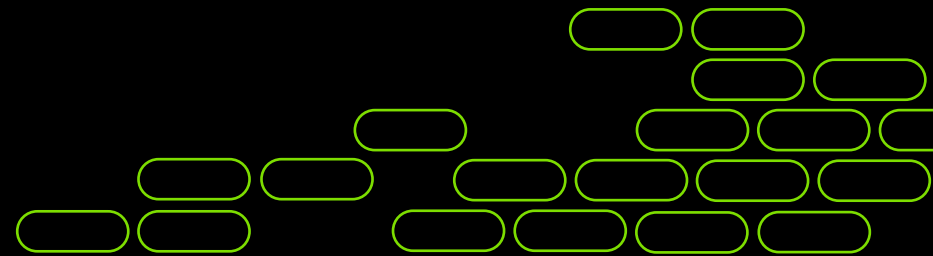




Aprenda com quem faz

< MBA EM ENGENHARIA DE DADOS />

EMENTA



Conteúdo_

> Sobre a XP EDUCAÇÃO	03
> Conteúdo	04-11
> Projeto Aplicado	12
> Contato	13

Sobre a XP EDUCAÇÃO

A XP Educação integra a XP Inc., um dos maiores grupos empresariais do país. A empresa que simplificou a vida financeira de milhares de pessoas agora vai transformar vidas e carreiras por meio da educação.

A Faculdade XP Educação traz toda a sua potência disruptiva e transformadora para oferecer uma formação profissional alinhada às demandas da nova economia digital, em um ecossistema educacional único.

Nossos compromissos são:

● Melhorar a vida das pessoas

A educação é a base de transformação da sociedade e contribui para desenvolver visão crítica e consciência cidadã.

● Formar novos talentos digitais

Cursos acessíveis e de qualidade para reduzir o déficit de talentos digitais no Brasil formando profissionais para o mercado de trabalho.

● Inovar para educar e transformar

Novas experiências educacionais tornam o aprendizado cada vez mais leve e atrativo para os alunos.

Bootcamp Engenheiro(a) de Dados_

● Objetivo principal:

O objetivo do Bootcamp é apresentar técnicas e tecnologias usadas na construção do pipeline de dados (data flow) que envolvem processos de coleta, preparação, armazenamento e acesso a dados, além da integração e combinação de diferentes fontes de dados.

● Pré-requisitos

- Conhecimento em banco de dados, modelagem, linguagem SQL, lógica de programação, programação orientada a objetos, linguagens de programação (preferencialmente Python e R).
- Inglês para leitura, Cartão de crédito para criação das contas free nas plataformas em nuvem.

Módulos_

● Fundamentos em Engenharia de Dados

- Conceitos fundamentais: dados, fontes de dados, Big Data, tipos de dados;
- Visão geral das atividades do engenheiro de dados;
- Modelagem de dados;
- Linguagem SQL;
- Fundamentos de Modelagem Dimensional;
- Tipos de arquitetura e etapas de modelagem dimensional;
- Data hubs, Data Lakes, Data Swamps e Data Ponds.

Duração: 32 horas

● Linguagem Python Aplicada à Engenharia de Dados

- Introdução à linguagem Python;
- Orientação a objetos, classes e herança na prática;
- Tipos de dados, variáveis e estruturas de armazenamento de dados;
- Estruturas condicionais e de repetição;
- Funções, Funções Built-in, Função Lambda;
- Map-reduce, filter, zip, list e enumerate na prática;
- Tratamento de erros;
- Principais bibliotecas: Pandas e Numpy.

Duração: 32 horas

● Armazenamento de Dados

- SGBDs SQL(relacionais), NoSQL, NewSQL: o que é, diferenças e vantagens, ACID versus teorema CAP / BASE, Categorias de SGBD NoSQL, Tecnologias existentes;
- Sistemas de arquivo: o que é, vantagens, tecnologias existentes;
- Armazenamento de dados em nuvem: AWS S3, GCS (Google Cloud) e Azure Blob Storage.
- Conceitos de Data Lake, Data Warehouse e Data Lakehouse.

Duração: 32 horas

● Pipeline de Dados

- Pipeline de Ciência de dados;
- Extração e Transformação de Dados;
- Ingestão de Dados;
- Soluções de ETL: Pentaho, Apache Nifi, Apache Airflow, KubeFlow, Prefect;
- Fluxo de dados: Data Flow com Airflow.

Duração: 32 horas

● Desafio Final

Conclusão da aplicação final.

Duração: 16 horas

Bootcamp Engenheiro(a) de Dados Cloud_

● Objetivo principal:

Formar profissionais capazes de projetar e implementar soluções e plataformas de Big Data, escaláveis, confiáveis e robustas em ambiente em nuvem, para uso analítico ou operacional.

● Pré-requisitos

- Conhecimento de SGBDs relacionais e NoSQL, linguagem SQL, linguagem de programação estruturada (preferencialmente Python);
- Inglês para leitura;
- Cartão de crédito para contas nas plataformas em nuvem.

Módulos_

● Fundamentos em Arquitetura de Dados e Soluções em Nuvem

- Fundamentos em arquitetura de dados
 - Componentes/Elementos de uma arquitetura de dados - o Well Architected Framework
- Modelos de Arquiteturas de Dados em Nuvem
 - Bancos de Dados gerenciados (relacional e não relacional);
 - Projetos de repositórios de dados: Data Lake e DW.
- AWS RDS;
- Data Lake com AWS S3;
- Data LakeHouse com Delta Lake e EMR;
- Extração de Dados com AWS DMS;
- Processamento de Big Data com AWS EMR e AWS Glue;
- Streaming de Dados com AWS Kinesis;
- Disponibilização de Dados com AWS Athena;
- IaC (Infraestrutura como Código): Terraform + Github.

Duração: 32 horas

● Tecnologias de Big Data - Processamento de Dados Massivos

- Computação distribuída, conceitos básicos;
- Introdução ao Ecossistema Apache Hadoop, Arquitetura e principais componentes;
- Apache Spark, conceitos, instalação e configuração;
- Dataframes;
- Spark SQL;
- Fluxos de Dados Contínuos com Spark Streaming;
- Spark ML, Spark GraphX;
- Spark na Nuvem.

Duração: 32 horas

● Desenho de Arquiteturas de Dados Escaláveis

- Arquiteturas escaláveis
 - Conceito e tipos de escalabilidade
- Arquitetura de microsserviços, conceitos e aplicações
- Tecnologias para escalabilidade de soluções em nuvem (Load Balancers, Auto-scaling);
- Virtualização e Containers;
- Docker e Kubernetes na prática;
- Kubernetes - implantação de soluções utilizando manifests .yaml, helm charts e operadores;
- Monitoramento de soluções escaláveis com Prometheus e Grafana.

Duração: 32 horas

● Processamento de Fluxos Contínuos de Dados

- Stream Processing Applications;
- Arquitetura de sistemas de stream processing;
- Padrões de projetos de fluxos contínuos de dados;
- Arquiteturas de projetos de dados (Arquitetura streaming, arquitetura lambda, arquitetura kappa e arquitetura Unifield);
- Coleta, armazenamento e processamento de dados de fluxos contínuos:
 - Conexão a fontes de dados com Kafka Connect
 - Processamento em Tempo Real com ksqlDB e Spark Streaming
 - Entrega de dados em solução de armazenamento (Data Lake, Data Warehouse, DBs modernos)
- Arquitetura e pipelines de dados para tempo real usando o Kafka e novos storages para armazenamento (Apache Druid, Apache Pinot, Hive, Elasticsearch, Neo4j).

Duração: 32 horas

● Desafio Final

Conclusão da aplicação final.

Duração: 16 horas

Projeto Aplicado_

● Sobre o P.A:

A partir do aprendizado nos bootcamps que compõem a sua Trilha de Pós-Graduação, é hora de começar seu Projeto Aplicado.

É a última etapa da sua especialização, momento em que você será, mais do que nunca, protagonista da sua jornada de conhecimento. No PA, você será apresentado para problemas com alto grau de incerteza e o desafio será propor ideias inovadoras para solucioná-los.

O projeto substitui o tradicional Trabalho de Conclusão de Curso, com uma diferença fundamental: é focado em entrega de valor, aplicabilidade, praticidade e com situações atuais que ocorrem no mercado de trabalho.

Ao longo do PA, você terá dinâmicas de grupo e orientação personalizada para auxiliá-lo(a) nesta etapa.

Ao final dos dois bootcamps e do PA, **você receberá seu título de pós-graduação chancelado pelo MEC** e terá ainda mais habilidades técnicas para alavancar sua carreira.



Aprenda com quem faz

www.xpeducacao.com.br

