Readme

We used the same implementation to sort CSVs, and used socket to communicate between clients and the server. When a client first contacts the server, it sends '>' followed by the column it wants to sort. The server will then record information and assign the client a session id. Once the client received the session id, it searches through directory and sends out all CSVs found with the session id and file size, so that the server knows which client it is and how much to read. We implemented a socket pool for the client using semaphores. The default maximum number of available sockets is 1, but it can be changed by the "-s" argument. After sending all CSVs, the client then sends '<' with its session id as the dump request to the server. The server would then merge, send back, and free all the data of that client. We assume that the command line arguments are in correct format, and the directory has at least one csv file.

We did both extra credit 1 and 2, implementations are stated above.

Please grade the EC2 version first. The professor said we can turn them in separately. If something crashes really bad, please use the original one. Sorry for the inconvenience.