Data Dictionary

1. Id – unique id for each record
2. Diagnosis – whether Benign(B) or Malignant(M)

*Ten real-valued features are computed for each cell nucleus*:

1. Radius - mean of distances from center to points on the perimeter
2. Texture - standard deviation of gray-scale values
3. Perimeter – distance around the nuclear border
4. Area – no. of pixels in the interior of the nuclear border & adding one half of the pixels on the perimeter
5. Smoothness - mean of the local variation in radius lengths
6. Compactness - perimeter^2 / area - 1.0
7. Concavity - severity of concave portions of the contour
8. Concave.points - number of concave portions of the contour
9. Symmetry - relative difference in length between line segments perpendicular to and on either side of the major axis.
10. Fractal\_dimension - “coastline approximation” – nuclear border irregularity

**NOTE:**

Se – standard error

Worst – worst cell/largest cell

Mean – simple mean