BU495Q

Project Sheet

Please forward questions to Prof. Michael Pavlin (mpavlin@wlu.ca)

The project will be an important part of the course.

Goal of the project:

1. To acquire hands on experience applying advanced analytic techniques to a business problem you are interested in.
2. To acquire experience formulating a business problem in a quantitative manner.
3. To clearly and persuasively use the results of a quantitative analysis to argue for a particular course of action.

Groups:

The project will be done alone or in groups of 2 (group size will be taken into account in grading).

Audience:

Outside of appendices, all writing should be directed towards a non-technical person that is well versed in the business problem (the person who wrote the HBR "Keeping up with your quants" article is who you should keep in mind).

Deliverables:

The project will be divided into two parts. The proposal and the final report.

1. **The proposal -- 15% of final grade**
   Up to 2 clearly written pages.
   You need to convince the reader that (1) there is a problem and (2) that it is feasible to solve this problem.
   Definition of the Business Problem: This should include a short description of the context
   Expected Results:
   Description of Data: What data do you have access to? Where does it come from?
   Does the data exist to support your conjectures? What do you need to do to put into a format that can be analyzed? Is the quality of data sufficient? How will you assess the quality of the data?
   Risks: What could go wrong? What don't you know about the data? What don't you know about the methods you are using (e.g. do they scale to the size of the dataset you want to use)?

2. **Final Report -- 30% of final grade**
   Up to 4 concise and well argued written pages plus Appendix consisting of up to 4 figures (if required the appendix may include up to a two page technical discussion).
   Describe **briefly** problem, motivation and dataset.
   Describe **actual** analysis and why you chose any techniques (if the course of action differed from the one proposed then explain why).

Describe Results referencing all included figures.

Give recommendations!

All code and (if possible) datasets should be provided in a separate zip or tar.gz file

Formatting:

All documents should have greater than 1inch margins, 11 pt font and be **double spaced**.

Datasets:

- My Data
  - Electricity market data from Midcontinent ISO
- WRDS – financial data,
  - Accounting data (compustat)
  - Stock trading data (CRSP)
  - and more: Analyst statements etc.
- Build your own
  - We will discuss retrieving Tweets using R
- Other sources:
  - marketing data: Dominick's Database http://research.chicagobooth.edu/kilts/marketing-databases/dominicks
  - e.g. statscan, U.S. DOE, municipal governments, academic studies
  - icpsr is a great place to find datasets http://www.icpsr.umich.edu/

If you can't find a particular type of data, let me know!

Finding a question:

Please look for a dataset that is of interest to you and were its analysis may be of interest to future employers! There are two types of question that I would suggest:

a. A data question: the answer to your question lies within your data (see questions below).
b. A methodological question: you are interested in finding a methodology to answer a particular type of question. In this case you may worry more about how a particular algorithm performed. You should compare at least two algorithms/methodologies.

Examples of Potential Questions:

- The COO at Blackberry wants to understand how social media has responded to the release of the Blackberry Classic.
- As a bankrupty lawyer you would like to a head start on finding clients. Could accounting data be used to automatically determine the potential for insolvency in the next year or two?
- Boss wants to do a promotion on a particular SKU, operations department wants to know what that is going to do future demand.

- *Example data question:* You are tasked with finding potential locations for a natural gas peaking plant using pricing data.
- *Example methodological question:* You are a partner in a "green" investment firm devise an automated method for identifying "green" companies.