Worksheet 4 - regular expressions I

Jo Hardin

2023-01-11

Your Name: _____

Names of people you worked with: _____

- Introduce yourself. What is the most challenging thing you have to do in the next 10 days?

- What classes are you taking in the spring? How many of them are SDS?

Task:

The Jewish Festival of Lights is written as חנוכה in Hebrew and translated to many different spellings in English. The correct English spelling of the holiday can be given in any of the following thirteen variants, according to https://www.holidays.net/chanukah/spelling.htm.

```
lights <- c("Chanuka", "Chanukah", "Chanukkah", "Channukah", "Hanukah",
            "Hannukah", "Hanukkah", "Hanuka", "Hanukka", "Hanaka", "Haneka",
            "Hanika", "Khanukkah")

lights
```

```
 [1] "Chanuka"   "Chanukah"  "Chanukkah" "Channukah" "Hanukah"   "Hannukah"
 [7] "Hanukkah"  "Hanuka"    "Hanukka"   "Hanaka"    "Haneka"    "Hanika"
[13] "Khanukkah"
```

Provide a regular expression which will capture all thirteen variants.

Solution:

[CHK]h?ann?[aeiu]kk?ah?

```
grepl("[CHK]h?ann?[aeiu]kk?ah?", lights)
```

```
[1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```