

# Bootcamp Engenheiro(a) de Dados

---

EMENTA DO CURSO



# Olá!

Neste Bootcamp você conhecerá as técnicas e tecnologias usada na construção do pipeline de dados (data flow) que envolvem processos de coleta, preparação, processamento, armazenamento e acesso a dados.



# Bootcamp Engenheiro(a) de Dados

## Conteúdo

Sobre o IGTI	04
IGTI em números	05
Sobre o Bootcamp	06
Engenheiro(a) de Dados	07
Módulos	08
Contato	11

... IGTI é exemplo de programas a distância (EAD)”...

—— Valor Econômico

Valor  
ECONÔMICO

## Sobre o IGTI

O Instituto de Gestão e Tecnologia da Informação, fundado em 2006, é uma instituição de ensino superior credenciada pelo MEC. É referência nacional na formação profissional em TI e Tecnologias Emergentes, e possui um modelo educacional a distância que prioriza a excelência acadêmica dos seus alunos.

**Conecte-se com o futuro! Estude a distância com aulas semanais por videoconferência.**



## IGTI em números

Muitos são os motivos para investir no seu futuro profissional através do IGTI.



### 85% de Satisfação

Percentual de alunos que avaliam as suas disciplinas com 4 ou 5 estrelas.



### 98% de Satisfação no Atendimento

A educação do IGTI é centrada na experiência e satisfação do aluno.

## O que estão falando sobre o IGTI?

O IGTI foi contemplado em 2020, pelo segundo ano consecutivo, como uma das 3 empresas mais inovadoras no uso da TI no Brasil pelo IT Forum, na categoria Educação.



### 100% Interativo

As aulas semanais ao vivo permitem esclarecer dúvidas diretamente com o professor.



### 93% de Retenção

Taxa de alunos que concluem o curso muito superior à média de mercado para EAD.

Formação intensiva, prática e imersiva em habilidades para atuação no mercado de tecnologia.

# Sobre os Bootcamps

## Benefícios

- Domine as técnicas e ferramentas mais utilizadas no mercado.
- Adquira a experiência necessária ao mercado com aprendizado prático.
- Garanta o acesso às melhores oportunidades de carreira.

## Metodologia

Imersivo: Uma abordagem hands-on de alto impacto para formar skills técnicos muito mais rápido que o ensino tradicional.

Prático: Um programa de ensino focado nas habilidades técnicas para atuar imediatamente no mercado de trabalho.

Interativo: Aulas semanais por videoconferência para solução de dúvidas, aprofundamento nos conteúdos, execução / correção de atividades práticas e orientações e debates sobre os desafios propostos.

## Informações Complementares

**Carga horária** 148 horas

**Duração** 2 meses

# 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.

## Objetivos específicos:

- Aprenda a construir e automatizar seu pipeline de dados.

- Conduza iniciativas envolvendo o planejamento de soluções integradas.

## 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

---



## Aquecimento e Regras do Jogo

- Visão geral da dinâmica do curso, conteúdos e ferramentas utilizadas.

**Duração: 4 horas**



## 1 Fundamentos

- Conceitos fundamentais: Dados, fontes de dados, Big Data, tipos de dados;
- Visão geral do pipeline de ciência de dados: coleta, preparação, armazenamento, processamento/análise, visualização;
- Atividades de engenharia do Pipeline de Dados;

- Fundamentos de modelagem de dados;
- Linguagem SQL;
- Visão geral dos processos de coleta, armazenamento e preparação de dados
- Exemplo prático de coleta, preparação e armazenamento de dados;
- Visão geral em Cloud computing.

**Duração: 32 horas**



## 2 Coleta e Preparação de Dados

- Fundamentos da coleta de dados;
- Particularidades dos processos aquisição, extração, ingestão, coleta;
- Coleta de dados em bases de dados (Relacional e NoSQL);



- Coleta de dados via API (Twitter, site de notícias, dados governamentais);
- Coleta de dados via crawler e scraper  
Manipulando dados desalinhados, inconsistentes e não padronizados;
- Transformação, limpeza, enriquecimento de dados;
- Integração de dados, Tratamento de dados ausentes;
- Categorização e discretização;
- Técnicas de anonimização de dados (LGPD);
- Combinando dados de múltiplas fontes;
- Normalização de Dados;
- Transformação e transferência de dados;
- Ferramentas para preparação de dados;
- Exemplos e práticas de preparação em dados estruturados, semiestruturados e não estruturados;

- Preparação de dados para mineração de texto (n-gramas, remoção de stopwords e normalização morfológica (stemming);
- Estudos de caso;

**Duração: 32 horas**

### **3 Novos Paradigmas em Armazenamento de Dados**

- SGBDs SQL, NoSQL, NewSQL. ACID versus Teorema CAP / BASE, Categorias de SGBD NoSQL, Tecnologias existentes;
- Sistemas de arquivo;

- Armazenamento de dados em nuvem: Conceitos de Data Lake, Data Warehouse e Data Lakehouse;
- Soluções de Data Lake - AWS S3, GCS (Google Cloud) e Azure Blob Storage;
- Soluções de DW - AWS Redshift, Google Big Query e Azure SQL Datawarehouse;
- Engines de Data Lake: Dremio, Trino, Presto;
- Soluções de Armazenamento escalável - Apache Druid, Apache Pinot, Elasticsearch.

**Duração: 32 horas**

## 4 Pipelines de Dados

- Atividades do pipeline de dados: aquisição, transformações, ingestão;

- Visão geral de soluções para extração, ingestão, transformação, armazenamento e análise de dados;
- Fluxo de dados ETL/ELT;
- Implementação e automatização de ETL/ELT;
- Pipeline de Dados com Apache Airflow;
- Pipeline de Dados com Apache Nifi;
- Exemplos de Pipeline de Dados em tempo real com Apache Kafka.

**Duração: 32 horas**

## 5 Desafio Final

Conclusão da aplicação final.

**Duração: 16 horas**



## Contato

Rua Roma 561, Santa Lúcia.  
CEP 30.360-680  
Belo Horizonte, Minas Gerais

[www.igti.com.br](http://www.igti.com.br)

0800 200 4488 | (31) 3047-3612

[Inscreva-se no programa](#)