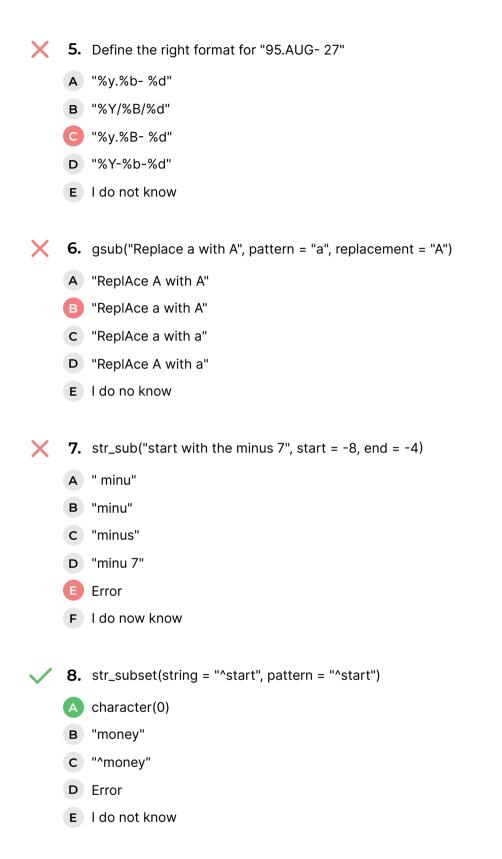


DM-Quiz-2020-Q9

18.18% (2/11)

- 1. (geners ← c("Action, Adventure, Comedy", "Comedy", "Comedy, Drama, Drama, Romance", "Crime, Drama, History")) Which of the functions has the following output? [1] NA NA "Drama" "Drama"
 - A str_extract(string = geners, pattern = "Drama")
 - **B** str_count(string = geners, pattern = "Drama")
 - str_subset(string = geners, pattern = "Drama")
 - D str_detect(string = geners, pattern = "Drama")
 - E I do not know.
- 2. str_subset(string = c("abc", "abbc", "abbbc", "ac"), pattern = "ab?c")
 - A [1] "abc" "ac"
 - B [1] "abc" "abbc" "abbbc" "ac"
 - **c** [1] "abc"
 - [1] "abc" "abbc" "abbbc"
 - E I do not know
- **3.** str_subset(string = c(" space(1)", "space (1)", "nospace", " ", " "), pattern = "\\S+")
 - A " space(1)", "space (1)", "nospace"
 - B " space(1)" , "space (1)", " " , " "
 - **c** "space (1)", " ", " "
 - **D** I do not know
- **4.** str_subset(string = c("0L","L1", "LL", "Li", "."), pattern = "[aeuio]")
 - A "Li"
 - B "OL","L1", "LL", "Li", "."
 - c "0L","L1", "LL", "Li"
 - D "LL", "Li"
 - E I do not know.



- **9.** Regular Expressions are
 - A set of commands used to match a family of text (alphanumeric, digits, words) to detect string sequences in a large text data
 - B series of functions used to extract information from text variables.
 - c meaning and interpretation of words, signs, and sentence structure
 - D collections of documents containing (natural language) text
 - E I do not know
- **10.** String manipulation is
 - A set of commands used to match a family of text (alphanumeric, digits, words) to detect string sequences in a large text data
 - B series of functions used to extract information from text variables.
 - c meaning and interpretation of words, signs, and sentence structure
 - D collections of documents containing (natural language) text
 - E I do not know
- ✓ 11. Sparsity is
 - A number of terms with more than one occurrence
 - B number of term with 0 occurrence in all documents
 - percentage of element with value 0 in total number of elements in document term matrix
 - D I do not know