Project Problems for Bus Q600 (Fall 2013)

- See the Instructions at the Project link <`http://profs.degroote.mcmaster.ca/ads/parlar/courses/q600/Project/index.html`>

- **Submissions of the project by email WILL NOT BE ACCEPTED!**

- **Late submissions of the project after the final exam date WILL NOT BE ACCEPTED!**

The Excel sheets mentioned below are part of the Excel file you will find in the Data Set section of the Project link <`http://profs.degroote.mcmaster.ca/ads/parlar/courses/q600/Project/index.html`>.

Problems	Maximum Mark
1	15
2	20
3	20
4	15
5	20
6	10
Total	**100**

Problem 1 (Income Distribution) [15 marks]

Refer to the Excel sheet named <`1. Income Distribution`>.

This sheet contains the (last year's) total incomes of 1,000 people aged 25. This group is assumed to constitute the population under study.

(a) [5 marks] Describe the distribution of the incomes of these 1,000 people. Include the mean, population standard deviation, minimum, maximum, range and a box plot.

(b) [5 marks] Choose five samples each of size 30 as follows: (i) Sample 1: ID 1–30, (ii) Sample 2: ID: 101-130, (iii) Sample 3: ID: 201–230, (iv) Sample 4: ID: 501-530 and (v) Sample 5: ID: 701-730. (We assume that these are random samples.)

- Calculate the sample mean $\bar{x}$ for each of these samples.

- Find the mean of all sample means and compare it to its theoretical value. Is there a discrepancy between the two? Why?

(c) [5 marks] The government has decided to provide financial aid to all 1,000 individuals in this study using a "lottery." If the average (mean) income of a randomly selected 36 individuals is less than \$20,000, the 1,000 individuals will each receive a cheque for \$400. What is the probability of this outcome?

Problem 2 (Cost Centres) [20 marks]

Refer to the Excel sheet named <2. Cost Centres>.

A facility (such as a hospital) has three cost centres for ordering the supplies it needs. Each cost centre has its own monthly budget that must be adhered to. Furthermore, you cannot order supplies under another cost centre if you have exceeded your supply budget.

The budgets for each of the three cost centres and the actual monthly expenditures that occurred for the 2012 calendar year are given in the Excel spreadsheet.

(a) [15 marks] We would like to know if the budgets were appropriate. If not, what budget levels would you recommend for each cost centre? If your recommendation is implemented, there should be a "small" chance of the true (but unknown) average monthly expenditures exceeding the budget.

(b) [5 marks] In your analysis in Part (a) above, how large should the sample size be to ensure that your margin of error is \$1,000?

Problem 3 (Travel Expenditures) [20 marks]

Refer to the Excel sheets named <3. Travel Expenditures-A & B> (for parts a and b below) and <3. Travel Expenditures-C> (for part c below).

Do you agree with the following statements on European and North American travel expenditures based on the data for the period from 1988 to 2010?

(a) [6 marks] Based on the European Leisure Travel and Business Travel expenditures, European tourism industry has been primarily dependent on leisure travellers.

(b) [6 marks] Based on the North American Leisure Travel and Business Travel expenditures, North American tourism industry has been primarily dependent on business travellers.

(c) [8 marks] On average, Europe gets more Visitor Exports (international tourist expenditures) than North America gets from Domestic Travel and Tourism expenditures.

Problem 4 (Healthy vs. Failed Companies) [15 marks]

Refer to the Excel sheet <4. Healthy vs. Failed>.

How do the companies that fail differ from those that continue to do business? One study compared several characteristics of 68 healthy and 33 failed firms, where one variable was the ratio of current assets to current liabilities (columns named "ratio"). This is, roughly speaking, what the firm is worth divided by what it owes.

Take Group 1 as the healthy firms and Group 2 as those that failed. The question is whether or not the mean ratio of current assets to current liabilities is different for the two groups.

(a) [5 marks] Analyze the data for each group and determine whether or not there are outliers or strong departures from normality.

(b) [5 marks] Do mean asset/liability ratios differ for the two groups? Explain.

(c) [5 marks] What is the difference in mean asset/liability ratios? Which interval contains the true (but unknown) difference in mean asset/liability ratios?

Problem 5 (OPEC Countries) [20 marks]

Refer to the Excel sheet named <5. OPEC Countries>.

OPEC countries derive their main income from oil and natural gas. Consider the data for the last 40 years given in the Excel spreadsheet (some with missing data).

(a) [10 marks] Consider the Oil Rents data: It is claimed that the average percentage oil rents (as % of GDP) for the OPEC countries listed is the same. Is this statement true? Explain!

(b) [10 marks] Consider the Gas Rents data: It is also claimed that the average percentage natural gas rents (as % of GDP) of the OPEC countries listed are the same. Is this statement true? Explain!

Problem 6 (DJIA) [10 marks]

Refer to the Excel sheet named <6. DJIA>.

This question is concerned with some characteristics (assets, sales and profits) of the 30 stocks in the Dow Jones Industrial Average (DJIA). (All figures in billions of dollars.) We would like to examine how the profits are related to assets and sales.

1. [3 marks] Consider assets. How many extreme outliers do you see? What is the mean, standard deviation and median? Is the distribution skewed? In what direction? Now consider the sales and perform the same analysis as you did for assets. Are the companies that are outliers in assets also outliers in sales? Are you surprised that Wal-Mart higher sales relative to assets?

2. [2 marks] Consider profits. How many extreme outliers do you see? What is the mean, standard deviation and median? Is the distribution skewed? In what direction?

3. [5 marks] How do you predict profits from, (i) assets only, (ii) sales only, and (iii) assets and sales? For each case, explain why or why not your model is significant.

   (a) For (i), let Assets = 50 and predict Profits.
   (b) For (ii), let Sales = 40 and predict Profits.
   (c) For (iii), let Assets = 50 and Sales = 40, and predict Profits. Which model gives you more accurate results?

4