Tutorial 2

STAT 3013/4027/8027

---

1. Answer the following questions from SI: 2.1, 2.2, 2.4.
2. A text file named 'gdp.txt' is on the course Wattle site and provides the 2013 Gross Domestic Product (GDP) values for 214 nations from the World Bank (http://data.worldbank.org). The World Bank states:

   > GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in current U.S. dollars. Dollar figures for GDP are converted from domestic currencies using single year official exchange rates. For a few countries where the official exchange rate does not reflect the rate effectively applied to actual foreign exchange transactions, an alternative conversion factor is used.

- Consider the following:
  a. Visually display the data and discuss. Try taking the natural log of the data (when statisticians say "log" they mean natural log).
  b. Compute a six number summary of the data.
  c. Based on the "box plot rule", determine if there are any outliers. Which countries are outliers? To use the rule examine the following: Are any values in the data below the 1st Quartile - 1.5 IQR? Are any values in the data above the 3rd Quartile + 1.5 IQR? IQR is the inter-quartile range.
  d. Let $Y = \log(\text{GDP})$. Suppose $Y \sim \text{normal}(\mu, \sigma^2)$. What is your best guess for μ and σ^2 as functions of Y (call these T_1 and T_2)? What are the means (expected values) of T_1 and T_2?