

Problemset 2

Your task is to evaluate and demonstrate for a given social media platform or other web platform (1) options to access interesting data, and (2) analyse the suitability of the data for social media analytics purposes. Your submissions will be distributed to all other students, as an input and inspiration for the capstone project.

Pick yourself a social media platform or other web platform of your choice.

1. Describe briefly the platform and its main features.
2. Describe in a structured way what sort of data is available on the platform, and to which category of data it belongs (e.g. text, network, images, metadata, ...)
3. Investigate options to access data from the platform. Provide demo code to showcase feasibility of data access. This should not at all be exhaustive. Only provide a minimalistic code snippet that exposes the basic method of data access. Do NOT scrape large amounts of data. Print or display the first few rows of the data you have accessed.
4. Briefly summarise your insights in a few sentences: What are challenges or limitations of accessing data from this platform? Do you have any recommendations that facilitate data access? Have you come across alternative ways to access data that you have not implemented in your demo code?
5. Brainstorm ideas on how the data could be used for social media analytics purpose: What kind of questions could be answered, or problems be solved with the data? Who would be the target audience for such analysis? Hint: Think about how the data could be related to one of the course topics (e.g. network analysis, topic modelling, machine learning for text data, sentiment analysis, retrieval augmented generation), based on your current understanding of these topics.