Jo Daviess	2020-03-22	18	8	0	12%	NA
Knox	2020-03-29	11	1	0	0%	NA
LaSalle	2020-03-18	22	17	2	13%	3%
Lawrence	2020-04-07	2	2	0	35%	NA
Lee	2020-04-04	5	6	0	36%	NA
Livingston	2020-03-22	18	10	0	13%	NA
Logan	2020-03-25	15	3	0	7%	NA
Marion	2020-03-30	10	4	0	14%	NA
Mason	2020-03-31	9	3	0	12%	NA
Massac	2020-04-01	8	2	0	9%	NA
Montgomery	2020-03-29	11	4	1	13%	0
Morgan	2020-03-25	15	8	1	14%	0
Moultrie	2020-04-02	7	2	0	10%	NA
Ogle	2020-03-31	9	12	0	28%	NA
Pike	2020-04-04	5	1	0	0%	NA
Randolph	2020-03-30	10	31	0	34%	NA
Richland	2020-04-07	2	1	0	0%	NA
Saline	2020-03-30	10	2	0	7%	NA
Shelby	2020-04-07	2	3	0	55%	NA
Stephenson	2020-03-22	18	6	0	10%	NA
Wabash	2020-04-06	3	1	0	0%	NA
Washington	2020-03-18	22	2	0	3%	NA
Whiteside	2020-03-15	25	21	1	12%	0
MEDIAN		10	4	0	12%	0%

(ii) Metro Counties

County	First Case Reported on...	No of Days	Number of Cases	Number of Deaths	CGR, Cases	CGR, Deaths
Bond	2020-03-29	11	3	0	10%	NA
Boone	2020-04-05	4	7	1	49%	0%
Calhoun	2020-04-05	4	1	0	0%	NA
Champaign	2020-03-22	18	69	1	24%	0%
Clinton	2020-03-15	25	13	0	10%	NA
Cook	2020-01-24	76	10520	312	12%	8%
De Witt	2020-04-03	6	3	0	18%	NA
DeKalb	2020-03-21	19	32	1	18%	0%
DuPage	2020-03-14	26	875	30	26%	13%
Ford	2020-03-31	9	5	1	18%	0%
Grundy	2020-03-24	16	12	0	16%	NA
Henry	2020-03-27	13	16	0	21%	NA
Jackson	2020-03-18	22	22	1	14%	0%
Jersey	2020-04-03	6	5	0	27%	NA
Kane	2020-03-10	30	296	19	19%	10%
Kankakee	2020-03-18	22	145	6	23%	8%
Kendall	2020-03-18	22	79	2	20%	3%
Lake	2020-03-11	29	1041	23	24%	11%
Macon	2020-03-28	12	13	1	21%	0%
Macoupin	2020-04-02	7	10	0	33%	NA
Madison	2020-03-18	22	82	2	20%	3%
Marshall	2020-03-25	15	1	0	0%	NA
McHenry	2020-03-10	30	180	5	17%	5%
McLean	2020-03-20	20	65	2	21%	3%
Menard	2020-03-29	11	2	0	6%	NA
Mercer	2020-04-02	7	2	0	10%	NA
Monroe	2020-03-23	17	20	1	18%	0%
Peoria	2020-03-16	24	16	2	12%	3%
Piatt	2020-04-02	7	4	0	20%	NA
Rock Island	2020-03-22	18	67	1	23%	0%
Sangamon	2020-03-14	26	35	2	14%	3%
St. Clair	2020-03-14	26	127	3	19%	4%
Stark	2020-04-08	1	1	0	0%	NA
Tazewell	2020-03-26	14	17	2	20%	5%
Unknown	2020-03-23	17	62	0	24%	NA
Vermilion	2020-04-01	8	5	0	20%	NA
Will	2020-03-16	24	914	35	28%	15%
Williamson	2020-03-18	22	9	0	10%	NA
Winnebago	2020-03-15	25	63	3	17%	4%
Woodford	2020-03-14	26	7	0	7%	NA
MEDIAN		18	16.5	1	18%	3%

Appendix 2: Case-Fatality Rate: Metro and Nonmetro Counties

Rural County	Case-Fatality Rate
Carroll	20%
Christian	17%
LaSalle	12%
Montgomery	25%
Morgan	13%
Whiteside	5%
Metro County	Case-Fatality Rate
Boone	14%
Champaign	1%
Cook	3%
DeKalb	3%
DuPage	3%
Ford	20%
Jackson	5%
Kane	6%
Kankakee	4%
Kendall	3%
Lake	2%
Macon	8%
Madison	2%
McHenry	3%
McLean	3%
Monroe	5%
Peoria	13%
Rock Island	1%
Sangamon	6%
St. Clair	2%
Tazewell	12%
Will	4%
Winnebago	5%

Note: Only counties reporting Covid-19 fatalities are listed.