Data Processing

1. Deal with NaN
2. Text Data?
3. Binarization
4. Scaling/Normalization
5. Scatterplots/Histograms
   1. Clustering
   2. Informs feature extraction and model selection
6. Feature extraction (x^2, xy, xy^2)
7. Model fitting
   1. Model selection (machine learning)
8. Model testing
   1. Cross validation
   2. Test sets
9. Done or go to 5