

Week 3 - Notes

STRSPLIT

```
mydata <- c("teste", "teste.2", "teste3")
mydata <- strsplit(mydata, "\\.")
firstElement <- function(x){x[1]}
results <- sapply(mydata, firstElement)
print(results)
```

```
## [1] "teste" "teste" "teste3"
```

```
# GSUB
```

```
mydata <- c("teste", "teste.2", "teste3")
mydata2 <- c("1,2", "1,3", "1,4")
mydata <- cbind(mydata, mydata2)
print(mydata)
```

```
##      mydata  mydata2
## [1,] "teste"   "1,2"
## [2,] "teste.2" "1,3"
## [3,] "teste3"  "1,4"
```

```
mydata[,2] <- gsub(",", ".", mydata[,2])
print(mydata)
```

```
##      mydata  mydata2
## [1,] "teste"   "1.2"
## [2,] "teste.2" "1.3"
## [3,] "teste3"  "1.4"
```

```
# GREP
```

```
mydata <- c("luiz", "fernando", "luizfernando")
grep("fer", mydata)
```

```
## [1] 2 3
```

```
# GREPL
```

```
grepl("fer", mydata)
```

```
## [1] FALSE TRUE TRUE
```

```
mydata3 <- mydata[grepl("fer", mydata)]
print(mydata3)
```

```
## [1] "fernando"      "luizfernando"
```

Regular expressions I

¹ ^ - start of line \$ - end of line [Bb][Uu][Ss][Hs] - looks for any bush word
² am - looks for “I am” or “i am”
[a-zA-Z] - any line that starts with number followed by character [^?.\$] - not ending by ? or .

Regular expressions II

9.11 - looks for any character between 9 and 11 flood|fire - looks for both flood and fire ^([Gg]ood|[Bb]ad)
[Gg]eorge ([Ww].)? [Bb]ush - looks for any george bush with w in the middle or not (the . is necessary because
. is a special character) (.*) - any character repeated

¹li
²0-9