# Data Challenge #1 - Breast Cancer Cell Detection

Please limit yourself to 4 hours time!  
  
Your task is to develop a model that predicts whether a biopsied breast cell is benign (not harmful) or malignant (cancerous), given a set of attributes about the cell.  
  
Even though this seems like a fairly straightforward task, there are so many ways you can explore, visualize, engineer your features, and tell a story with this data. It will be interesting to see the different approaches people take! Some important questions to think about are: how can you best explore the data? Why did you select your particular model? How did you validate your model? What insights can you derive from your results?  
  
Follow data challenge best practices. Pay particular attention to how you present your findings - communicate your critical thinking, tell a data story  
  
Please code and annotate your analysis in a Jupyter notebook. Please place your submission in the submission folder.  
  
The dataset consists of 699 cells for which you have the following features:  
  
Sample code number: id number  
Clump Thickness: 1 - 10  
Uniformity of Cell Size: 1 - 10  
Uniformity of Cell Shape: 1 - 10  
Marginal Adhesion: 1 - 10  
Single Epithelial Cell Size: 1 - 10  
Bare Nuclei: 1 - 10  
Bland Chromatin: 1 - 10  
Normal Nucleoli: 1 - 10  
Mitoses: 1 - 10  
Class: (2 for benign, 4 for malignant)  
The dataset is also available here: https://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/breast-cancer-wisconsin.data  
  
Please push your submissions to the Submissions sub-folder with the naming convention: <fname>\_<lname>\_DC1.