Huang, Jonathan
CS5785: Modern Analytics
September 9, 2014

CS5785: Homework 0

1 Setting Up Python

I signed up for Kaggle, installed Python (Canopy), and got it working with my preferred IDE (Sublime Text). On an interesting note, it turns out that getting Sublime to play nice with Canopy is a little annoying.

2 Iris Flowers

1 In this dataset, there are 4 features/attributes per sample (in addition to the specific class for each entry), 3 different iris classes, and 50 samples of each class recorded.

2 The code for this problem can be found in HW0.py (which should be attached to this submission). Specifically, the parsing code is found on lines 5 (for the N x p array containing attributes) and 6 (for the N vector containing species).

3 The plot for this problem is on the following page. The code used to generate this plot can be found in HW0.py. Specifically the code used was lines 14 to 38.
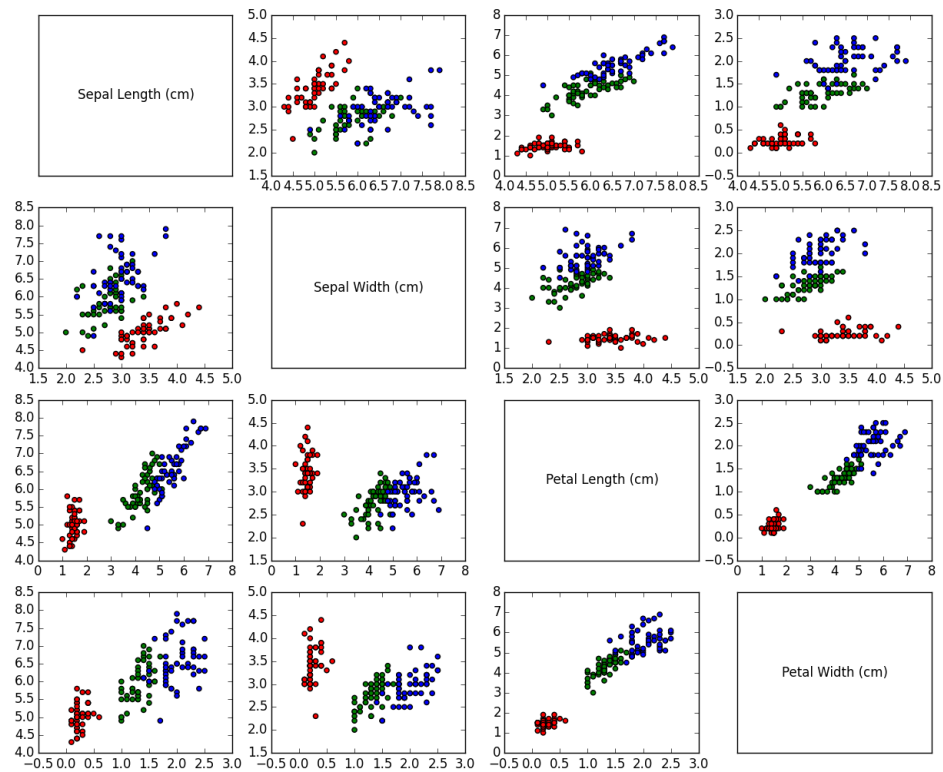
Figure 1: Plot 2.3: Iris Data (red = setosa, green = versicolor, blue = virginica)