Team Project Assignment

Obtain a data set from websites like 'Kaggle' or 'KDNuggets', 'data.gov' or from one of your team members that relates to an interesting business or administrative question. If no data are available from administrative settings, then something of interest to the team in a broader context could also make a good topic; for example, national or global economic data, the environment, medical services, and so on.

A good data set will have the following features:

1) The data will contain multiple variables.

2) Prediction will be an appropriate approach for both the question you want to answer and the data that are available.

3) It is not essential, but it would be beneficial if there are a mixture of data types among the variables; that is, both categorical and interval/ratio data, so that you can experiment with coding and interpreting different types.

Project Proposal: Your project proposal can be a brief (one-page) description of the proposed work, the data source and the intended analytical method(s). Your final work may end up requiring quite different methods. Separate guidelines for the proposal will be posted.

Final Report and Presentation: Conduct an analysis of your data and prepare a report on your findings. You have limited time to carry out your analysis and prepare a report, so you may choose to treat this as a preliminary study. Studies like this can become quite complicated. It is not necessary to explore all possible models and designs – just explore a few interesting ones and leave others for recommended future work. You may use any statistical package with which you are comfortable.

Your report should be concise – essentially an extended summary describing your objectives, the analyses, conclusions and recommendations. A maximum of *five pages* plus attachments should be sufficient, but shorter is better! Detailed mathematical details and computer printouts should only be included if they are directly relevant to the report, and they should generally be placed in appendices as exhibits that are cited in the report. Your intended readers are managers who may have little recent knowledge or experience with the technical details of analytics. If necessary, you may assume that one of your readers is a manager with some analytics experience (perhaps your boss!) who will want a record of some of the details. Such details should be in the appendix. Part of the grade for this assignment will be based on the clarity and appropriate level of the writing and professional formatting of the report.

Your submission should include the report (.PDF format please) and an electronic copy of the data set along with output files, if appropriate. Please let me know if the data are confidential, because they may be interesting teaching cases for use in other classes!

Your team will give a short presentation on your project in the final class. You may choose one member of your team to make the presentation or any combination of members, but the result should be clear and professional. The presentation should take no more than fifteen minutes, with an additional five minutes for questions.