

Data Science & Sociology – Proposal Form (your project)

Please use this form, and enter text using a computer on this form in the space that is provided and then print out the form with the newly entered text into it.

Name(s) of students	KU Student Number

Project Title:

Changes in Topics During COVID-19 Within The Korean Sociology Studies

Research question:

During the COVID-19 period, did topics of Korean sociology appear different from other periods?

During the COVID-19 period, was changes of topics of Korean sociology affected by confirmed cases of COVID-19?

1. State the problem (dependent variable(s)) you are working on in one sentence. For example, you could write: “I am/we are intending to study the relationship between X and Y”, or “I am/we are testing the proposition that X causes Y”, or “I am investigating whether when in the presence of Z, X behaves differently than Y”, and so on.

I am intending to study the relationship between whether or not it is during COVID-19 period, the number of the COVID-19 confirmed cases, and the change in topics of Korean sociology studies. (blue statements are independent variable and red one is dependent.)

2. State in one or two sentences what the theoretical issue is that is at stake in your study, that is, the general issue that your study relates to. It is important that this is not the same as what you are working on. If you are working on the proposition “X causes Y”, then the theory you are addressing is a more general version of the study you are conducting.

Studies in the “Korean Journal of Sociology(한국 사회학)” had developed in two directions since 1964 until 2017: ‘specialization’ and ‘convergence’. (Seol et al., 2018)

Korean sociology was not negatively affected by COVID-19 in terms of the number of papers or the number of collaborators.(Kang & Kwon, 2021)

Further details are in item 4.

3. Describe the data and data source that you will be using to address the problem you are considering. Describe the sample for your study and the design of your study.

I'll use the data that describes the number of confirmed cases of COVID-19 during 2020 and 2021, which is provided by WHO(World Health Organization). Since this data includes confirmed cases of COVID-19 over two hundreds nations by day, I extracted data about Republic of Korea and grouped data by month.

In addition, I'll use the data that refers to KCI(Korea Citation Index)-listed papers. This data includes titles, topics, keywords of papers, and so on. I will only use data belonging to “Social Science in General(사회과학일반)” and “Sociology(사회학)”

These data do not need to be sampled, because it is the population itself. However, I conceptualized Korean sociology studies into KCI-listed papers that is included in “Social Science in General” and “Sociology” field, so there is some filtering process, which means there are some studies that have been omitted but can be considered sociology papers. And if there is some bias, this data may be reduced to only the data of the “*Korean Journal of Sociology*”.

4. Description of Project: This section identifies the problem you consider, the data that you will use, the methods that you will employ, and identifies at least two social science* sources that speak to the issue you consider and provide the theoretical motivation for the study. Use proper bibliographic referencing for your sources (APA style), including the relevant page numbers.

According to Seol et al.(2018, pp. 199-200), studies in the “*Korean Journal of Sociology*” are specialized. During 1964-1979, main topic is about urban and rural. And during 1980-1999, the topic is changed to social structure, and at last, 2000-2017, topics are divided into culture, gender, immigration, and so on. At the same time, there is tendency to convergence. So studies have used common concepts and topics, making clusters of themselves.

Additionally, there is study about impact of COVID-19 on social science study. Kang & Kwon(2021, p. 196) proved that the number of social science papers and collaborators had not decreased during COVID-19 period, unlike that of natural science and humanities.

However, I suspect the social sciences have been affected in some way by COVID-19. Therefore, I'll look in topics and keywords, not the numbers. In this case, I'll prove that COVID-19 affects on topics of Korean sociology studies by demonstrating difference in topics during COVID-19 period(2020-2021) and non-COVID-19 period(2018-2019). Moreover, I'll also check the relationship between the number of confirmed cases of COVID-19 and changes in topic.

In the terms of methodology and data, there are several possible ways to analyze topic. First, there is “Subject Area” column in the KCI data, therefore I can classify studies. However, this column is not assigned well, as almost values are “Sociology” or “Social Science in General”, which is same as main category I filtered. Second, there is NLP(Nature Language Processing) algorithm like topic-modeling that can analyze the main text and the abstract, which is same method Seol et al.(2018) used. But I can't use this method, since individual user can't automatically get all of paper main texts and abstracts. Finally, I can use NLP on “Keyword” column, which is composed of multiple keywords. This method is most difficult but still possible and effective way.

In this case, I have to analyze vectors of words, using word-embedding methods. However, this process needs document or sentence, therefore I'll consider each record as sentences, since one record has multiple keywords and keywords in same record has similar meanings. And if needed, I'll use more period KCI paper data to train the model. Additionally, if this process is insufficient to train model(this can be happen because there are too many different keywords), I'll divide keywords in corpus, However, this can be an unnecessary or intrusive process since keyword itself is a word that is semantically divided enough.

REFERENCES

- Kang, J. H. & Kwon, E. R. (2021). The Effect of COVID-19 Pandemic on Research Productivity in South Korea: A Comparative Analysis of Korean Journal Articles across Academic Fields. *Korean Journal of Sociology*, 55(1), 179-199.
- Seol, D. H., Ko, J. H. & Yoo, S. H. (2018). Korean Sociological Association and Sociological Research: Changes in the Areas of Sociology in Korea 1964-2017. *Korean Journal of Sociology*, 52(1), 153-213.