EXPLORE YOUR COMPUTER

GOALS

- Learn the different types of files
- Browse the file system
- Create/delete/move folders and files from the command line

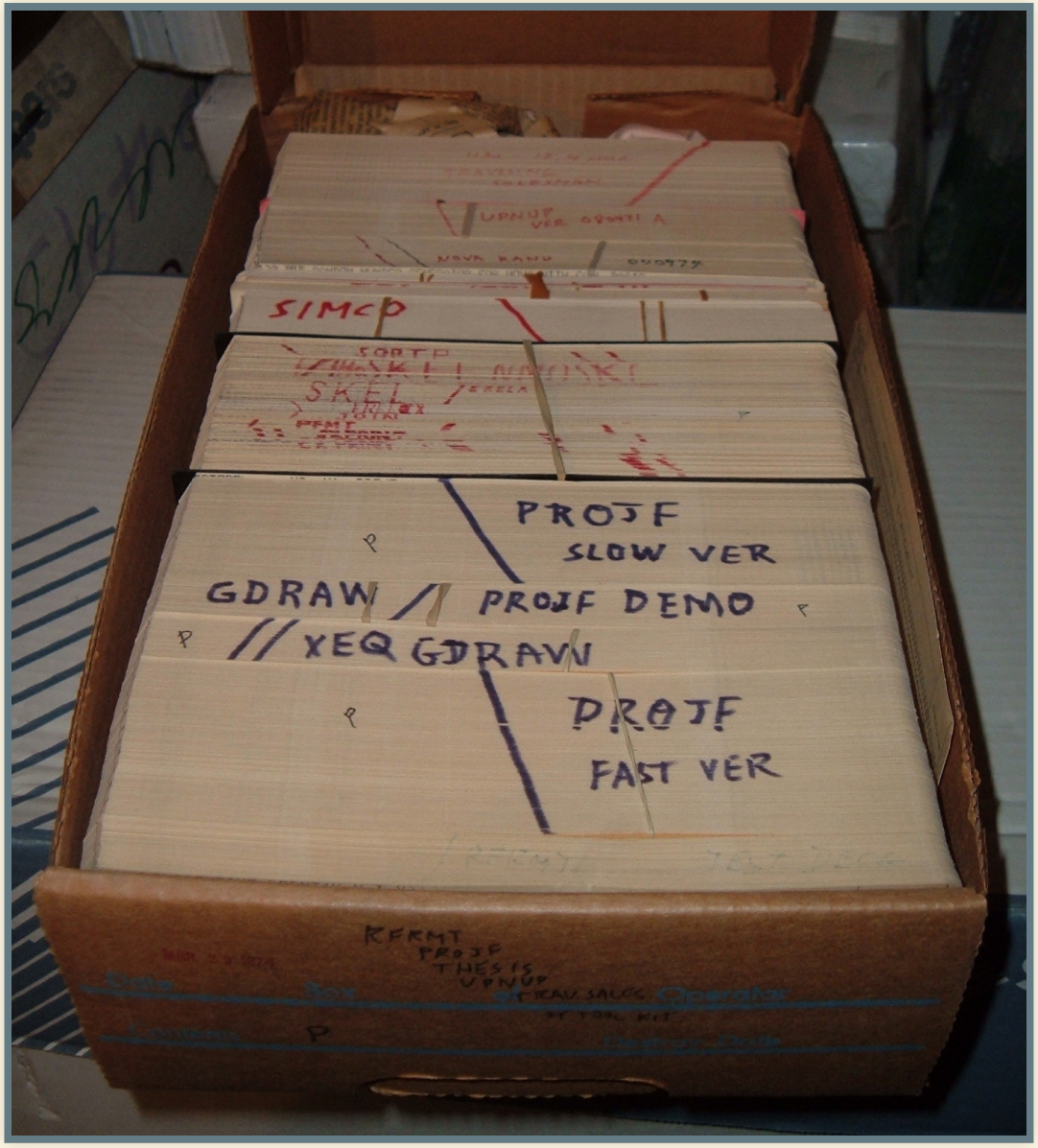
FILES, FORMATS AND CHARACTER ENCODING

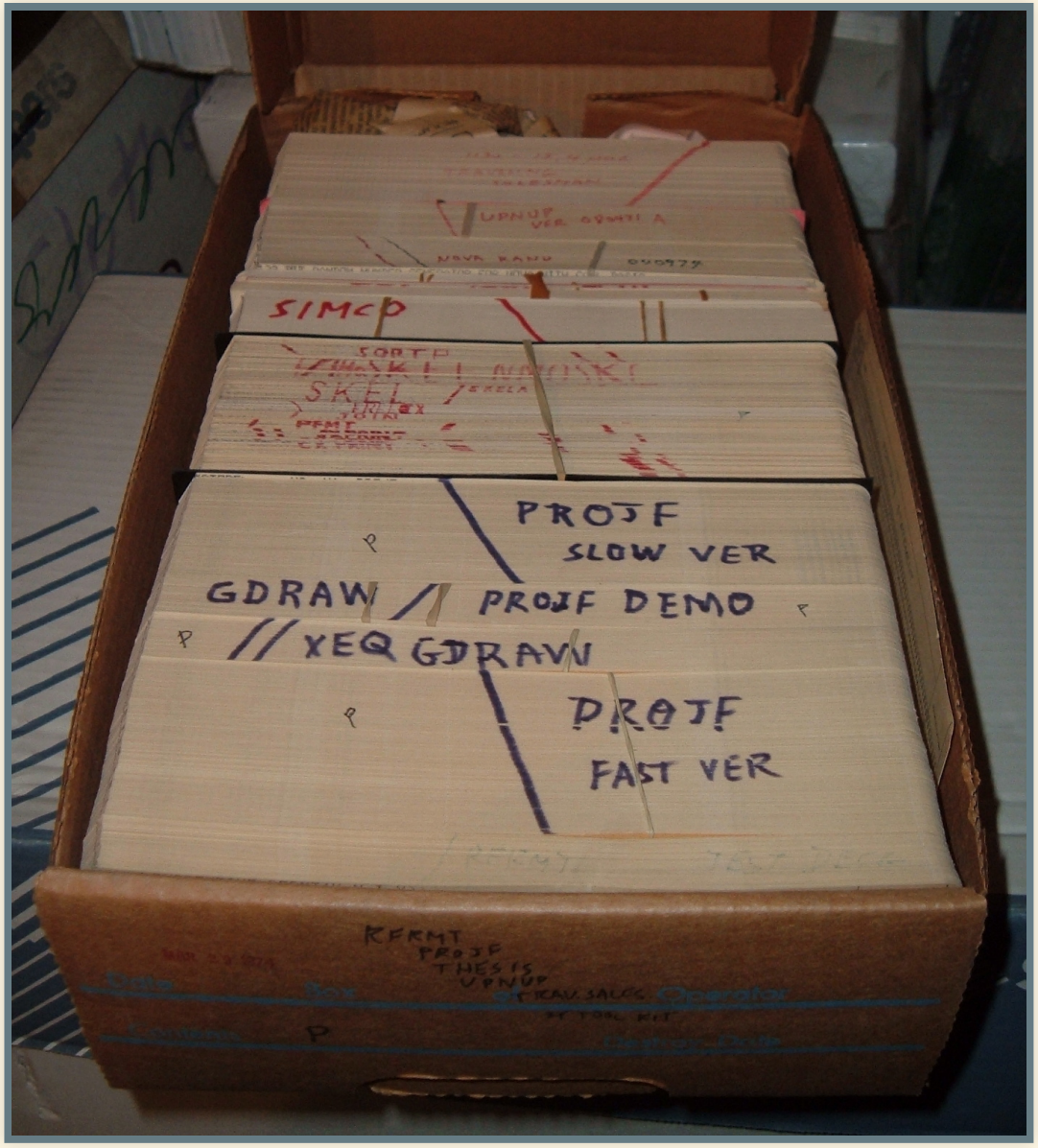
FILE TYPES

- **executable files**: they contain programs
- **data files**: they contain different types of information: text, image, audio, video.

FILE TYPES (ANOTHER TAXONOMY)

- **text files**: data is stored using electronic text (characters). They are human-readable.
- **binary files**: data is stored in binary format. They are computer-readable but not human-readable.


SIMCP
PROJF
SLOW VER
GDRAW // PROJF DEMO
// XEQ GDRAW
PROJF
FAST VER
KERMT
PROJF
THESIS
UPNUP



Punched card decks by ArnoldReinhold

FORMATS

- The file extension can be used as a hint to the format
- The format tells us the specific software we need to read it and/or manipulate it (e.g. data files) or run it (i.e. executable files)

CHARACTER ENCODING

- Any “text” is created from **characters**.
- Characters that are needed for a specific purpose are grouped into a **character set**.
- To refer to a character in an unambiguous way, each character is associated to a **code point** (a number).
- Characteres are stored in the computers as one of more **bytes**.
- A **character encoding** is a set of mappings between the bytes and the character set.

WHAT CHARACTER ENCODING?

UTF-8

WHAT CHARACTER ENCODING?

UTF-8

Your machine might be using a different encoding by default (eg. Windows-1252 or CP-1252)

IN THIS COURSE WE WILL WORK WITH:

- **`.txt`** plain text
- **`.csv`** tabular plain text (comma-separated values)
- **`.json`** JavaScript Object Notation. It stores and trasmits data consisting of attribute-value pairs and collections of elements
- **`.py`** source code in the script language Python
- **`.ipynb`** notebook documents used by Jupyter Notebook, interactive computational environment to work with Python

THE FILE SYSTEM

THE TERMINAL

- Also known as console, command-line interface
- It processes commands to a computer program in the form of lines of text

TO OPEN THE TERMINAL

- Linux: press **`CTRL + ALT + T`**
- Windows: press **`Windows + X`**; select **`Windows PowerShell`** or open the start menu and type "Powershell"
- Mac: in the Finder, open **`/Applications/Utilities`** (or use Launchpad and type "Terminal")

COMMAND

It is a request the user sends to the OS for it to execute it. Elements:

1. Name of a executable program (e.g. **`python`**)
2. A set of options whose name is preceded by one or two hyphens (e.g. **`-i`** or **`--help`**)
3. A set of arguments to define the file(s) or the data that must be used

COMMAND EXAMPLE

```
python my_wonderful_collation_script.py
```

FILE/DIRECTORY NAMING CONVENTIONS

Do not use whitespaces!

(and avoid special characters as well)

PATHS

- A path specifies a unique location in a file system. It points to a location by following the directory tree hierarchy, thus it looks like a slash-separated list of directory names followed by either a directory name or a file name. Eg.:
  - macOS: **`/Users/`*`username`*`/Desktop`**
  - Linux: **`/home/`*`username`*`/Desktop`**
  - Windows: **`C:\Users\`*`username`*`\Desktop`**

PATH REPRESENTATION (CONVENTIONS)

- root directory
  - UNIX: `/`
  - Windows: `\` (relative to current working directory root) or ***`drive_letter`*****`:\`**
- directory separator
  - UNIX: `/`
  - Windows: `\`

PATH REPRESENTATION (CONVENTIONS)

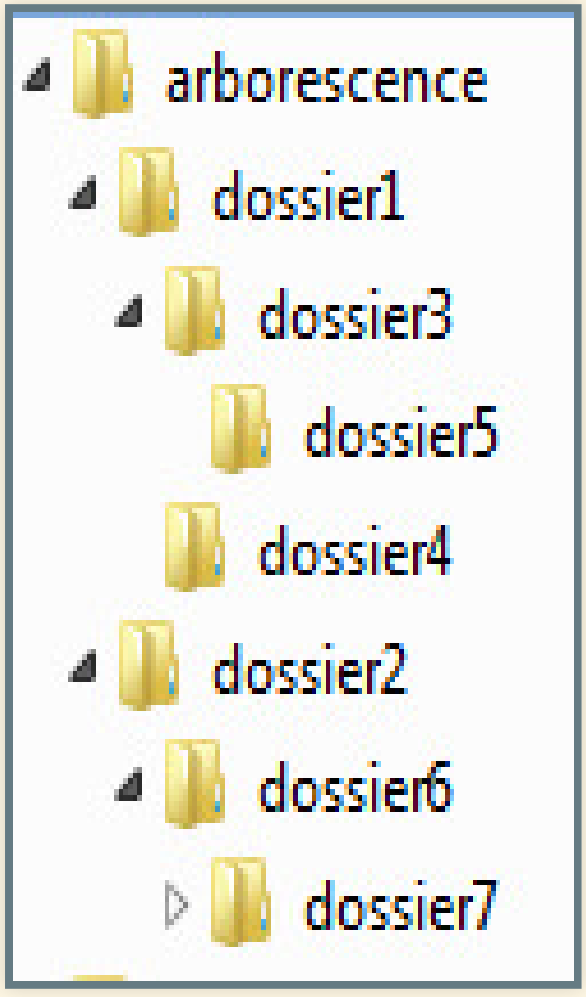
- current directory: `.`
- parent directory: `..`

TYPES OF PATH

- relative: path in relation to the current location
- absolute: path in relation to the disk root

TO CHANGE FOLDER

- UNIX: **`cd` *`directory_name`***
- Windows: **`Set-Location` *`directory_name`*** or **`cd` *`directory_name`*** or **`chdir` *`directory_name`***

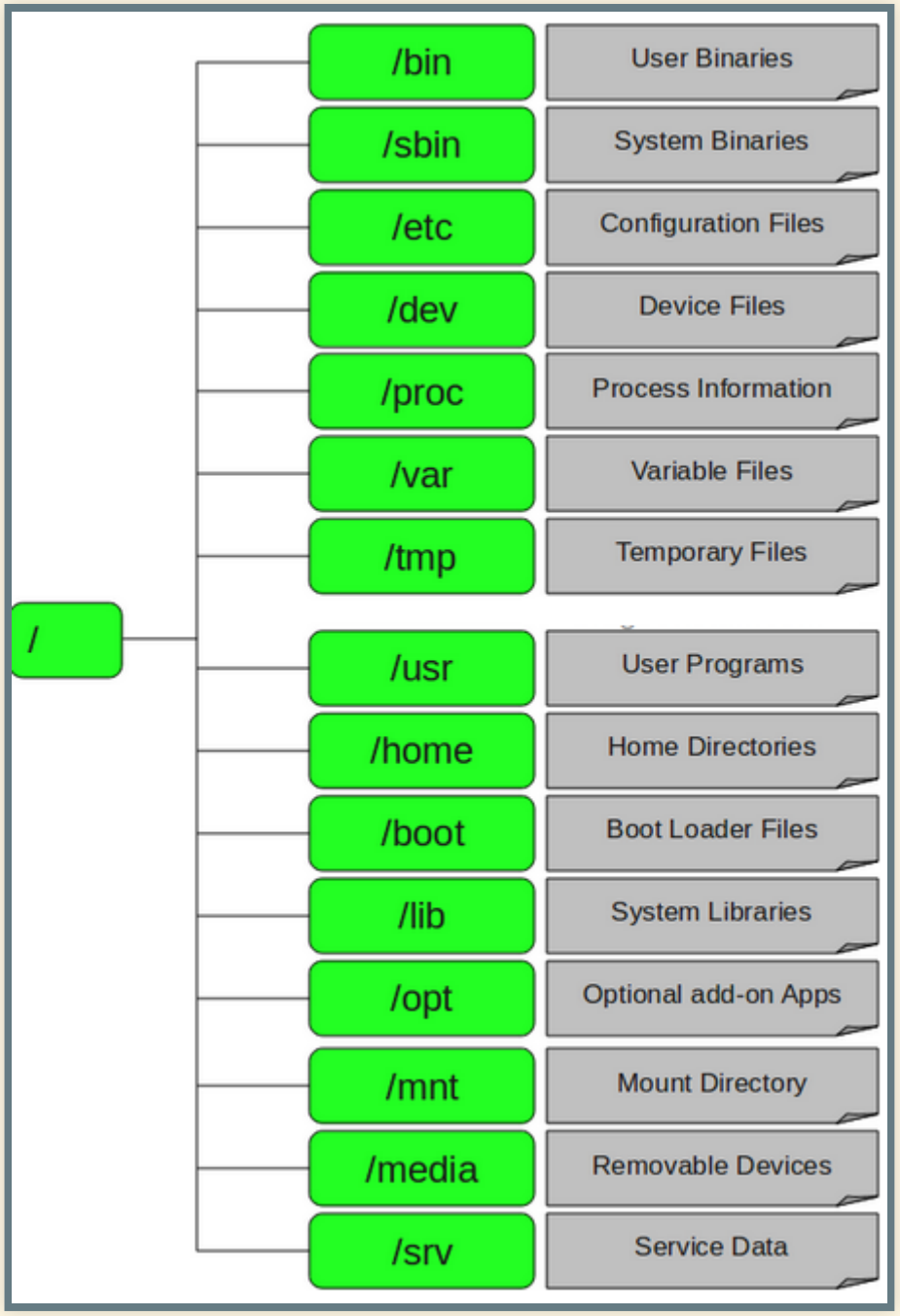


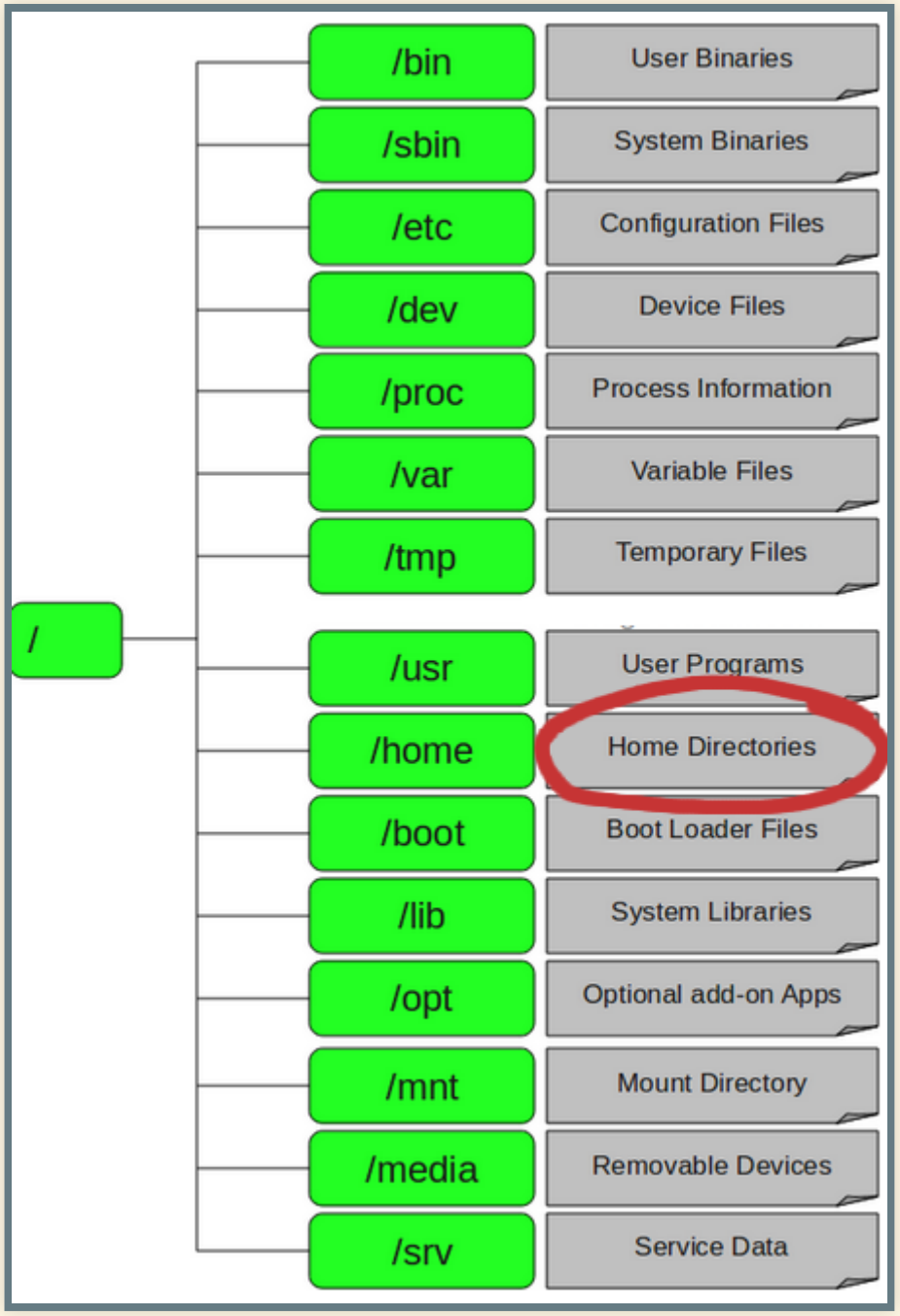
We are in “dossier3”

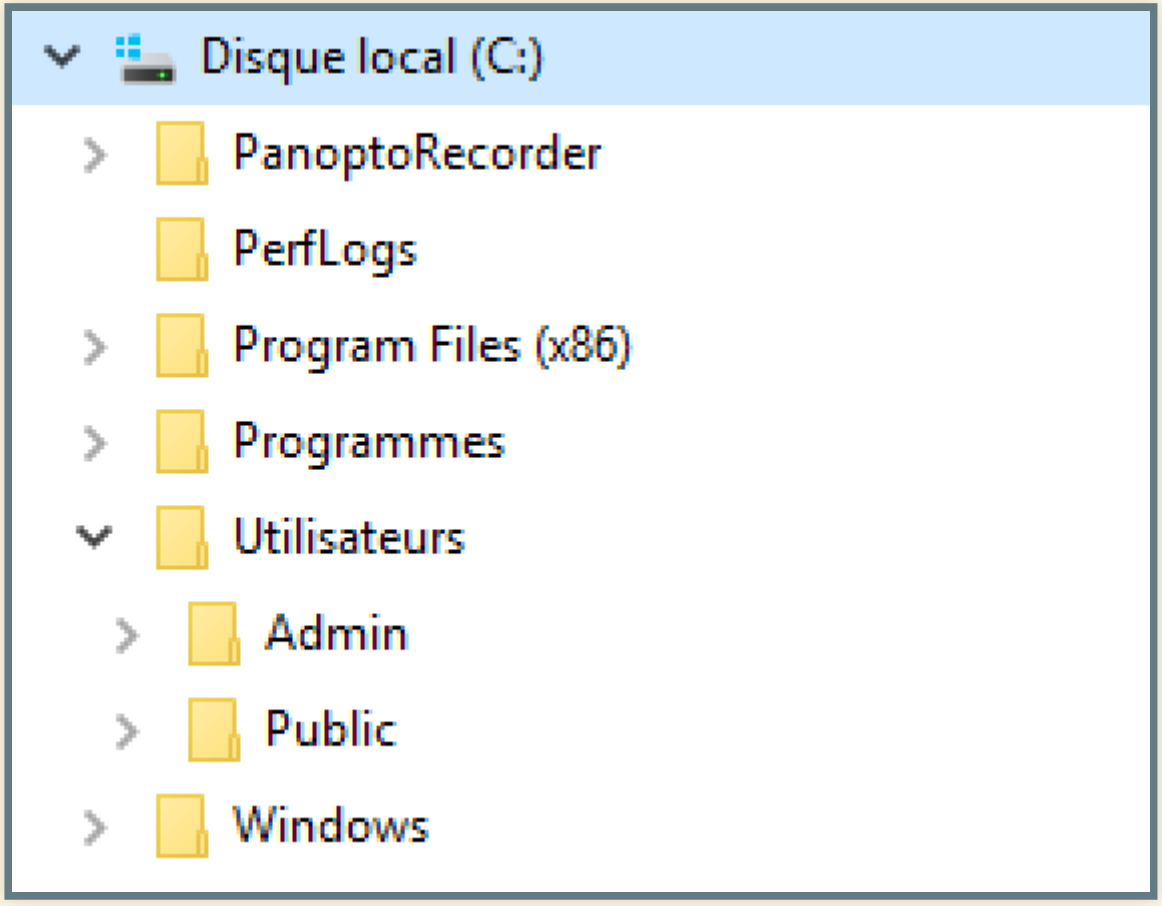
- “my” absolute path: **`/arborescence/dossier1/dos`**
- to move to dossier5 (relative path): **`c`** **`dossier5`**
- to move to dossier1 (relative path): **`c`**
- to move to dossier4 (relative path): **`c`** **`../dossier4`**
- to move to root: **`cd /`**

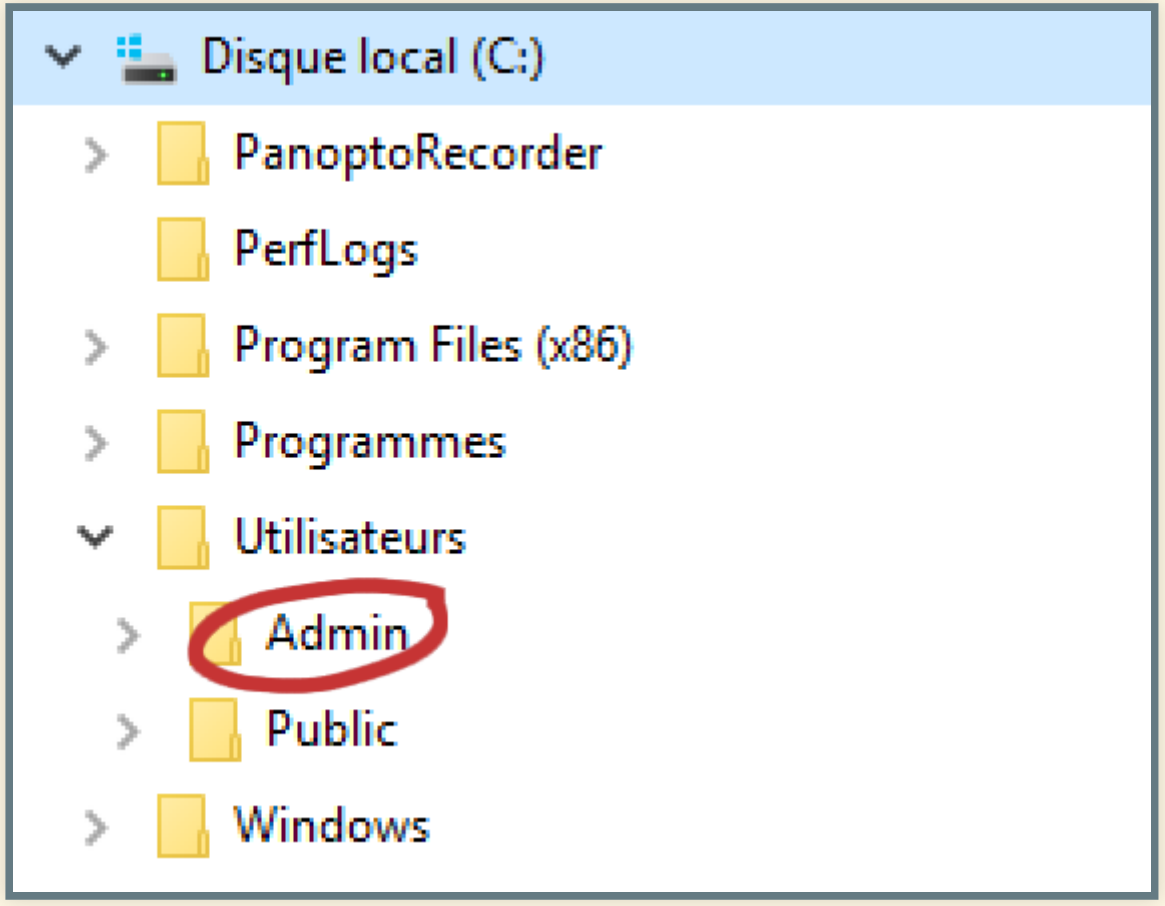
DIRECTORIES

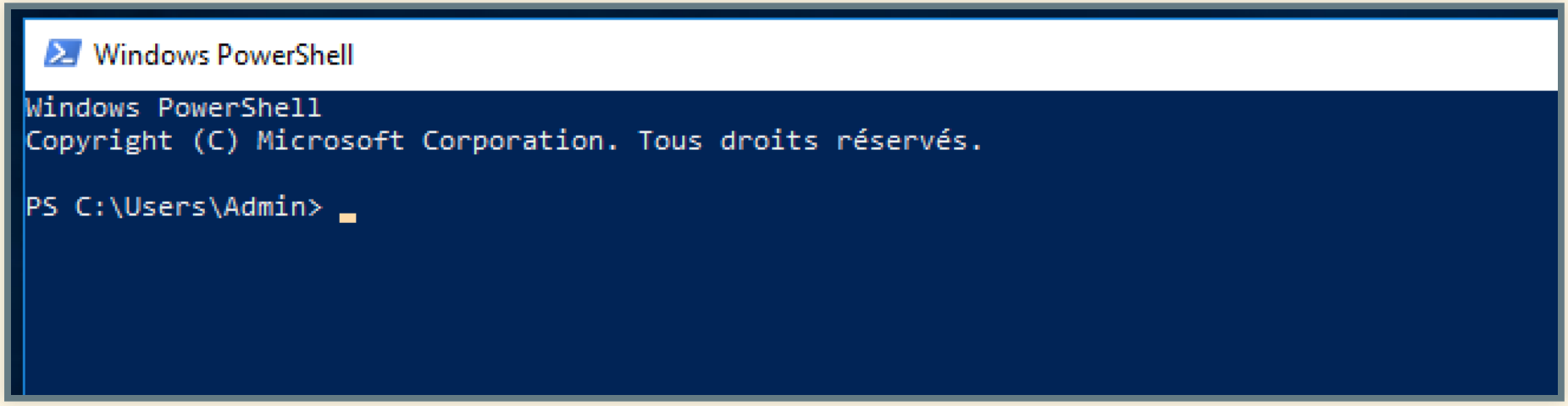
When we open the terminal, we are at the root of our personal folder

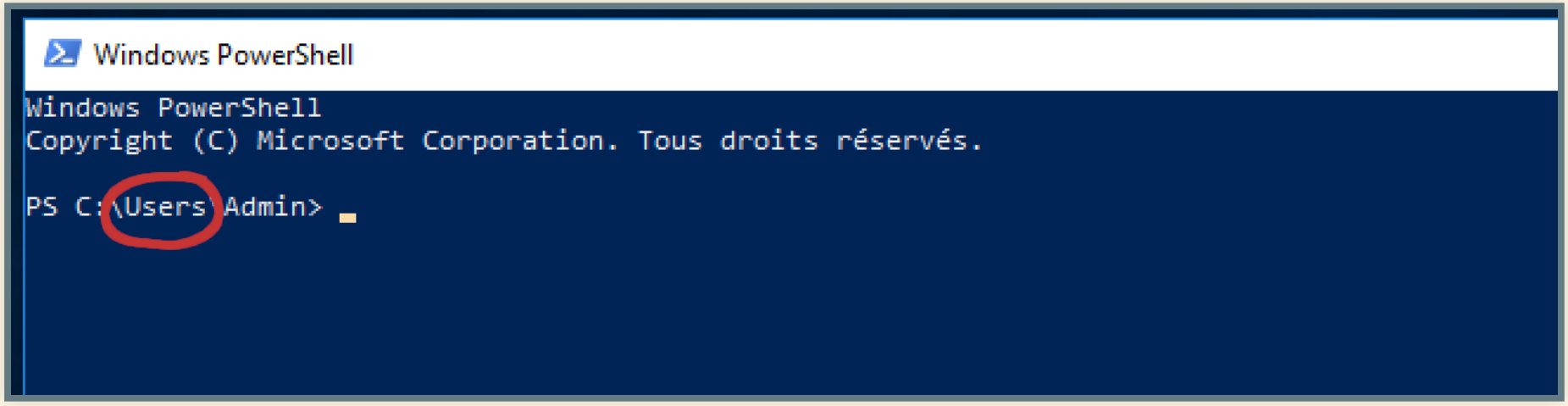

/
/bin
User Binaries
/sbin
System Binaries
/etc
Configuration Files
/dev
Device Files
/proc
Process Information
/var
Variable Files
/tmp
Temporary Files
/usr
User Programs
/home
Home Directories
/boot
Boot Loader Files
/lib
System Libraries
/opt
Optional add-on Apps
/mnt
Mount Directory
/media
Removable Devices
/srv
Service Data

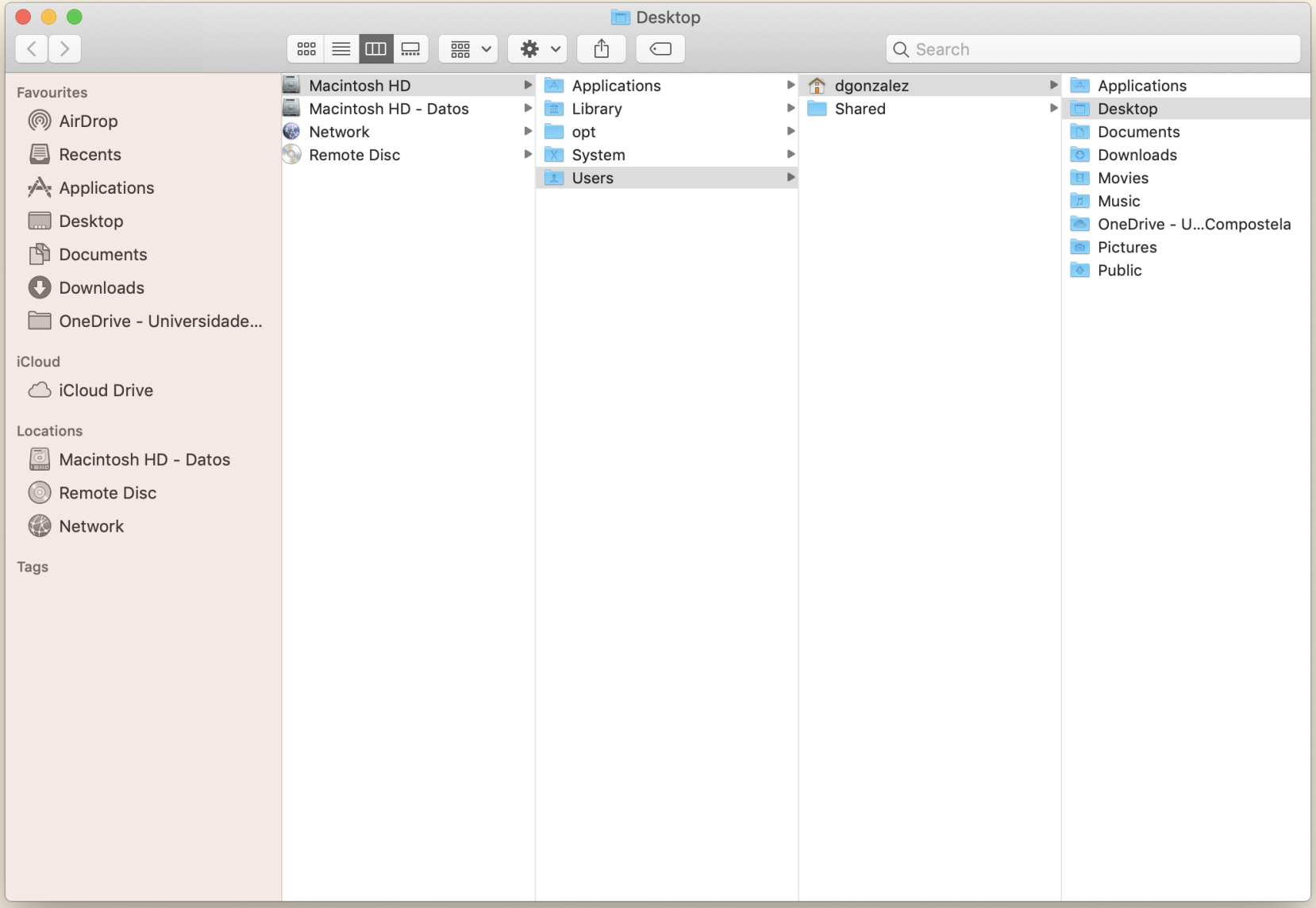

/
/bin
User Binaries
/sbin
System Binaries
/etc
Configuration Files
/dev
Device Files
/proc
Process Information
/var
Variable Files
/tmp
Temporary Files
/usr
User Programs
/home
Home Directories
/boot
Boot Loader Files
/lib
System Libraries
/opt
Optional add-on Apps
/mnt
Mount Directory
/media
Removable Devices
/srv
Service Data


Disque local (C:)
PanoptoRecorder
PerfLogs
Program Files (x86)
Programmes
Utilisateurs
Admin
Public
Windows


Disque local (C:)
PanoptoRecorder
PerfLogs
Program Files (x86)
Programmes
Utilisateurs
Admin
Public
Windows


Windows PowerShell
Windows PowerShell
Copyright (C) Microsoft Corporation. Tous droits réservés.
PS C:\Users\Admin>


Windows PowerShell
Windows PowerShell
Copyright (C) Microsoft Corporation. Tous droits réservés.
PS C:\Users\Admin>


Desktop
Search
Favourites
AirDrop
Recents
Applications
Desktop
Documents
Downloads
OneDrive - Universidade...
iCloud
iCloud Drive
Locations
Macintosh HD - Datos
Remote Disc
Network
Tags
Macintosh HD
Macintosh HD - Datos
Network
Remote Disc
Applications
Library
opt
System
Users
dgonzalez
Shared
Applications
Desktop
Documents
Downloads
Movies
Music
OneDrive - U...Compostela
Pictures
Public

LIST THE CONTENTS OF A DIRECTORY

- UNIX: `ls`
- Windows: `ls` or `dir`

EXERCISE 1

1. Open the terminal
2. List all files and directories
3. Go to your desktop

MODIFYING THE FILE SYSTEM

CREATE A DIRECTORY

- UNIX: **`mkdir`** ***`directory_name`***
- Windows: **`mkdir`** ***`directory_name`*** or **`New-Item`** ***`directory_name`***

COPY A FILE

`cp file_name_source file_name_output`

MOVE A FILE

`mv file_name_source file_name_output`

DELETE A FILE

`rm file_name`

DELETE A DIRECTORY

`rm -r directory_name`

EXERCISE 2

Using the command line:

1. Open the terminal and go to the folder where you downloaded the GitHub repo of this course
2. Create the folder “my_data”
3. Bonus (optional): Copy or move at least one of the files with which you want to work in this course

ADDITIONAL MATERIALS

OS: WINDOWS

- Dawson, Ted: "Introduction to the Windows Command Line with PowerShell". *The Programming Historian*.

OS: ALL

- Milligan, Ian & Baker, James: “Introduction to the Bash Command Line”. *The Programming Historian*.
- W3C: “Character encodings for beginners”.
- Priceton University Library: “File naming and structure”. Research Data Management at Princeton.

RESOURCES

- Encoding Explorer (by Webatic)
- FileInfo (database with file formats and extensions)