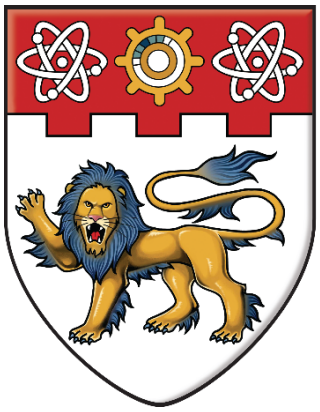

NANYANG TECHNOLOGICAL UNIVERSITY

CS2400 Individual Assignment
Multiple and Polynomial Regression
March 2025

Wee Kim Wee School of Communication and Information
Nanyang Technological University
Singapore

CS2400 Individual Assignment (15%)

Deadline: 14 April 2025

This assignment requires you to perform multiple regression and polynomial regression on a dataset. Study the dataset that has been uploaded to NTULearn. The dataset comprises the death rate (the dependent variable) and fifteen independent variables:

Part A (Multiple Regression)

Find a multiple regression model to fit the relationship between two independent variables (select two out of the fifteen independent variables available) and the death rate (the dependent variable). Justify the selection of the independent variables. The regression model to be used is:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \epsilon$$

Part B (Polynomial Regression)

Find a polynomial regression model to fit the relationship between one independent variables (select one out of the fifteen independent variables available) and the death rate (the dependent variable). Use the following regression model:

$$y = b_0 + b_1 x + b_2 x^2 + \epsilon$$

Report

Write up your findings in a short 100- to 150-word report, presenting:

I the primary equations (also called the normal equations) for your regression (in both equation and matrix formats);

Remember that both the equation format and the matrix format of the primary equations are two ways of expressing the same information. One is not better than the other. The matrix format is preferred only because Cramer's Rule (which is typically used to solve the simultaneous equations) requires the equations to be expressed in the matrix format.

II an explanation for your choice of variables; and

III the results of your regression analysis (the final equation) for Parts A and B.

For submission, upload two files to NTULearn:

I A Microsoft Excel file containing your calculations. Cramer's Rule should be used to solve the simultaneous equations, and the residual (ϵ_i) for each point should be calculated. A cover page that clearly indicates your name (as reflected in the matriculation card), and your matriculation number should be included.

II A Microsoft Word file containing (1) the short 100- to 150-word report summarising your findings; and (2) the 350- to 400-word memo for your employers, an agency of the United Nations.

In your memo, suggest three variables that could improve the prediction of death rates. Death rates is a critical parameter for countries that are ageing. If the number of deaths in a population outnumber the number of births, the population will decline. This has serious knock-on effects on the economy (there will be fewer workers in the economy), the tax collection (fewer workers means less tax collected), and even the defence of the country (fewer able-bodied men of the right age to defend the country). This is why being able to predict death rates is important, if a little morbid.

One variable that you can suggest is access to healthcare, indicated by the number of doctors per thousand population[i]. A smaller number of doctors indicates that each doctor must attend to more people, a parameter that gains importance as a population ages. Think of **two more variables** that will be useful in predicting the death rate but not available in the dataset.

A frequently asked question

The same question asked in several ways:

- Can I "reuse" an independent variable used in Part A (Multiple Regression) again in Part B?
- Can I use an independent variable twice (in Part A and Part B)?
- Does the independent variable in Part B have to be different (from the ones used in Part A)?

Yes, you can reuse the same independent variable in Part A and Part B.

Uploading Instructions

1. Zip your Microsoft Word and Microsoft Excel files together into a single file.
2. Rename the zipped file. Use the name on your matriculation card as the filename, e.g., `Tan_Ah_Kow.zip`.
3. Upload the renamed zipped file to the correct folder.

Important Note

Remember to use the name reflected in your matriculation card for zip file. This is important to ensure that the correct marks are keyed in.

[i] The World Bank frequently reports this. Looks at this page.