

HW3

NZ

2025-02-02

REGULAR EXPRESSIONS

1

```
find: [ \t]{2, }  
replace: ,
```

Capture everything with two or more spaces. Replace with comma

2

```
find: (\w+), (\w+), (.*)  
replace: \2 \1 \(\3\)
```

Capture every word that is separated by commas and any character except after the comma and the newline. change the order of the words (1 and 2) and then add parentheses to the last characters (3 corresponding to the affiliation)

3

```
find: ^\s*(\d{4} .+?\s*.mp3)\s*  
replace: \1\n
```

Capture all the words that we want in the first line. Replace for a line break for each

4

```
find: ^\s*(\d{4})\s*(.+?)\s*(.mp3)  
replace: \2_\1\3
```

Choose the four digit number for each line. Move to the position that we want, before .mp3

5

```
find: (\w)\w+?,(\w+),[\d\.]+,(\d+) #replace: \1_\2,\3
```

Capture the first letter of the genus name, then the species name and the numerical value. Replace all the words of the genus word except for the first letter, and delete the numeric values with points.

6

```
find: (\w)\w+?,(\w{4})\w+?,[\d\.]+\,(\d+)  
replace: \1_\2,\3
```

7

```
find: ^(\w{3})\w*,(\w{3})\w*,([\d\.]+\,)(\d+)$  
replace: \1\2, \4, \3
```

Capture the first four letters of the species name

8

```
find: NA  
replace: 0
```

```
find: [^\w\s/0-9\t\n]  
replace:(just replace)
```

First, select the word NA and replace for 0. Then, select characters other than: letters, digits 0 to 9 and / (to avoid replacing these symbols in the date column) and just replace.