print "Istanbul"

**datafile='Istanbul.csv'**

**with util.get\_learner\_data\_file(datafile) as f:**

**alldata = np.genfromtxt(f,delimiter=',')**

**# Skip the date column and header row if we're working on Istanbul data**

**if datafile == 'Istanbul.csv':**

**alldata = alldata[1:,1:]**

data=alldata[1:,1:]

#print data.shape[0]

**data=np.random.permutation(data)**

easiest way imo, initialize lists to collect leaf size, outrmse, inrmse

2 nested loops first for I in range 1-51 to iterate over leaf size

Second loop is to run 20 times at each leaf size (only recommended for rt and bagging to smooth out lines

They can put into a data frame and average by leaf size or do another list to collect the averaged after the 2nd loop

After putting into a df, they can plot (might be best to make the leaf size the index

They will need to collect the data for their experiments. They should definitely do dt, bagging and rt