Regex

Alyssa Green

2023-02-08

GitHub Repository (https://github.com/alyssagreen02/BIOL432)

Load Library

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

1. Import Data

```
MData=read.csv("measurements.csv")
```

2. Add new column

```
MData=MData%>%mutate(Sp=gsub("^(.).*\\s(\\w+)","\\1\\. \\2",Organism))

head(MData)
```

```
##              Organism Limb_width Limb_length Observer    Volume            Sp
## 1   Specious peoples          6           0    Sam P    0.0000     S. peoples
## 2       Orng samples          7           7 Alyssa G  269.3916     O. samples
## 3 Pinkus grafiarous          8           4   Josh G  201.0619 P. grafiarous
## 4 Pinkus grafiarous          9           3 Alyssa G  190.8518 P. grafiarous
## 5   Specious peoples          6           4   Josh G  113.0973     S. peoples
## 6       Orng samples          6           6   Josh G  169.6460     O. samples
```

New Data Frame

```
MeasShort=MData%>%select(Sp, Limb_width, Limb_length, Observer, Volume)
head(MeasShort)
```

```
##               Sp Limb_width Limb_length Observer    Volume
## 1    S. peoples          6           0    Sam P    0.0000
## 2    O. samples          7           7 Alyssa G  269.3916
## 3 P. grafiarous          8           4   Josh G  201.0619
## 4 P. grafiarous          9           3 Alyssa G  190.8518
## 5    S. peoples          6           4   Josh G  113.0973
## 6    O. samples          6           6   Josh G  169.6460
```

Save Data Set

```
write.csv(MeasShort, file = "MeasShort.csv", row.names = F)
```