

# DM-Quiz-2020-Q9

## 11 Questions

1. `(genres <- 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"

2/8 ☒ A `str_extract(string = genres, pattern = "Drama")`  
0/8 ☐ B `str_count(string = genres, pattern = "Drama")`  
2/8 ☐ C `str_subset(string = genres, pattern = "Drama")`  
2/8 ☐ D `str_detect(string = genres, pattern = "Drama")`  
2/8 ☐ E I do not know.

2. `str_subset(string = c("abc", "abbc", "abbbc", "ac"), pattern = "ab?c")`

1/8 ☒ A `[1] "abc" "ac"`  
2/8 ☐ B `[1] "abc" "abbc" "abbbc" "ac"`  
0/8 ☐ C `[1] "abc"`  
3/8 ☐ D `[1] "abc" "abbc" "abbbc"`  
2/8 ☐ E I do not know

3. `str_subset(string = c(" space(1)", "space (1)", "nospace", " ", " "), pattern = "\\S+")`

2/8 ☒ A `" space(1)" , "space (1)" , "nospace"`  
2/8 ☐ B `" space(1)" , "space (1)", " " , " "`  
0/8 ☐ C `"space (1)", " " , " "`  
4/8 ☐ D I do not know

4. `str_subset(string = c("OL", "L1", "LL", "Li", ":"), pattern = "[aeuio]")`

3/8 ☒ A `"Li"`  
2/8 ☐ B `"OL", "L1", "LL", "Li", ":"`  
0/8 ☐ C `"OL", "L1", "LL", "Li"`  
0/8 ☐ D `"LL", "Li"`  
3/8 ☐ E I do not know.

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

4/8 ☒ A "%y.%b- %d"

0/8 ☐ B "%Y/%B/%d"

3/8 ☐ C "%y.%B- %d"

0/8 ☐ D "%Y-%b-%d"

1/8 ☐ E I do not know

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

7/8 ☒ A "ReplAce A with A"

1/8 ☐ B "ReplAce a with A"

0/8 ☐ C "ReplAce a with a"

0/8 ☐ D "ReplAce A with a"

0/8 ☐ E I do no know

7. str\_sub("start with the minus 7", start = -8, end = -4)

4/8 ☒ A " minu"

1/8 ☐ B "minu"

0/8 ☐ C "minus"

0/8 ☐ D "minu 7"

1/8 ☐ E Error

2/8 ☐ F I do now know

8. str\_subset(string = "^start", pattern = "^start")

3/8 ☒ A character(0)

0/8 ☐ B "money"

0/8 ☐ C "^money"

1/8 ☐ D Error

4/8 ☐ E I do not know

**9. Regular Expressions are**

- 3/8 **A** set of commands used to match a family of text (alphanumeric, digits, words) to detect string sequences in a large text data
- 2/8 **B** series of functions used to extract information from text variables.
- 3/8 **C** meaning and interpretation of words, signs, and sentence structure
- 0/8 **D** collections of documents containing (natural language) text
- 0/8 **E** I do not know

**10. String manipulation is**

- 5/8 **A** set of commands used to match a family of text (alphanumeric, digits, words) to detect string sequences in a large text data
- 3/8 **B** series of functions used to extract information from text variables.
- 0/8 **C** meaning and interpretation of words, signs, and sentence structure
- 0/8 **D** collections of documents containing (natural language) text
- 0/8 **E** I do not know

**11. Sparsity is**

- 2/9 **A** number of terms with more than one occurrence
- 1/9 **B** number of term with 0 occurrence in all documents
- 3/9 **C** percentage of element with value 0 in total number of elements in document term matrix
- 3/9 **D** I do not know