* Provide a course-specific introduction to yourself. What has been your experience with database systems and data preparation to date? What databases have you used? What issues have you encountered with databases and data preparation in the past?

* Social media sites generate continuous stream of massive amount of data (primarily unstructured data) every minute, in your opinion, would you prefer to use a DBMS engine or use file-processing to store/query/update the big-data collected?

My name is Zeeshan Latifi. I work for a healthcare company in the Chicago area as a Senior BI Analyst. My experience with DB’s is somewhat limited. I have some SQL experience from undergrad as well as some experience doing some relationship building in BI tools at work. I’ve used Teradata and SQL server. One thing I have encountered when working with databases and data preparation is that sometimes setting up the relationships can be confusing. Primary keys and secondary keys is what I struggle with the most. However, some tools make the relationship building between your data easier; others are more difficult.

When using social media as your data source, using a file processing system would work best. Something like Hadoop. However, it mainly depends on the type of data you’re collecting. If it is simply text and you understand the structure, you can pull the data and format in a standard DBMS engine. However, if you’re working with a variety of data (images, sounds, etc.), I think it’d be best to use a file processing approach.