Worksheet 8 - Regular Expressions I
Monday, September 30, 2024

Jo Hardin

Your Name: _____

Names of people you worked with: _____

What is the most challenging thing you have to do in the next 10 days?

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"
 [6] "Hannukah"  "Hanukkah"  "Hanuka"    "Hanukka"   "Hanaka"
[11] "Haneka"    "Hanika"    "Khanukkah"
```

Provide a regular expression which will capture all thirteen variants but will not capture the string "Christmas".

Solution:

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

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

```
 [1]  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
[11]  TRUE  TRUE  TRUE FALSE
```