A lot of data is being generated online as a by-product of human behavior. These data can give interesting insights for social scientists. However, since these data are often stored in large \textit{raw} log data tables, collecting and analyzing the data is not straightforward for those inexperienced in handling such files. In the current chapter, we outline a procedure for pre-processing raw data into an analyzable format. As our example, we use data from students in a learning management system. Data from are often used to gain insight in students’ online learning behavior (\textit{learning analytics}). The chapter provides a hands-on manual for pre-processing data from Moodle, using the open-source software, scripts are provided for every pre-processing step, and we highlight decisions that have to be made during the process. The whole process is designed in such a way that researchers with little background in advanced data manipulation techniques are able to transform their own data with relatively little own input. As a proof of concept, the chapter ends by using the pre-processed data for a common analysis in the field of learning analytics: predicting student performance.