File Systems

Nick Ulle

Motivation

In order to read a data set,

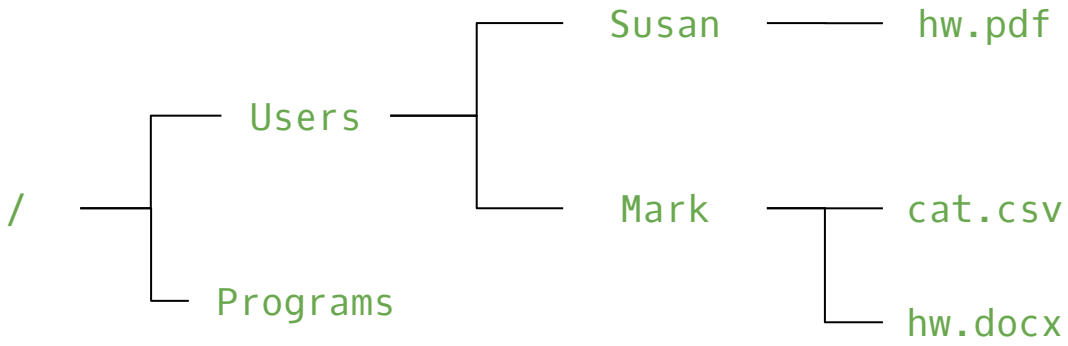
first tell R where the data is.

File Systems

Your computer's files are organized into a **file system**.

The file system looks like an upside-down tree.

Each **directory** (or "folder") is a branch of the tree.



The Root

The **root** is the beginning: the top of your file system.

The name of the root depends on your operating system and computer:

- Mac OS X & Linux: `/`
- Windows: usually `C:/`, but sometimes `D:/`, `E:/`, etc…

Paths

A **path** is a list of directories that lead to a directory or file.

Separate directories with forward slashes `/` (not commas and spaces).

Examples:

```
/Users/Susan/hw.pdf

/home/nick

/snacks/chocolate/
```

Paths

A **path** is a list of directories that lead to a directory or file.

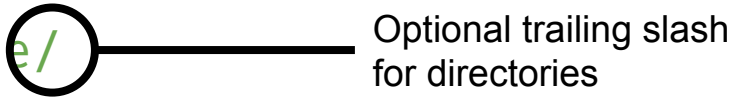
Separate directories with forward slashes `/` (not commas and spaces).

Examples:

```
/Users/Susan/hw.pdf

/home/nick

/snacks/chocolate/
```



Paths on Windows

Windows uses backslashes `\` instead of forward slashes:

```
C:\Users\Susan\hw.pdf
```

```
C:\home\nick\
```

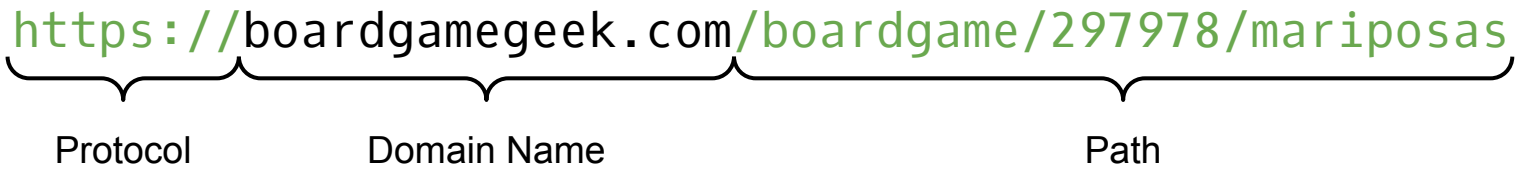
```
C:\snacks\chocolate\
```

But even on Windows, R uses forward slashes in paths:

```
C:/Users/Susan/hw.pdf
```

URLs

Website URLs (uniform resource locators) are a generalization of paths:



For local files, only the path needs to be specified.

Relative Paths, Part 1

An **absolute path** begins from the root.

A **relative path** begins from somewhere else.

An absolute path: `/snacks/chocolate/meiji/hello-panda`

A relative path starting from the `/snacks/chocolate` directory:

`meiji/hello-panda`

Relative Paths, Part 2

Relative paths have two advantages:

1. Shorter
2. Name and location of starting point can vary

Say I keep my snacks in `/junkfood`, not `/snacks`.

The path `chocolate/meiji` still works.

The path `/snacks/chocolate/meiji` doesn't.

Path Shortcuts

A few special characters are shortcuts:

- `.` is the current directory
- `..` is the directory above
- `~` is the home directory (more about this later)

So:

`./chocolate/meiji` is the same as `chocolate/meiji`

`/Users/Mark/..` is the same as `/Users`

The R Working Directory

Nick Ulle

Data Frames

Nick Ulle

Factors

Nick Ulle

File Formats

Nick Ulle