

The terminal, for data people.

D8A.ACADEMY/RESOURCES/TERMINAL-SHEET

MOVE AROUND

<code>pwd</code>	where am I
<code>ls -lah</code>	sizes + hidden files
<code>cd -</code>	back to last dir
<code>mkdir -p a/b</code>	nested dirs

SLICE & COUNT

<code>cut -d, -f1 f.csv</code>	column 1
<code>sort uniq -c</code>	value counts
<code>awk -F, '{print \$2}'</code>	column 2
<code>tr ',' '\t'</code>	swap delimiters

PIPES

<code>a b</code>	output feeds input
<code>> f.txt</code>	write (overwrites)
<code>>> f.txt</code>	append
<code>cmd1 && cmd2</code>	2 only if 1 worked

LOOK AT DATA

<code>head -5 f.csv</code>	first rows
<code>tail -f app.log</code>	follow live
<code>wc -l f.csv</code>	row count
<code>less f.csv</code>	scroll, q quits

SEARCH

<code>grep -r 'TODO' .</code>	find in files
<code>grep -c error app.log</code>	count matches
<code>find . -name '*.csv'</code>	find files
<code>history grep ssh</code>	what did I run

SURVIVE

<code>ctrl+c · ctrl+r</code>	kill · search history
<code>man cmd</code>	the manual
<code>chmod +x run.sh</code>	make it runnable
<code>rm</code>	← no trash bin. think first.

WORTH MEMORIZING

01 Value counts, no pandas

The one-liner that answers "what is in this column" on any machine.

```
cut -d, -f3 sales.csv -- take column 3
| sort                -- group same values
| uniq -c             -- count each
| sort -rn            -- biggest first
| head                -- top 10
```

→ Works on a 10GB file the laptop cannot open, over ssh, with nothing installed.

02 Trap: the terminal forgives nothing

Three classics that hurt exactly once.

YOU TYPED	WHAT HAPPENS	THE HABIT
<code>rm -rf \$DIR/</code>	empty var = deletes everything	quote it, check with <code>echo first</code>
<code>> data.csv</code>	file wiped before you look	<code>>></code> appends, <code>></code> overwrites
<code>cat 10gb.csv</code>	terminal drowns	<code>head, less, wc -l</code> instead

→ There is no undo and no trash. The pros are not braver, they just echo before they run.