

Big Data – Data Engineering

Tema da aula
GIT - Versionamento de Código

FIA Business School

NOSSOS DIFERENCIAIS | QUEM SOMOS



BUSINESS SCHOOL

Graduação, pós-graduação, MBA, Pós-MBA, Mestrado Profissional, Curso In Company e EAD



CONSULTING

Consultoria personalizada que oferece soluções baseadas em seu problema de negócio



RESEARCH

Atualização dos conhecimentos e do material didático oferecidos nas atividades de ensino



Líder em Educação Executiva, referência de ensino nos cursos de graduação, pós-graduação e MBA, tendo excelência nos programas de educação. Uma das principais **escolas de negócio do mundo**, possuindo convênios internacionais com Universidades nos EUA, Europa e Ásia. +8.000 **projetos de consultorias** em organizações públicas e privadas.



Único curso de graduação em administração a receber as notas máximas



A primeira escola brasileira a ser finalista da maior competição de MBA do mundo



Única *Business School* brasileira a figurar no *ranking* LATAM



Signatária do Pacto Global da ONU



Membro fundador da ANAMBA - Associação Nacional MBAs



Credenciada pela AMBA - Association of MBAs



Credenciada ao Executive MBA Council



Filiada a AACSB - Association to Advance Collegiate Schools of Business



Filiada a EFMD - European Foundation for Management Development



Referência em cursos de MBA nas principais mídias de circulação



LABDATA FIA

NOSSOS DIFERENCIAIS | QUEM SOMOS

O **Laboratório de Análise de Dados** – LABDATA é um Centro de Excelência que atua nas áreas de ensino, pesquisa e consultoria em análise de informação utilizando técnicas de *Big Data*, *Analytics* e *Inteligência Artificial*.



Profª Drª Alessandra Montini

O LABDATA é um dos pioneiros no lançamento dos cursos de *Big Data* e *Analytics* no Brasil. Os diretores foram professores de grandes especialistas do mercado.

+10 anos de atuação.

+9000 alunos formados.

Docentes

- Sólida formação acadêmica: doutores e mestres em sua maioria;
- Larga experiência de mercado na resolução de *cases*;
- Participação em congressos nacionais e internacionais;
- Professor assistente que acompanha o aluno durante todo o curso.

Estrutura

- 100% das aulas realizadas em laboratórios;
- Computadores para uso individual durante as aulas;
- 5 laboratórios de alta qualidade (investimento +R\$2MM);
- 2 unidades próximas à estação de metrô (com estacionamento).

Corpo Diretivo

COORDENADORES DO LABDATA | ATUAÇÃO ACADÊMICA E PROFISSIONAL



PROFA. DRA. ALESSANDRA DE ÁVILA MONTINI

Diretora do LABDATA-FIA, apaixonada por dados e pela arte de lecionar. Tem muito orgulho de ter criado na FIA cinco laboratórios para as aulas de Big Data e Inteligência Artificial. Possui mais de 20 anos de trajetória nas áreas de Data Mining, Big Data, Inteligência Artificial e Analytics. Cientista de dados com carreira realizada na Universidade de São Paulo. Graduada e mestra em Estatística Aplicada pelo IME-USP e doutora pela FEA-USP. Com muita dedicação chegou ao cargo de professora e pesquisadora na FEA-USP, ganhou mais de 30 prêmios de excelência acadêmica pela FEA-USP e mais de 30 prêmios de excelência acadêmica como professora dos cursos de MBA da FIA. Orienta alunos de mestrado e de doutorado na FEA-USP. Parecerista da FAPESP e colunista de grandes portais de tecnologia.



Conteúdo Programático



ANALYTICS



ANALYTICS



ESTATÍSTICA
APLICADA



MACHINE
LEARNING



DEEP LEARNING

TECNOLOGIA



CLOUD



HADOOP



SPARK



INGESTÃO DE
DADOS



NoSQL

LINGUAGEM



PYTHON

PROJETO



PROJETO ANALYTICS



PROJETO FINAL



Conteúdo da Aula

- 1. Introdução
- 2. Conceito de GIT
 - i. GIT
 - ii. Fundamentos GIT
 - iii. Comandos do GIT
 - iv. GIT HUB
 - v. Acessos
 - vi. Ambiente
 - vii. Ferramentas



GIT

1. INTRODUÇÃO | BIG DATA

GIT

Git é um sistema de controle de versão distribuído, utilizado para gerenciar e versionar projetos de software, dados e outras formas de arquivos. Algumas das principais características do Git incluem comandos que veremos no decorrer da aula.





Repositório

Repositório: um repositório é onde o Git armazena todos os arquivos e histórico de versão de um projeto. Cada repositório é independente, permitindo que vários desenvolvedores trabalhem em diferentes partes do projeto simultaneamente.



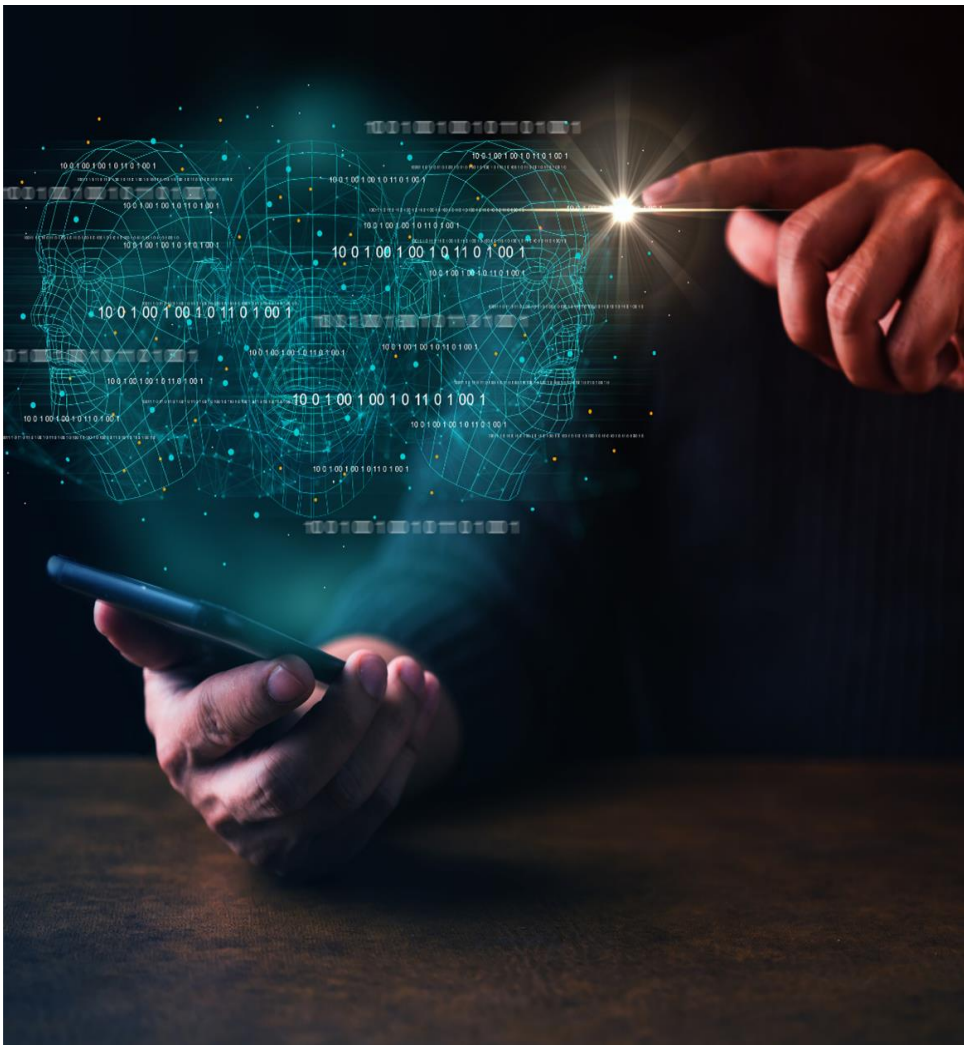
1. Conceitos

[illegible]

Commits: um commit é uma versão específica do projeto que é registrada no histórico do Git. Cada commit inclui um conjunto de mudanças que foram feitas desde a última versão registrada.

GIT

1. Conceitos



Merge

Merge: o processo de mesclagem de um branch é a incorporação das alterações feitas em um branch para o branch principal do projeto.

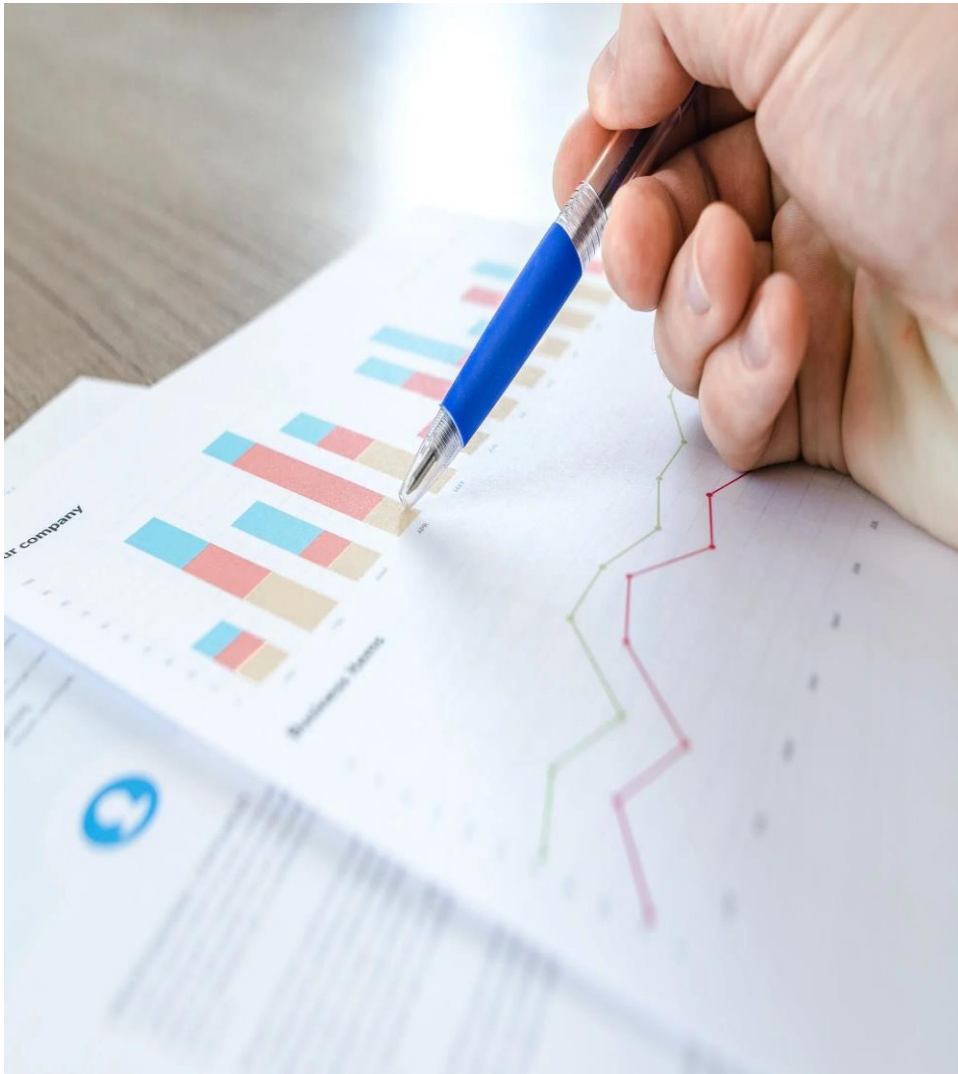




Pull requests (PR)

Pull requests: um pull request é uma solicitação para que as alterações em um branch sejam mescladas com o branch principal do projeto. Isso permite que outros desenvolvedores revisem as alterações antes de serem incorporadas ao projeto.





Clone

Clone: um clone é uma cópia completa de um repositório, incluindo todo o histórico de versão, criada para permitir que outro desenvolvedor trabalhe no projeto sem afetar o repositório original.



GIT

1. Conceitos

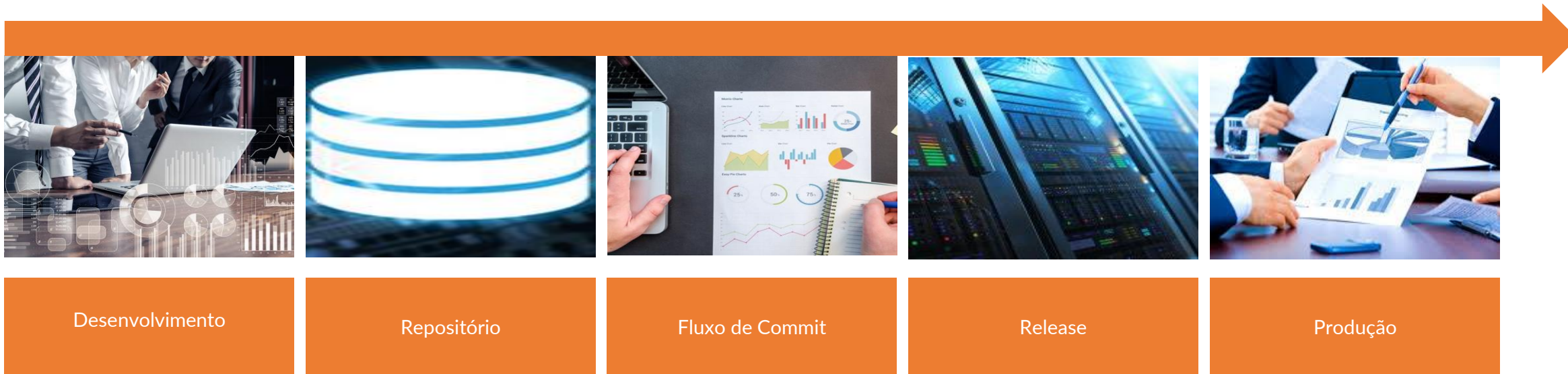


Fork

Fork: um fork é uma cópia completa de um repositório, criada por outro desenvolvedor para trabalhar em suas próprias alterações no projeto sem afetar o repositório original.



Fluxo de trabalho GIT



Esses são apenas alguns dos conceitos básicos do Git. Há muitos outros recursos avançados e comandos que podem ser usados para gerenciar e versionar projetos com o Git



GIT

1. Conceitos



COMANDOS

git init: cria um novo repositório Git vazio.

git clone: faz uma cópia completa de um repositório existente.

git add: adiciona alterações feitas em arquivos ao índice do Git para prepará-los para um commit.

git commit: cria um novo commit contendo as alterações adicionadas ao índice.

git status: exibe o status atual do repositório, incluindo quais arquivos foram modificados, adicionados ou excluídos.

git log: exibe o histórico de commits do repositório.

git branch: exibe todas as ramificações (branches) do projeto.

git checkout: alterna para uma ramificação ou commit específico.

git merge: mescla alterações de um branch para outro.

git pull: puxa as alterações mais recentes do repositório remoto e mescla-as com o branch atual.

git push: envia as alterações locais para um repositório remoto.

git remote: gerencia conexões com repositórios remotos.



GIT

1. Conceitos



GIT

1. Conceitos



GIT HUB

GitHub é uma plataforma web de hospedagem de repositórios Git, que permite que desenvolvedores compartilhem, colaborem e controlem versões de seus projetos de software e outros arquivos.

O GitHub é amplamente utilizado por desenvolvedores de todo o mundo para hospedar e colaborar em projetos de código aberto e proprietários. Ele se tornou uma parte essencial do ecossistema de desenvolvimento de software e dados moderno.





Principais Características do GitHub

- Repositórios públicos e privados: os desenvolvedores podem hospedar seus projetos em repositórios públicos, que são visíveis para todos os usuários do GitHub, ou em repositórios privados, que são visíveis apenas para os colaboradores especificados.
- Colaboração: os desenvolvedores podem colaborar em projetos por meio de ramificações, pull requests e revisões de código.
- Ferramentas de gerenciamento de projetos: o GitHub fornece ferramentas para gerenciar e acompanhar o progresso de projetos, incluindo quadros de tarefas (project boards), issues e milestones.
- Integrações: o GitHub se integra a muitas outras ferramentas de desenvolvimento, como IDEs, ferramentas de integração contínua e entrega contínua (CI/CD), serviços de hospedagem de documentação e outros.



CRIAÇÃO DE CONTA NO GIT HUB



GIT

1. Conceitos

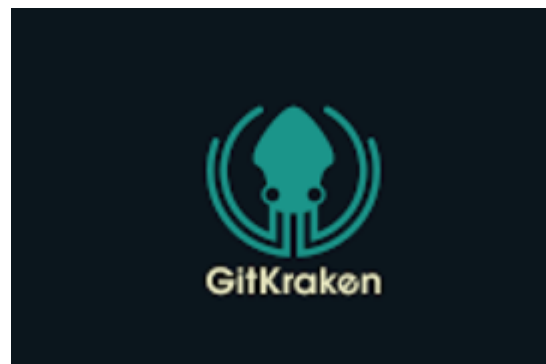
21



GIT

1. Conceitos

22



PREPARAÇÃO DO AMBIENTE

