Question 1

Primarily, Data Science is a process rather than an event. A process in which one attempts to validate a hypothesis or a model, revealing hidden trends and insights that lie within the data, explaining these acumens through storytelling with the goal of generating insight companies and organizations may use to make strategic decisions with favorable outcomes. Starting with curiosity and data, Data Science explores, manipulates and analyzes data to garnish critical answers using algorithms on structured and unstructured data from widely varied sources, in an attempt to add knowledge and insight, providing irreplaceable value to business.

The Data Scientist themselves is a curious and well-rounded individual who demonstrates specific subject matter expertise, understanding of software programming and analytics, with the ability to transform the results of the data science process into an engaging story that companies will need to understand to navigate the ever-changing digital world in which we live. A Data Scientist is well versed in assembling structured and unstructured data, cleaning it, organizing it, and using the tools available to analyze it. All with the purpose of gaining clear and concise insights to develop solutions for varied enterprises.

Question 2

Having worked and trained in risk and compliance for many years, I am confident that a deep knowledge of Data Science would equip me to add greater value and expertise to this field. During my employment at Digital Outsource Services, an outsourced auxiliary to Micro Gaming, I underwent extensive training in risk and compliance, working primarily – but not exclusively – for the Australian market. I utilized these skills when working for Lio-Global, managing risk and compliance, performing due-diligence and KYC (know your client) for the company, and drafting financial risk reports for foreign investors. I have also worked with GoVerify and Alternative Investments, fulfilling similar roles in assisting corporations and private individuals to safeguard their wealth and assets, assuring all investments and purchases were compliant and all risk protocols followed. Data Science would be invaluable to the anti-money laundering and Know Your Client sectors, allowing the data scientist to access, analyze and manipulate the data available to better understand the client base, their needs and preferences, as well as, flag transactions and the flow of money that would indicate illegal attempts to transfer, hide or clean wealth. The use of data science tools would prevent backlogs within the risk and compliance sectors, and do away with manual attempts to harness important data, which leaves room for human error and relies heavily on the risk management team. Instead, these teams would be able to flag and action events based on data science protocols and algorithms, streamlining their efforts and help mitigating the shortfalls of relying on human-based data mining, while extending the reach of the company to access structured and unstructured data. With machine-learning capabilities, the reduced the time needed to process the data would allow AML and KYC teams to focus on high priority matters, rather than the process of data gathering, manipulating and structuring. My interest in data science comes from my experience in these diverse settings, requiring risk and compliance expertise. I am enthusiastic about the application of machine-learning and data science in improving the anti-financial crime

sector and look forward to utilizing the skills I am learning to improve processes and actionable outcomes.

Question 3

The structure will largely be based on the length of the document. A brief report (5 or fewer pages) will be concise and take the form of a summary, swiftly presenting the key findings. An extended documents (of more than 100 pages in length) will build progressively develop the argument, contain details of other relevant work, research methodology, data sources and intermediate findings, concluding with the main results.

Cover Page: The cover page should include the title of the report, names of the authors, their affiliations, contacts, the name of the publisher, and the date of publication. These contact details on the cover page makes it easier for the reader to contact the author(s).

Table of Contents: The table of contents contains the main headings, lists of tables and figures the document will contain.

Abstract: An abstract - or "executive summary" – elucidates the core of the arguments in three paragraphs or less for a brief report, or slightly longer for larger documents.

Introduction: This section presents the conundrum for the reader who may be viewing the topic for the first time, gradually introducing them to the subject matter and the purpose of the intricate details to follow.

Literature Review: Depending on the breadth of the subject matter the literature review may be brief, citing only the most influential authors. Alternatively, if the subject matter is highly nuanced then one should cite the pertinent research, offering adequate context before the analysis. Herein lies the opportunity to highlight any gaps in existing knowledge the analysis will attempt to fill. In this section, the document will formally introduce the research question and hypothesis.

Methodology: Here the document will introduce the research methods and data sources utilized for the analysis. Any new data collection exercise needs comprehensive explanation. Reference to the literature review will bolster choice for variables, data, and methods, ultimately explaining how the research has helped answer your research question.

Results: Here one presents their empirical findings, starting with descriptive statistics and illustrative graphics, culminating in the formal testing of your hypothesis. One may also report empirical techniques used that fall under data mining.

Discussion Section: In this section, one will rely on the influence of narrative to communicate their thesis to the reader(s) and make the numbers comprehensible. Reference to the research question and any knowledge gaps will be identified to the reader(s) and ultimately you will highlight the findings of your analysis.

Conclusion: This generalized section will highlight your specific findings, taking on a marketing approach to promote these findings to the reader so they are not left stranded in the caveats outlined earlier in the document. One may also identify possible developments in research and subsequent applications resulting from the research.

Appendices: If needed, this section will include references and acknowledgments of sources and authors used in the research.