a)

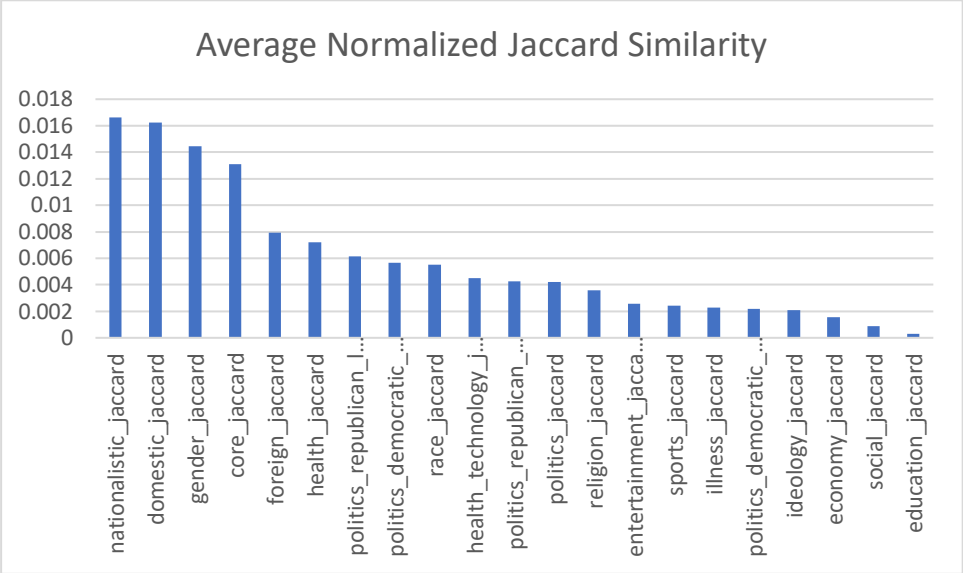
Ohio has experienced a significant impact from the COVID-19 pandemic, with the state reporting a total of 3,426,822 cases and 139,540 deaths as of April 2023. The first case of COVID-19 was reported on March 9, 2020, and Ohio experienced a peak in cases in late April and early May 2020, with over 1,000 new cases per day. The state implemented measures such as school closures, social distancing, and a stay-at-home order to slow the spread of the virus.

Ohio's experience has been similar to that of other US states. The state has reported a significant number of cases and deaths, with the pandemic affecting both urban and rural areas. And the state also experienced the peak of cases during December of 2020 and during December 2021 and January of 2022, which is also similar to the trend that other states experienced. And also, the most death cases of Ohio state are in people over 60 years old which is also a similar pattern with other states.

Prior to the state confirming any cases, Governor DeWine made the decision to cancel Arnold Classic. And, on March 9, DeWine asked colleges and universities to go to online courses. And on March 12, DeWine announced that all schools from K-12 would close for a 3-week break. Except the school closures, Ohio state also instituted a ban on gatherings of more than 100 and ordered the closure of all bars and restaurants. On March 22, Ohio state issued a statewide stay-at-home order, requiring the closure on nonessential businesses. Until April 14, DeWine announced Phase 1 of the state's reopening plan. And on May 12, 2021, DeWine announced that all COVID health orders, including the mask mandate, would be lifted on June 2.

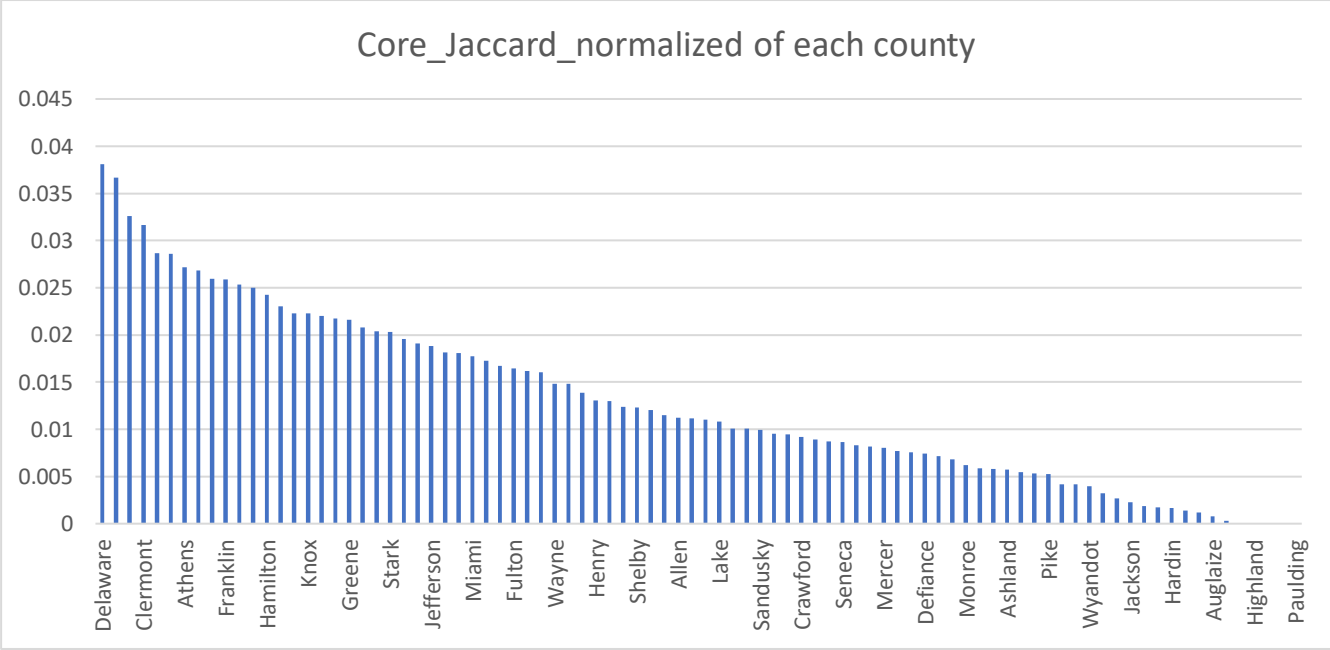
As for whether Ohio has dealt with the pandemic successfully, Wikipedia does not make a definitive statement. However, according to the website, because of the early intervention in the pandemic by its government and medical community, Ohio had decreased the impact of the state's healthcare system by 50 to 75%. And the Governor DeWine was one of the first state governors to "sound the alarm" about the coronavirus threat, taking action before Ohio had any confirmed cases. The Hill said he'd "been one of the most aggressive governors in responding to the pandemic". From all these statements, I believe that Ohio has dealt with Covid-19 successfully.

b)



From the graph, I found that the most topics that related to Covid-19 measured by Jaccard similarity is the nationalistic topic. The similarity of domestic topic is nearly equal to that of the nationalistic topic. And the top four topic is nationalistic, domestic, gender and core. The Jaccard Similarity of all those four topic are above 0.01, and have a large lead compared with other topics. And the similarity of education topic is the minimum, which is nearly 0.

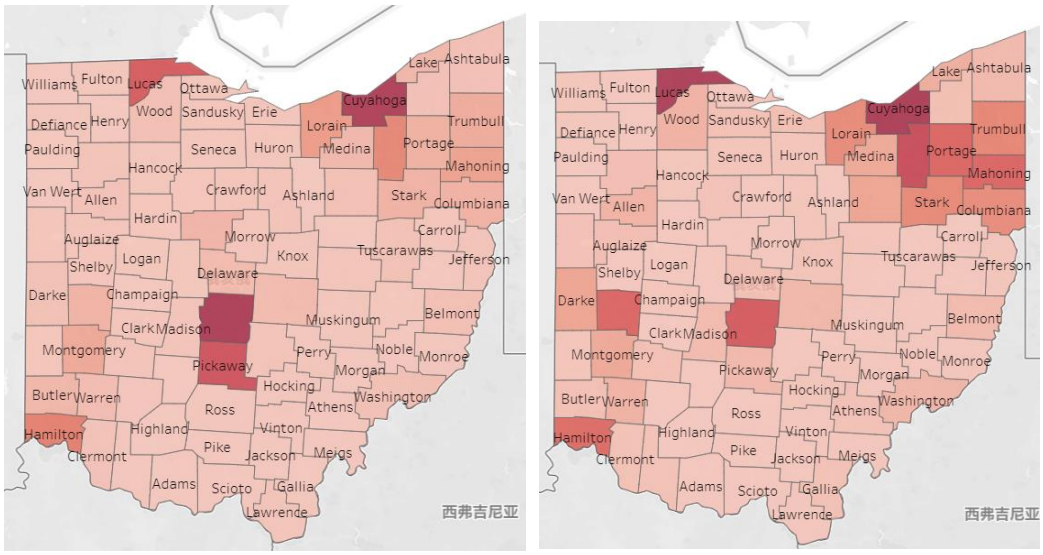
c)



From the graph, I found that Delaware county published the most tweets that are related to Covid-19. And compared to the number of cases of each county, I found that, for the counties that with the top-15 average awareness value of core-Jaccard-normalized, those counties also has the most Covid-19 cases, which shows

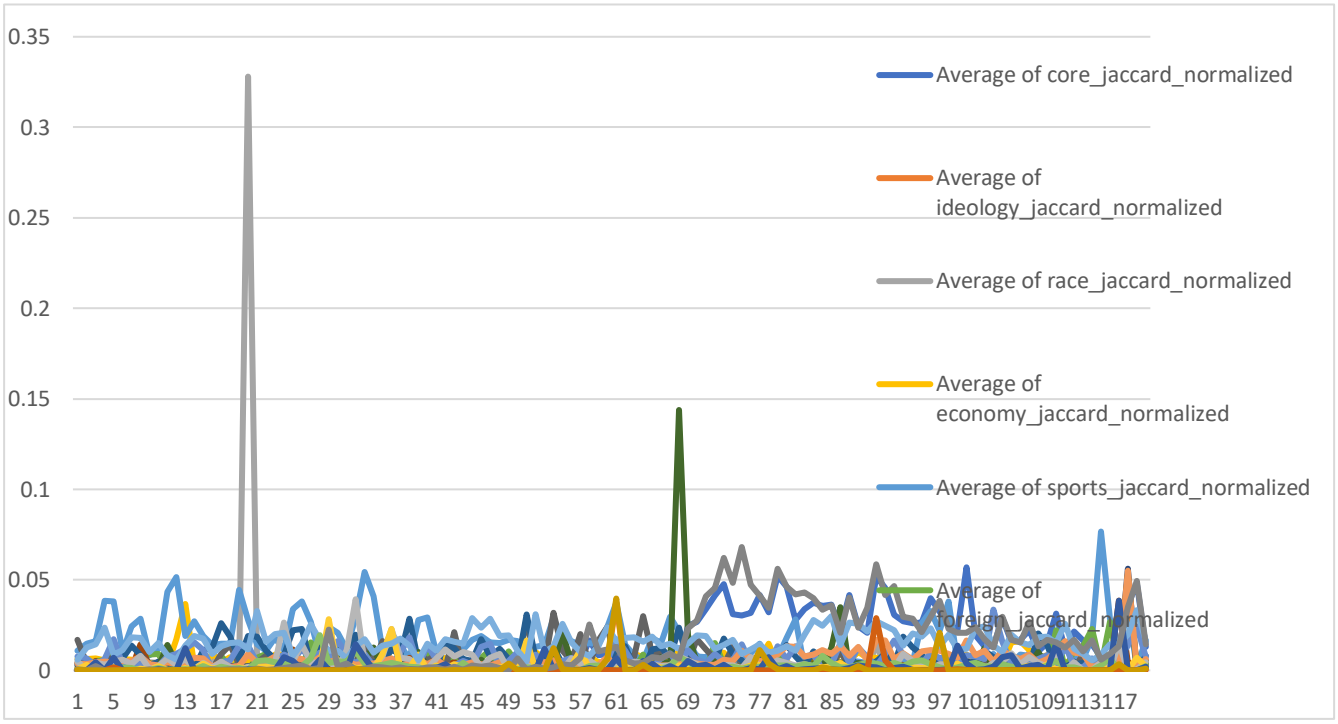
that those two variables have correlation and the Core_Jaccard_Normalized can be one of the most important attributes that we used to predict the number of cases.

d)



The left-side graph shows Number of Average Covid-19 Cases per Capita, and the right-side graph shows Number of Average Covid-19 Deaths per Capita. From the graph, the top-5 counties with high number of per capita cases are Franklin, Cuyahoga, Pickaway, Lucas and Hamilton and the top-5 counties with high number of per-capita deaths are Cuyahoga, Lucas, Summit, Franklin, and Mahoning. The comparison between number of percapita cases and number of per-capita deaths shows that there is a strong correlation between cases and deaths. Therefore, "death" is an important variable to help predict the number of cases.

e)



In the graph, X-axis represents "Day" variable, and Y-axis represents the level of awareness. The overall data values are all under 0.05, and on some special days, there is a substantial increase in the AWARENESS level for certain topics. This may be directly related to the hot events or special activities of the day that led to a heated discussion on a topic and the increasing in the number of tweets on the subject.