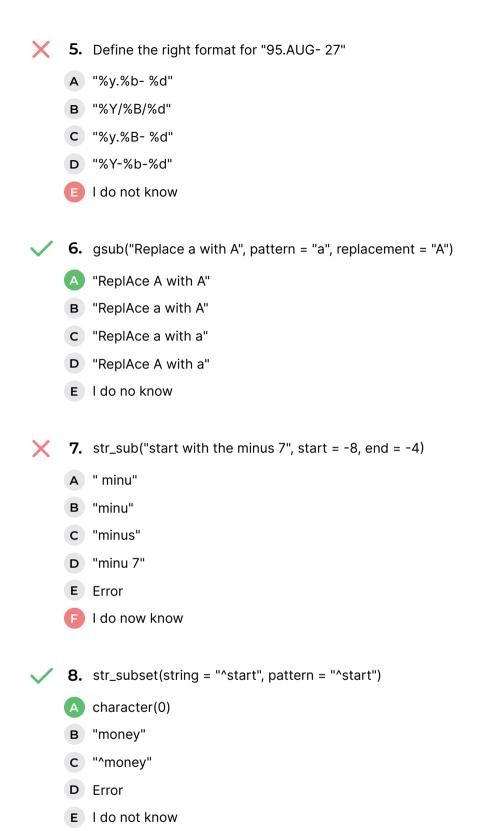# DM-Quiz-2020-Q9

27.27% (3/11)

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")

C 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"

D [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 "0L","L1", "LL", "Li", "."

C "0L","L1", "LL", "Li"

D "LL", "Li"

E I do not know.

5. Define the right format for "95.AUG- 27"

A. "%y.%b- %d"

B. "%Y/%B/%d"

C. "%y.%B- %d"

D. "%Y-%b-%d"

E. I do not know

6. gsub("Replace a with A", pattern = "a", replacement = "A")

A. "ReplAce A with A"

B. "ReplAce a with A"

C. "ReplAce a with a"

D. "ReplAce A with a"

E. I do no know

7. str_sub("start with the minus 7", start = -8, end = -4)

A. " minu"

B. "minu"

C. "minus"

D. "minu 7"

E. Error

F. I do now know

8. str_subset(string = "^start", pattern = "^start")

A. character(0)

B. "money"

C. "^money"

D. Error

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

C percentage of element with value 0 in total number of elements in document term matrix

D I do not know