**Media Marketing data**

This dataset (“weekly\_media\_sample.csv”) contains a sample of weekly media marketing data (spanning a total of 208 weeks), recording revenues generated from various spending on three media. Other variables such as competitors and newsletter subscription are also included.

A brief description of the data variables is as follows:

*X* – Week number;

*DATE* – Date of the week;

*revenue* - Revenue (ROI);

*media1\_S* – Spend on Media 1;

*media2\_S-* Spend on Media 2;

*media3\_S*- Spend on Media 3;

*competitor\_sales* – Competitors sales;

*newsletter* – Number of newsletter subscription (note that this is regarded as a marketing activity with no media spend).

## **Part 1: Classical Marketing Data Modelling (40%)** 1a: Modelling

You are to use the data collected between 2012/8/6 and 2016/5/30 (inclusive) to train your model.

* Build a statistical regression model to determine the effectiveness of marketing strategy in generating revenues. You must carefully justify the choice of your model and the variables used. ***[10 marks]***
* Interpret the results obtained from your model. ***[5 marks]***
* Discuss the limitations of the model used, in the context of media marketing. ***[5 marks]***

## 1b: Prediction

Use the data between the test period ,2016/6/6 and 2016/7/25 (inclusive), to assess your model.

* Use the model built in 1a to predict the marketing revenues for the test period. Plot your answers on a graph, including 95% intervals. ***[5 marks]***
* Assess the predictive performance of your model quantitatively. ***[5 marks]***

## 1c: Temporal effects

A colleague argued that it is important to account for temporal effects in any media marketing dataset.

* Do you agree with this? Explain. ***[2 marks]***
* Explore the inclusion of temporal components in your model. You are free to use any model for this, e.g. time series models such as ARIMA. [***5 marks]***
* Provide a similar analysis as above to check if the model proposed here is suitable. ***[3 marks]***

**Part 2: Bayesian methods (60%)**

## 2a: Bayesian model estimation

Pick a model from Part 1.

* Fit the model in a Bayesian framework using non-informative priors. You must also answer the following questions:  
  - Write down your priors. Why are they non-informative?   
  - Are your priors conjugate?  
  - Which Markov chain Monte Carlo (MCMC) algorithm did you use? Why?  
   ***[15 marks]***
* Have your MCMC chains converged? Provide evidence (e.g. trace plots, ACF plots). ***[5 marks]***
* Produce a summary of the parameters estimated. ***[5 marks]***
* Compare your results with those from Part 1. Are they similar? Explain. ***[5 marks]***

## 2b: Prior choice

* Suppose an expert claimed that Media 3 has no impact in generating marketing revenues. The expert is rather certain about his personal belief, but do not want to remove the variable from the model. Instead, he wished to combine his belief with information learned from the dataset. How would you modify your prior specification to account for this? State and justify your decision carefully. ***[5 marks]***
* Rerun the Bayesian estimation using the prior you proposed. What do you notice from the parameters estimated? ***[5 marks]***
* Additionally, you also wanted to incorporate the following opinions from another expert:   
  - “There is a strictly positive relationship between newsletter and marketing revenues.”  
  - “I am fairly certain that one unit invested in Media 1 will result in two units of marketing revenues returned.”   
  - “I believe the impact of Media 2 is four times that of what I specified for Media 1.”

- “For every unit increase in competitor\_sales, the change on revenues returned must range between 0 and 0.3.”  
  
Modify your prior specification carefully with justifications. ***[10 marks]***

* Re-fit your Bayesian model and report your results. ***[10 marks]***