Analytics of Business Intelligence

Chapter # 2b - Grep Questions

Dr. Wajahat Gilani

Rutgers Business School

February 22, 2020

Replacing Patterns

The question was about replacing characters (or numbers) for a column where data messed up a bit. Here is an example of a small table:

```
1    a = c('4x','6a','234aax','abc56')
2    b = c('rty1a','cc6b','2h3g4','abc56')
3    dt = data.table(a,b)
```

a	b
4x	rty1a
6a	cc6b
234aax	2h3g4
abc56	abc56

That creates the following table:

Using the gsub() Function

There is another function that works like grep, but instead of finding patterns that meet a certain criteria, it takes it one step further and changes values that meet those criteria.

```
1 dt[,a:=gsub('[a-z]','',a)]
2 dt[,b:=gsub('[a-z]','',b)]
```

That creates the following table. Notice how the gsub() function is in the **column** section not the row section. This is because we are changing the values.

a	b
4	1
6	6
234	234
56	56