

2. Operating System and User Interface

Operating System (OS) is an interface between user and the computer applications that handles the input output (I/O) operations in the computer. An operating system is a program that acts as an intermediary between the application programs and the computer hardware. You cannot directly use computer applications (or programs) with computer hardware without a translation system between the hardware and the applications. This translation system is called the OS. The Windows or Mac OS (IOS) works behind-the-scenes to run your computer (i.e. the software and the hardware). It tells the computer what to do when it starts up (the process is called booting) and keeps track of your documents, files, and other software. It also provides the standard user interface component (like menus and the desktop) that you see when you look at your computer screen. Both the Windows and the Mac OS operating systems use a graphical interface (pictures or icons instead of text) that allow you to immerse yourself in multitasking (accessing multiple applications and files simultaneously). While working with windows OS you play with a series of boxes and icons, which can be opened and closed as needed.

Operating System Components

Process management

Networking

Memory management

System Protection

File management

Command-interpreter system

Secondary storage management

The Operating System Layer



Figure 3: Operating System Layers

Functions of the OS

1. Resource sharing: manages resources by ensuring proper and effective use. □ Provision of virtual machine: Hiding the hardware details from the user.

2. Memory management: It manages the effective usage of internal memory, RAM while running multi- programs. ② Protector and error handling.
3. File management, saving, opening the saved file, updating, copying and deleting.
4. Facilitates booting (loading the OS from disk to RAM).
5. Manages multitasking: determines which application should run, in what order, and how much time should be allowed for each application before giving another application access to run.

Classification of OS

This could be done based on the number of program they can handle at the same time or the number of user(s) that can be accommodated on the system simultaneously. Hence we have single tasking and multitasking and, single user and multi-user OS.

Single user OS is multitasking but can only allow only one user at a time e.g. PC-DOS, MS-DOS, CP/M, OS/2. Multi-user OS is multitasking and at the same time multi-user (See the computer Network section for the architecture). This is done when a number of computers (workstation) connect to a central computer so that all the other computers can use the work on the central computer e.g. UNIX, XENIX, PC-MOS, windows NT, Linux and Novell. These OSs are called Network OS.

Distributed operating System

Distributed systems use multiple central processors to serve multiple real time application and multiple users. Data processing jobs are distributed among the processors accordingly to which one can perform each job most efficiently. The processors communicate with one another through various communication lines (such as high-speed buses or telephone lines). These are referred as loosely coupled systems or distributed systems. Processors in a distributed system may vary in size and function. These processors are referred as sites, nodes, and computers and so on.

The advantages of distributed systems are following:

1. With resource sharing facility user at one site may be able to use the resources available at another node.
2. Speedup the exchange of data with one another via electronic mail.
3. If one site fails in a distributed system, the remaining sites can potentially continue operating.
4. Better service to the customers.
5. Reduction of the load on the host computer.
6. Reduction of delays in data processing

Network operating System

Network Operating System runs on a server and provides server the capability to manage data, users, groups, security, applications, and other networking functions. The primary purpose of the network operating system is to allow shared file and printer access among multiple computers in a network, typically a local area network (LAN), a private network or to other networks. Examples of network operating systems are Microsoft Windows Server 2003, Microsoft Windows Server 2008, UNIX, Linux, Mac OS X, Novell NetWare, and BSD.

The advantages of network operating systems are following:

1. Centralized servers are highly stable.
2. Security is server managed.
3. Upgrades to new technologies and hardware can be easily integrated into the system.
4. Remote access to servers is possible from different locations and types of systems.

The disadvantages of network operating systems are following.

1. High cost of buying and running a server.
2. Dependency on a central location for most operations.
3. Regular maintenance and updates are required.

Computer Networks

A computer network is a set of computers sharing resources located on or provided by network nodes. Computers use common communication protocols over digital interconnections to communicate with each other. These interconnections are made up of telecommunication network technologies based on physically wired, optical, and wireless radio-frequency methods that may be arranged in a variety of network topologies. Computer networks support many applications and services, such as access to the World Wide Web, digital video and audio, shared use of application and storage servers, printers and fax machines, and use of email and instant messaging applications.

The nodes of a computer network can include personal computers, servers, networking hardware, or other specialized or general-purpose hosts. They are identified by network addresses and may have hostnames. Hostnames serve as memorable labels for the nodes and are rarely changed after initial assignment. Network addresses serve for locating and identifying the nodes by communication protocols such as the Internet Protocol.

Computer networks may be classified by many criteria, including the transmission medium used to carry signals, bandwidth, communications protocols to organize network traffic, the network size, the topology, traffic control mechanisms, and organizational intent.

LAN: Local Area Network is small network setup within the office premises to share the resources from Server to clients (individual PCs).

WAN: Wide Area Network is setup within the organization in a given location to share the resources within organization.

WWW: World Wide Web is the largest Network of computers using Internet Protocol (IP) to share the resources worldwide. See the following figure to get an idea of how a modern day organization utilizes the LAN, WAN and WWW network to accomplish the computing needs.

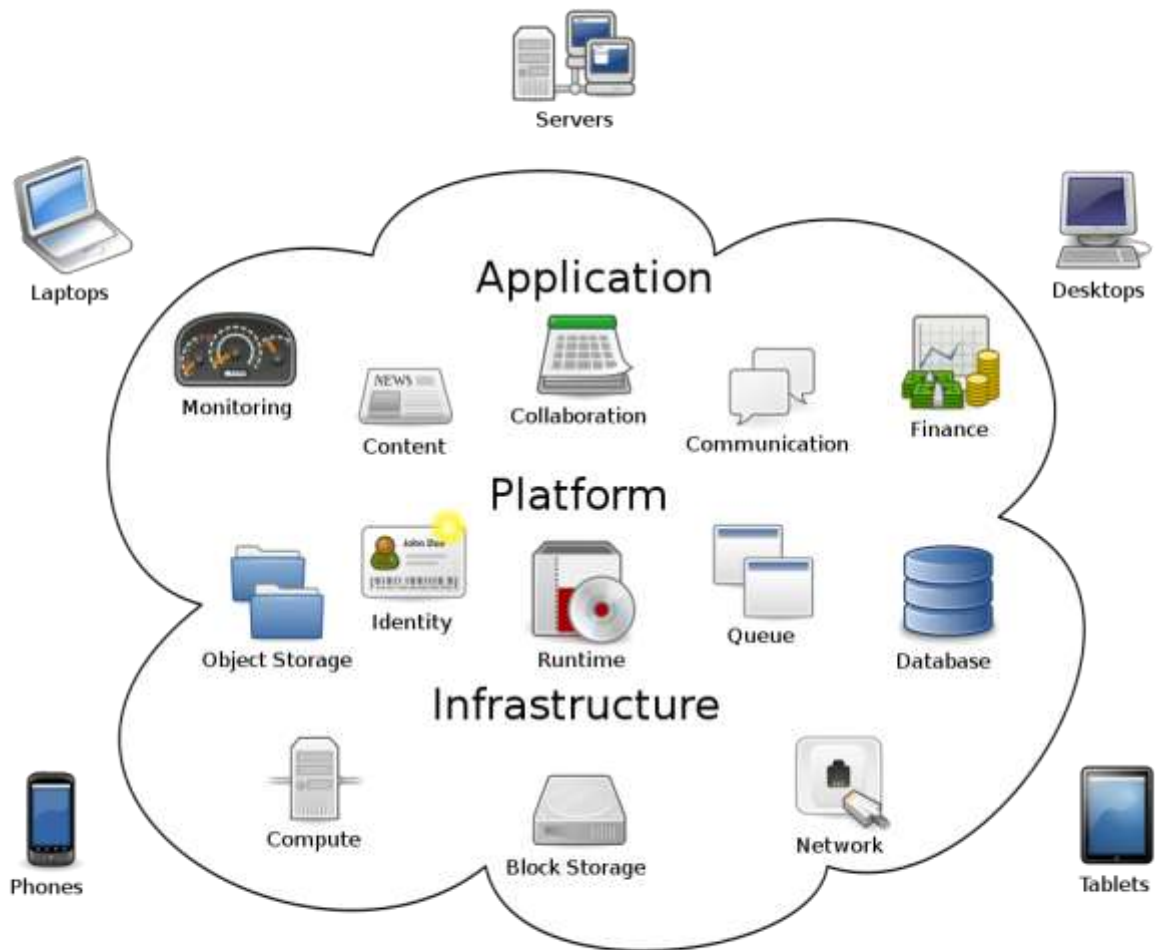


A typical network setup in the modern-day computing environment

Cloud Computing

Cloud computing is the on-demand availability of computer system resources, especially data storage (cloud storage) and computing power, without direct active management by the user. Large clouds often have functions distributed over multiple locations, each of which is a data center. Cloud computing relies on sharing of resources to achieve coherence and typically uses a pay-as-you-go model, which can help in reducing capital expenses but may also lead to unexpected operating expenses for users. The advantages of using cloud services are lower implementation costs, more secure platform to store and retrieve the resources and availability of latest technology to use. However, the downside is the exposure of sensitive to data to third party service provider and the possibility of infringement of property rights.

The big player in cloud services are: Azure from Microsoft, AWS from Amazon, Google Cloud, and Oracle.



Cloud computing metaphor: the group of networked elements providing services does not need to be addressed or managed individually by users; instead, the entire provider-managed suite of hardware and software can be thought of as an amorphous cloud.

Virtual Machines

A Virtual Machine (VM) is a compute resource that uses software instead of a physical computer to run programs and deploy apps. One or more virtual “guest” machines run on a physical “host” machine. Each virtual machine runs its own operating system and functions separately from the other VMs, even when they are all running on the same host. This means that, for example, a virtual MacOS virtual machine can run on a physical PC.

Virtual machine technology is used for many use cases across on-premises and cloud environments. More recently, public cloud services are using virtual machines to provide virtual application resources to multiple users at once, for even more cost efficient and flexible compute.

What are virtual machines used for?

VMs allow a business to run an operating system that behaves like a completely separate computer in an app window on a desktop. VMs may be deployed to accommodate different levels of processing power needs, to run software that requires a different operating system, or to test applications in a safe, sandboxed environment.

Virtual machines have historically been used for server virtualization, which enables IT teams to consolidate their computing resources and improve efficiency. Additionally, virtual machines can perform specific tasks considered too risky to carry out in a host environment, such as accessing virus-infected data or testing operating systems. Since the virtual machine is separated from the rest of the system, the software inside the virtual machine cannot tamper with the host computer.

Advantages of virtual machines

Virtual machines are easy to manage and maintain, and they offer several advantages over physical machines:

VMs can run multiple operating system environments on a single physical computer, saving physical space, time and management costs.

Virtual machines support legacy applications, reducing the cost of migrating to a new operating system. For example, a Linux virtual machine running a distribution of Linux as the guest operating system can exist on a host server that is running a non-Linux operating system, such as Windows.

VMs can also provide integrated disaster recovery and application provisioning options.

Disadvantages of virtual machines

While virtual machines have several advantages over physical machines, there are also some potential disadvantages:

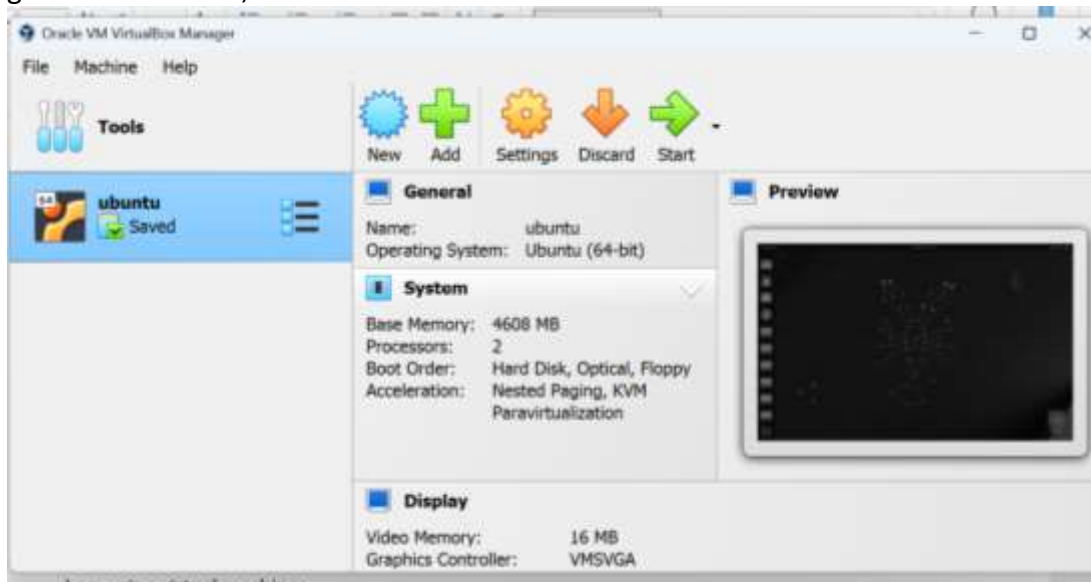
Running multiple virtual machines on one physical machine can result in unstable performance if infrastructure requirements are not met.

Virtual machines are less efficient and run slower than a full physical computer. Most enterprises use a combination of physical and virtual infrastructure to balance the corresponding advantages and disadvantages.

Linux OS and File Systems

Unix was created at Bell Labs in 1970 written in the C programming language, which was developed at the same time. It supports large numbers of simultaneous users, runs with few alterations on many hardware platforms (provides some platform independence) and of course it was and is a simple, elegant, and easy to use (at least compared to its predecessors). Linux is a relatively new UNIX flavor derived from the work of Linus Torvalds, who was interested to develop a UNIX for academic use. Linux is one of the many versions of the UNIX operating system. Working on Linux means working on one of the flavors of UNIX. The main advantage that Linux is absolutely free, you need not spend even the cost CD it can be entirely free downloadable from the Internet. (No registration fees, no costs per user, free updates, and freely available Operating System and Networking Lab source code). It is portable (means can be configured on any processor like Intel, Solaris, etc.), dual-bootable, fast, reliable, secure and versatile. These properties make it popular among the System Administrators. While working on it you may realize many more important features and advantages of Linux. Most of the exercises are command line based but similar exercises you can try on your GUI based Linux.

I am using virtual machines called Oracle VM with ubuntu linux OS. By clicking on Start icon, we get the ubuntu on, the host is Windows OS.



We can run OS commands either via command line interface (CLI) or graphical user interface (GUI). Ubuntu provides nice GUI to explore the file systems and you can create a folder/directory by right clicking on the folder icon as shown below:

- `date` show date and time
- `history` list of previously executed commands
- `pine` send or receive mail messages
- `msgs` display system messages
- `man` show on-line documentation by program name
- `info` on-line documentation for GNU programs
- `w, who` who is on the system and what are they doing
- `who am i` who is logged onto this terminal
- `top` show system status and top CPU-using processes
- `uptime` show one line summary of system status
- `finger` find out info about a user@system

File Management

- `cat` combine files
- `cp` copy files
- `ls` list files in a directory and their attributes
- `mv` change file name or directory location
- `rm` remove files
- `ln` create another link (name) to a file
- `chmod` set file permissions
- `des` encrypt a data file with a private key
- `find` find files that match specified criteria

Display Contents of Files

- `cat` copy file to display device
- `vi` screen editor for modifying text files
- `more` show text file on display terminal with paging control
- `head` show first few lines of a file(s)

42

- `tail` show last few lines of a file; or reverse line order
- `grep` display lines that match a pattern
- `lpr` send file to line printer
- `pr` format file with page headers, multiple columns etc.
- `diff` compare two files and show differences
- `cmp` compare two binary files and report if different
- `od` display binary file as equivalent octal/hex codes
- `file` examine file(s) and tell you whether text, data, etc.
- `wc` count characters, words, and lines in a file

Operat

Directories

- `cd` change to new directory
- `mkdir` create new directory
- `rmdir` remove empty directory (remove files first)
- `mv` change name of directory
- `pwd` show current directory

Description of Commonly Used UNIX Commands

The description for the most commonly used UNIX commands is given below in an alphabetic order.

a) cat

cat allows you to read multiple files and then print them out. You can combine files by using the > operator and append files by using >>.

Syntax: cat [argument] [specific file]

Example:

cat abc.txt

If you want to append three files (abc.txt, def.txt, xyz.txt), give the command as,

cat abc.txt def.txt xyz.txt > all

b) cd, chdir

cd (or chdir) stands for "change directory". This command is the key command to move around your file structure.

Syntax: cd [name of directory you want to move to]

When changing directories, start with / and then type the complete file path, like

cd /vvs/abc/def

in this eg. You are moving from vvs directory to abc directory then move to def directory .

You can also move relative to the current directory by typing

cd vvs/abc/def

To move relative to the parent directory of your current directory, use

Cd ../vvs/abc/def

c) chmod

chmod (which stands for "change mode") changes who can access a particular file. A "mode" is created by combining the various options from who, opcode, and permission.

Syntax: chmod [option] mode file

If you look at a list of files using the long list command ls -l, you'll see the permissions, owner, file size, modification time, and filename. The first column of the list shows who can read, write, and execute the files or directories, in other words, the permissions. It basically shows who has permission to do what to a given file or directory. r stands for "read" and means that you're allowed to read the file or directory. w stands for "write" and gives permission to edit or change the file as well as create, move, rename, or remove a directory. x stands for "execute" which gives permission

to run a file or search a directory. Every file or directory has four sets of rwx permissions. The first set represents the user (u), the second set represents the group (g), the third set represents other (o), and the fourth set represents all (a). The column will look like this: rwxrwxrwx

Each set of rwx represents user, group, and other respectively. Only the owner of a file or a privileged user may change the permissions on a file. There are two ways to change permissions on a file or directory, either numerically or by using lettered commands. Both ways use the command `chmod`. To add permissions to a file, you use `+`, to remove permissions you use `-`.

For example, take a file:

```
-rw-r--r-- 1 yash mony 476 Apr 14 17:13 vvs.txt
```

To allow a group (mony, in this case) "write" access, you would type:

```
chmod g+w vvs.txt
```

If you wanted to remove "read" ability from "other" you would type:

```
chmod o-r vvs.txt
```

It is also possible to specify permissions using a three-digit sequence. This is a more efficient way to change permissions (or at least it requires less typing), so use this method if it doesn't confuse you. Each type of permission is given an octal value. Read is given the value of 4, write is given the value of 2, and execute is given the value of 1. These values are added together for each user category. The permissions are changed by using a three-digit sequence with the first digit representing owner permission, the second digit representing group permission, and the third digit representing other permission. For example, if you wanted to make `vvs.txt` readable, writable, and executable for the user, readable and writable for the group, and readable for other, you would type: `chmod 764 vvs.txt`

The first digit means readable and writable for the user (4+2+1), the second digit means readable and writable for the group (4+2+0), and the third digit means readable for other (4+0+0). if you want to change the permissions on a directory tree use the `-R` option. `chmod -R` will recursively change the permissions of directories and their contents.

d) cp

The `cp` command copies files or directories from one place to another. You can copy a set of files to another file, or copy one or more files under the same name in a directory. If the destination of the file you want to copy is an existing file, then the existing file is overwritten. If the destination is an existing directory, then the file is copied into that directory.

Syntax: `cp [options] file1 file2` If you want to copy the file `favourites.html` into the directory called `laksh`, you give

the command as:

```
cp favourites.html /vvs/laksh/
```

A handy option to use with `cp` is `-r`. This recursively copies a particular directory and all of its contents to the specified directory, so you won't have to copy one file at a time.

e) `date`

The `date` command can be used to display the date or to set a date. In unix the term date includes the time as well.

Syntax: `date [option] [+format]`

```
date [options] [string]
```

The first structure shows how `date` can be used to display the current date. A certain format can be specified in which the date should be displayed. Check the Unix manual for specific formats and options. The second structure allows you to set the date by supplying a numeric string. Only privileged users will be able to use this second command structure.

f) `diff`

`diff` displays the lines that differ between two given files.

Syntax: `diff [options] [directory options] file1 file2`

`diff` can be an extremely valuable tool for both checking errors and building new pages. If you run a `diff` between two files, you'll be shown what differences the files have line by line. The lines referring to `file1` are marked with the `<` symbol. The lines referring to `file2` are marked by the `>` symbol. If the file is a directory, `diff` will list the file in the directory that has the same name as `file2`. If both of the files are directories, `diff` will list all the lines differing between all files that have the same name. If you have a file that is not working properly, it can be a great help to check it against a similar file that is working. It will often quickly alert you to a line of code that's missing. A handy option to use if you want to generally compare two files without noting the complex differences between them is the `-h` option (`h` stands for half-hearted). Using `-i` as an option will ignore differences in uppercase and lowercase characters between files, and `-b` will ignore repeating blanks and line breaks.

g) exit

The exit command allows you to terminate a process that is currently occurring. For example, if you wanted to leave a remote host that you were logged onto (see rlogin also), you should type exit. This would return you to your home host.

h) find

find searches through directory trees beginning with each pathname and finds the files that match the specified condition(s). You must specify at least one pathname and one condition.

Syntax: find pathname(s) condition(s)

There are several handy conditions you can use to find exactly what you want. The -name condition will find files whose names match a specified pattern. The structure for the name condition is: find pathname -name pattern

The condition -print will print the matching files to the pathname specified. -printcan also be used in conjunction with other conditions to print the output. If you wanted to find all the files named favorites.html in the directory Ram, then

you'd do this:

```
find /Ram -name favorites.html -print
```

This looks through the directory Ram and finds all the files in that directory that contain favorites.html, then prints them to the screen. Your output would look like this:

```
/Ram/sixteen_candles/favorites.html
```

```
/Ram/favorites.html
```

```
/Ram/breakfast_club/favorites.html
```

All meta-characters (!, *, ., etc.) used with -name should be escaped (place a \ before the character) or quoted. Meta-characters come in handy when you are searching for a pattern and only know part of the pattern or need to find several similar patterns.

For example, if you are searching for a file that contains the word "favorite", then use the meta-character * to represent matching zero or more of the preceding characters.

This will show you all files which contain favorite.

```
find /Ram -name '*favorite*' -print
```

This looks through the directory Ram and finds all the files in that directory that contain the word "favorite". The output would look like this:

```
/Ram/sixteen_candles/favorites.html
```

```
/Ram/favorites.html
```

```
/Ram/least_favorites.html
```

```
/Ram/breakfast_club/favorites.html
```

```
/Ram/favorite_line.html
```

The -user condition finds files belonging to a particular user ID or name.

i) **grep**

The grep command searches a file or files for lines that match a provided regular expression ("grep" comes from a command meaning to globally search for a regular expression and then print the found matches).

Syntax: `grep [options] regular expression [files]`

To exit this command, type 0 if lines have matched, 1 if no lines match, and 2 for errors. This is very useful if you need to match things in several files. If you wanted to find out which files in our vvs directory contained the word "bca" you could use grep to search the directory and match those files with that word. All that you have to do is give the command as shown:

```
grep 'bca' /vvs/*
```

The * used in this example is called a meta-character, and it represents matching zero or more of the preceding characters. In this example, it is used to mean "all files and directories in this directory". So, grep will search all the files and directories in vvs and tell you which files contain "bca".

j) **head**

Displays the first ten lines of a file, unless otherwise stated. Syntax: `head [-n] [files]`

For example, the following command will display the first 15 lines of favourites.html.

```
head -15 favourites.html
```

k) **kill**

kill ends the execution of one or more process ID's. In order to do this you must own the process or be designated a privileged user. To find the process ID of a certain job give the command ps.

Syntax: kill [options] PIDs

There are different levels of intensity to the kill command, and these can be represented either numerically or symbolically. kill -1 or HUP makes a request to the server to terminate the process, while kill -9 or kill KILL forces a process to terminate absolutely. Most politely, UNIX users will attempt to kill a process using -1 first before forcing a process to die.

l) ls

ls will list all the files in the current directory. If one or more files are given, ls will display the files contained within "name" or list all the files with the same name as "name". The files can be displayed in a variety of formats using various options.

Syntax: ls [options] [names]

ls is a command you'll end up using all the time. It simply stands for list. If you are in a directory and you want to know what files and directories are inside that directory, type ls. Sometimes the list of files is very long and it flies past your screen so quickly. You miss the file you want. To overcome this problem give the command as shown below:

ls | more

The character | (called pipe) is typed by using shift and the \ key. | more will show as many files as will fit on your screen, and then display a highlighted "more" at the bottom. If you want to see the next screen, hit enter (for moving one line at a time) or the spacebar (to move a screen at a time). | more can be used anytime you wish to view the output of a command in this way. A useful option to use with ls command is -l. This will list the files and directories in a long format. This means it will display the permissions (see chmod), owners, group, size, date and time the file was last modified, and the filename.

```
drwxrwxr-x vvs staff 512 Apr 5 09:34 sridhar.txt
```

```
-rwx-rw-r-- vvs staff 4233 Apr 1 10:20 resume.txt
```

```
-rwx-r--r-- vvs staff 4122 Apr 1 12:01 favourites.html
```

There are several other options that can be used to modify the ls command, and many of these options can be combined. -a will list all files in a directory, including those files normally hidden. -F will flag filenames by putting / on directories, @ on symbolic links, and * on executable files.

m) man

The man command can be used to view information in the online Unix manual.

Syntax: `man [options] [[section] subjects]`

`man` searches for information about a file, command, or directory and then displays it on your screen. Each command is a subject in the manual. If no subject is specified, you must give either a keyword or a file. You can also search for commands that serve a similar purpose. For example, if you want more information about the `chmod` command, you should type:

```
man chmod
```

A screen will then appear with information about `chmod`. Type `q` to quit.

n) `mkdir`

`mkdir` creates a new directory.

Syntax: `mkdir [options] directory name`

For example, to create a directory called `prakhayath` in the present working directory, give the command as,

```
mkdir prakhayath
```

o) `more`

`more` displays the contents of files on your screen.

Syntax: `more [options] [files]`

To have the next line displayed, hit the return key, otherwise press the spacebar to bring up the next screen. Press `h` for assistance with other commands, `n` to move to the next file, or `q` to quit.

p) `less`

`less` is similar to `more` in that it displays the contents of files on your screen. Unlike `more`, `less` allows backward and forward movement within the file. It does not read the whole file before displaying its contents, so with large files `less` displays faster than `more`. Press `h` for assistance with other commands or `q` to quit.

Syntax: `less [options] [files]`

q) `mv`

`mv` moves files and directories. It can also be used to rename files or directories.

Syntax: `mv [options] source target`

If you wanted to rename `vvs.txt` to `vsv.txt`, you should give the command as:

Similarly, we can run OS commands using CLI or GUI in Windows as seen in the classroom. Just type `cmd` in the search box and press Enter to go to command prompt as shown below:

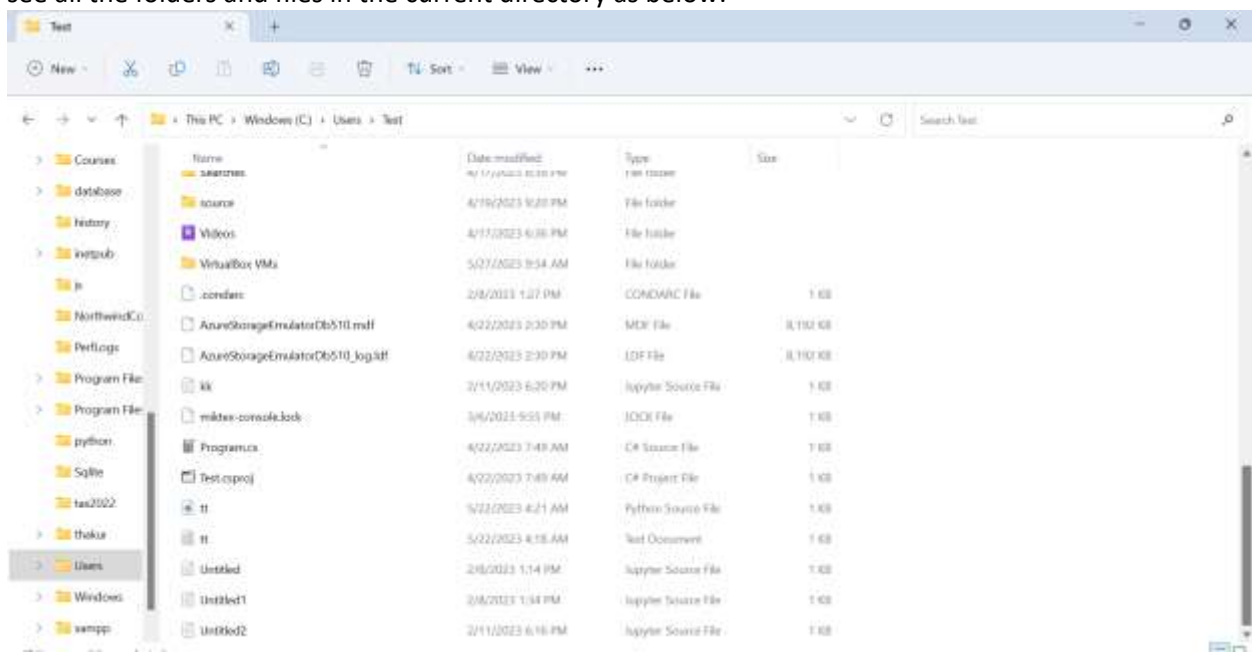

```
Command Prompt
04/22/2023 02:30 PM      8,388,688 AzureStorageEmulatorOb510.mdf
04/22/2023 02:30 PM      8,388,688 AzureStorageEmulatorOb510_log.ldf
04/22/2023 08:07 AM      <DIR>      bin
04/17/2023 06:36 PM      <DIR>      Contacts
04/17/2023 06:36 PM      <DIR>      Desktop
05/05/2023 03:53 AM      <DIR>      Documents
05/29/2023 08:42 PM      <DIR>      Downloads
04/17/2023 06:36 PM      <DIR>      Favorites
02/11/2023 07:20 PM      891 Wk.ipynb
04/17/2023 06:36 PM      <DIR>      Links
03/06/2023 10:55 PM      22 miktex-console.lock
04/17/2023 06:36 PM      <DIR>      Music
04/22/2023 08:07 AM      <DIR>      obj
04/17/2023 06:36 PM      <DIR>      OneDrive
04/17/2023 06:36 PM      <DIR>      Pictures
04/22/2023 07:49 AM      186 Program.cs
02/08/2023 02:40 PM      <DIR>      PythonProject
04/17/2023 06:36 PM      <DIR>      Saved Games
04/17/2023 06:36 PM      <DIR>      Searches
04/19/2023 09:20 PM      <DIR>      source
04/22/2023 07:49 AM      171 Test.csproj
05/22/2023 04:21 AM      190 tt.py
05/22/2023 04:18 AM      185 tt.txt
02/08/2023 02:14 PM      589 Untitled.ipynb
02/08/2023 02:34 PM      1,005 Untitled1.ipynb
02/11/2023 07:16 PM      965 Untitled2.ipynb
04/17/2023 06:36 PM      <DIR>      Videos
05/27/2023 09:54 AM      <DIR>      VirtualBox VMs
12 File(s)      16,781,445 bytes
35 Dir(s)      338,758,316,032 bytes free

C:\Users\Test>
```

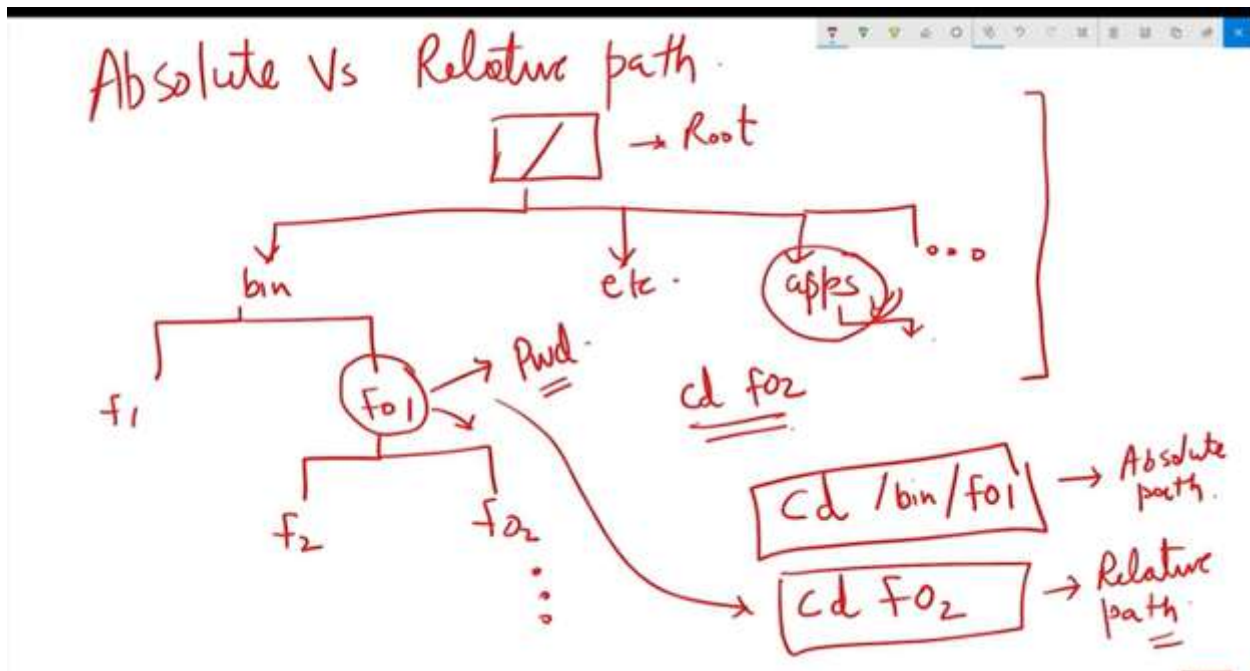
After invoking dir command we can see all the files and folder in the current directory named c:\Users\Test>

All the folders are indicated by <DIR> and the files are shown with file size and the filenames with extension. E.f., tt.py indicates this file is of python program, extension .py indicates so, and the file size is 190 bytes. File extension gives us some idea about the type of information contained in that file.

The same information can be seen using GUI in Windows. Just click on the disk icon on the task bar, and see all the folders and files in the current directory as below:



You can right click the mouse and create a new folder in the current directory easily. We must have a separate folder to manage the similar/related files. The file name should also be informative so that we can open the right file for the right information in the future when we need to work on that file.



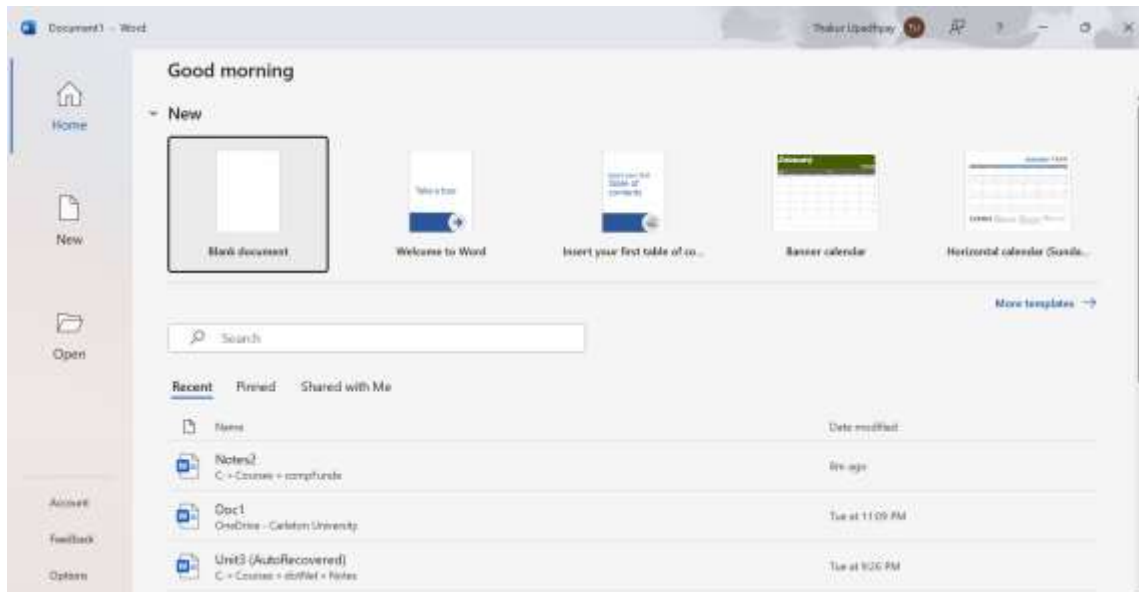
The above figure shows a typical folder structure of a disk, where / indicates the root directory and from there you can have many branches of folders and files as leaves, like a tree. You can run tree command to see the structure of the file systems in your disk drive.

Moreover, you need to keep track of the relative file path to locate the particular file in the given folder. E.g. If you are at bin folder and want to get to fo2 folder, then invoke `cd fo1/fo2`, and from fo2 to apps then you have to move to the root (\) first and start the journey from there. The command will be `cd\apps`.

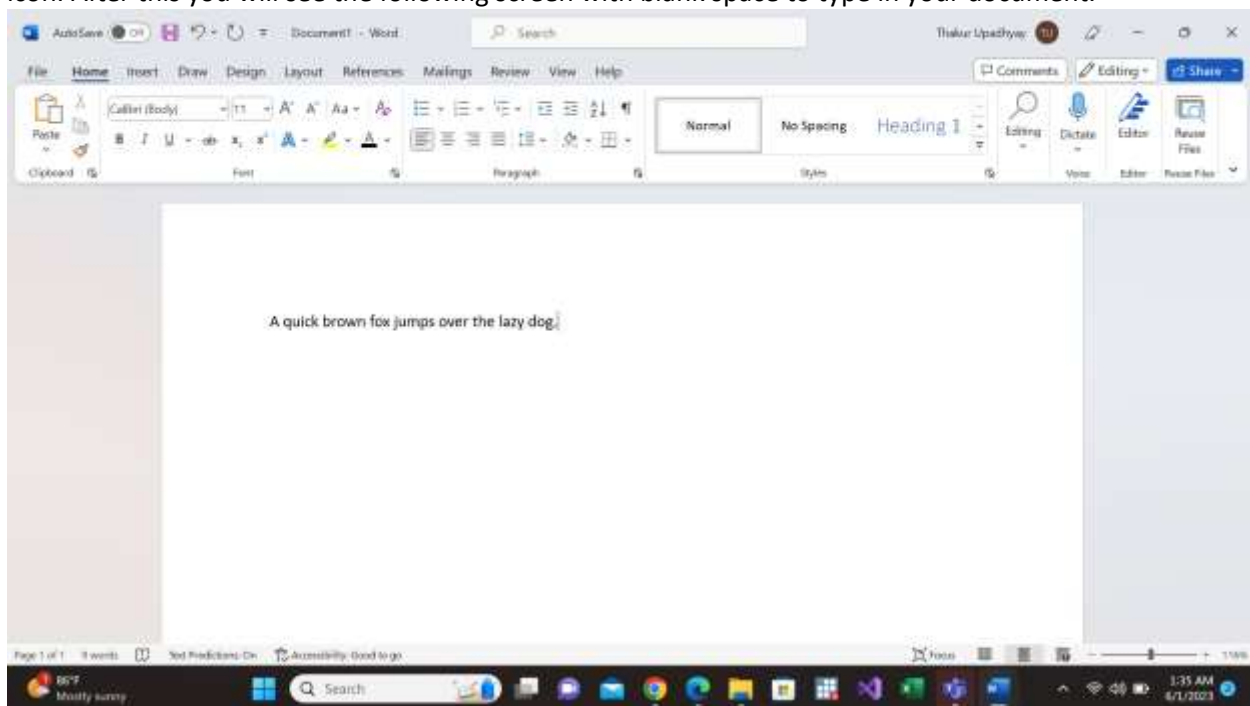
Office packages: Brief introduction to Word processor, Spreadsheet, power point presentation

Microsoft Inc has three popular applications software for managing the daily office work. We have MS Word for word processing and document creation and editing, and MS Excel for spreadsheet calculations, where we can solve the computations of different transactions or records in a flat file or a Table. MS PPT is designed for preparing beautiful presentations of different reports and project works.

To run the Word application, just type word in the search box or click on the Word icon on your window, then you will see the following interface.



You can see the recently opened files (e.g. Notes1, Doc1 etc.) and a box saying Blank coument as caption. If you want to work on the previously saved document then just click on the file and you will see the document open on the window. If you want to start a new document then click on the Blank document icon. After this you will see the following screen with blank space to type in your document.



Here you can type your document, format it as appropriate and save it. To save the file you just click on File menu and Save. You need to save your file in the right folder and the file name should be informative. The etension of Word file will be .docx by default. Following are useful short cut keys to be used while working on the Word document.

To do this	Press
Open a document.	Ctrl+O
Create a new document.	Ctrl+N
Save the document.	Ctrl+S
Close the document.	Ctrl+W
Cut the selected content to the Clipboard.	Ctrl+X
Copy the selected content to the Clipboard.	Ctrl+C
Paste the contents of the Clipboard.	Ctrl+V
Select all document content.	Ctrl+A
Apply bold formatting to text.	Ctrl+B
Apply italic formatting to text.	Ctrl+I
Apply underline formatting to text.	Ctrl+U
Decrease the font size by 1 point.	Ctrl+Left bracket ([)
Increase the font size by 1 point.	Ctrl+Right bracket (])
Center the text.	Ctrl+E
Align the text to the left.	Ctrl+L
Align the text to the right.	Ctrl+R
Cancel a command.	Esc
Undo the previous action.	Ctrl+Z
Redo the previous action, if possible.	Ctrl+Y
Adjust the zoom magnification.	Alt+W, Q, then use the Tab key in the Zoom dialog box to go to the value you want.
Split the document window.	Ctrl+Alt+S

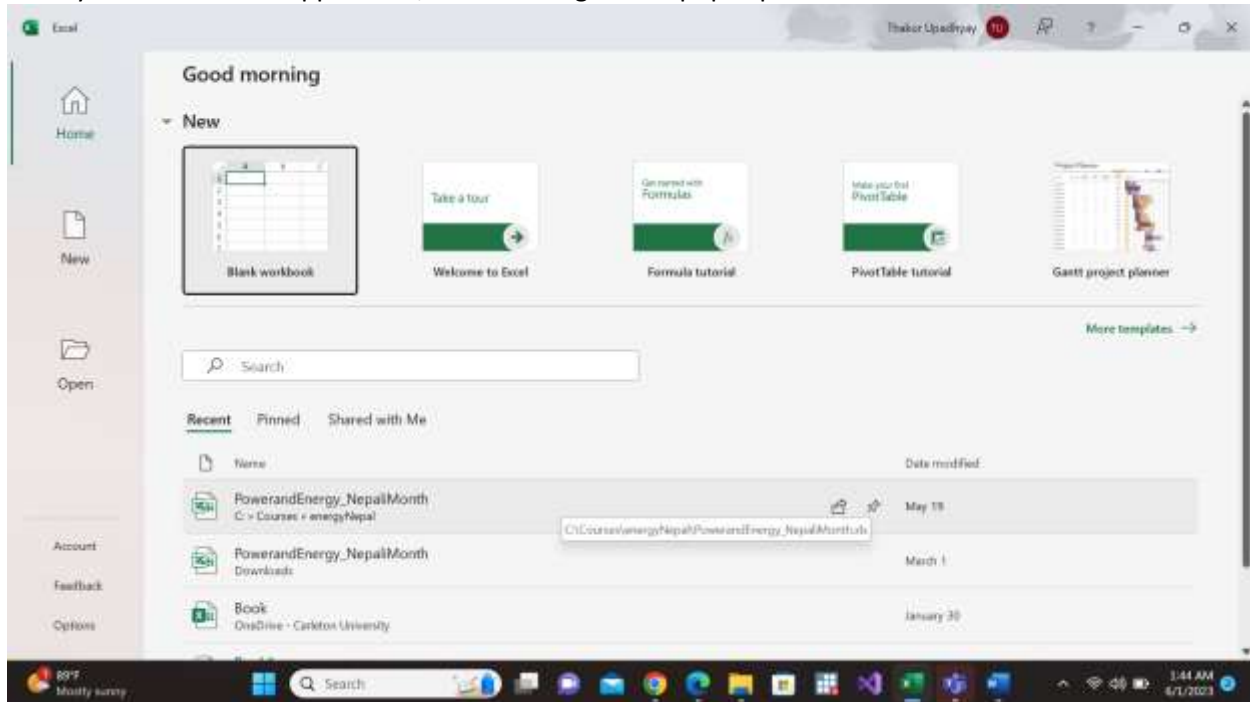
These short cuts are also applicable to Excel and PPT as per the needs.

We can copy and paste the part of the document from one apps to another when needed. E.g., we can make a nice table with some calculations in excel and bring that table to our word document for preparing the report. Also, some of these contents can be copied and pasted in the PPT for preparing the presentation. You will be practicing more functions to format the document in MS Word in lab.

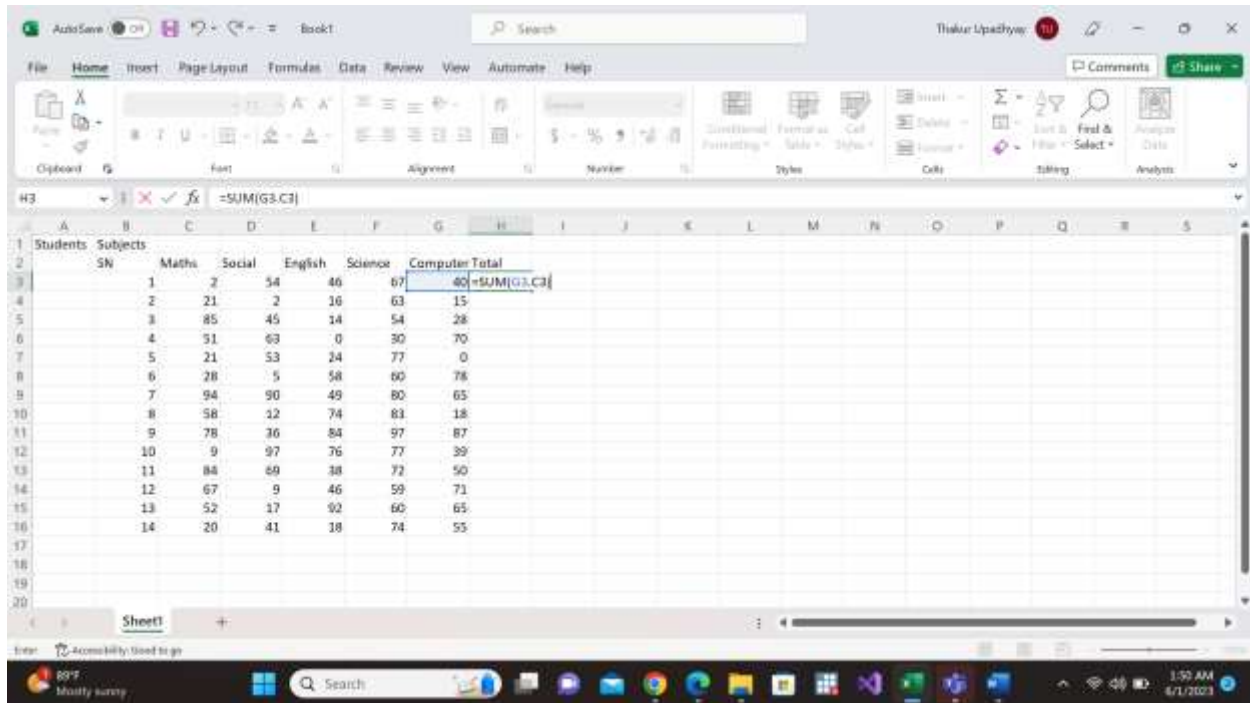
MS Excel

MS Excel is another application that deals with calculations, visualization and some statistical functions in a spreadsheet. A spreadsheet is divided into cells with corresponding cell references that starts with column name and ends with row number. E.g. A1 indicates the first cell with column A and row number 1. All calculations and function computations are based on these references.

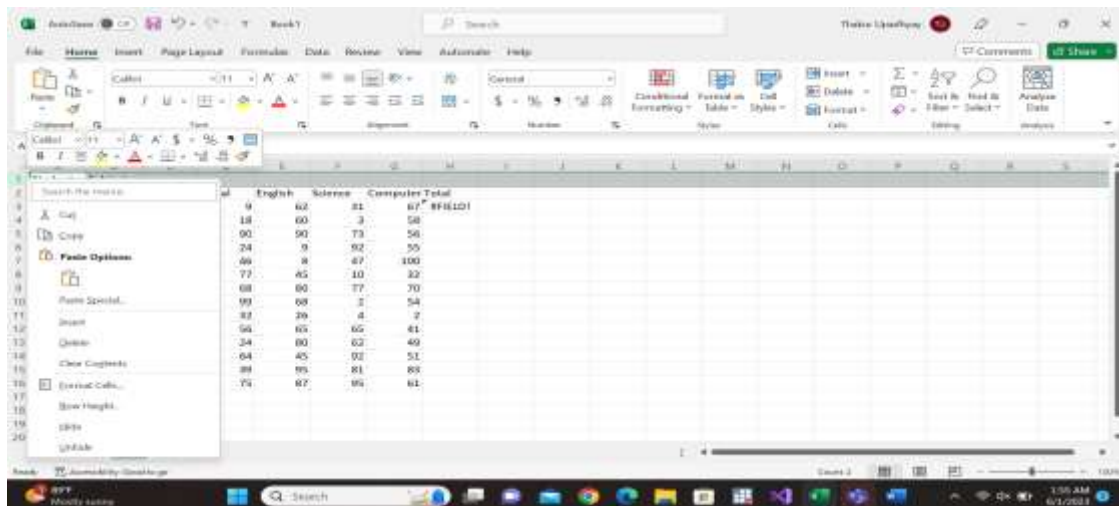
To run the excel application, just type excel on the search box or click on Excel icon from application lists. Once you run the excel application, the following screen pops up.



Here also, you can see the recently opened excel files and the blank workbook. To work a new, just click on Blank Workbook, then you will see the following screen.



Here, I have prepared a marksheets of 14 students in a class with five subjects. An excel sheet (Sheet1 here) contains cells formed by intersestion of a row and a column. The cursor in the screen is at the locaiton H3 with a formula to calculate the sum of the marks for student number 1. Always, use = sign first to type any excel formula. Excel has many built in formula liks sum(), average(), etc. and you can type your own formula to compute the desired expression. You can instert the row by right clicking on the row and clicking on Insert menu from the menu list as shown below. Similarly, a column can be inserted by right clicking on the column and pressing Insert. The short cuts for Word are also applicable here.



After, wrting the sum function for student 1 on H3 cell, we just copy that formula and pase that to all the cells below so that we can get the total marks for all the students. This type of cell referencing is call relative reference because we copied the H3 cell which has sum for the cells ranging from C3 to G3, and when we pasted that formula below the cells corresponding rows increases by one automatically.

We can now compute the percentage of of the totals for each student as below:

AutoSave Book1

Thakur Upadhyay

File Home Insert Page Layout Formulas Data Review View Automate Help

Comments Share

Clipboard Font Alignment Number Styles Cells

Conditional Formatting Format as Table Cell Styles

Insert Delete Format

Sort & Filter Find & Select Analyze Data

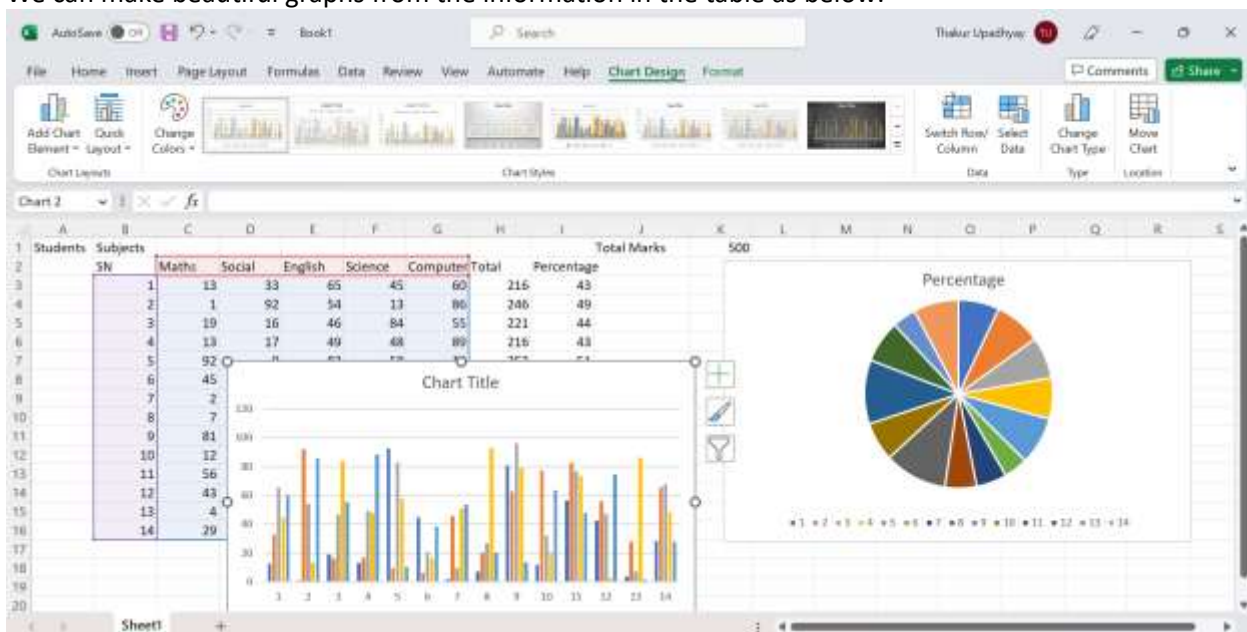
Formulas

$=H3/$K$1*100$

Students	Subjects	Maths	Social	English	Science	Computer	Total	Percentage
1	96	28	33	33	65	254	51	
2	20	36	33	8	65	161		
3	46	43	8	13	23	129		
4	18	30	63	17	88	236		
5	15	7	90	58	6	176		
6	18	58	54	81	26	236		
7	76	26	87	9	29	227		
8	69	3	97	60	24	253		
9	26	70	86	65	33	280		
10	55	18	50	56	50	229		
11	20	93	24	84	76	297		
12	93	40	38	13	7	191		
13	53	89	81	56	27	308		
14	16	71	18	3	87	196		

The percentage is calculated using the expression $H3/\$K\$1*100$, i.e. total divided by Total marks of 500, which resides in K1 location. The K1's value is same for all the students, so we need to make that reference absolute by pressing F4, you can see the formula on the formula box. Once we copy and paste the formula we get the percentages for all students calculated easily.

We can make beautiful graphs from the information in the table as below:

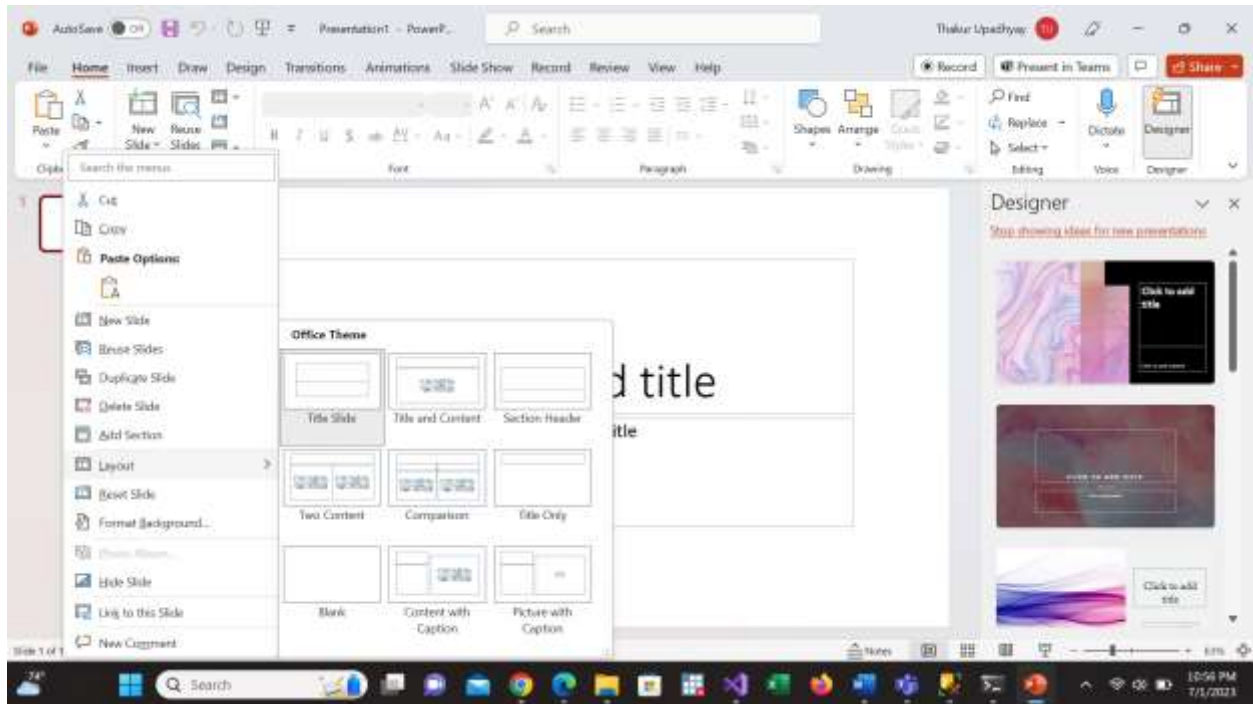


To make the graphs, first select the data, e.g. , the percentage column and click on the insert menu and select the pie chart then you get the pie chart as shown on the right. More practice in Lab.

MS Powerpoint

MS Powerpoint is another application in the office suite that deals with slide preparation and presentation for an effective communication of project/report, etc.

To run the PPT, just type ppt on the search box or ppt icon on the desktop, and then just click on the blank ppt and right click on the layout, you will get the following gui.



You can select different layout for different purposes. You can copy paste the desire part of the text/image/graphs etc. from MS Word or Excel or any other source onto the slide for presentation.

E.g.

Business Analytics

Creating Value from the *NEW OIL*
(bigdata)

Thakur Upadhyay, PhD
Department of Economics
Carleton University, Canada

Outlines

- Background
- Dawn of bigdata era
- Analytics ecosystem
- Dashboard with Power BI
- Machine learning example
- Future research agenda

Defining Business Analytics



BA is the use of:

(big)data,
information technology,
statistical and computer science algorithms,
quantitative methods, and
mathematical models

to help managers gain improved insight about their business operations and make better, factbased decisions.

- Traditionally also called Business Intelligence (BI)

You can choose different font sizes and color scheme etc. You can add animations and other presentation styles as per need. Be concise while preparing the PPT, don't put lots of text in one slide. This will distract the audience's attention. More practive in the Lab.

Database: Very brief introduction to flat File system, RDBMS, NOSQL

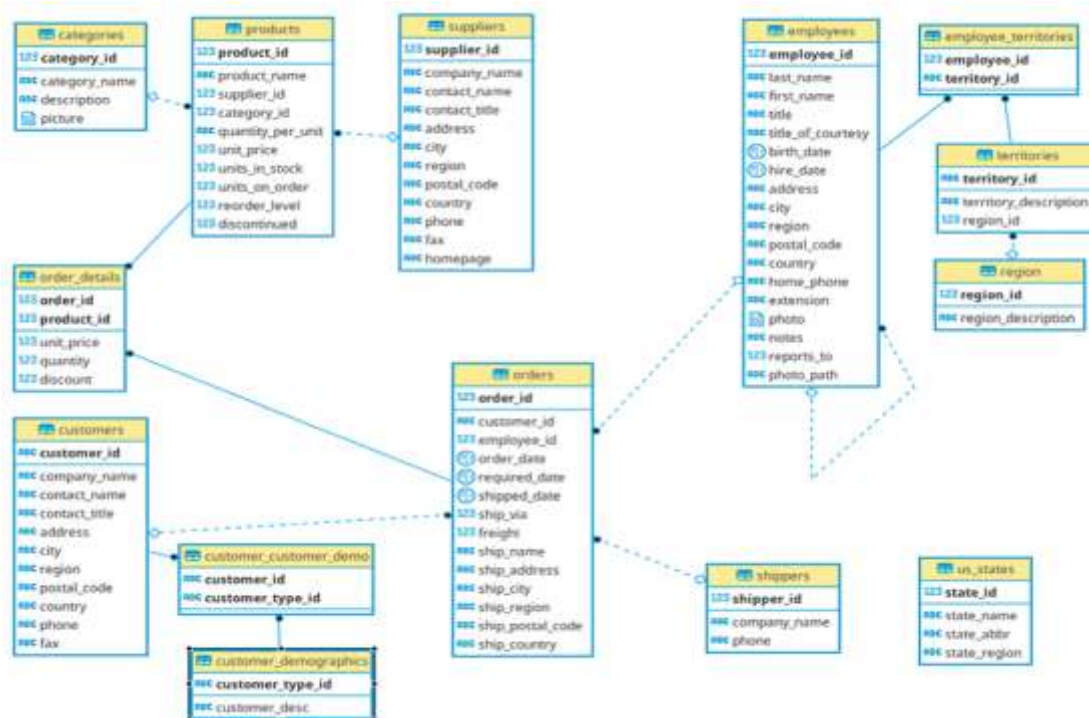
Database

- A database is an organized collection of structured information, or data, typically stored electronically in a computer system.
- A database is usually controlled by a database management system (DBMS).
- The main purpose of the database is to operate a large amount of information by storing, retrieving, and managing data.
- Data within the most common types of databases in operation today is typically modeled in rows and columns in a series of tables to make processing and data querying efficient.
- Most databases use Structured Query Language (SQL) for writing and querying data
- There are many databases available like SQL Server, Oracle, MySQL, MongoDB, PostgreSQL, Sybase, Informix, etc.

Before the invention of Relational Database Management System, all the columns of all the tables in a business organization used to store in a single file, called a flat file as shown below:

	OrderID	ProductID	UnitPrice	Quantity	Discount	OrderID	CustomerID	EmployeeID	OrderDate	RequiredDate	ShippedDate	ShipVia	Freight	ShipName	ShipAddress
1	10248	11	14	12	0	10248	VINET	5	7/4/1996	8/1/1996	7/16/1996	3	32.38	Vins et alcools Chevalier	59 rue de l'Abbaye
2	10248	42	9.8	10	0	10248	VINET	5	7/4/1996	8/1/1996	7/16/1996	3	32.38	Vins et alcools Chevalier	59 rue de l'Abbaye
3	10248	72	34.8	5	0	10248	VINET	5	7/4/1996	8/1/1996	7/16/1996	3	32.38	Vins et alcools Chevalier	59 rue de l'Abbaye
4	10249	14	16.6	9	0	10249	TOMSP	6	7/5/1996	8/16/1996	7/10/1996	1	11.61	Torne Specialitäten	Luisenstr. 48
5	10249	51	42.4	40	0	10249	TOMSP	6	7/5/1996	8/16/1996	7/10/1996	1	11.61	Torne Specialitäten	Luisenstr. 48
6	10250	41	7.7	10	0	10250	HANAI	4	7/6/1996	8/5/1996	7/12/1996	2	65.83	Honari Carnes	Rua do Paço, 67
7	10250	51	42.4	35	0.15	10250	HANAI	4	7/6/1996	8/5/1996	7/12/1996	2	65.83	Honari Carnes	Rua do Paço, 67
8	10250	65	16.8	15	0.15	10250	HANAI	4	7/6/1996	8/5/1996	7/12/1996	2	65.83	Honari Carnes	Rua do Paço, 67
9	10251	22	16.8	6	0.05	10251	VICTE	3	7/6/1996	8/5/1996	7/15/1996	1	41.34	Victualles en stock	2, rue du Commerce
10	10251	57	15.6	15	0.05	10251	VICTE	3	7/6/1996	8/5/1996	7/15/1996	1	41.34	Victualles en stock	2, rue du Commerce
11	10251	65	16.8	20	0	10251	VICTE	3	7/6/1996	8/5/1996	7/15/1996	1	41.34	Victualles en stock	2, rue du Commerce
12	10252	20	64.6	40	0.05	10252	SUPRO	4	7/9/1996	8/6/1996	7/11/1996	2	51.3	Suprêmes délices	Boulevard Tirou, 255
13	10252	33	2	25	0.05	10252	SUPRO	4	7/9/1996	8/6/1996	7/11/1996	2	51.3	Suprêmes délices	Boulevard Tirou, 255
14	10252	60	27.2	40	0	10252	SUPRO	4	7/9/1996	8/6/1996	7/11/1996	2	51.3	Suprêmes délices	Boulevard Tirou, 255
15	10253	31	10	20	0	10253	HANAI	3	7/10/1996	7/24/1996	7/16/1996	2	58.17	Honari Carnes	Rua do Paço, 67
16	10253	30	14.4	42	0	10253	HANAI	3	7/10/1996	7/24/1996	7/16/1996	2	58.17	Honari Carnes	Rua do Paço, 67
17	10253	49	16	40	0	10253	HANAI	3	7/10/1996	7/24/1996	7/16/1996	2	58.17	Honari Carnes	Rua do Paço, 67
18	10254	24	3.6	15	0.15	10254	CHOPS	5	7/11/1996	8/6/1996	7/23/1996	2	22.98	Chop-suey Chinese	Hauptstr. 31
19	10254	55	19.2	21	0.15	10254	CHOPS	5	7/11/1996	8/6/1996	7/23/1996	2	22.98	Chop-suey Chinese	Hauptstr. 31

This way of storing information in a single flat file was easy to implement but it would create lots of complication when the data size grows. In order to solve this problem, a RDBMS application software is used in practice. For this, the information in this flat file are broken down into a various tables (this process is called normalization) and later they get connected via relational algebra. The Tables's relations look like this.



Organizing information in this structure gives lots of advantages in managing huge amounts of database. A SQL database programming language helps us in different manipulation and query of the database. e.g. To get all the fields of the tables, we run the following SQL script.

```

SELECT *
FROM [Order Details] as ord
INNER JOIN Orders
ON ord.OrderID = Orders.OrderID
  
```

```
INNER JOIN Products as pr
ON ord.ProductID = Pr.ProductID
INNER JOIN Customers as cu
ON Orders.CustomerID = cu.CustomerID
INNER JOIN Employees as emp
ON Orders.EmployeeID = emp.EmployeeID
```

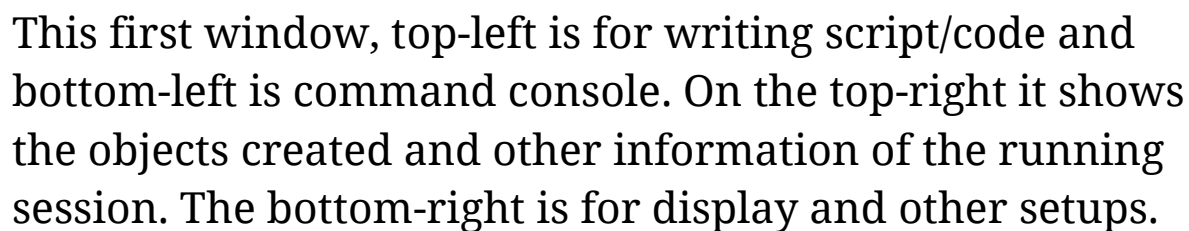
Nowadays, due to widespread information availability on the net, a semi structured database management system, NoSQL, is also getting popular. NoSQL, also referred to as “not only SQL”, “non-SQL”, is an approach to database design that enables the storage and querying of data outside the traditional structures found in relational databases. While it can still store data found within relational database management systems (RDBMS), it just stores it differently compared to an RDBMS. The decision to use a relational database versus a non-relational database is largely contextual, and it varies depending on the use case.

Instead of the typical tabular structure of a relational database, NoSQL databases, house data within one data structure, such as JSON document. Since this non-relational database design does not require a schema, it offers rapid scalability to manage large and typically unstructured data sets.

Data analysis: Very brief introduction to SPSS, R, Python

Managing database with RDBMS or NoSQL is not enough, we need to analyse those data in order get some business insights. Generally, the analysis consists of descriptive statistive (mean, median, mode, range variance/SD, kurtosis and skewness), predictive (regression and classification) and prescriptive (optimization) models. We need a specialized software for data analysis and SPSS, R and Python are widely used software packages for this purpose. SPSS stands for Statistical Package for Social Sciences and menu/GUI based easy to use statistical software for data analysis. Python, which we learned in detail while learning programming

First, we need to install R software and R Studio for an IDE to write code for data analysis. This R studio looks like:



In order to analyse the data in R, first we need to save our dataset in Excel or other RDBMS package and import them into R. See the R-script for further details of the analysis part.

