Lab Exercises

1. Given is the case of Breast Cancer in our folder ***Cancer***(Cancer.csv). There are two classes of breast cancer namely recurrence-events and no-recurrence-events. Apply appropriate model to the data and find the accuracy, sensitivity and specificity of it.

2. Given is the Bank Marketing in the folder Bank. The following are the input and output variables respectively:

Input variables:	
# bank client data:	
1	age (numeric)
2	job : type of job (categorical: "admin.","unknown","unemployed","management","housemaid","entrepreneur","student","blue collar","selfemployed","retired","technician","services")
3	marital : marital status (categorical: "married","divorced","single"; note: "divorced" means divorced or widowed)
4	education (categorical: "unknown","secondary","primary","tertiary")
5	default: has credit in default? (binary: "yes","no")
6	balance: average yearly balance, in euros (numeric)
7	housing: has housing loan? (binary: "yes","no")
8	loan: has personal loan? (binary: "yes","no")
# related with the last contact of the current campaign:	
9	contact: contact communication type (categorical: "unknown","telephone","cellular")
10	day: last contact day of the month (numeric)
11	month: last contact month of year (categorical: "jan", "feb", "mar", ..., "nov", "dec")
12	duration: last contact duration, in seconds (numeric)
# other attributes:	
13	campaign: number of contacts performed during this campaign and for this client (numeric, includes last contact)
14	pdays: number of days that passed by after the client was last contacted from a previous campaign (numeric,
15	previous: number of contacts performed before this campaign and for this client (numeric)
16	poutcome: outcome of the previous marketing campaign (categorical: "unknown","other","failure","success")
Output variable (desired target):	
17	y :has the client subscribed a term deposit? (binary: "yes","no")

Use the dataset **bank-full.csv**. You can choose the breaks in the cut() function to be applied for numerical variables which need to be converted to Categorical variables. Find the accuracy, sensitivity and specificity. Also draw the ROC curve stating the AUC.