Data Analysis Challenge

The aim of the data analysis challenge is to see how you carry out the analysis of a novel set of data and how you report your findings. There is no "right answer" to the challenge, we just want to see how you approach analysing a novel data set. We'd expect you to spend a maximum of 2-3 hours doing the challenge.

Dr Louisa West has given you one time-series dataset (dat1.csv). She says it has 105032 observations measured in 5 different ways (A,B,C,D,E) and a column labelled "id" which represents the time points in ascending order. Unfortunately, Dr James sent you the data from Base Camp in Antarctica then immediately set off for a three-week field trip staying in a tent on the ice and now can only make contact in emergencies. Sadly, she forgot to give you any of the metadata, including the data type in each column and the time interval that observations were made on!

Although you don't know much about where the data came from or what they are, you decide to do some preliminary analysis to understand the properties of the data (problems, trends, correlations and so on) so that you can ask Dr West some questions about the system studied when you are back in touch.

We'd like you to do this preliminary analysis using your favourite tools and techniques and write a short report on how you did this and what you find. The report should include your code and any plots you make and should also include the questions you would ask Dr West on her return. You might choose to make your report available through your github account or other version control system, alternatively, email it to James Abbott (j.abbott@dundee.ac.uk) or let us know where we can download it from by Sunday 16th October.