

### **Some Questions**

- 1) What will be the things to look out for when collecting the appropriate tweets for a certain event?
- 2) Will there be any hardware or software issues that will arise when collecting these types of data?
- 3) How will I initially pull data using an API so that I will run into the least amount of resistance?
- 4) In the event that the data provider runs into a conflict, what will be my role in making sure the dataset is pure?
- 5) Given a certain amount of data at one point in time, what is the best way to make sure it is safe from potentially harmful changes?
- 6) What measures must be taken in order for my API key to be kept safe? In the event that it is compromised, how shall I proceed with the data already collected?
- 7) If any part of the data shall be omitted for the purpose of running a less but more efficient model, then how should the foregone data be interpreted? If context is lost, then what should be the excuse?