Introduce yourself and share with your peers your background and any experience you have with data science. (1 mark)

Hello all,

My name is Austin. I'm a self-taught python programmer who's trying to transition into the field of Artificial Intelligence and Deep Learning. My dream is to become an AI Engineer, but I figured that learning about data science and statistics as a whole would prove beneficial to my endeavors, so here I am. Other than basic analysis of small data sets in python, I have no prior experience with data science.

Based on the videos and the reading material, how would you define a data scientist and data science? (3 marks)

Although there seems to be some debate on exactly what constitutes "data science", I like the definition provided by the author in one of our prior readings - "Data science is what data scientists do". I particularly like this definition because, as the author mentions, it does not limit data science to those with PhD's in statistics and programming. Although I come from a computer science background myself, I like the emphasis on accepting new scientists from all different disciplines. I think that this interdisciplinary acceptance is what truly makes a field excel, and those that stand to bar entry because "the data is too small" or the scientist doesn't have the "correct" background, only do so because they fear diversity may disrupt the status quo. We saw it over and over in the videos - curiosity, passion for the subject, and the ability to tell a good story. These were the traits the professor said made a great data scientist. The necessary math, programming, and relational database knowledge can all be learned. What really matters is that you can take the data, and glean from it a wonderful, compelling story. After all, the chief job of data science is to present their findings to those in charge of practice and policy.

As discussed in the videos and the reading material, data science can be applied to problems across different industries. What industry are you passionate about and would like to pursue a data science career in? (1 mark)

The field that I'm most excited about going into is law enforcement. Since I could read, I gobbled up everything crime related that I could get my hands on. I know that crime prediction is a prime example of data analysis, but this isn't what interests me. It's important, yes, but it's just not what I'm excited about. What am excited about, however, is the possibility of using data science to turn the old, unstructured data of cold cases into new, workable databases that could potentially generate new leads using new machine learning algorithms. Plenty of local police departments don't even have a cold case unit, and often, detectives are forced to use their days off to work on them. This makes working on the cases agonizingly slow. Remember, these cases were unsolvable even when they were fresh, when detectives could devote all of their time to working them. Imagine how little progress you'd make working something this difficult only in your free time. However, I think now is the time to breath new life into these cases. Over the years, new technology has consistently emerged to reopen new cases. DNA testing is a prime example of this. Cases that had long since gone cold were suddenly given new life with the introduction of DNA testing, and with each incremental improvement we make on the process, more and more cases can be reopened and investigated. I think data science could be the next step in this process. With the vast amount of data collected as evidence, I believe this sector is ripe for innovation.

Based on the videos and the reading material, what are the ten main components of a report that would be delivered at the end of a data science project? (5 marks)

1. Cover Page
2. Table of Contents
3. Abstract
4. Introductory
5. Literature Review
6. Methodology
7. Results
8. Discussion
9. Conclusion
10. Housekeeping (references, acknowledgments, appendices)