Instructions on how to run the scripts:

1- Install the required packages and libraries.
2- Run the list of functions in DataProcessing.r file.
3- Read the .csv file, using readCSV().
4- EventLogs is the dataset that is used as the input from the beginning of the script and the output of each function is the input for the next one.
5- Clean the column headers from spaces and other special characters, using the cleanHeaders().
6- Select the columns that are required for the analysis using the selectColumns().
7- Filter the rows according to the values of the different attributes/columns, this will depend on the analysis. For example, removing events with certain values or the empty values.
8- Remove rows with low frequency using removeEventsLowFrequency(), if these records will not add value to the analysis.
9- Remove the cases with few events using deleteTraceLengthLessThan(), for example if a case have only 3 events and if these cases do not add value to the analysis.
10- Remove incomplete traces using deleteTruncatedTracesStart() and deleteTruncatedTracesEnd().
11- Remove all events with a certain values using deleteAllEvents(), specifying the name of the event. This function can be used several times depending on the analysis questions.
12- Concatenate two columns into one column using concatenateColumns(), for example concatenate two columns which will result in a new column that will be used in the analysis.
13- Keep only the first occurrence of an event using keepFirstEvent().
14- Keep only the last occurrence of an event using keepLastEvent().
15- Depending on the type of the analysis and the questions that the analysis want to answer, some of these functions can be executed multiple times.