If a dataset is larger than the size of your RAM, you might still be able to analyze the data on a single computer. By default, the Python pandas library will read in an entire dataset from disk into memory. If the dataset is larger than your computer's memory, the program won't work.

However, the Python pandas library can read in a file in smaller chunks. Thus, if you were going to calculate summary statistics about the dataset such as a sum or count, you could read in a part of the dataset at a time and accumulate the sum or count.

Here is an example of how this works.