Week 11 assignment.

Write 3 scripts, each of which generates a result file. Turn in 7 files (3 scripts, 3 results file, and csv file for part 2)

Part 1, Numpy (use numpy operations, i.e. no loops).

Write a script that:

- Creates a 2D array (10 rows, 7 columns) with random floating point values between 0 and 10 and print the integer parts of the array (there are several ways to do this)
  - Prints this to the results file.
- Use np.arange() to make a 1D array with 12 elements, then use "in place" conditions to change the array so that all elements greater than 3 and less than 9 are replaced by their negative values.
  - Print the array to the results file.

Part 2. Pandas (use pandas operations, i.e. no loops)

- Find a data set in .csv or .tsv format with a first row containing headers
- Write a script that loads the data into a dataframe
- Write the shape of the dataframe and the names of the columns to a results file.
- Write out the datatypes for each column to a file.
- Calculate using the pandas.DataFrame.corr() method the correlation between all of the pairs of columns that contain numerical data.   Write the correlations to the results file. (if some columns are mostly numeric, but there is missing data, then .corr() may not work. In this case you need to convert missing data to NaNs).

Part 3 more pandas

Write a script that loads ENS_Genic_Codon_Branch_ADJ_Residuals.tsv using '\t' as the separator

- This contains the branch lengths for a lot of genes on a tree of 17 species of mammals.
- Replace all the 0's with NaNs (do this in place, read about nans)
- Build a new dataframe with one row per gene, and one column that contains the mean branch lengths.
- Print the shape of this dataframe and the mean of all the mean branch lengths, across genes, to a file.