**Description of Breast Cancer Dataset**

1. Number of Instances: 699 (as of 15 July 1992)  
  
2. Number of Attributes: 10 plus the class attribute  
  
3. Attribute Information: (class attribute has been moved to last column)  
  
 # 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)  
  
4. Missing attribute values: 16  
  
 There are 16 instances in Groups 1 to 6 that contain a single missing   
 (i.e., unavailable) attribute value, now denoted by "?".   
  
5. Class distribution:  
   
 Benign: 458 (65.5%)  
 Malignant: 241 (34.5%)