# List of String Manipulation Functions in R

|  |  |
| --- | --- |
| **Functions** | **Description** |
| nchar() | It counts the number of characters in a string or vector. In the stringr package, it's substitute function is str\_length() |
| tolower() | It converts a string to the lower case. Alternatively, you can also use the str\_to\_lower() function |
| toupper() | It converts a string to the upper case. Alternatively, you can also use the str\_to\_upper() function |
| chartr() | It is used to replace each character in a string. Alternatively, you can use str\_replace() function to replace a complete string |
| substr() | It is used to extract parts of a string. Start and end positions need to be specified. Alternatively, you can use the str\_sub() function |
| setdiff() | It is used to determine the difference between two vectors |
| setequal() | It is used to check if the two vectors have the same string values |
| abbreviate() | It is used to abbreviate strings. The length of abbreviated string needs to be specified |
| strsplit() | It is used to split a string based on a criterion. It returns a list. Alternatively, you can use the str\_split() function. This function lets you convert your list output to a character matrix |
| sub() | It is used to find and replace the first match in a string |
| gsub() | It is used to find and replace all the matches in a string / vector. Alternatively, you can use the str\_replace() function |

# List of Regular Expression Commands in R

|  |  |
| --- | --- |
| **Function** | **Description** |
| grep | returns the index or value of the matched string |
| grepl | returns the Boolean value (True or False) of the matched string |
| regexpr | return the index of the first match |
| gregexpr | returns the index of all matches |
| regexec | is a hybrid of regexpr and gregexpr |
| regmatches | returns the matched string at a specified index. It is used in conjunction with regexpr and gregexpr. |