LESSON 8

report.awk

```
gsub(/,/, "", $2)
toupper($1)
substr($4, 1, 4)
```

AWK

String Functions



A FREE MYWEBUNIVERSITY COURSE

AWK SCRIPTING

LESSON 8

String Functions

length, substr, index, sub, gsub, match, toupper and tolower

length, substr and index

Call	Result	Note
`length("Kabul")`	5	characters
`substr("2026-10-24", 1, 4)`	2026	start at 1
`substr("Kabul", 3)`	bul	to the end
`index("Herat", "rat")`	3	position
`index("Herat", "x")`	0	not found

- length returns the number of characters. substr of a string, start, length cuts out a ...
- index returns where one string first appears inside another, or zero if it is not there.

01

sub and gsub

clean.awk

```
{
    gsub(/,/, "", $2)              # 1,250 -> 1250
    sub(/^ +/, "")                 # trim leading spaces
    gsub(/[ \t]+$/, "")            # trim trailing spaces
    n = gsub(/error/, "ERROR")      # n = number replaced
    print
}
```

- sub replaces the first match of a regular expression, and gsub replaces every match. Both …
- gsub of slash comma slash with nothing removes all commas from dollar two, which is how you …

02

toupper, tolower and match

match.awk

```
tolower($1) == "mariam" { print "found" }

match($0, /[0-9]+\.[0-9]+\.[0-9]+\.[0-9]+/) {
    ip = substr($0, RSTART, RLENGTH)
    print ip
}
```

- toupper and tolower change case, which is useful before comparing names typed in different …
- match finds a regular expression and sets RSTART and RLENGTH, so you can cut out exactly …

03

Building Strings with sprintf

ids.awk

```
{
    id  = sprintf("STU-%04d", NR)        # STU-0001
    avg = sprintf("%.1f", $2 / 3)
    key = sprintf("%s|%s", $3, $1)        # Kabul|Mariam
    print id, key, avg
}
```

- sprintf formats a string just like printf, but stores the result instead of printing it.
- Use it to build keys, IDs with leading zeros, or rounded numbers that you keep for later.

04

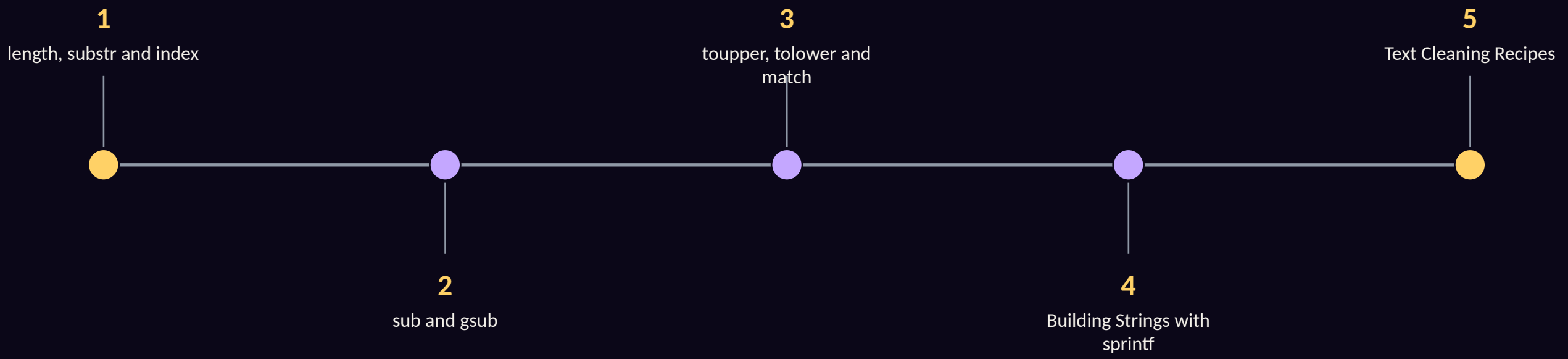
Text Cleaning Recipes

Code	Does
`gsub(/^[ \t]+|[ \t]+$/, "")`	trim both ends
`gsub(/ +/, " ")`	squeeze spaces
`gsub(/"/, "")`	remove quotes
`$1 = toupper(substr($1,1,1)) substr($1,2)`	capitalize

- Here are recipes you will use again and again: trim spaces, normalize case, remove quotes, …
- Save them in a small file of AWK functions, which we will learn to write in the next lesson.

05

Lesson Map


1
length, substr and index
2
sub and gsub
3
toupper, tolower and match
4
Building Strings with sprintf
5
Text Cleaning Recipes

Sources and Further Reading

Call	Result	Note
`length("Kabul")`	5	characters
`substr("2026-10-24", 1, 4)`	2026	start at 1
`substr("Kabul", 3)`	bul	to the end
`index("Herat", "rat")`	3	position
`index("Herat", "x")`	0	not found

AWK SCRIPTING · 8

match.awk

```
tolower($1) == "mariam" { print "found" }

match($0, /[0-9]+\.[0-9]+\.[0-9]+\.[0-9]+/) {
    ip = substr($0, RSTART, RLENGTH)
    print ip
}
```

Thank you for watching

MyWebUniversity.com

Free technical education since 2000 — English and Dari

@MyWebUniversity · @MyWebUniversity-Dari