**Problem:** Multiple .csv files of data was generated by the software for a commonly used lab test. Due to the nature of the test,

**Solution:** A script and web application was written in R and R Shiny, respectively, that utilized parsing and text mining principles (via the *grep()* and *gsub()* functions in the *stringr* package) to extract the useful data, and then produce the desired calculations and statistical analysis on the data set.

The user is allowed to enter the number of replicates per sample treatment group, which facilitates statistical calculations. The app accounts for two different formats found in the .txt files, which were generated by two different test machines (“Left” and “Right”). The user interface (UI) is seen in Figs. 2 and 3.

**Benefit:** A .csv file was produced as output. The .csv file, shown as Fig. 4, represents the important data and calculations for all samples in a functional and readable format. The resulting output eliminated the need for manual data entry by lab technicians, significantly reducing both the time spent and risk of errors in data preparation for lab reports.

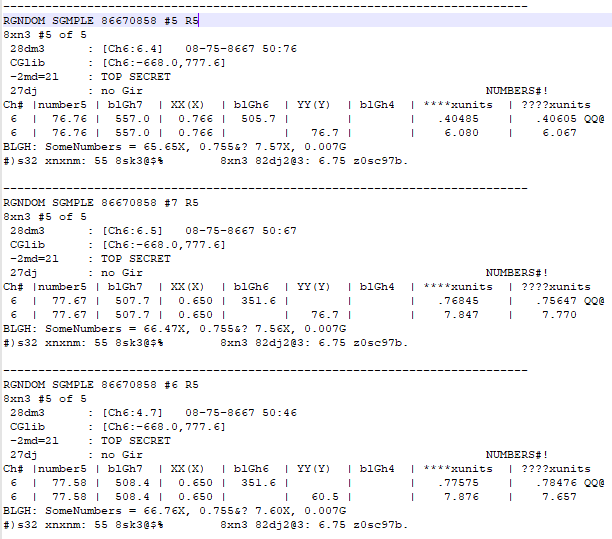


Figure 1: Original form of test data in .txt file, from which data extraction was quite unenjoyable.

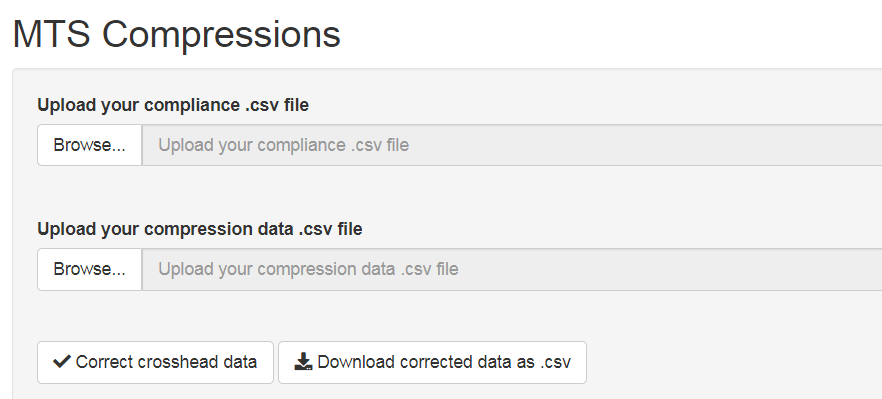


Figure 2: User interface of application. User uploads two .csv files, a compliance data file and compression data file. Once both files are uploaded, selects the machine from which the data was generated, and selects the appropriate number of replicates via slider.

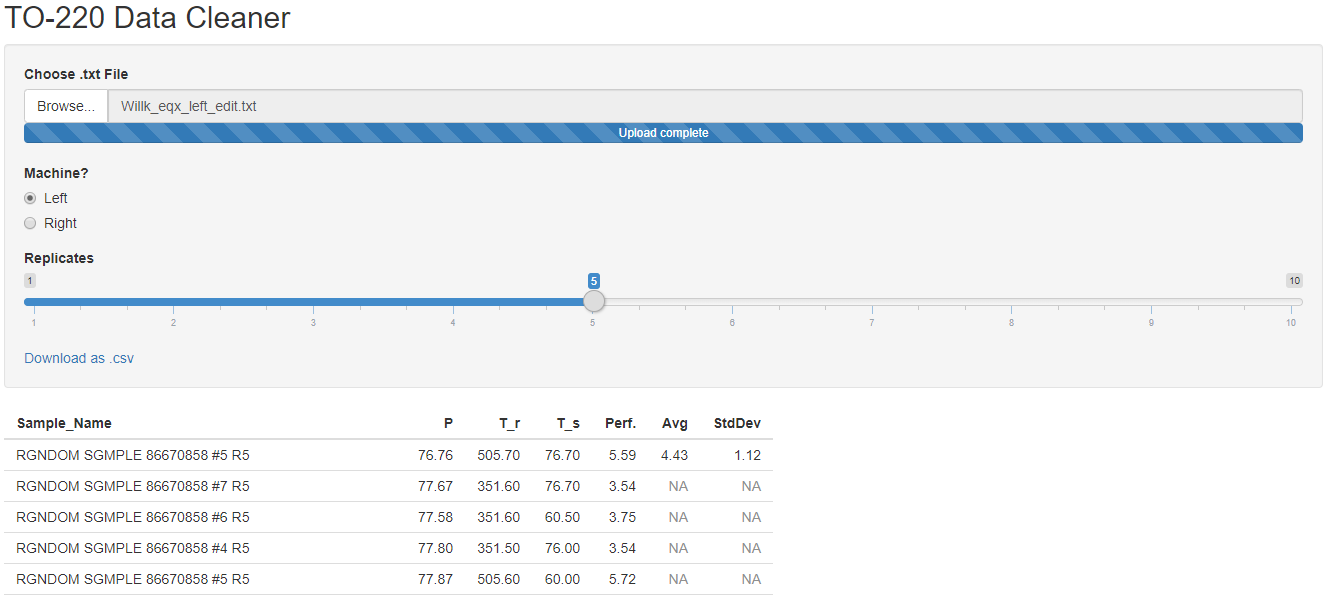


Figure 3: Once file is uploaded, a preview of the parsed data is shown below the user interface. The preview adjusts itself as the number of replicates varies. When satisfied, the user is able to download a .csv of the clean data.

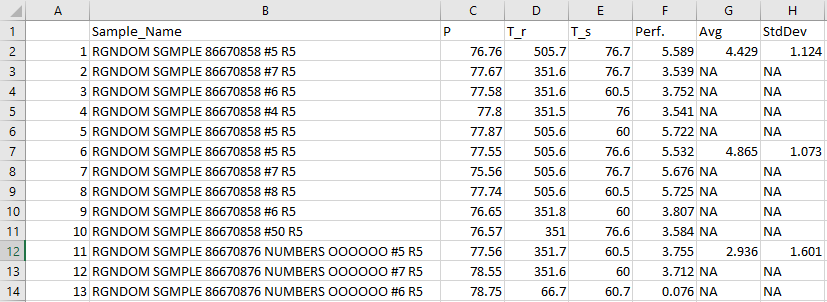


Figure 4: Output .csv in Excel which allowed readable record-keeping. Data includes calculations of interest and basic descriptive statistics from which Excel graphs were easily generated.