Dong Liang 1399853

I am from MGI and I already finished “Geoscripting” and “Geo-information Tools” courses. So, I have experience in Python, R and GIS. But I have not applied my programming skills to data science projects except for those in the GIS/RS domain. So, I chose this course to enrich my experience in programming combined with environmental science and social science. And I am also interested in how to handle different data sources and combine these data in one project using Python and GIS software. After completing this course, I hope I will be able to combine complex data sources and analyze the data in the environmental science domains.

**Personal learning goals**

Data:

“To be able to acquire and process data from social media using Python.” Currently, I have only used Python in geographical domain. At the end of this course, I would like to be able to acquire data from social media using web scrapers and analyze these data by Python. Therefore, I will produce an executable Python Jupyter notebook which shows how a web crawler acquires data from social media websites and analyze the data to know about the perception of air pollution in Poland.

Information:

“To be able to do spatial regression analysis between potential factors and air pollution.” Currently, I do not have any experience in spatial regression analysis. Also, I have never used Geoda before. At the end of this course, I want to be able to do spatial regression analysis using Geoda or Python. Therefore, I will produce text document or Python Jupyter notebook to show the steps of spatial regression analysis between factors and air pollution.

Knowledge(pitfalls of using the machine learning to analyze limited data)

“Understand the pitfalls of combining different temporal resolutions and how it would influence the quality of data.” Currently, I only know the basic concept of temporal resolution but at the end of this course, I will be able to explain the pitfalls of combining different temporal resolution and how it affects the quality of data and the analysis result. I will demonstrate this in a small essay in which I will provide a literature review and an analysis of our project procedure to show how the quality of data is affected by different temporal resolutions.