

# Project Description & Requirement

#	Attribute	Domain
-----		
1.	Sample code number	id number
2.	Clump Thickness	1 - 10
3.	Uniformity of Cell Size	1 - 10
4.	Uniformity of Cell Shape	1 - 10
5.	Marginal Adhesion	1 - 10
6.	Single Epithelial Cell Size	1 - 10
7.	Bare Nuclei	1 - 10
8.	Bland Chromatin	1 - 10
9.	Normal Nucleoli	1 - 10
10.	Mitoses	1 - 10
11.	Class:	(2 for benign, 4 for malignant)

Note: Missing attribute values: 16

There are 16 instances that contain a single missing (i.e., unavailable) attribute value, now denoted by "?". You may just delete those instances.

Note: 11 attribute "Class" is our response.

Project Requirements:

Write a report on your findings and try to answer the following questions:

1. Is the full model (including all 10 features) good? Is there any non-important feature? If so, how do you remove them?
2. Propose the best model in your mind. (hint: there could be many criteria like prediction error, goodness of fit and etc.) Remember there may not be unique solution for this question.