a. Analyse activity and language on the forum over time. Some starting points:

• Describe your data: How active are participants, and are there periods where this increases

or decreases? Is there a trend over time?

• Looking at the linguistic variables, do these change over time? Is there a relationship

between variables?

b. Analyse the language used by groups. Some starting points:

• Threads indicate groups of participants communicating on the same topic. Describe the

threads present in your data.

• By analysing the linguistic variables for all or some of the threads, is it possible to see a

difference in the language used by different groups?

• Does the language used within threads (or between threads) change over time? How

consistent or variable is the language used within threads?

c. Challenge: Social networks online. We can think of participants posting to the same thread

at similar times (for example during the same month) as forming a social network. When

these participants also post to other threads over the same period, their social network

extends.

• Can you define, graph and describe the social network that exists at a particular point in

time, for example over one month? How does this change in the following months?

• Note: you only need to analyse a small portion of the social network over a short time

period. We will cover social network analysis in Lecture 5.

d. Reflection on your investigation. What did you first investigate? How did you then modify

your research based on the results of your first investigation?

• Using one of the data science methodologies in Lecture 4, illustrate your research process