Assignment 1

In this assignment you will practice using R for data analytics. Please write R code to analyze a cancer patient dataset and answer the following questions.

1. **load the following data set**
   https://stats.idre.ucla.edu/stat/data/rdm/patient_pt2_dm.csv
   (Hint: read.cvs() function, please see textbook page 23 for explanations and examples.)
2. **display the summary of the data**
   (Hint: summary() function, please see textbook page 23, 52, 55 and 56 for examples)
3. **display the first six observations**
   (Hint: head() function, please see textbook page 55 and 56 for examples)
4. **display observations where the patients' ages are greater than 50**.
   (Hints: subset() function, please see textbook page 29 and 30 for explanations and examples)
5. **display only two columns of data: tumorsize and cancerstage**
   (Hints: data frame indexing, please see textbook page 25 for explanations)
6. **computer the mean of age**
   (Hint: mean() function, please see textbook page 31 and 52 for examples.)
7. **computer the correlation between tumorsize and age**
   (Hint: cor() function, please see textbook page 31 for explanations)
8. **plot a histogram for bmi**
   (Hint: hist() function, please see textbook page 52 and 55 for examples)
9. **plot a scatterplot for age and tumorsize**
   (Hint: plot() function, please see textbook page 39 for explanations)
10. **display frequency table for ntumors**
    (Hint: table() function, please see textbook page 29 and 33 for examples)

Note that you can find hints for all questions, including function names or textbook page numbers. If you don't know how to use a function, please try the following

(1) use HELP() or ? to get details of the function in R
(2) search the function in your textbook.
(3) Google it
(4) search it in the online books posted on Blackboard "R Learning Resources" section.
(5) ask about it on Blackboard General Help forum

Please submit your answers in a MS Word file generated by R Markdown, which includes questions, R code, and the results.