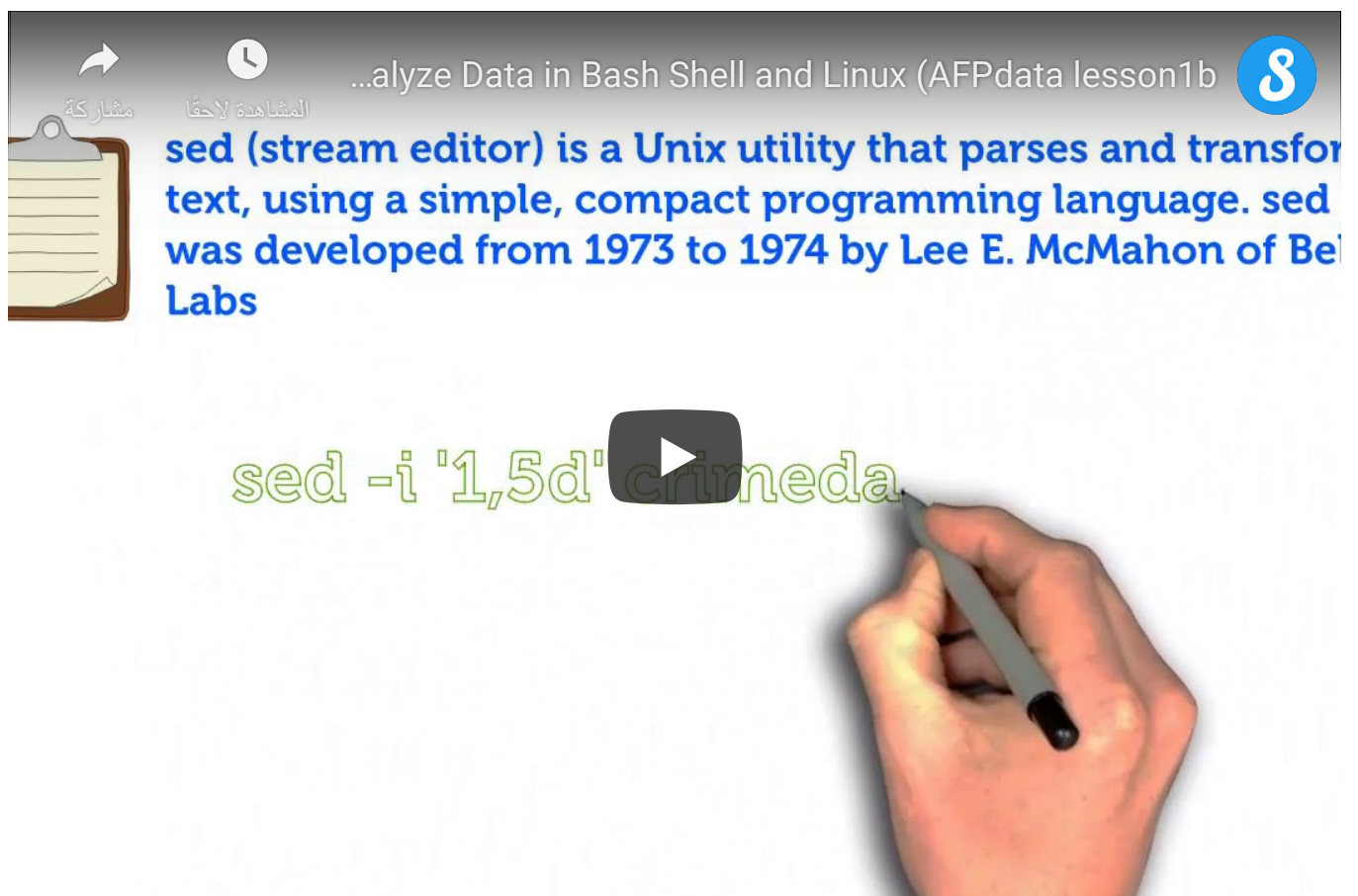


# Data Preview

Assuming you have saved your downloaded file as `tmpB2ZP9eAFP---Crime-incidents-data.xls` (from the [data.gov.au](https://data.gov.au)), the first thing you want to perform is to convert the data into the `.csv` format. Type the following command in the terminal:

```
libreoffice --headless --convert-to csv *.xls
```



Video lecture: Data xls to csv conversion and data preview

This command, using command line interface of the LibreOffice (which comes free with most of the Linux systems) converts all Excel `(*.xls)` files in the current folder to `.csv`. Let's also simplify the name of the file with `mv` command:

```
mv tmpB2ZP9eAFP---Crime-incidents-data.xls afpdata1.csv crimedata-au.csv
```



Awesome! next what we'll do is get a sneak peak of our data. As a quick reminder, the `cat` command shows us all the lines in a file.

```
cat crimedata-au.csv | csvlook
```

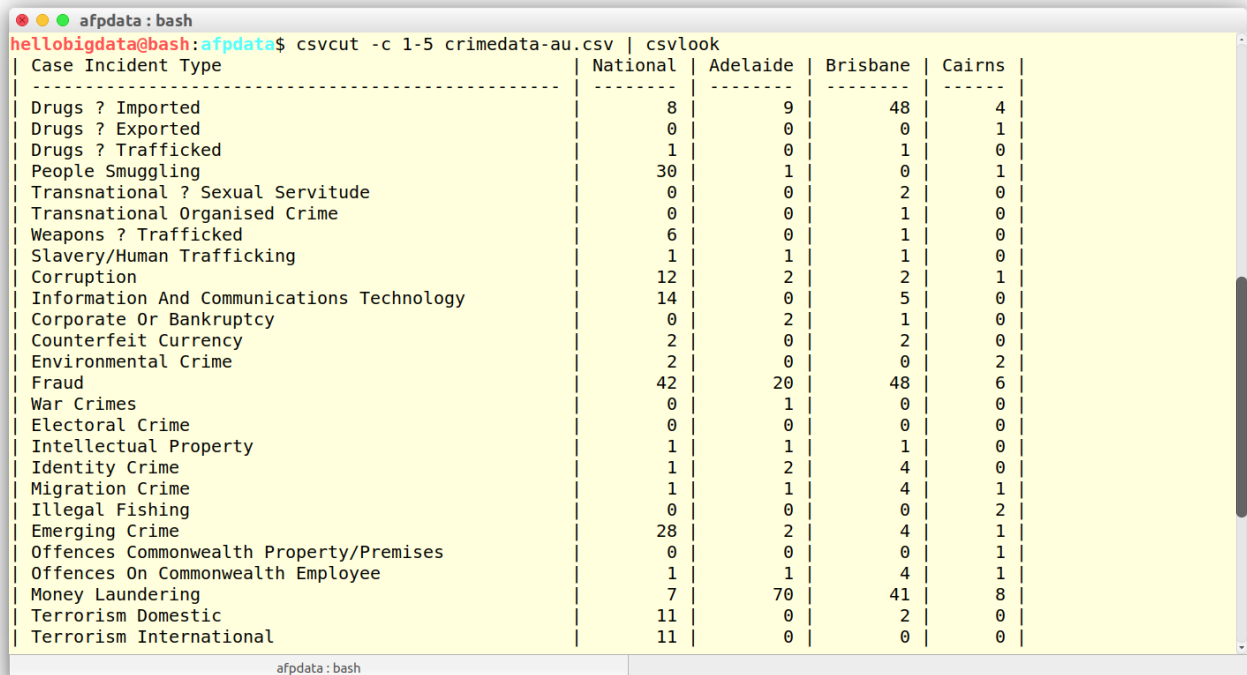


However, before we show the data again, we want to install a tool called `csvkit`. Once installed correctly, we should be able to run the following command:

```
csvcut -c 1-5 crimedata-au.csv | csvlook
```



The first command (`csvcut`) cuts out the first five columns and the pipes the output to the `csvlook` command:



```
afpdata : bash
hellobigdata@bash:afpdata$ csvcut -c 1-5 crimedata-au.csv | csvlook
```

Case Incident Type	National	Adelaide	Brisbane	Cairns
Drugs ? Imported	8	9	48	4
Drugs ? Exported	0	0	0	1
Drugs ? Trafficked	1	0	1	0
People Smuggling	30	1	0	1
Transnational ? Sexual Servitude	0	0	2	0
Transnational Organised Crime	0	0	1	0
Weapons ? Trafficked	6	0	1	0
Slavery/Human Trafficking	1	1	1	0
Corruption	12	2	2	1
Information And Communications Technology	14	0	5	0
Corporate Or Bankruptcy	0	2	1	0
Counterfeit Currency	2	0	2	0
Environmental Crime	2	0	0	2
Fraud	42	20	48	6
War Crimes	0	1	0	0
Electoral Crime	0	0	0	0
Intellectual Property	1	1	1	0
Identity Crime	1	2	4	0
Migration Crime	1	1	4	1
Illegal Fishing	0	0	0	2
Emerging Crime	28	2	4	1
Offences Commonwealth Property/Premises	0	0	0	1
Offences On Commonwealth Employee	1	1	4	1
Money Laundering	7	70	41	8
Terrorism Domestic	11	0	2	0
Terrorism International	11	0	0	0

afpdata : bash

Formatted first five lines of the crimedata-au.csv

