Project Proposal:

Names:
Eric Youn
Kevin Wang
Zakaria Alhomsy
Zhengde Song
Siyi Gao

Research Question: Is there a positive correlation between COVID-19 deaths and an increase in cryptocurrency investment in the US.

(2-3 Paragraphs & 2 References)
Background & Prior Work:
As panic and lack of confidence increased in centralized organizations, such as the government or banks, people may have begun to take interest in alternative forms of decentralized stores of value. The market of cryptocurrencies boomed during the one year of COVID-19.which could be seen from the price increase in Bitcoin by around 640% (from $7,300 to more than $46,800 in March, 2021). Online cryptocurrencies enjoy many benefits over traditional currencies such as being easily accessible from within people's homes while they were quarantined and not facing fee transfer problems, something which over 1.7 billion people faced with traditional currencies that was resolved using cryptocurrencies. Online cryptocurrencies are also decentralized in nature and not tied to the stability of governments, many of which took a great hit during the pandemic.

Our group noticed the staggering difference in the behaviour of the cryptocurrencies market and decided to dig deeper into the reason for the contrast.

Knowing the fact that there are many variables which might affect people's decision on the cryptocurrencies market, our group decides to pick the variable of the number of COVID-19 deaths and conduct a research to find out if there is a correlation between it and an increase in cryptocurrency investment.

References:
https://clsbluesky.law.columbia.edu/2021/03/26/how-the-covid-19-pandemic-affected-the-cryptocurrency-market/

https://www.frontiersin.org/articles/10.3389/fpsyg.2021.647691/full

Hypothesis: We hypothesize that there will be a positive correlation between the onset of COVID-19 and an increase in public interest in cryptocurrencies in the US. Because of at the time waning confidence in the government's ability to deal with COVID-19, we predict that people began to increase investments into cryptocurrencies instead of traditional investments.

Data: An ideal dataset would include historical prices and open interest in the various cryptocurrencies, such as Bitcoin, Ethereum, DogeCoin, etc. that we are tracking before and during the COVID-19 pandemic. These prices would ideally be recorded per day which would give us ample data to determine if there were noticeable trends and if there were significant differences pre-COVID and during the pandemic. We would also need data on the number of cases and deaths due to COVID-19 around the world each week. This would ideally also be stored on a daily scale so that we can correlate it with our cryptocurrency price action. We would also like to measure interest in cryptocurrencies through historical search term data from Google Trends.

Ethics & Privacy:

For data collection, there is no worry of getting informed consent because all of the data that we will be using for the project will be public data that is free for everyone's use. The bitcoin dataset will contain no personal identifiable information, as that data will simply track bitcoin's price movements day to day, and the COVID-19 data set will also contain no personal identifiable information because the identities are already given as a unique identifier instead of a name or other identifier that can be linked to a specific person.

Because of the public nature of our datasets, data security isn't a major concern. We will be using a private github repository to store the data, and the data will only be shared with the group members, the professor, and TAs. If a person requests for their information to be removed, we will be able to remove that row from the dataset. There is no data retention plan to delete the data after it is no longer needed because it is a public dataset that is available to everyone. In terms of equitable impact, there is likely to be no impact as this is an observational study that won't change the behaviors or circumstances of those with COVID-19 or those buying bitcoin.

The bitcoin dataset is very unlikely to contain bias, as it is simply stating the previous events that occurred in bitcoin's price. The COVID-19 dataset, however, could be biased in that the sample of COVID patients is suffering from selection bias in the way that the patients were selected. This would mean that they were not randomly sampled from the total population of COVID patients. To solve this we would have a disclaimer that our analysis is only relevant to the original population that it was sampled from, such as say the country that the dataset was collected from. To detect this bias we could run a statistical test with a similar COVID dataset and see if there is an appreciable difference between them.

Team Expectations: We expect each team member to be cooperative and do their fair share of work. We will plan to meet at a minimum of once a week to discuss and divide the work needed for the project. If a team member is being uncooperative(not attending meetings or doing their assigned part), then we will try to reach out to them through email, if there is no improvement from the team member, then we will contact the Professor and inform them about the issue.

Project Timeline Proposal:

- Week 4: Deciding on the research topic and finishing project proposal,

- Week 5: Find datasets, do background research on topic
- Week 6: Wrangle Data, discuss which things to chart
- Week 7: Begin Analysis,
- Week 8: Composing the Ethics & Privacy section.
- Week 9: Composing the conclusion and finalize the project
- Week 10: Turn in Final Project