wellcome
connecting
science

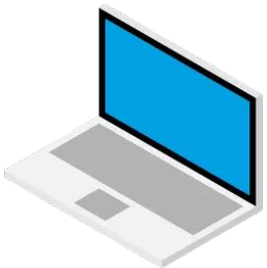
Introduction to the Virtual Machine

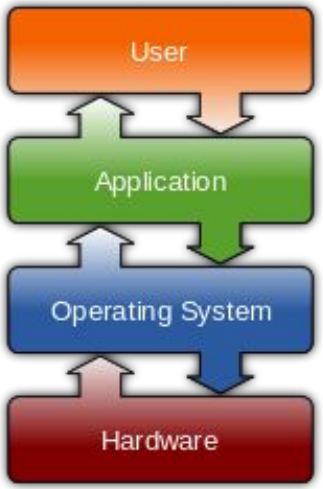
Dr. Jacqui Keane

@drjkeane
drjkeane@gmail.com

Operating System (OS)

- Software that supports the computer's basic functions
  - Manages computer hardware (screen, mouse, keyboard)
  - Provides tools for managing files, running software
  - Provides a way via software applications to interact with the computer



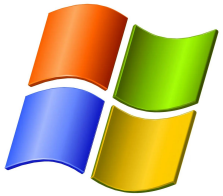


Operating System (OS)

- Examples?

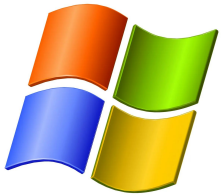
Operating System (OS)

- Examples?



Operating System (OS)

- Examples?



Virtual Machine (VM)

- VM is a computer environment that can be run on any computer
  - OS, data, software applications

- Allows you to run one OS (Linux) on another OS (Windows)

Virtual Machine (VM)

- VM is a computer environment that can be run on any computer
  - OS, data, software applications

- Allows you to run one OS (Linux) on another OS (Windows)

Virtual Machine (VM)

- Created a VM for this course
  - Linux OS
  - Data for practicals
  - Bioinformatics software (fastqc, samtools, ariba etc.)
- Continue to use it after the course (with limitations!)

Questions?

wellcome
connecting
science

Linux for Bioinformatics

Dr. Jacqui Keane

@drjkeane
drjkeane@gmail.com

Unix/Linux

- What is Unix?
  - Standard operating system (alternative to MS Windows, Mac OS)
  - Provides a way for you to interact with the computer (through commands and scripts)
  - Many ‘flavours’ of Unix, using Linux

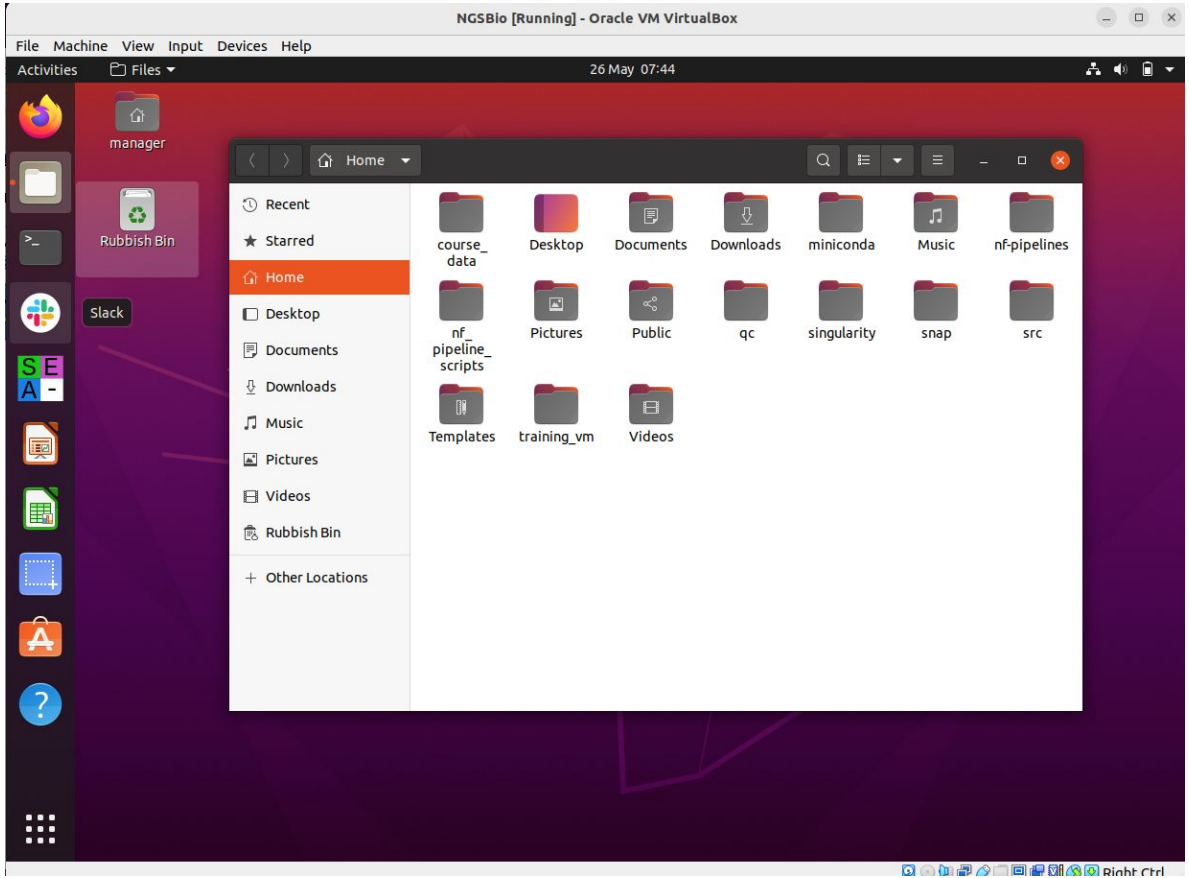
Unix/Linux

- What is Unix?
  - Standard operating system (alternative to MS Windows, Mac OS)
  - Provides a way for you to interact with the computer (through commands and scripts)
  - Many ‘flavours’ of Unix, using Linux

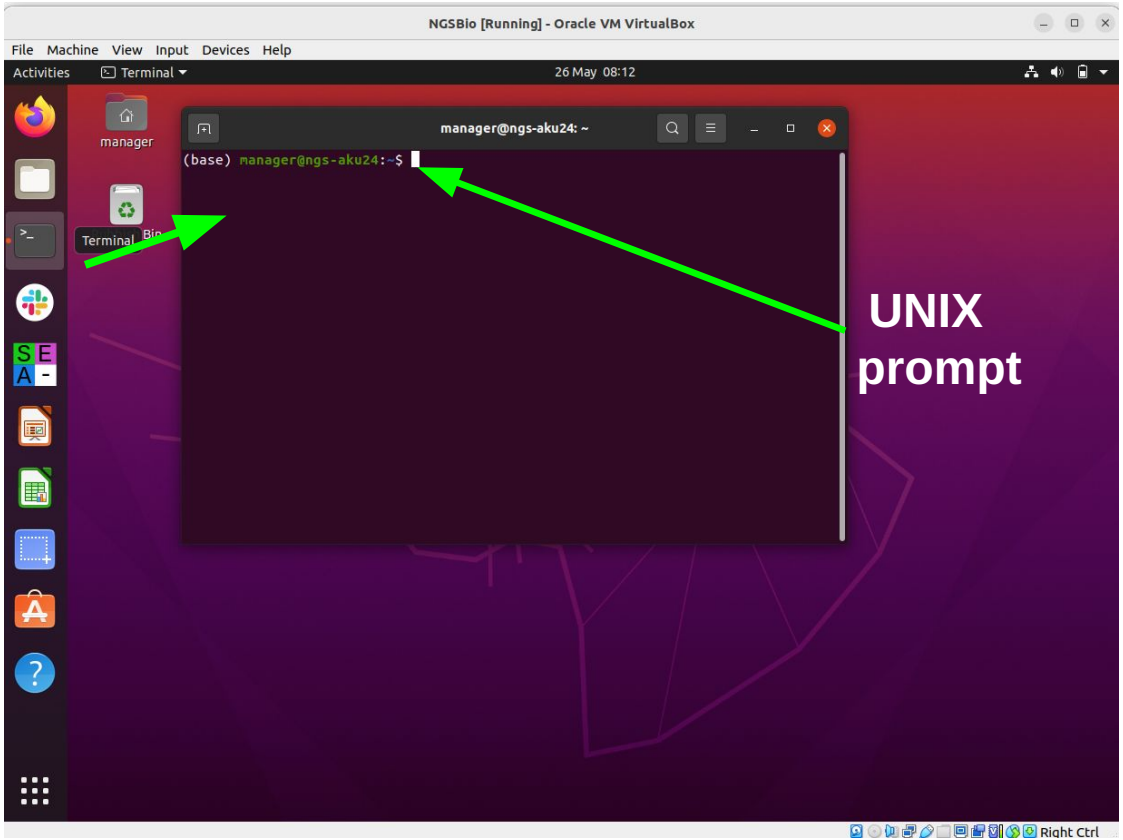
- Why use Unix/Linux?
  - Output of lots of biological research exists in large text files
  - Very suitable for working with such files and widely used in scientific community
  - Powerful and flexible commands for processing large text files
  - Saves you time, automating repetitive tasks
  - Powerful and stable operating system, compatibility with most bioinformatics tools
  - Flexibility and customisation, package management, open source

Using Linux

File Explorer



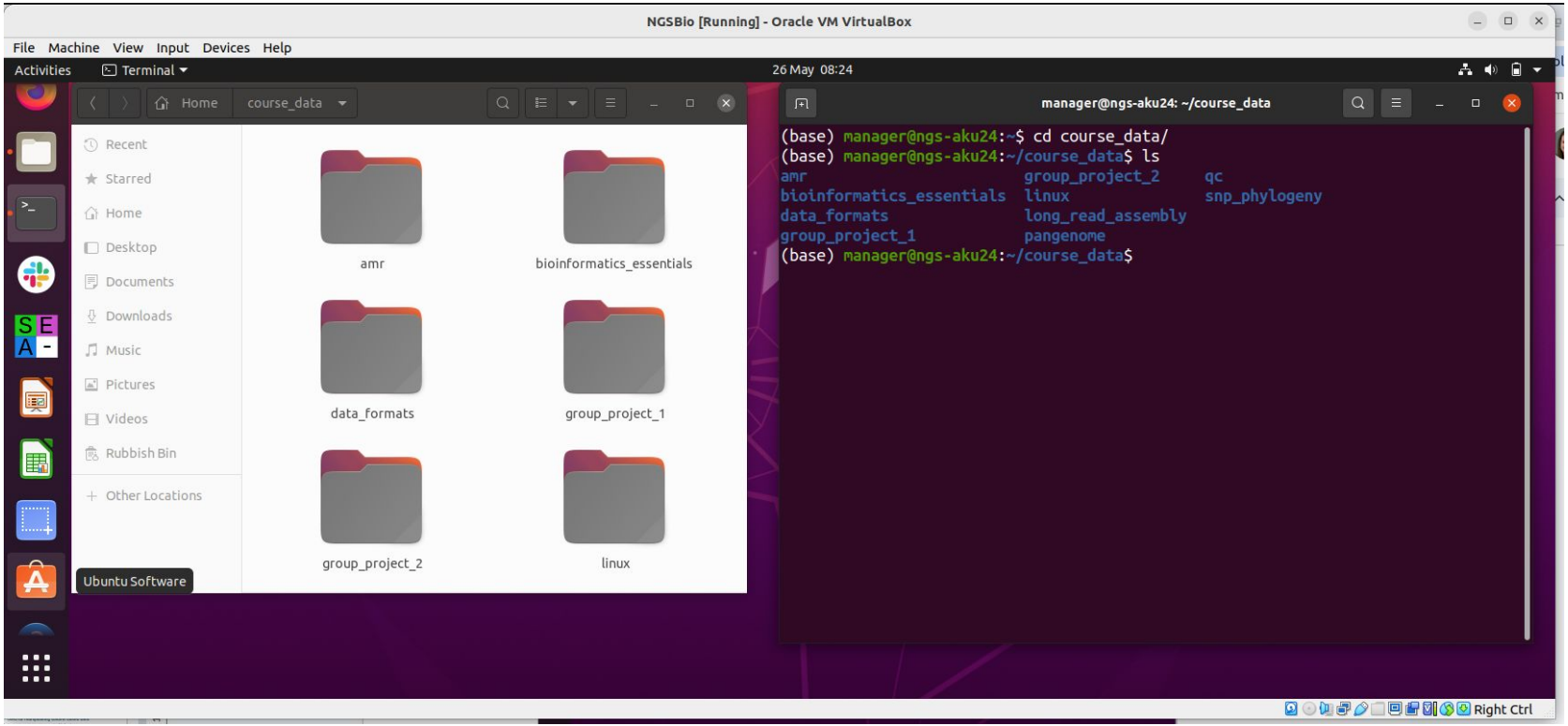
Terminals and Commandline



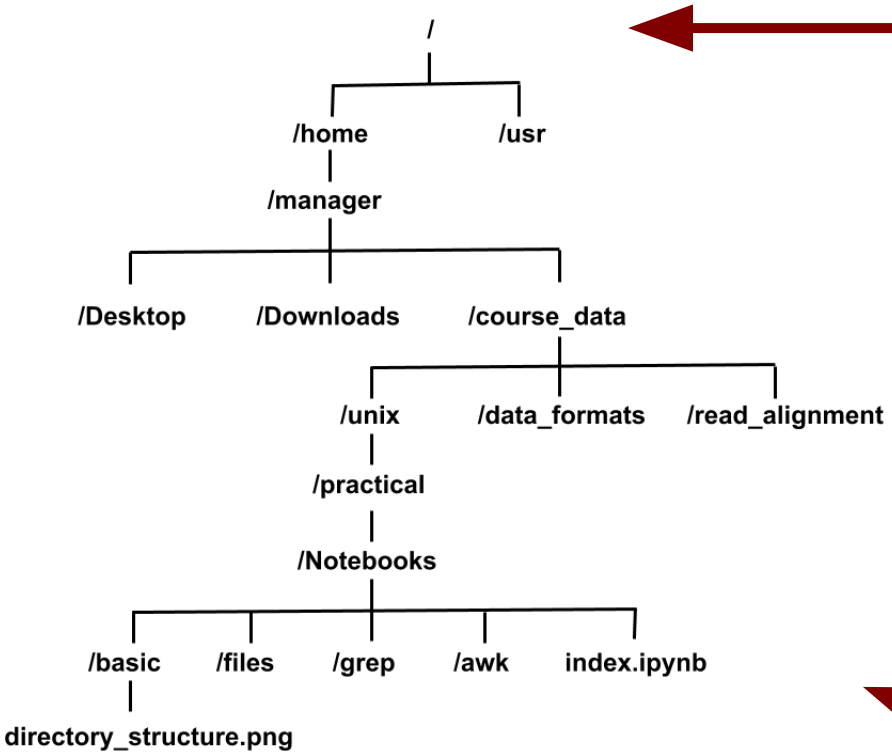
Linux Commands

Command	What it does
`ls`	List the contents of the current directory/folder
`cd`	Changes a directory/folder
`mv`	Moves a file
`cp`	Copies a file
`rm`	Remove a file
`less`	Displays the contents of a file
`head`	Displays the first ten lines of a file
`tail`	Displays the last ten lines of a file
`cat`	Concatenate files together
`pwd`	Print working directory
`mkdir`	Make a new directory

`ls` command


NGSBio [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Activities
Terminal
26 May 08:24
Home
course_data
Recent
Starred
Home
Desktop
Documents
Downloads
Music
Pictures
Videos
Rubbish Bin
Other Locations
amr
bioinformatics_essentials
data_formats
group_project_1
group_project_2
linux
Ubuntu Software
manager@ngs-aku24: ~/course_data
(base) manager@ngs-aku24:~$ cd course_data/
(base) manager@ngs-aku24:~/course_data$ ls
amr
bioinformatics_essentials
data_formats
group_project_1
group_project_2
linux
long_read_assembly
pangenome
qc
snp_phylogeny
(base) manager@ngs-aku24:~/course_data$
Right Ctrl

Directory Structure



root directory

/home/manager/course_data/unix/practical/Notebooks/index.ipynb

Linux Tips & Tricks

- Linux is case sensitive
  - Typing LS is NOT the same as typing ls
- You need to put spaces between
  - a command
  - the values passed to the command
    - **mkdir new_dir will create a new directory**
    - **mkdirnew_dir will just give an error!**
- Linux is not psychic! If you misspell the name of command or a file it will not understand you

Linux Practical

- basic linux - directory structure
- grep - search for occurrences that match a specific pattern
- awk - used for manipulating column based data
- For loops - repeat a process multiple times
- Bash scripting - record a sequence of commands (use later)
    - Basic bash script
    - Where to store your scripts
- Practical exercises are found on the VM (unix.pdf)

Questions?

Getting Started