Coursera: IBM Data Science Professional Certificate

Course 1: What is data science?

Final Peer-graded Assignment

In this Assignment, you will demonstrate your understanding of the videos and the readings by answering open-ended questions, defining data science and data scientist, and describing the different sections comprising a final deliverable of a data science project.

Q1: Based on the videos and the reading material, how would you define a data scientist and data science?

A data scientist is a problem solver who works on data in an analytical fashion in order to explore, discover and further communicate the hidden relations or phenomena behind.

Data science is the practice of collecting, clearing, storing, mining and interpreting data with the ultimate goal of providing policy or strategy recommendations.

Q2: 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?

I studied law and linguistics. I am particularly interested in Natural Language Processing (NLP) in a multilingual environment and I am curious how it may be extended to the legal field.

Q3: 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?

A well-structured data science report should comprise ten main sections as follows: 1) cover page, 2) table of contents, 3) abstract/executive summary, 4) introduction, 5) literature review, 6) methodology, 7) results, 8) discussion, 9) conclusion, and 10) references, acknowledgement, and appendices (if needed).