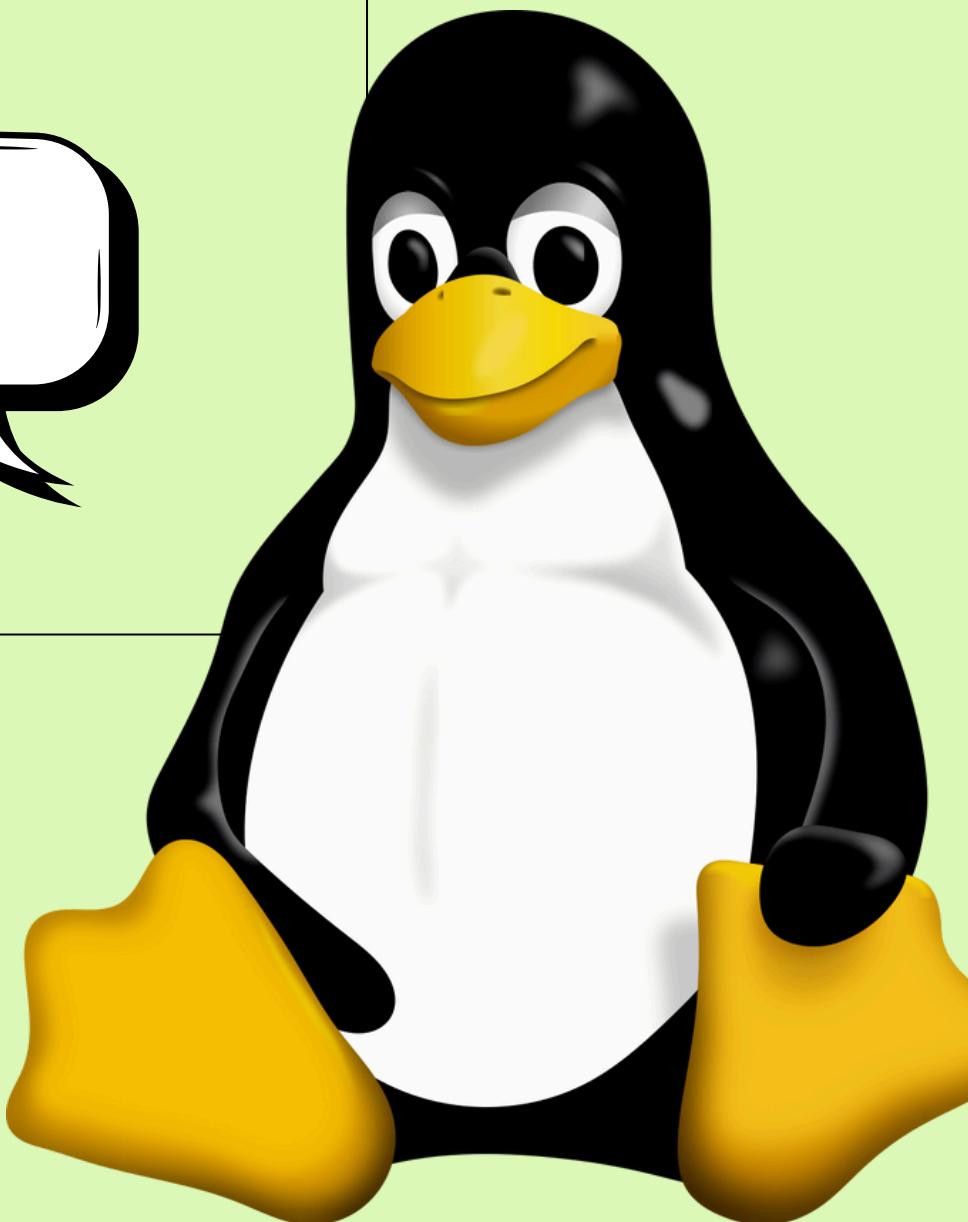


# Using Linux

- Introduction of Linux and Bash for Data Engineering
- Linux Terminal in Cloud Environment
- Bash Statement(Shebang, chmod, permission)
- Basic Bash Command
- Basic Pipeline shell
- SSH (Connecting to VSCode and Cloud)

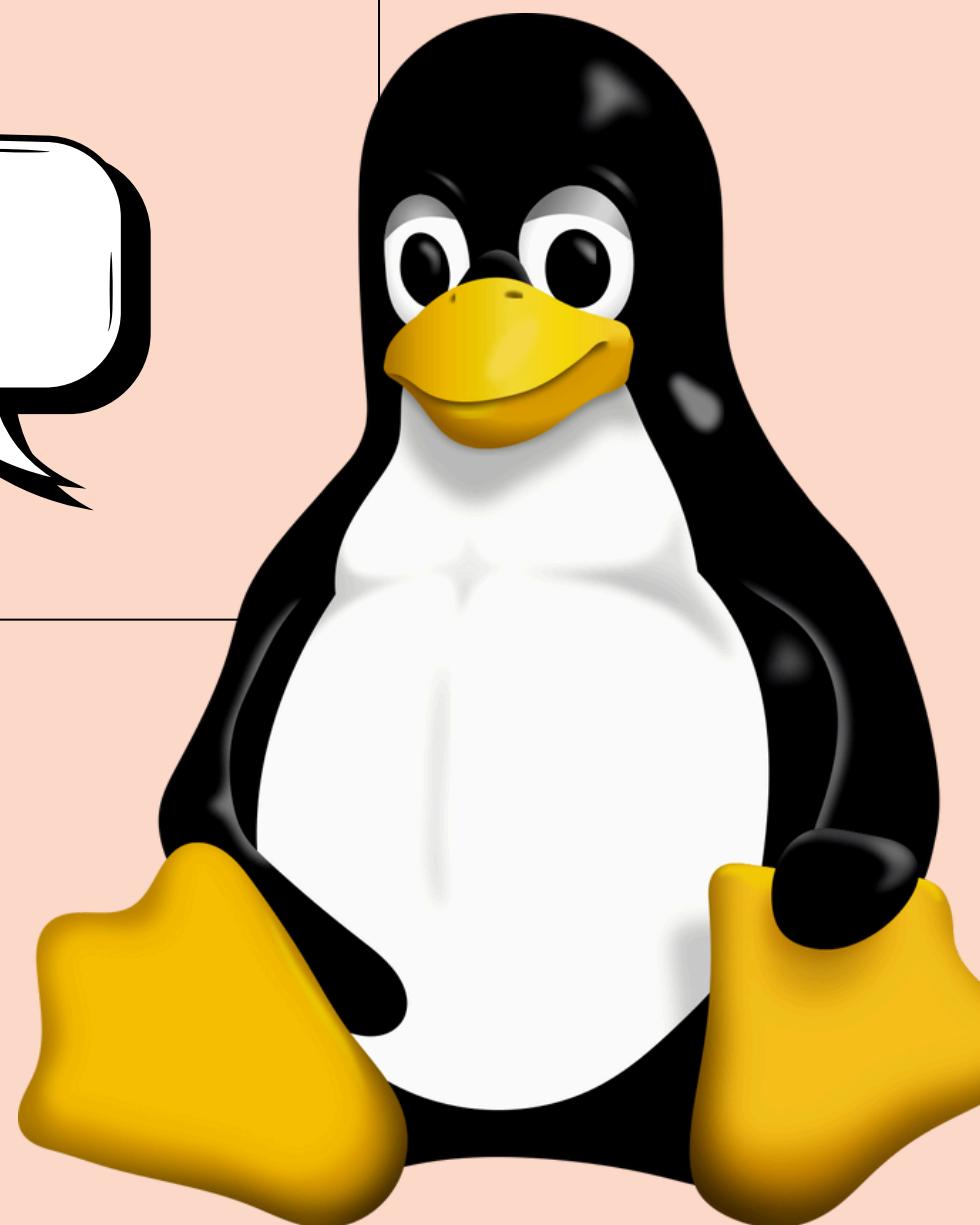
Hi! I'm linux



# Introduction of Linux and Bash for Data Engineering

- What is Linux/ Bash?
- Why is Linux?
- Terminal in Linux
- Show code example

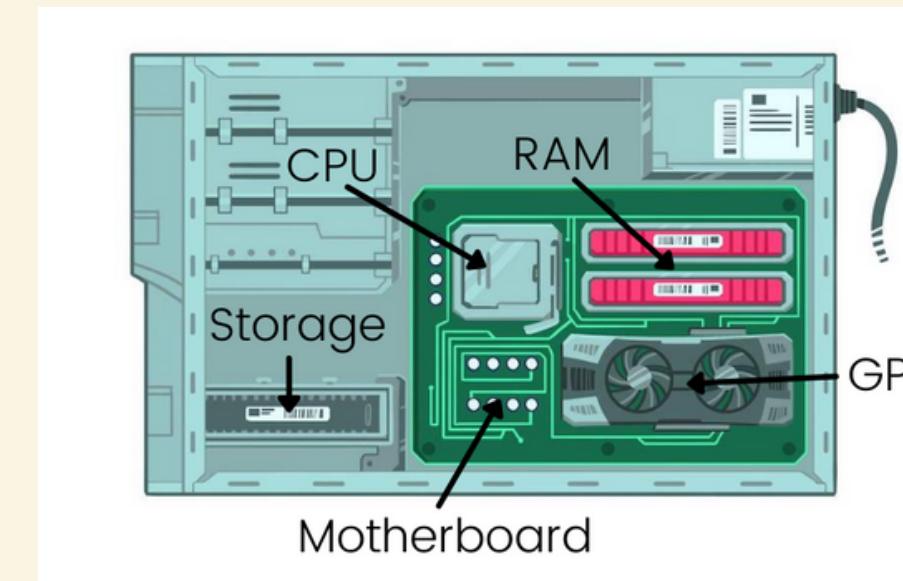
Hi! I'm linux



# What is Linux?

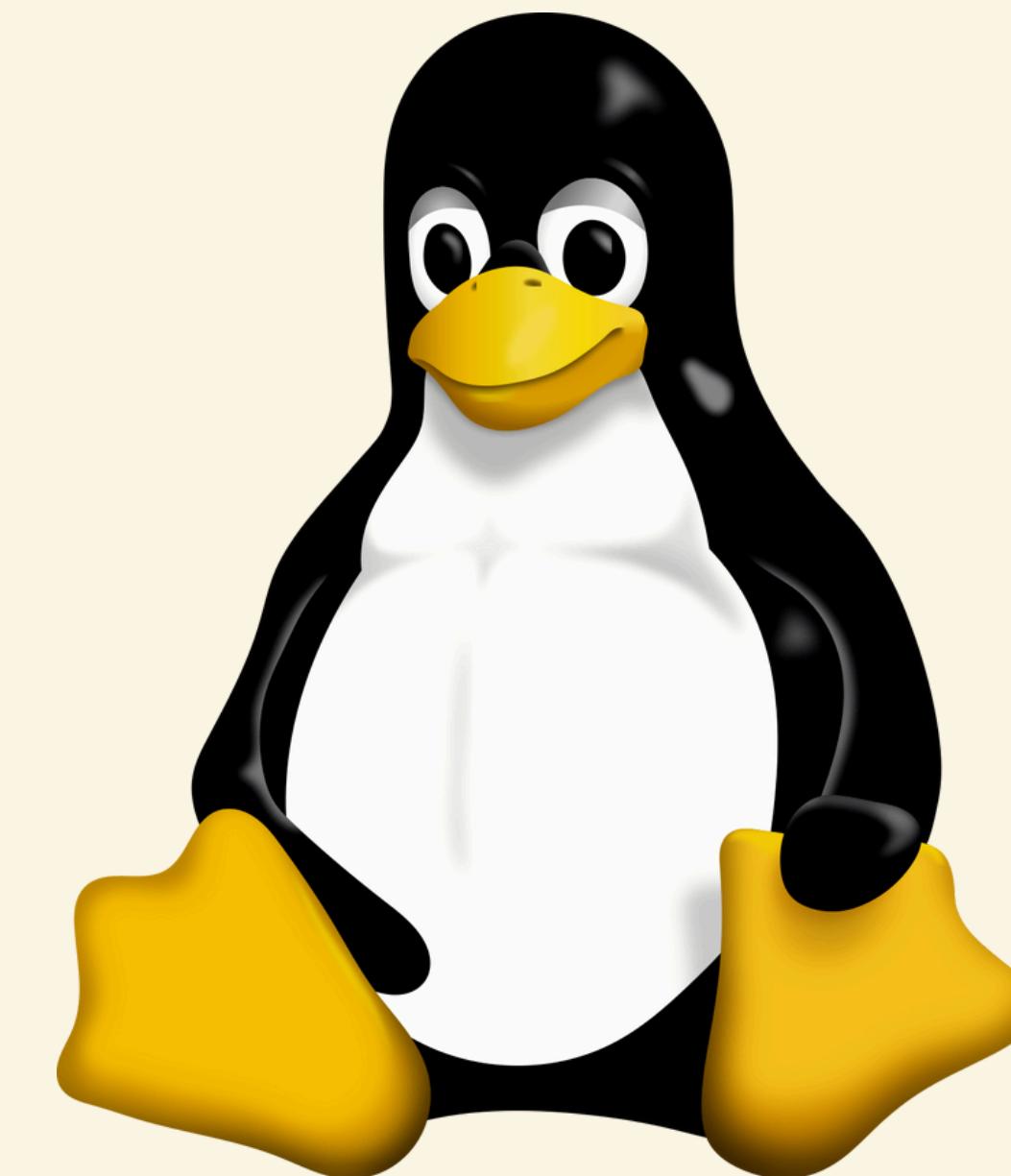
Operation system

Open Source



# What is Bash?

command language

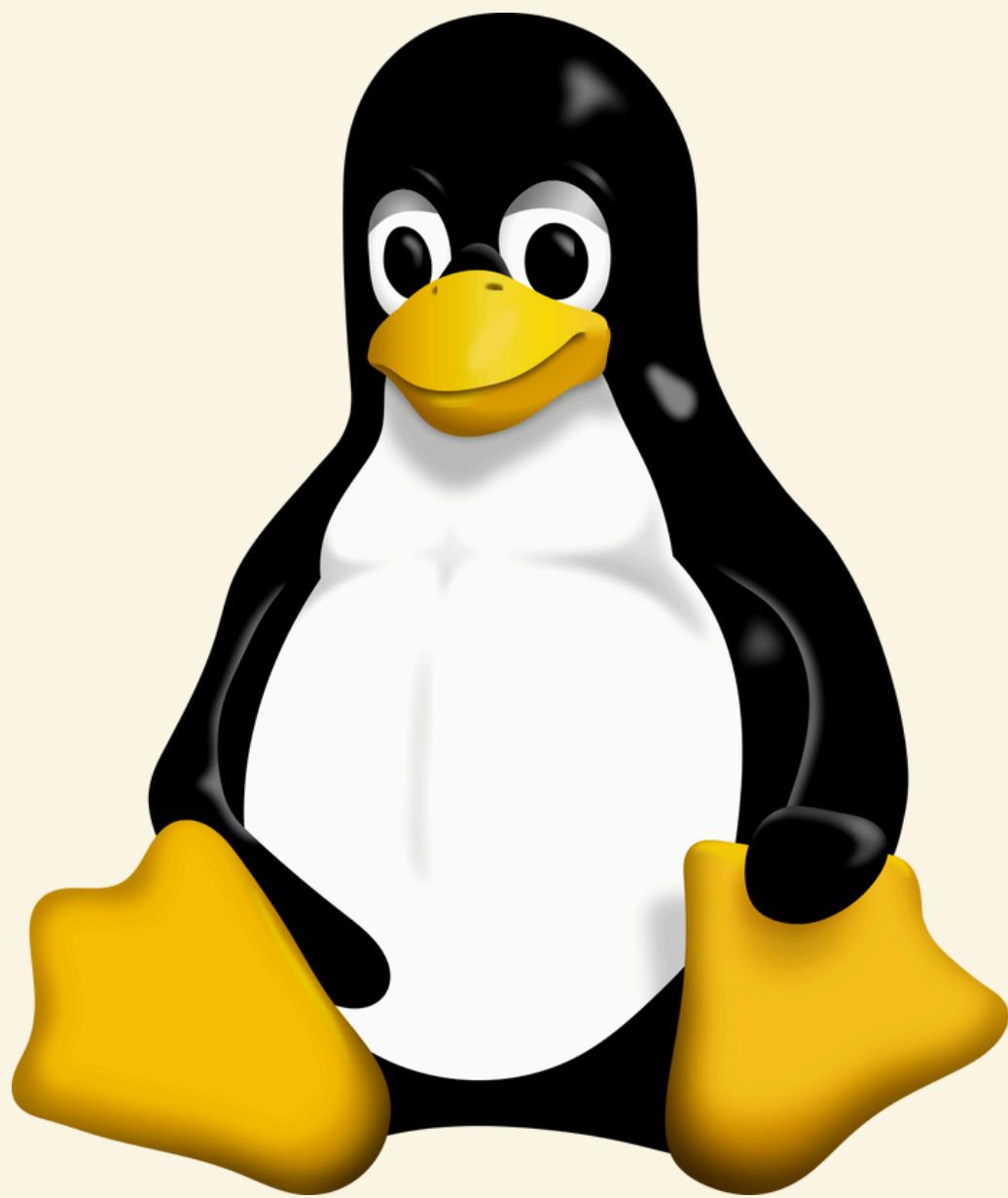


```
kb@phoenixNAP:~$ bash script.sh
[sudo] password for kb:
Hit:1 http://repo.mysql.com/apt/ubuntu focal InRelease
Hit:2 http://rs.archive.ubuntu.com/ubuntu focal InRelease
Hit:3 http://security.ubuntu.com/ubuntu focal-security InRelease
Hit:4 http://rs.archive.ubuntu.com/ubuntu focal-updates InRelease
Hit:5 http://rs.archive.ubuntu.com/ubuntu focal-backports InRelease
Reading package lists... Done
Building dependency tree
Reading state information... Done
All packages are up to date.
Reading package lists... Done
Building dependency tree
Reading state information... Done
Calculating upgrade... Done
0 upgraded, 0 newly installed, 0 to remove and 0 not upgraded.
Done! ←
kb@phoenixNAP:~$
```

# Why is Linux?

Fully Control

Security



Work anywhere in cloud!



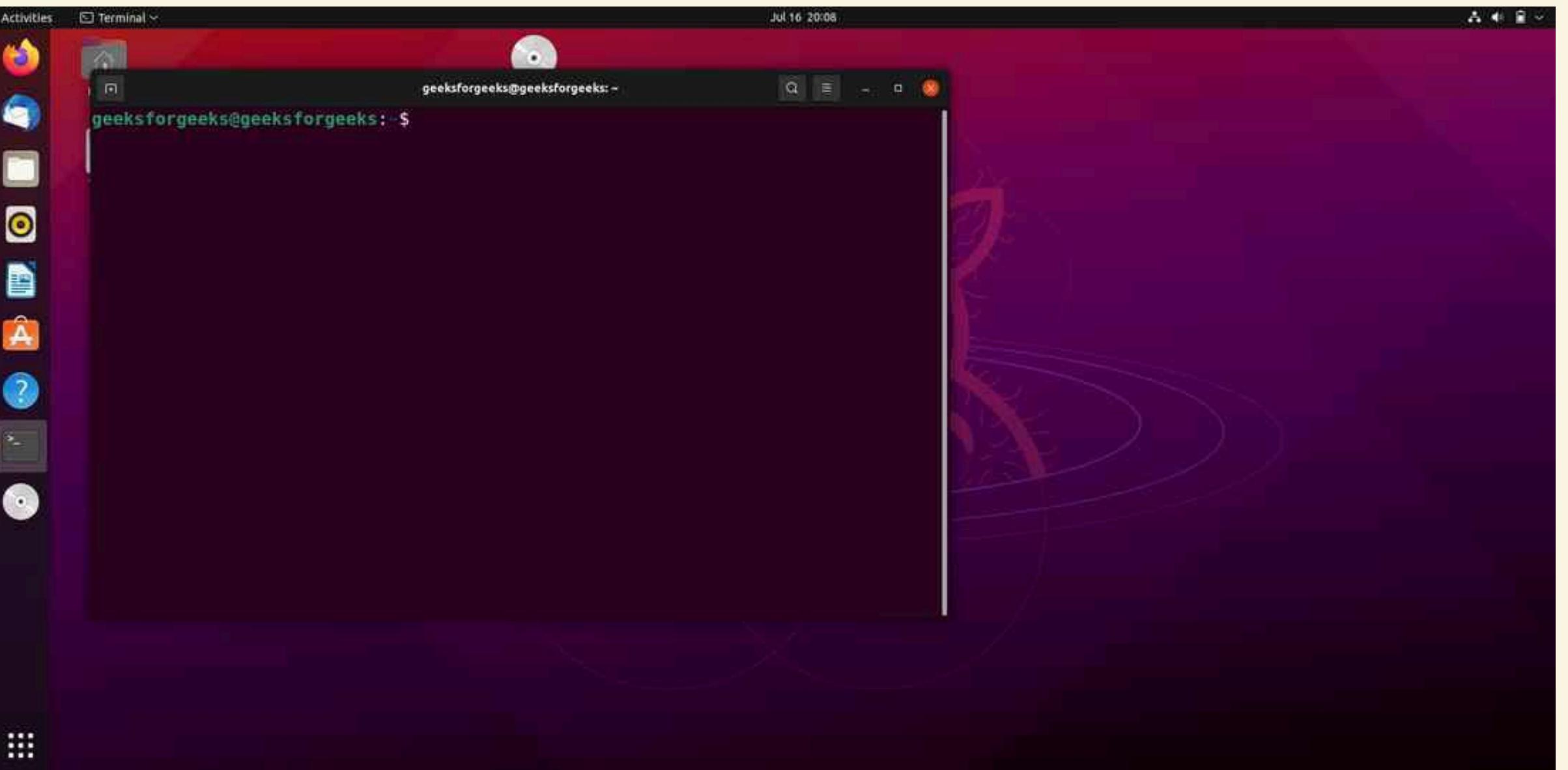
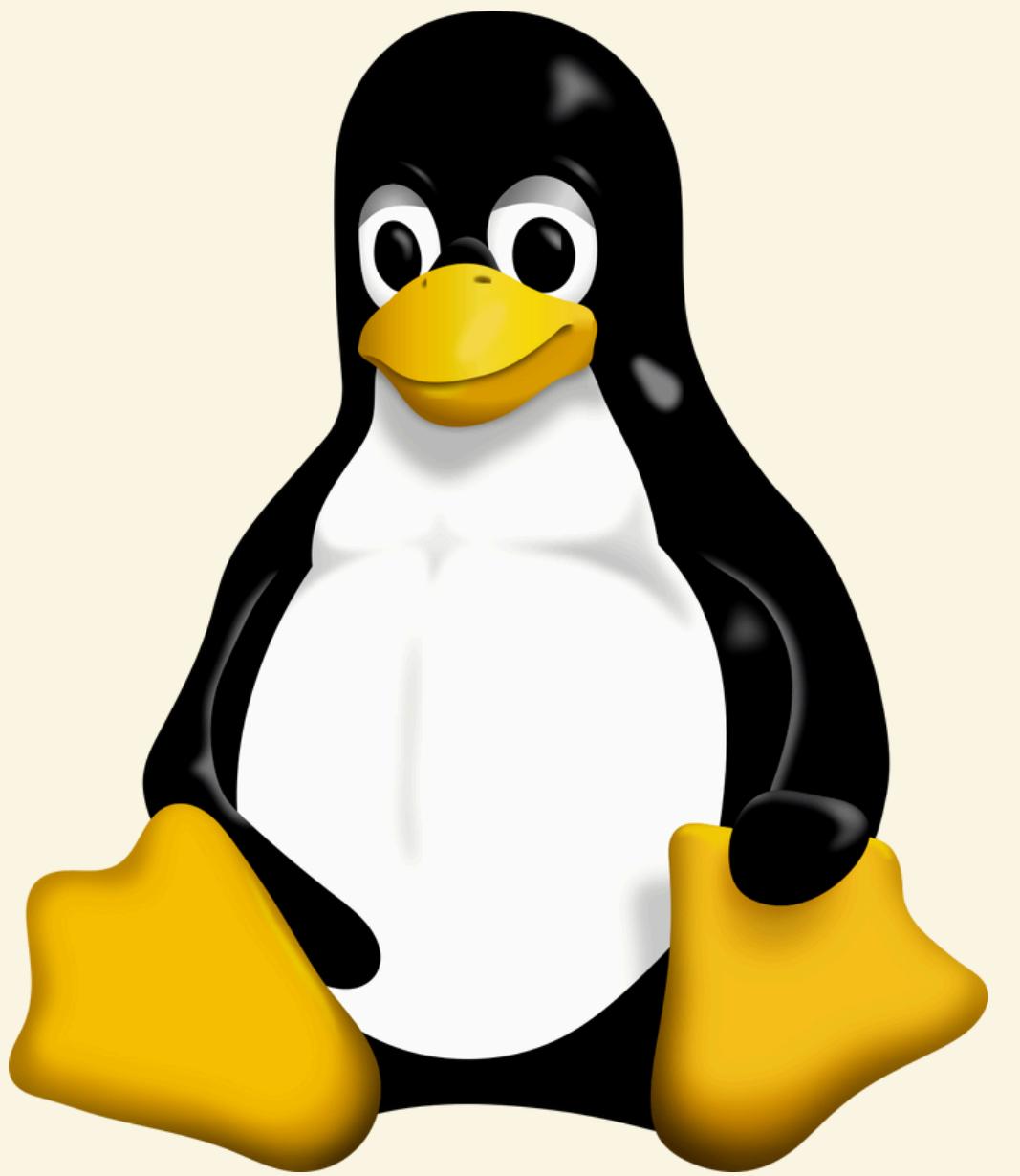
Google Cloud



Multi-task Faster

No Virus

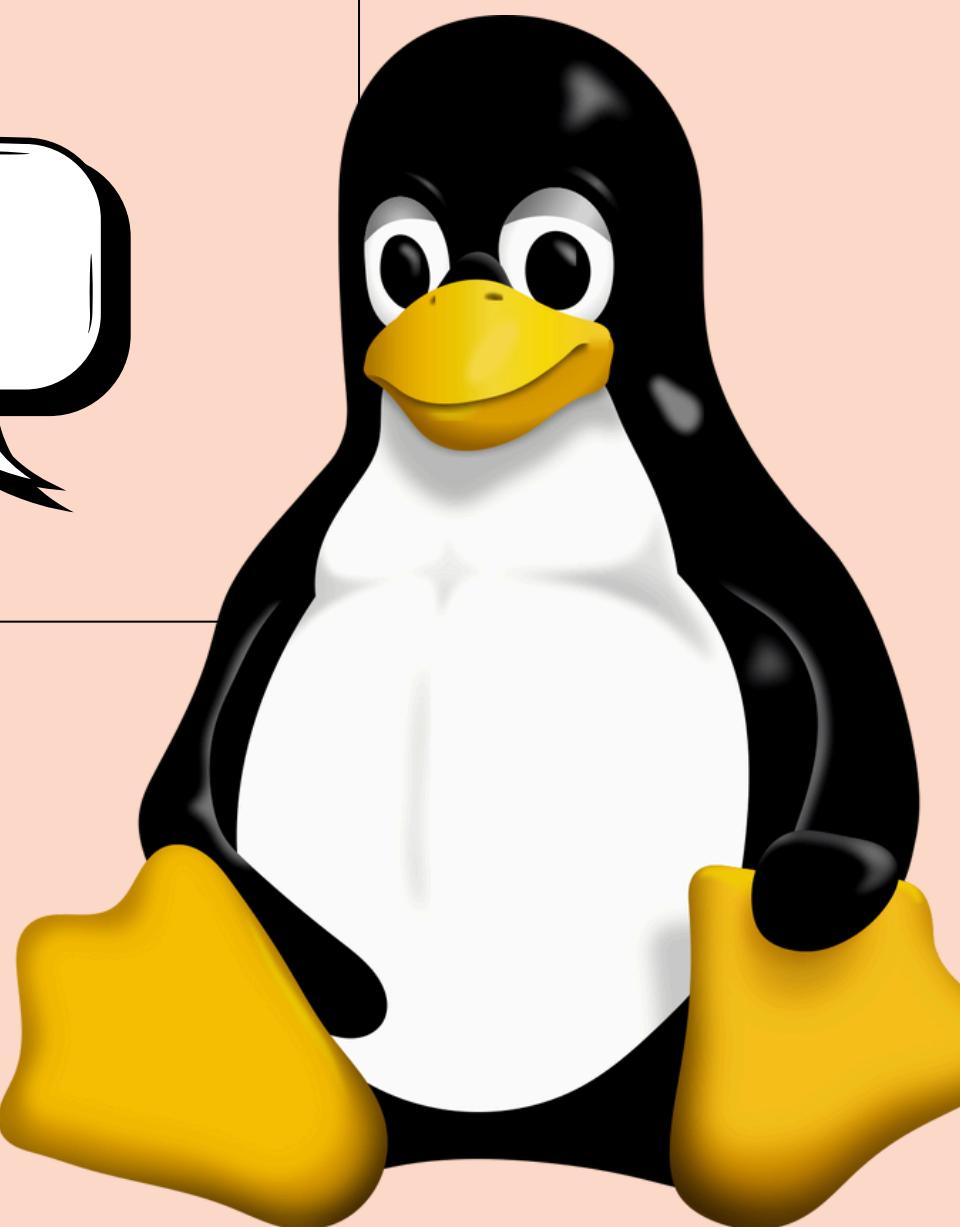
# What am I look like?



# Linux Terminal in Cloud Environment

- AWS
- GitHub
- Azure
- Google Cloud Platform (GCP)

Linux Cloud



The screenshot shows the AWS Cloud9 interface. At the top, there's a navigation bar with the AWS logo, a 'Services' dropdown, a search bar, and user account information for 'smith' in 'Sydney'. A sidebar on the left is titled 'Developer Tools' and contains a 'CloudShell' section for the 'ap-southeast-2' region. The main content area features a large heading 'AWS Cloud9' and a sub-heading 'A cloud IDE for writing, running, and debugging code'. To the right of this text is a white callout box with the heading 'New AWS Cloud9 environment' and a prominent orange 'Create environment' button. Below the main content, there's a terminal window showing a prompt '[cloudshell-user@ip-10-130-32-212 ~]\$'. The bottom of the screen includes standard footer links for 'CloudShell', 'Feedback', 'Privacy', 'Terms', and 'Cookie preferences', along with a copyright notice for 2024.

aws | Services | Search [Alt+S]

smith | Sydney | Developer Tools

# AWS Cloud9

A cloud IDE for writing, running, and debugging code

New AWS Cloud9 environment

Create environment

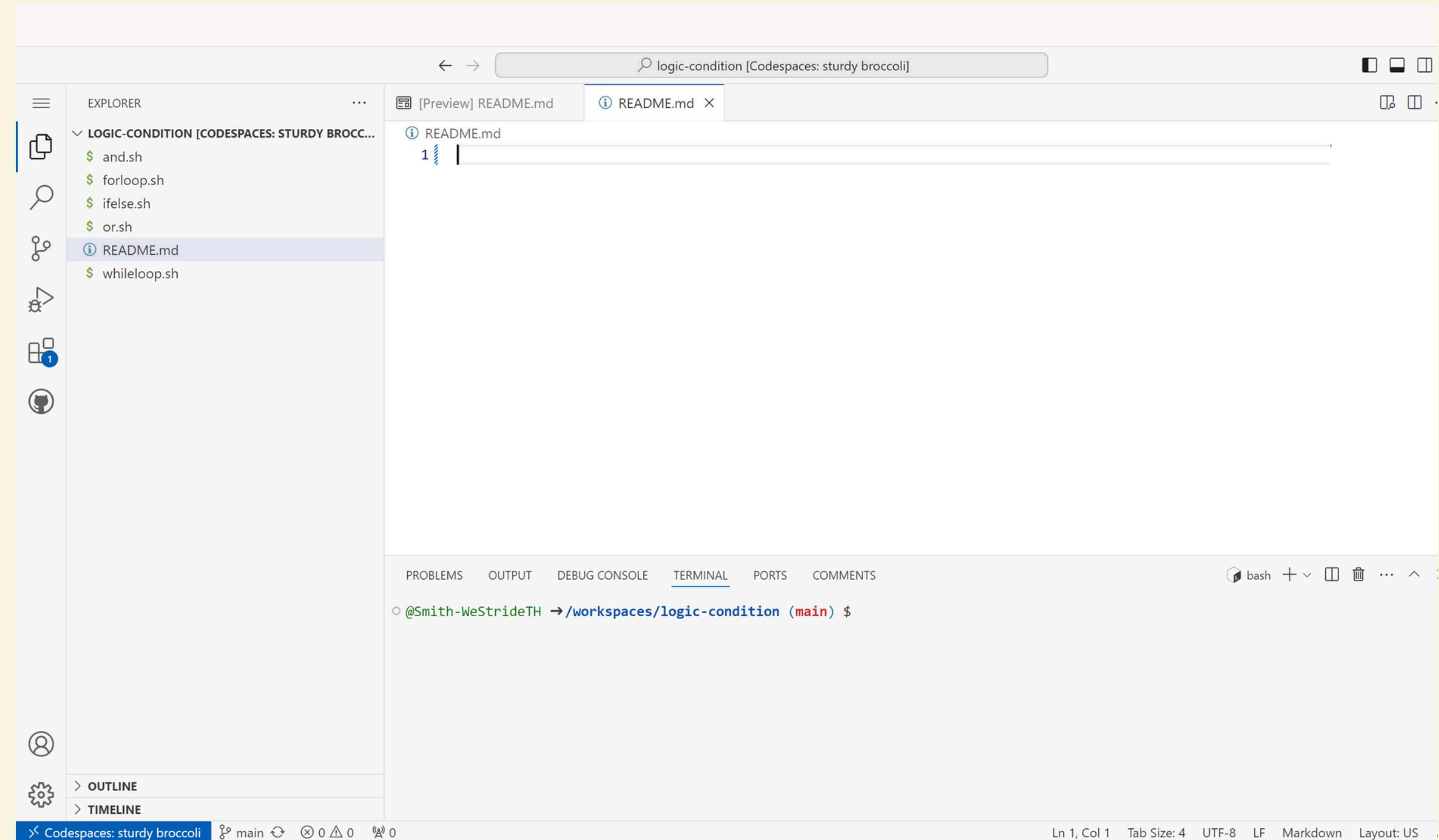
CloudShell

ap-southeast-2

[cloudshell-user@ip-10-130-32-212 ~]\$

© 2024, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

# GitHub



# Azure

The screenshot shows the Microsoft Azure portal interface. At the top, there is a blue header bar with the "Microsoft Azure" logo, a search bar containing "Search resources, services, and docs (G+)", and several icons for notifications, settings, and help.

The main area is titled "Azure services" and features a grid of service icons:

- Create a resource
- Storage accounts
- All resources
- Azure AI services
- Azure Cosmos DB
- Virtual machines
- Quickstart Center
- Azure OpenAI
- Kubernetes services
- More

Below this is a terminal window titled "Bash". The command prompt shows "smith\_aueng [ ~ ]\$".

# GCP

The screenshot shows the Google Cloud Platform (GCP) homepage. At the top, there is a banner with a gift icon and the text: "Start your Free Trial with \$300 in credit. Don't worry—you won't be charged if you run out of credits. [Learn more](#)". To the right of the banner are "DISMISS" and "START FREE" buttons. Below the banner is the Google Cloud logo and a "Select a project" dropdown menu. A search bar contains the placeholder "Search (/) for resources, docs, products, and more". To the right of the search bar are icons for "Search", "Star", "Project", "Bell", "Help", and a profile icon with a 'T'.

Welcome, Smith Tripornkanokrat

## Try Google Cloud with \$300 in free credits

- ✓ Access to Google Cloud products and services
- ✓ 90 days to spend your credits
- ✓ No billing during trial

Other options

Explore new models and experiment with different prompts using Google's largest AI model.

[TRY GEMINI](#)

Configure Google Cloud for scalable, production-ready enterprise workloads.

CLOUD SHELL

Terminal cloudshell [Open Editor](#) | - + X

```
Welcome to Cloud Shell! Type "help" to get started.  
To set your Cloud Platform project in this session use "gcloud config set project [PROJECT_ID]"  
smith@cloudshell:~$ 
```



Google Cloud

## WS Cloud9

```
# play with terminal
- list s3 bucket: (aws s3 ls)
- create new s3 bucket
- list s3 bucket: (aws s3 ls)
```

## Azure

```
# play with terminal
-vim ~/.bashrc
-create new echo
-source ~/.bashrc
```

## GCP

```
# play with terminal
-vim ~/.bashrc
- create new alias
- source ~/.bashrc
-alias
```

# Example

## Common

<b>root directory</b>	pwd
<b>change directory</b>	cd
<b>list</b>	ls
<b>display all infomation</b>	uname -a
<b>create file/folder</b>	touch
<b>look inside file</b>	cat

## Text

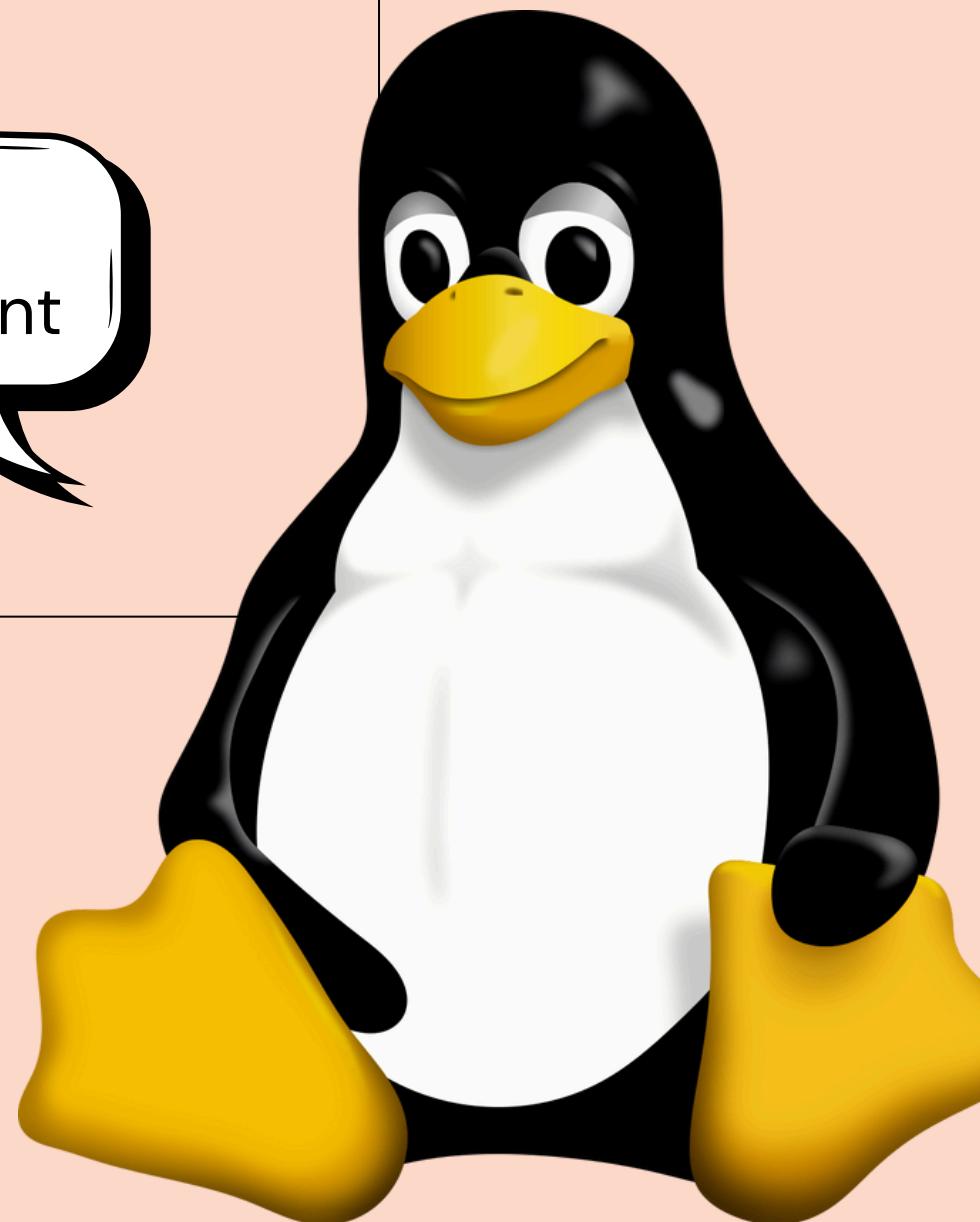
<b>a text editor</b>	vim
<b>reads and executes</b>	source
<b>write &amp; quit</b>	wq
<b>print</b>	echo
<b>alias name</b>	alias

# Introduce (shebang, chmod +x, permission)

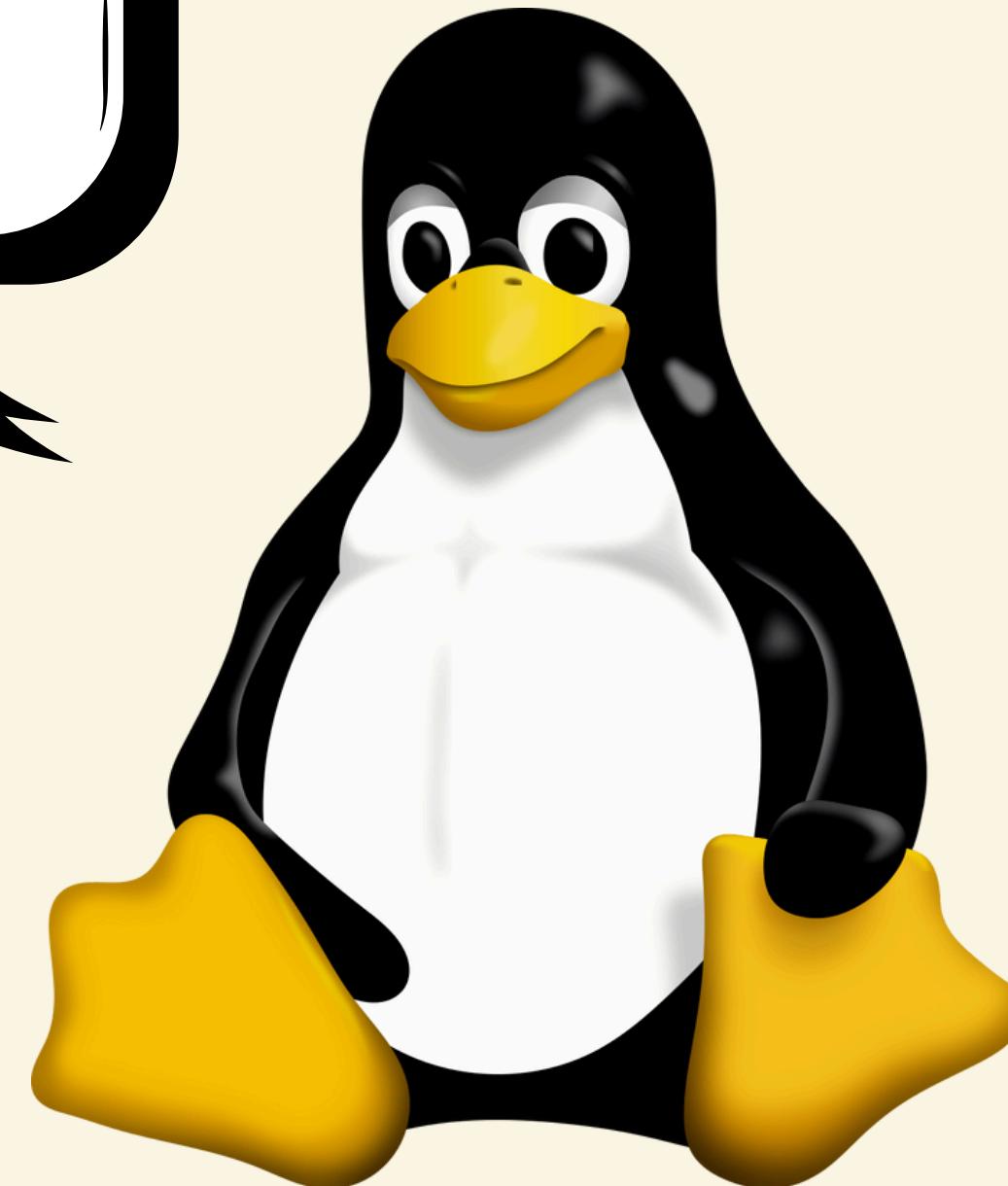
Introduce concept using GitHub:

- **shebang** : `#!/usr/bin/bash`
- **chmod +x**
- **permission**

Let's get  
bash statement



# shebang & chmod

A cartoon illustration of Tux, the official mascot of Linux. He is a black and white penguin with a yellow beak and feet, sitting upright with his hands clasped together.

Shebang: `#!/usr/bin/bash`

“interpreter path for terminal”

chmod +x

“change and modify permission”

# Permission types (user-group-other)

(user-group-other)

rw-rw-rw -> 666

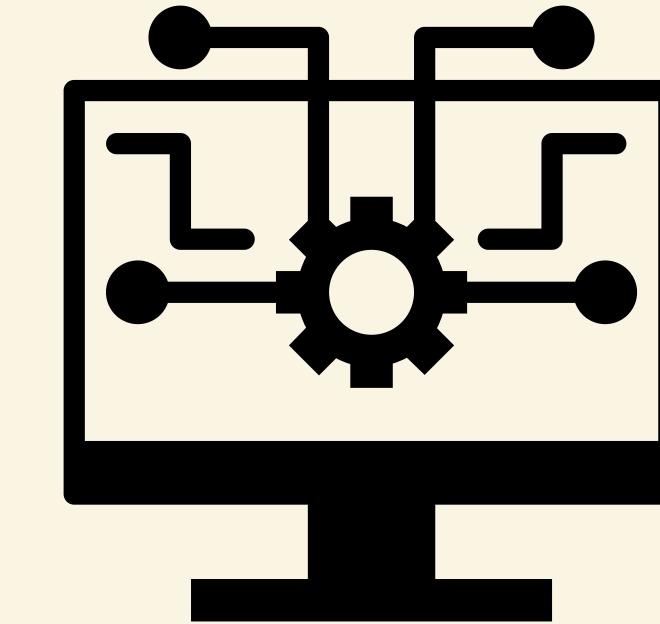
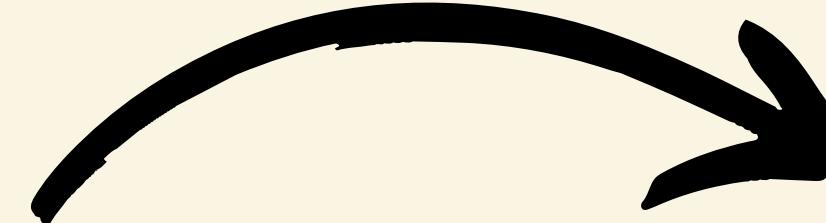
Number	Binary	Representation	Permissions
0	000	---	No Permission
1	001	--x	Execute
2	010	-w-	Write
3	011	-wx	Write, Execute
4	100	r--	Read
5	101	r-x	Read, Execute
6	110	rw-	Read, Write
7	111	rwx	Read, Write, Execute

chmod +x

rwX-rwX-rwX -> 777

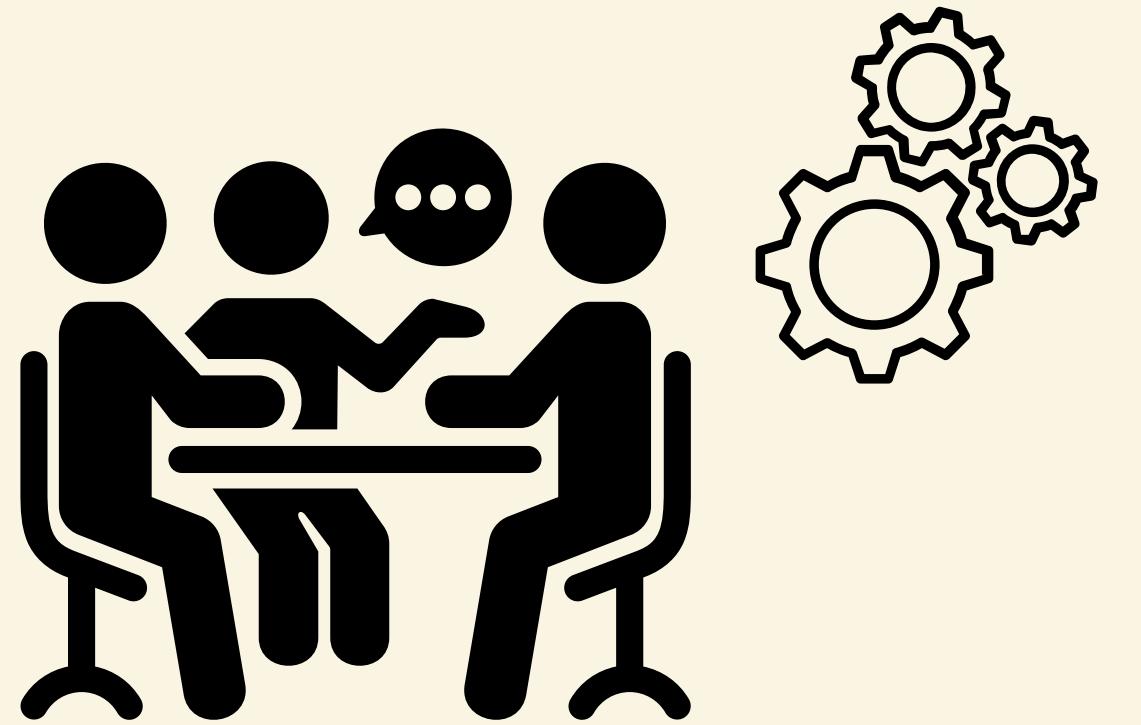
**(user-group-other)**

**User**

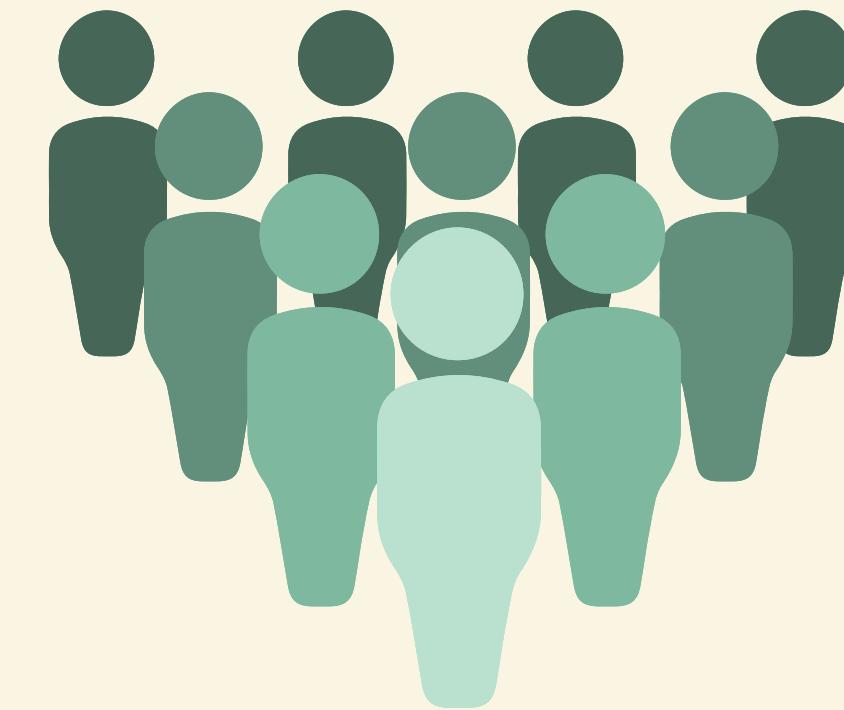


Installing / Modifying / updates/ maintenance

**Group**



**Other**



# Check permission

ls -l

ls -lah

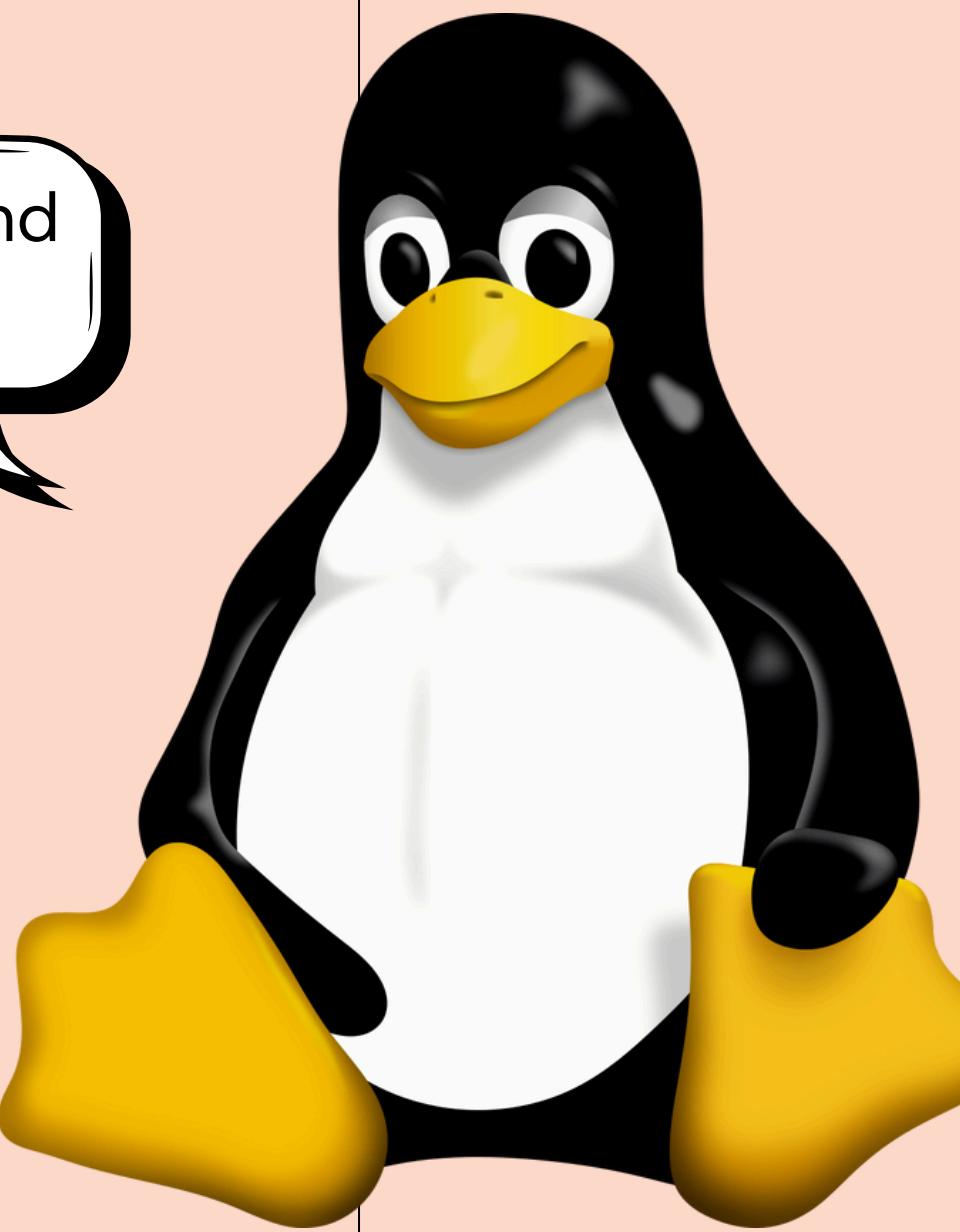


# Basic Linux-Bash Command

## Command :

- Linux Command Structure
- Calendar
- Disk size
- shell directory
- process
- stop/kill process
- Background/Foreground execution
- permission

Useful command



# Command Structure

[Cmd -Option Target]

---

ls -lah /usr/sbin

ls: list

l: listing format

a: all information

h: human readable

mkdir newfolder

mkdir: make directory

mkdir -p root/dir/sub\_dir

mkdir: make directory

p: parent

<https://explainshell.com/explain?cmd=fg>

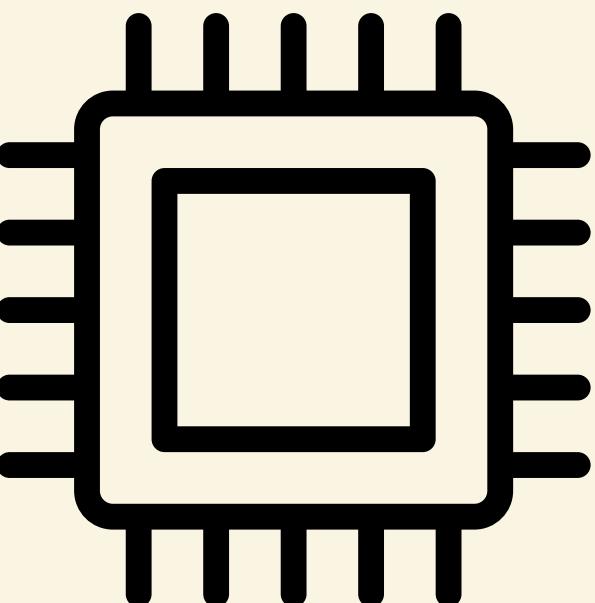
# Calendar – Date-Time

Calendar	cal
Date	date
location date directory	which date



# Disk Size

disk space usage	df-h
estimate file space usage	du-sh /usr/bin/*
display Linux tasks	top



# shell directory

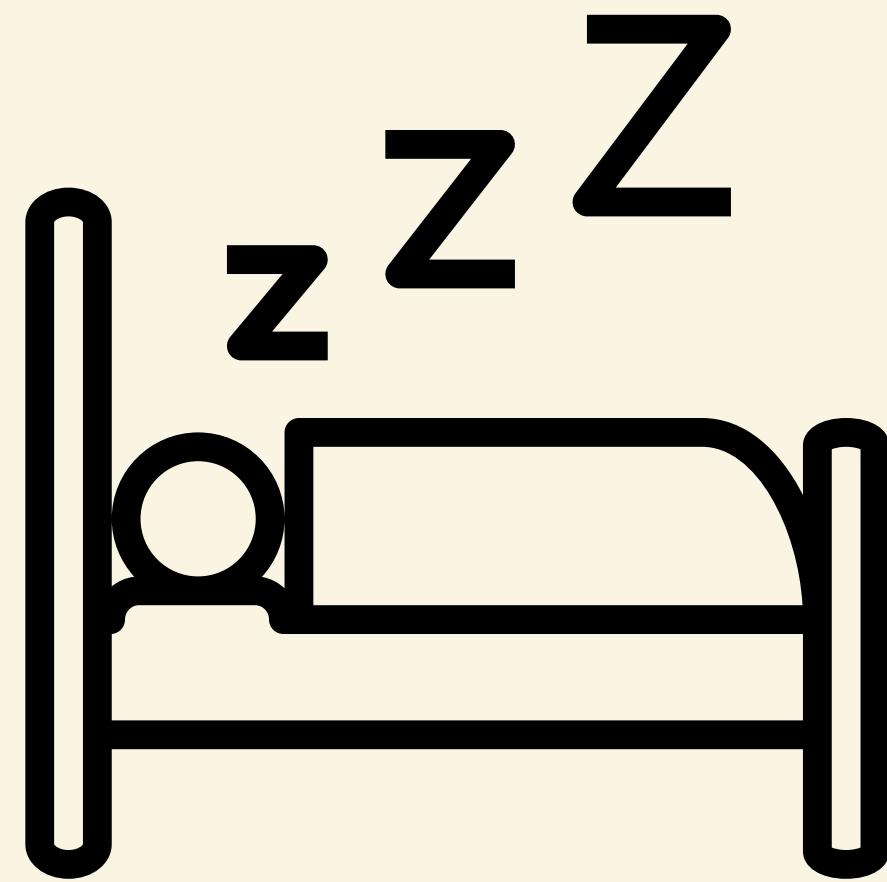
<b>change directory</b>	cd
Change to parent directory	cd ..
<b>print name of current/working directory</b>	pwd
Create new file	touch
<b>Print output</b>	cat
<b>make directories</b>	mkdir
<b>move (rename) files</b>	mv
<b>remove recursive/force</b>	rm -rf



# Process

report a snapshot of the current processes	ps
Stop Process	ctrl + z
Kill process	ctrl + c
foreground execution	fg
Background execution	./file.sh &

sleeper.sh



# Pipeline shell

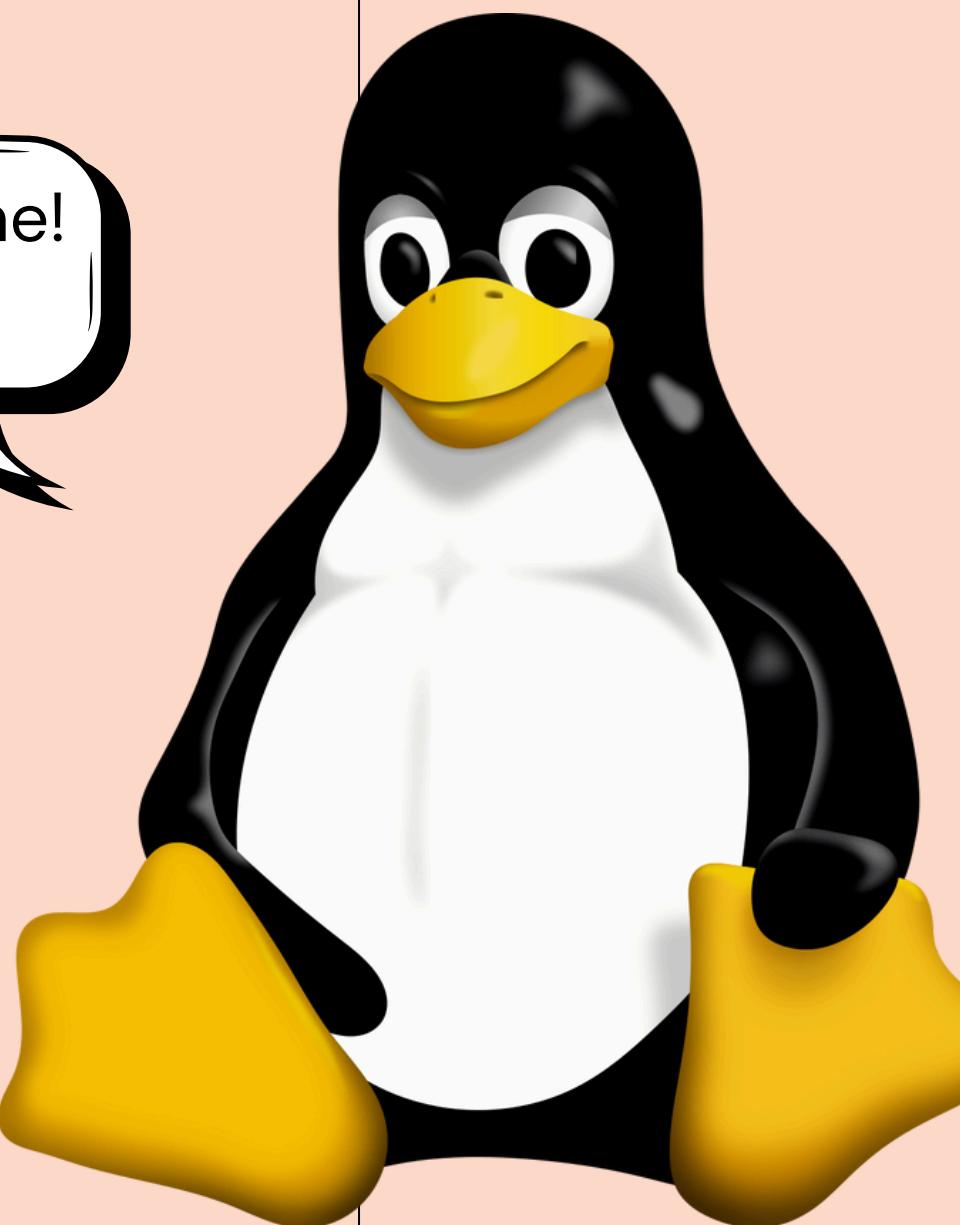
**Useful pipeline :**

- **Simple:** input | process > output
- **Condition:** one step fail then fail whole script

**Example:**

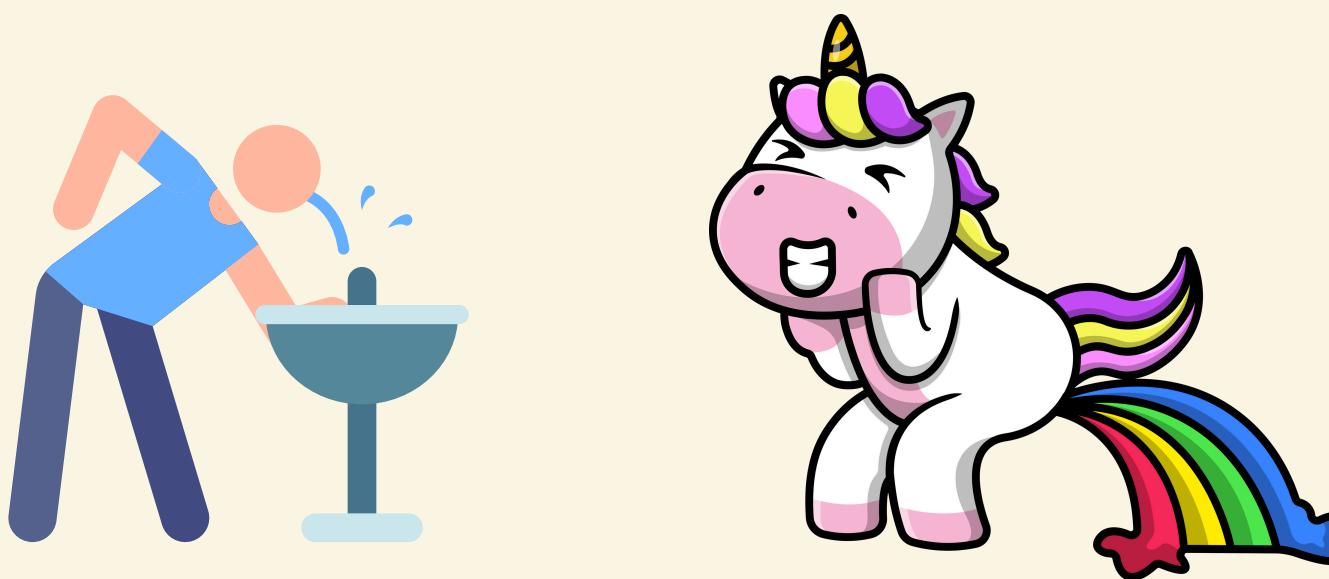
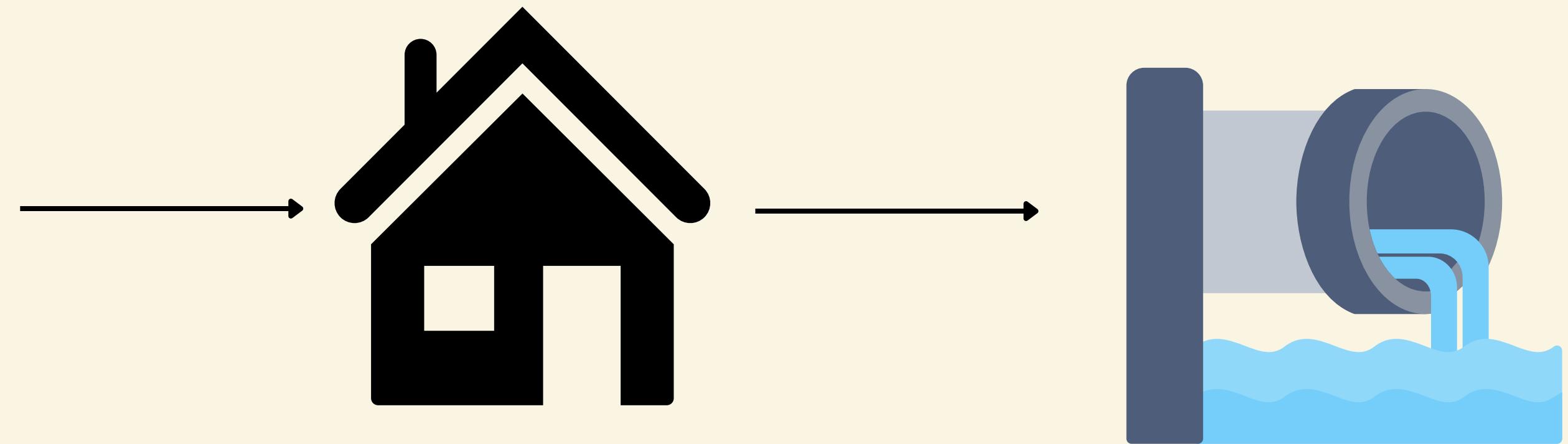
- **String:** sort and match shell directory
- **Append**
- **Output null**
- **History**

Pipeline! Pipeline!



# Pipeline

[Input | Process > output]



# Condition

[cmd && cmd && ...]

---



# SSH (Secure Shell)

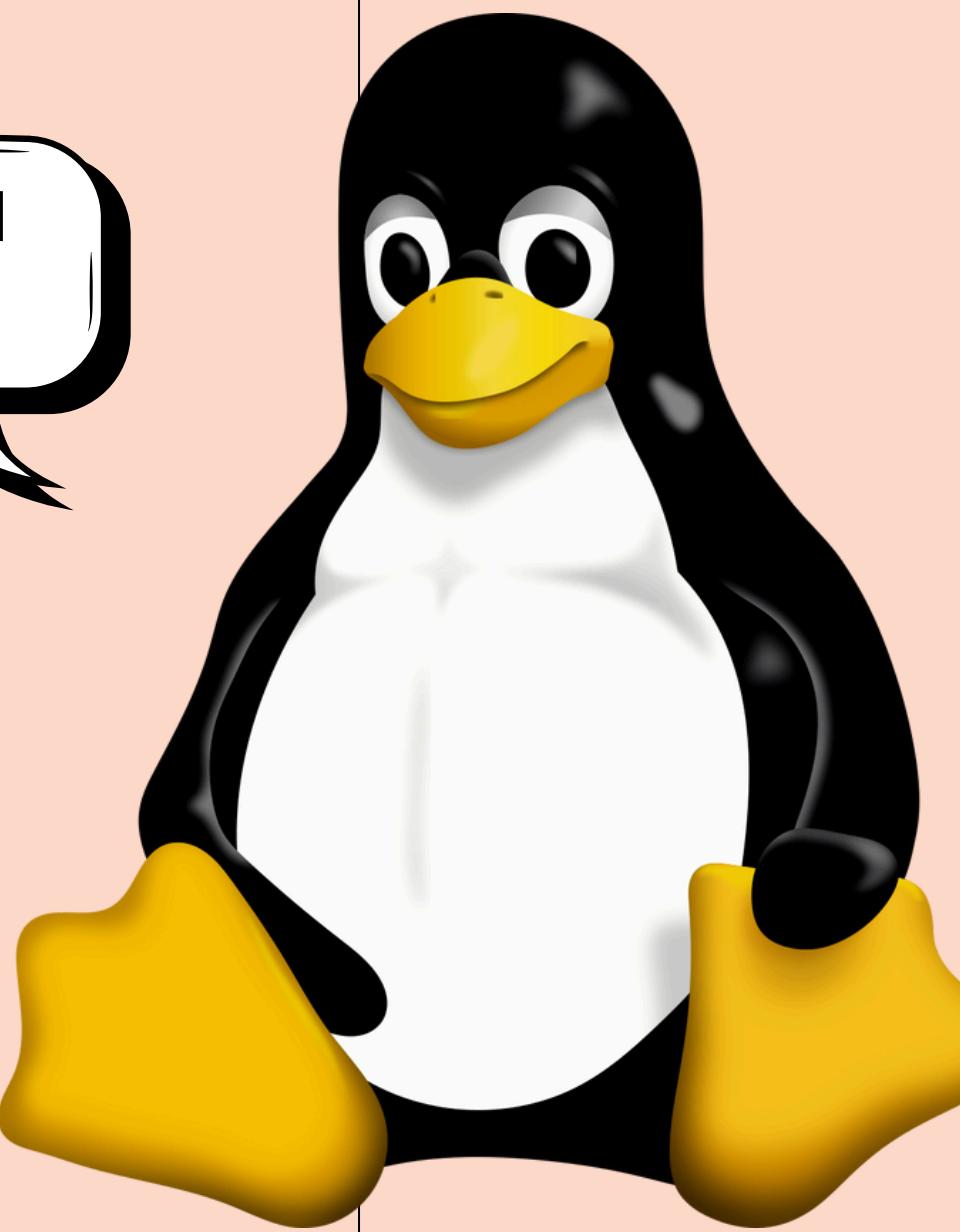
**Introduce to SSH :**

- **What is SSH?**
- **SSH Advantage**

**Example:**

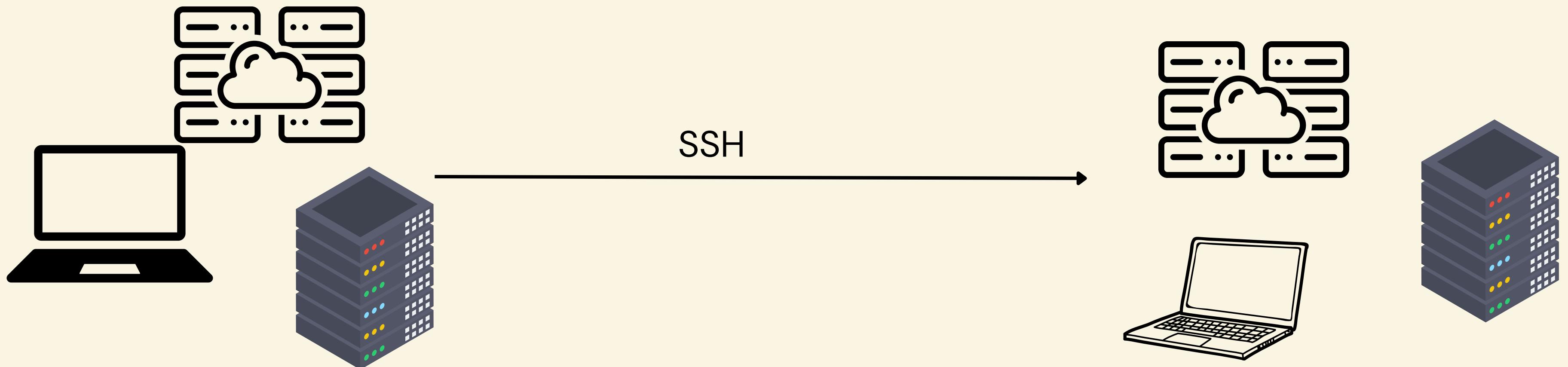
- **AWS connect to GitHub**
- **GitHub connect to AWS**

SSH appeared  
in 1995



# What is SSH and Advantage?

Client -> Host



**Client**

**Host**

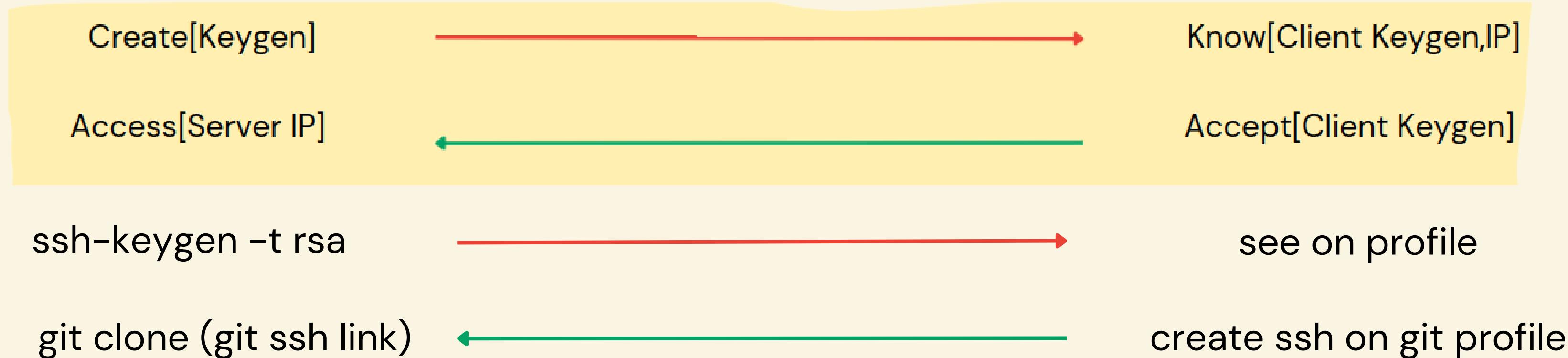
Create[Keygen]

Know[Client Keygen,IP]

Access[Server IP]

Accept[Client Keygen]

# AWS connect to GitHub



# GitHub connect to AWS

