**ECE 3822: Engineering Computation II**

# Homework No. 13: Python Programming

The goal of this homework is to reinforce your ability to read and parse data in Python using industry-standard file formats.

Write a Python script that creates four different Microsoft Excel files that have 3, 100, 1000 and 100,000 rows respectively. Each row contains the following entries: a random number, a valid URL, a first name, optional middle initial and a last name (e.g., “Mary A. Smith”, “John Doe”) and an address (“1949 North 12th Street”). A typical row in the spreadsheet would look like this:



Note that the creation of these files is an important part of this assignment. You will have to be creative in how you find this data and ingest it into your program. In your code in comments at the top of the main program file, briefly describe how you did this (where you got the data from, how you preprocessed it, etc.).

Next, write a Python program to read an arbitrary number of Excel files from the command line, load the data and sort it into alphabetical order based on the last name appearing in Column C. If two last names are identical, use the first name and the middle initial. Print the sorted list to stdout. Incorporate a –help option into your program.

Note that you are sorting all of the files into one list – not each file individually.

Display the amount of processing time required using Python’s built-in profiling tools. Comment on the amount of time used as a function of the total number of rows in a spreadsheet. Is the relationship linear? If so, how does Python achieve such efficiency? Add this information to the top of your main program as comments.

Submit your solution in this directory: /data/courses/ece\_3822/current/homework/hw\_13/<lastname>. The driver programs should have a name “my\_prog1.py” and “my\_prog2.py” respectively.