Marco Lewis - Reflective Report

As part of the group, I performed a variety of different tasks. During meetings, I acted as scribe, taking notes which would be shared later with the rest of my group. As part of the data exploration, I made code that would produce stacked bar charts of the data, giving us an initial idea of what features might be useful. For feature selection, I was involved in implementing a few of the methods, including the Chi Square association test, mutual information and multiple correspondence analysis (MCA), which I researched to get a good understanding. These methods were used to find the primary and secondary dependencies and reduce the number of dimensions of the features, which were used in our models. Further, I investigated different ways we could handle null data, but in the end we stuck with simply removing data subjects that had no data in the columns required. Early on, I had begun investigating whether implementing K-means could be used as one of our possible models. In the end the group agreed that K-means wouldn't be a good enough to use as a general model. I contributed to sections of the report that I felt I had a better understanding of, such as in data exploration and the section on MCA.

I feel I have learned a few things throughout the project. One thing I learned is how to deal with non-ideal situations, such as in this project where a lot of data was missing. While discussing with my group, we would discuss how best it would be to solve the problems that were given and new ones that arose. I also now see the importance in spending time exploring various aspects of the dataset and different approaches to selecting features. Further, researching MCA has shown, to me, the importance of researching unknown and new methods. Working in a larger group has been a new experience for me as well, since my degree has been spent working as an individual or in teams of size 2-4 people. This has given me the chance to develop my skills of working in a team and ensuring that my code is understandable.