The Hong Kong University of Science and Technology
The Department of ISOM

ISOM3530 Business Data Analytics
(2019 Spring)

Case study, due on **23:59, May 15, 2019**

You are a data scientist of a Portuguese banking institution. Your directors are interested in **predicting if the client will subscribe a term deposit** in marketing campaigns based on phone call. "bank_marking.csv" contains the data related to the above campaigns. Often, we need to call the clients more than once in order to access if the product would be or not subscribe.

The dataset has the following variables:

Variables	Description	Measurement level
Client data		
age	Age of the client	Numeric
job	Type of job ('admin.','blue-collar','entrepreneur','housemaid','management','retired','self-employed','services','student','technician','unemployed','unknown')	Categorical
marital	Marital status ('divorced','married','single','unknown'; note: 'divorced' means divorced or widowed)	Categorical
education	Education level ('basic.4y','basic.6y','basic.9y','high.school','illiterate','professional.course','university.degree','unknown')	Categorical
default	Do the client have credit default? ('no','yes','unknown')	Categorical
housing	Do the client have housing loan? ('no','yes','unknown')	Categorical
loan	Do the client have personal loan? ('no','yes','unknown')	Categorical
Last phone call related		
contact	Contact communication method ('cellular','telephone')	Categorical
month	Last contact month ('jan', 'feb', 'mar', ..., 'nov', 'dec')	Categorical
day_of_week	Last contact day of the week ('mon','tue','wed','thu','fri')	Categorical
duration	Last contact duration in second	Numeric
Other attributes		
campaign	Number of contacts performed during this campaign for this client	Numeric
pdays	Number of days passed by after the last contact	Numeric
previous	Number of contacts performed before this campaign for this client	Numeric
poutcome	Outcome of the previous campaign ('failure','nonexistent','success')	Categorical
Social and economic indicators		
emp.var.rate	Employment variation rate – quarterly indicator	Numeric
cons.price.idx	Consumer price index – monthly indicator	Numeric
cons.conf.idx	Consumer confidence index – monthly indicator	Numeric
euribor3m	Eurobor 3 month rate – daily indicator	Numeric
nr.employed	Number of employees – quarterly indicator	Numeric
Response		
y	Has the client subscribed a term deposit? ('yes','no')	Binary

Here are the concerns from your directors:

(1) Develop a better strategy for the further campaigns based on the data.
(2) Which features are important for the client to subscribe the product or not?
(3) Which groups of people will be likely to subscribe the product in a campaign?
(4) We need to present the results to the executive committee, so the model cannot be too complicated, and we need good graphics and outputs for this.

GUIDELINES OF REPORT

File Submission:

1. Upload your report, presentation slides, program code and data file on Canvas by the deadline.
2. Submit a hardcopy of report and also a copy of presentation slides right before you start the presentation.

Basic Report Formatting Rules:

1. Page limit: at most 10 pages main content (excluding the Appendices; at most 10 pages for appendices.
2. Font size: use 12-point type in the main text of the report.
3. Spacing: 8 points of lading (about 1 one-and-a-half spacing).
4. Margins: left and right margin should be l" (2.5 cm) on both sides.

HOW TO WRITE REPORTS

Here are some guidelines for you to write your reports.

1. **Project Outline**
    - Title Page: Must contain
        - Project title.
        - Name of group members and group number.
        - Date.
        - Executive Summary. (very short and to the point).
          The executive summary contains a brief account of your conclusion. Write the results not the description of the problem. Say: I found that the IBM price stock was at a higher level in October than in November, in October the mean price was .....

    - Introduction: If any, must be very short. Describe your project and give the background information. Try to be brief but at the same time do not leave out the relevant information. Here you include the basic descriptive statistics, graphs and summaries of the data. Please if you have several graphs or large tables put them in an appendix and refer to them on the text.

    - Analysis & Results: State your points or hypotheses and prove them or disprove them. Go point by point showing performing the corresponding hypothesis tests and interpreting the results. If you use a computer software that generates a sizable output file, put the output file as an appendix and refer to the output here for performing your hypothesis tests.

    - Conclusion: Like the executive summary but longer.

    - References: List of books or articles that are cited on the text. If you do not cite any then omit it (you are not required to have it).

- Appendix with computer output, computer generated or hand- made graphs. Try to underline and number the places in the output that are quoted in the above text, and use those numbers to refer to that piece of output in the text.

2. Writing your report

- Spelling and grammar are very important.

- Paragraph Structure: When you write a paragraph try to be organized. A way to do this is to start by saying the thesis, then follow with a list of arguments that prove it, quoting the output, performing hypothesis tests, etc. Finally reinforce the result.
    - Say the thesis.
    - Prove it.
    - Say it again.
- Another way to construct paragraphs is by using concatenation: from A prove B, from B prove C, and so on.

- The reports that you will write in this class will be no more than 3 or 4 pages of text plus a few more of output. The output should be organized in tables + charts, but never append a thick SAS printout. However, a reasonable report should be as brief as possible.

3. Graphs and charts are essential to display information

- Graphs and charts are essential to display information.