

## ECON 8320 – Tools for Data Analysis

### Assignment 6 [25 points]

---

Many of the functions that we perform in Python can be executed in parallel for increased performance. Let's start by re-writing one of the algorithms from Homework 2 using the `multiprocessing` library. The functions written in this assignment should check the number of available processor cores, and divide work evenly between the available cores. Parallelize the following functions:

1. Find the largest (or smallest) number in a list of any size
2. Use the `csv` module (documentation at <https://docs.python.org/3/library/csv.html#module-csv>) to read a csv file, and compute summary statistics for the third column. Summary statistics that should be calculated are minimum, maximum, mean, and standard deviation.

Note: If you wish to parallelize the computation of mean and standard deviation (as you would for a large dataset), this blog post explains how to update means and standard deviations when incorporating a new batch of data: <http://notmatthancock.github.io/2017/03/23/simple-batch-stat-updates.html>