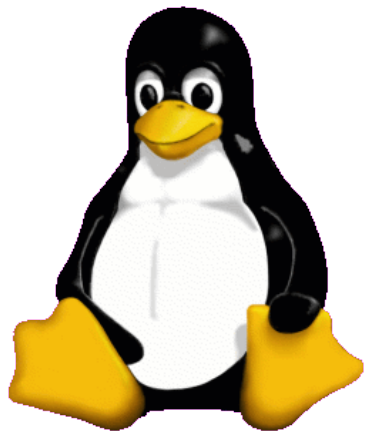# Unix102

BIOL647
Digital Biology

Rodolfo Aramayo

Unix102

System Information Commands

Date the computer was set to

```
> date
```

Unix102

System Information Commands

Reports the time the system is on

```
> uptime
```

Unix102

System Information Commands

Displays the users log into the system

```
> who

> w
```

Unix102
System Information Commands

Displays the system info

```
> uname

> uname -ap
```

Unix102

Disk Information Commands

Disk Free Space

```
> du
```

Unix102

Disk Information Commands

Humanized and size calculated as base2

```
> du –h
```

Unix102

Disk Information Commands

Humanized and all files

```
> du -ha
```

Unix102

Disk Information Commands

Depth option of one

```
> du -hd1
```

Unix102
Disk Information Commands

Allocation of a file versus size of a file
df Displays size set aside by the system on a Hard Drive (blocks)

```
> du -hd 1

> ls -alh
```

Unix102

Utility Programs

cal/ncal: Calendar

```
> cal

> cal 12 2021

> cal 2000

> ncal
```

Unix102

Utility Programs

bc: Bench Calculator

```
> bc

  1000/9

  scale=10

  1000/9

> bc -l # Precision
```

Unix102

Utility Programs

expr: Expression Evaluator

- Does Not work (entire argument)

```
> expr 1+1
```

- Does Not work (* is a wild–card and has special meaning)

```
> expr 1133 * 2245
```

- Does work (escaped characters)

```
> expr 1133 \* 2245
```

Unix102

Utility Programs

units: Unit Conversion

```
> man units
```

Unix102

Output of entire files

cat: Concatenate and write files

```
> cat file01 file02 > concatenated_file
```
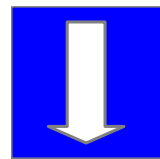
tac: Concatenate and write files in reverse

```
> tac file01 file02 > concatenated_file
```

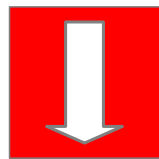# Unix102
Output of entire files
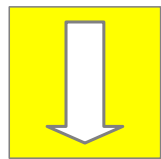
Given three Datasets A, B, and C

If Concatenated as … Result in …
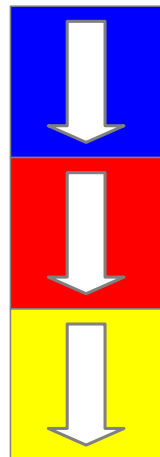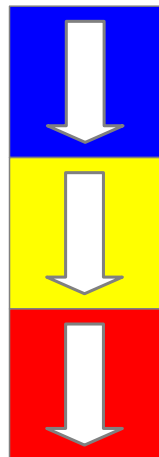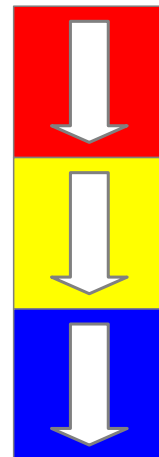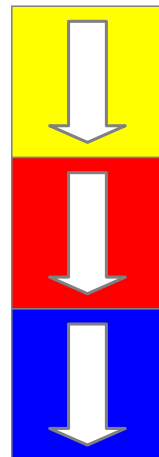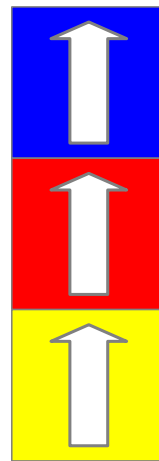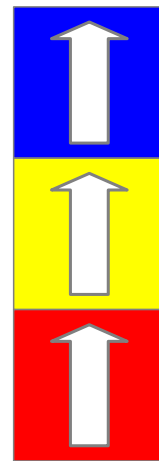
Unix102

Output of entire files

nl: Number lines and write files

```
> nl file01

> cat file01 file02 | nl

> cat -n file01 file02
```

Unix102
Introducing Tables

Table containing
12 fields (columns) and 15 records (rows or lines)

Fields Separators

Field 01
Record 01

Field 01
Record 08

1	2	3	4	5	6	7	8	9	10	11	12
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											

Unix102

Output of parts of files

head: Output the first part of files

```
> head -2 file01
```

tail: Output the last part of files

```
> tail -4 file01
```

Unix102
Output of parts of files

Table containing
12 fields (columns) and 15 records (rows or lines)

Fields Separators

Field 01
Record 01

1	2	3	4	5	6	7	8	9	10	11	12
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											

Head -2

Field 01
Record 08

Tail -4
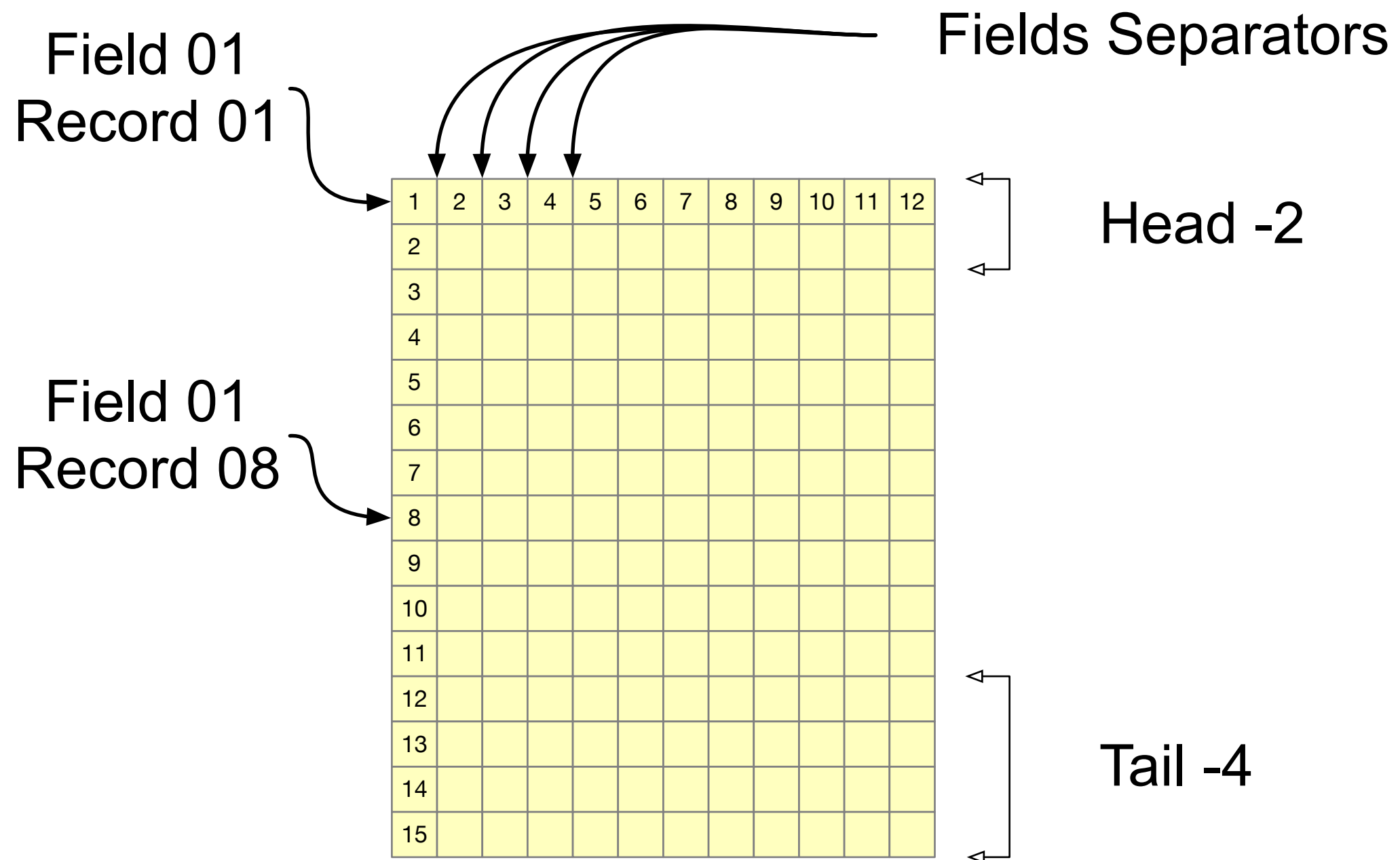
Unix102

Summarizing files

wc: Print newline, word, and byte counts

```
> wc -l

> wc -c
```

Unix102

Summarizing files

sum: Print checksum and block counts

```
> sum file01
```

cksum: Print and verify file checksums

```
> cksum file01
```

Unix102

Summarizing files

md5sum: Print or check MD5 digests

```
> md5sum file01
```

sha1sum: Print or check SHA-1 digests

```
> sha1sum file01
```

Unix102
Sorting files

Table containing
12 fields (columns) and 15 records (rows or lines)

Field 01
Record 01

Field 12
Record 01

Head -2

Field 01
Record 08

Tail -4

Unix102
Sorting files

Unix102
Sorting files

Sorting by Colors
On Field 04

Unix102
Sorting files

Field 01
Record 01

Field 12
Record 01

Head -2

Field 01
Record 08

Tail -4

Sorting by Colors
On Field 04

New Head -2

Previous Head

Previous Tail

Previous Head

New Tail -4

Previous Tail

Unix102
'Uniquifying' files

Removing Duplicate
Fields

1	2	3	4	5	6	7	8	9	10	11	12
3											
5											
2											
9											
4											
8											

Unix102

Operating on sorted files

sort: Sort text files

- Sort lines

```
> sort
```

- Reverse

```
> sort -r
```

- Numerical

```
> sort -n
```

- Ignore case

```
> Sort -f
```

- Unique output

```
> sort -u
```

Unix102

Operating on sorted files

uniq: Uniquify files

```
> uniq           # Filter in/out repeated lines

> uniq -d        # Returns repeated lines

> uniq -u        # Returns unduplicated lines
```

Unix102
Operating on sorted files

uniq: Uniquify files

- Filter in/out repeated lines

```
> uniq
```

- Returns repeated lines

```
> uniq -d
```

- Returns unduplicated lines

```
> uniq -u
```

Unix102

Operating on sorted files

comm: Compare two sorted files line by line

```
> comm sorted_file01 sorted_file02

> comm <(sort file01) <(sort file02)
```

Unix102

Comparing files

Checking Differences with diff, and sdiff

test01	test02	test03
apples	apples	oranges
oranges	oranges	walnuts
walnuts	grapes	chestnuts

The diff command displays the only line that differs between the two files. To understand the report, remember that diff is prescriptive, describing what changes need to be made to the first file to make it the same as the second file. This report specifies that only the third line is affected, exchanging walnuts for grapes

```
> echo -e "apples\noranges\nwalnuts" > test01

> echo -e "apples\noranges\ngrapes" > test02

> echo -e "oranges\nwalnuts\nchestnuts" > test03


> diff test01 test02

> sdiff test01 test02

> diff test01 test03

> sdiff test01 test03
```

Unix102
Operating on fields

cut: Print selected parts of lines

```
> ll / > root.txt
```

```
> cat root.txt | cut -c 2-10
```

```
> cat root.txt | cut -c 2-10,30-35
```

```
> cat root.txt | cut -c 2-10,30-35,45-
```

Unix102
Operating on fields
Translating Characters

```
> echo "a,b,c" | tr "," "-"

> echo "1234564321" | tr "123456" "EBGDAE"

> echo "12345643219" | tr "123456" "EBGDAE"

> echo "This is ROT-13 encrypted" | tr "A-Za-z" "N-ZA-Mn-za-m"

> echo "Guvf vf EBG-13 rapelcgrq" | tr "A-Za-z" "N-ZA-Mn-za-m"

> echo "already daytime" | tr "day" "night" <-Not Find and Replace

> echo "Ban Tan Cat Tac Dog Got" | tr "[:upper:]" "[:lower:]"

> echo "Ban Tan Cat Tac Dog Got" | tr "[A-Z]" "[a-z]"

> echo "Ban,Tan,Cat,Tac,Dog,Got" | tr "[,]" "[\t]"

> echo "Ban,Tan,Cat,Tac,Dog,Got" | tr "[,]" "[\t]" | cat -t

> echo "Ban,Tan,Cat,Tac,Dog,Got" | tr "[,]" "[\t]" | cat -te
```

Unix102
Operating on fields

Deleting and Squeezing Characters

```
| Option | Description                              |

| -d     | Delete characters                        |

| -s     | Squeeze characters                       |

| -c     | Use complement set                       |

| -dc    | Delete characters not in listed set      |

| -sc    | Squeeze characters not in listed set     |


# Delete Digits (not letters)

> echo "abc123333deee4567fg" | tr -d "[:digit:]"
```

Unix102
Operating on fields

Deleting and Squeezing Characters

```
| Option | Description                              |

| -d     | Delete characters                        |

| -s     | Squeeze characters                       |

| -c     | Use complement set                       |

| -dc    | Delete characters not in listed set      |

| -sc    | Squeeze characters not in listed set     |


# Delete Complement Digits (letters)

> echo "abc123333deee4567fg" | tr -dc "[:digit:]"
```

Unix102
Operating on fields
Deleting and Squeezing Characters

```
| Option | Description                                |

| -d     | Delete characters                          |

| -s     | Squeeze characters                         |

| -c     | Use complement set                         |

| -dc    | Delete characters not in listed set  |

| -sc    | Squeeze characters not in listed set |


# Squeeze digits

> echo "abc123333deee4567fg" | tr -s "[:digit:]"
```

Unix102

Operating on fields

Deleting and Squeezing Characters

```
| Option | Description                               |

| -d     | Delete characters                         |

| -s     | Squeeze characters                        |

| -c     | Use complement set                        |

| -dc    | Delete characters not in listed set       |

| -sc    | Squeeze characters not in listed set      |


# Squeeze Complement Digits

> echo "abc123333deee4567fg" | tr -sc "[:digit:]"
```

Unix102

Operating on fields

Deleting and Squeezing Characters

```
| Option | Description                                 |

| -d     | Delete characters                           |

| -s     | Squeeze characters                          |

| -c     | Use complement set                          |

| -dc    | Delete characters not in listed set   |

| -sc    | Squeeze characters not in listed set |


# Translate Digits to Letters

> echo "abc123333deee4567fg" | tr -ds "[:digit:]" "[:alpha:]"
```

Unix102
Operating on fields

Deleting and Squeezing Characters

```
| Option | Description                              |

| -d     | Delete characters                        |

| -s     | Squeeze characters                       |

| -c     | Use complement set                       |

| -dc    | Delete characters not in listed set  |

| -sc    | Squeeze characters not in listed set |


# Translate Squeeze Complement Digits to Letters

> echo "abc123333deee4567fg" | tr -dsc "[:digit:]" "[:alpha:]"
```

Unix102
Compiling software

Build from Source

- Note that username = Your Username
- Add the following to your .bashrc

```
> emacs .bashrc

export PATH=/home/DB2022_xx/Software/bin:$PATH

# To activate the .bashrc file

> source /home/DB2022_xx/.bashrc

> . /home/DB2022_xx/.bashrc
```

Unix102

Compiling software

Compiling Datamash

```
> wget http://ftp.gnu.org/gnu/datamash/datamash-1.3.tar.gz

> tar -xzf datamash-1.3.tar.gz

> cd datamash-1.3

> ./configure --prefix=/home/DB2022_xx/Software

> make

> make check

> make install
```

Unix102

BIOL647
Digital Biology

Rodolfo Aramayo