

Linux para Ciência de Dados



Universidade Estadual Paulista, Júlio de Mesquita Filho - UNESP

clayton.pereira@unesp.br

- ❑ [Dat Linux](#)
- ❑ [Linux for Data Science](#)
- ❑ [Linux para 2024: Uma Análise Comparativa](#)
- ❑ [Linux Journey](#)



Por que Linux é usado na Ciência de Dados?



Figure: Distros Linux

Por que Linux é amplamente utilizado na Ciência de Dados?

- **Ambientes de computação na nuvem:** Muitas plataformas de computação em nuvem, como AWS, Google Cloud e Microsoft Azure, utilizam Linux como sistema operacional padrão.

Por que Linux é amplamente utilizado na Ciência de Dados?

- **Ambientes de computação na nuvem:** Muitas plataformas de computação em nuvem, como AWS, Google Cloud e Microsoft Azure, utilizam Linux como sistema operacional padrão.
- **Facilidade na automação:** O Bash e outros interpretadores de comandos permitem automatizar processos de coleta, processamento e análise de dados.

Por que Linux é amplamente utilizado na Ciência de Dados?

- **Ambientes de computação na nuvem:** Muitas plataformas de computação em nuvem, como AWS, Google Cloud e Microsoft Azure, utilizam Linux como sistema operacional padrão.
- **Facilidade na automação:** O Bash e outros interpretadores de comandos permitem automatizar processos de coleta, processamento e análise de dados.
- **Compatibilidade com ferramentas de Ciência de Dados:** Muitas ferramentas como Python, R, Jupyter Notebook, Apache Spark, Docker e Hadoop são mais otimizadas e bem suportadas no Linux.

Por que Linux é amplamente utilizado na Ciência de Dados?

- ❑ **Ambientes de computação na nuvem:** Muitas plataformas de computação em nuvem, como AWS, Google Cloud e Microsoft Azure, utilizam Linux como sistema operacional padrão.
- ❑ **Facilidade na automação:** O Bash e outros interpretadores de comandos permitem automatizar processos de coleta, processamento e análise de dados.
- ❑ **Compatibilidade com ferramentas de Ciência de Dados:** Muitas ferramentas como Python, R, Jupyter Notebook, Apache Spark, Docker e Hadoop são mais otimizadas e bem suportadas no Linux.
- ❑ **Eficiência no processamento de dados:** Linux permite manipular grandes volumes de dados diretamente do terminal com comandos eficientes.

Diferenças entre Linux, Windows e macOS

Comparação de Desempenho, Linha de Comando e Compatibilidade

Característica	Linux	Windows	macOS
Desempenho	Alto, eficiente para servidores e computação em nuvem	Pode ser mais pesado devido a processos em segundo plano	Estável e otimizado para hardware Apple
Linha de Comando	Shell/Bash poderoso para automação	Prompt de Comando e PowerShell	Terminal Unix-like, similar ao Linux
Compatibilidade	Suporte nativo para Python, R, Hadoop, Docker e Spark	Algumas ferramentas exigem instalação manual	Boa compatibilidade, mas limitado a hardware Apple

Table: Comparação entre Linux, Windows e macOS - Parte 1

Comparação de Segurança e Custo

Característica	Linux	Windows	macOS
Segurança	Mais seguro devido ao sistema de permissões	Mais vulnerável a malwares	Seguro, mas depende do ecossistema Apple
Custo	Gratuito e open-source	Pago (licença do Windows)	Pago (inclusive no preço dos dispositivos Apple)

Table: Comparação entre Linux, Windows e macOS - Parte 2

Existem várias distribuições Linux, cada uma com propósitos específicos. Algumas das mais populares para Ciência de Dados incluem:

Ubuntu (mais popular)

- ☐ Fácil de usar, ampla compatibilidade com ferramentas de desenvolvimento.
- ☐ **Versão recomendada:** Ubuntu LTS (Long Term Support).

Debian

- ☐ Base do Ubuntu, mas mais estável e com foco em segurança.
- ☐ Ideal para servidores e ambientes corporativos.

CentOS (Red Hat-based)

- ☐ Utilizado em empresas e servidores.
- ☐ Mais seguro e confiável para aplicações corporativas.

Arch Linux

- ☐ Para usuários avançados, permite personalização total do sistema.

Kali Linux

- ☐ Focado em segurança e testes de penetração.

Comandos Básicos do Terminal

A linha de comando é uma ferramenta essencial no Linux e permite a execução rápida de tarefas sem a necessidade de interfaces gráficas.

Comandos Importantes

- ☐ `pwd` *# Mostra o diretório atual*
- ☐ `ls` *# Lista arquivos e pastas no diretório*
- ☐ `cd pasta/` *# Entra na pasta especificada*
- ☐ `cd ..` *# Volta um nível no diretório*

Comandos Básicos do Terminal

Manipulação de Arquivos

- ☐ `touch arquivo.txt` *# Cria um novo arquivo*
- ☐ `mkdir nova_pasta` *# Cria uma nova pasta*
- ☐ `cp arquivo.txt backup/` *# Copia arquivo para outra pasta*
- ☐ `mv arquivo.txt novo_nome.txt` *# Renomeia ou move um arquivo*
- ☐ `rm arquivo.txt` *# Remove um arquivo*
- ☐ `rmdir pasta_vazia` *# Remove uma pasta vazia*
- ☐ `rm -r pasta_cheia` *# Remove uma pasta com conteúdo*

Comandos Básicos do Terminal

Exibição e Edição de Arquivos

☐ `cat arquivo.txt`

Exibe conteúdo do arquivo

☐ `tail -n 10 arquivo.txt`

Exibe as últimas 10 linhas

☐ `head -n 10 arquivo.txt`

Exibe as primeiras 10 linhas

☐ `nano arquivo.txt`

Abre editor de texto no terminal

Comandos Básicos do Terminal

Outros Comandos Úteis

☐ whoami

Mostra o usuário atual

☐ clear

Limpa a tela do terminal

☐ history

Lista os comandos digitados



Vamos falar um pouco sobre Bash



Figure: Distros Linux

O que é o Bash?

Conhecendo

- ☐ Bash significa **B**ourne **A**gain **S**hell

O que é o Bash?

Conhecendo

- ☐ Bash significa **Bourne Again SHell**
- ☐ É um **interpretador de comandos** (shell) para sistemas Unix/Linux

O que é o Bash?

Conhecendo

- ☐ Bash significa **Bourne Again SHell**
- ☐ É um **interpretador de comandos** (shell) para sistemas Unix/Linux
- ☐ Utilizado para executar comandos, automatizar tarefas e criar scripts

O que é o Bash?

História do Bash

- ❑ Criado por **Brian Fox** em 1989 para o projeto GNU

O que é o Bash?

História do Bash

- ☐ Criado por **Brian Fox** em 1989 para o projeto GNU
- ☐ Substituiu o shell Bourne original (sh)

O que é o Bash?

História do Bash

- ☐ Criado por **Brian Fox** em 1989 para o projeto GNU
- ☐ Substituiu o shell Bourne original (sh)
- ☐ Tornou-se o shell padrão na maioria das distribuições Linux

O que é o Bash?

Características do Bash

- Linguagem interpretada

O que é o Bash?

Características do Bash

- ☐ Linguagem interpretada
- ☐ Forte integração com comandos do sistema

O que é o Bash?

Características do Bash

- ☐ Linguagem interpretada
- ☐ Forte integração com comandos do sistema
- ☐ Suporte a variáveis, estruturas de repetição, condicionais e funções

O que é o Bash?

Características do Bash

- ☐ Linguagem interpretada
- ☐ Forte integração com comandos do sistema
- ☐ Suporte a variáveis, estruturas de repetição, condicionais e funções
- ☐ Permite automação de tarefas com scripts `.sh`

O que é o Bash?

Por que aprender Bash?

- ☐ Domínio do terminal é essencial em ambientes Linux

O que é o Bash?

Por que aprender Bash?

- ☐ Domínio do terminal é essencial em ambientes Linux
- ☐ Scripts em Bash facilitam a repetição de tarefas

O que é o Bash?

Por que aprender Bash?

- ☐ Domínio do terminal é essencial em ambientes Linux
- ☐ Scripts em Bash facilitam a repetição de tarefas
- ☐ Base para automação de deploys, análises, e manutenção de servidores

O que é o Bash?

Comandos Básicos

□ pwd, ls, cd, mkdir, rm

O que é o Bash?

Comandos Básicos

- ☐ pwd, ls, cd, mkdir, rm
- ☐ cp, mv, cat, nano, echo

O que é o Bash?

Comandos Básicos

- ☐ pwd, ls, cd, mkdir, rm
- ☐ cp, mv, cat, nano, echo
- ☐ Comandos podem ser encadeados com `&&`, `||` e `;`

O que é o Bash?

Aplicações do Bash

- ☐ Administração de sistemas e servidores

O que é o Bash?

Aplicações do Bash

- ☐ Administração de sistemas e servidores
- ☐ Análises automatizadas e pré-processamento de dados

O que é o Bash?

Aplicações do Bash

- ☐ Administração de sistemas e servidores
- ☐ Análises automatizadas e pré-processamento de dados
- ☐ Execução de pipelines em ciência de dados

O que é o Bash?

Aplicações do Bash

- ☐ Administração de sistemas e servidores
- ☐ Análises automatizadas e pré-processamento de dados
- ☐ Execução de pipelines em ciência de dados
- ☐ Integração com Python e ferramentas de IA

Exemplo de Script Bash

```
#!/bin/bash  
echo " Digite - seu - nome : "  
read nome  
echo " Ola , - $nome ! - Bem-vindo - ao - Bash ! "
```

obrigado.