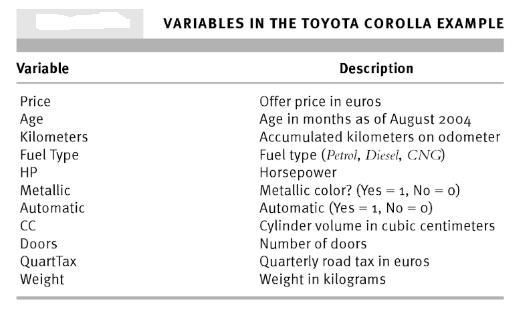
**Homework 2 (Due: Monday, Feb 25)**  
  
**Guideline:** **Please include the screen shot of R-syntax and the results as presented in Lab-3 for each part of the question except 2(a) for which you just provide R-syntax.**-------------------------------------------------------------------------------------------------------------------------------

1. The following represents the ages of the 50 richest people in the world in 2009.  
  
89, 89, 87, 86, 86, 85, 83, 83, 82, 81, 80, 78, 78, 77, 76, 73, 73, 73, 72, 69, 69, 68, 67, 66, 66, 65, 65, 64, 63, 61, 61, 60, 59, 58, 57, 56, 54, 54, 53, 53, 51, 51, 49, 47, 46, 44, 43, 42, 36, 35

a. Compute the measures of central tendency (mean, median and mode).  
  
b. Compute the measures of spreadness (range, variance & standard deviation).  
  
c. Find the quartiles.  
  
d. Find the five number summary.  
  
e. Make a box plot of the data with appropriate title, x-label, y-label. Describe the shape of the data distribution.

2. A large Toyota car dealership offers purchasers of new Toyota cars the option to buy their used car as part of a trade-in. In particular, a new promotion promises to pay high prices for used Toyota Corolla cars for purchasers of a new car. The dealer then sells the used cars for a small profit. To ensure a reasonable profit, the dealer needs to be able to predict the price that the dealership will get for the used cars. For that reason, data were collected on all previous sales of used Toyota Corollas at the dealership. The data include the sales price and other information on the car, such as its age, mileage, fuel type, engine size, etc. The description of the part of the variables is as follows  
  
  
The data file is **ToyotaCorolla.csv**. In this data set I would like you to perform the following tasks:  
  
a. Read the data file in R to create a SAS data set. Name it **TCor**.

b. Compute descriptive statistic (mean, median, mode, variance, standard deviation and quartiles) of Price variable.  
  
c. Compute descriptive statistic of **Price** variable by grouping variable **Fuel\_Type**.  
  
(**Hint:** you may want to google about this topic as: “descriptive statistic by grouping variable in R”)  
  
d. Construct a box-plot of Price variable.

e. Construct Box plots of Price variable by grouping variable **Fuel\_Type**.  
  
  
f. Create a subset (with name: **TcorSub**) of **Tcor** that contains data for Price (col=3), Age\_08\_04(col=4),

KM(col=7).  
  
g. Use **apply** to compute five number summary of all the variables in TcorSub.  
  
h. Create three data subsets (using Fuel\_Type as grouping variable). This means first subset contains data for on all variables for Fuel\_Type = Petrol, second one contas data on all variables for Fuel\_Type = Diesel and last one contains data for Fuel\_Type = GNG.

i. Determine five number summary of Price variable from each of the data set, you created in (h).