



**Fecomércio  
Sesc**

## **Data Science – Princípios e Técnicas**

Março  
2024



# Vamos nos conhecer



De onde eu vim e para aonde vamos?

# Quem sou eu?

# Data Science



---

Prof. Marco Mialaret

Doutorando em Ciência da Computação pelo CIN-UFPE.

Mestre em Matemática pela UFPB.

Licenciado em Matemática pela UFRPE e Tecnólogo em Banco de Dados pelo SENAC.

Atuo na área acadêmica superior desde 2011 e atuo como cientista de dados (~~quando sobra tempo~~).

# Data Science



---

Onde me encontrar:

<https://www.linkedin.com/in/marco-mialaret-junior/>

e

<https://github.com/MatmJr>

# A disciplina

# Data Science



## Ementa

Ciência de Dados, conceitos, abrangência e mercado de trabalho.

Estatística aplicada à Ciência de Dados.

Principais conceitos relacionados: BI, BIG Data, Data Mining, IA, Machine Learning e Deep Learning.

Definição de Problemas em DataScience.

Modelagem de Dados para projetos de Ciência de Dados.

Preparação, limpeza e tratamento de dados.

Explorando a linguagem R em seus fundamentos aplicados à Ciência de Dados.

Explorando a linguagem Python e bibliotecas disponíveis para IA e ML em seus fundamentos aplicados à Ciência de Dados.

Bancos de dados NoSql e armazenamento distribuído.

Processamento Distribuído e ferramentas. Princípios de Ciência de Dados associados à IA e Machine Learning.

# Data Science

---

## Objetivo Geral

Capacitar estudantes com conhecimentos introdutórios de Data Science, cobrindo desde os fundamentos teóricos até a aplicação prática de técnicas.



# Data Science



---

## Objetivos Específicos

- Introduzir os conceitos fundamentais de Data Science, sua importância no cenário atual de tecnologia e análise de dados, e explorar as principais áreas de aplicação.
- Explorar conceitos estatísticos essenciais aplicados à Ciência de Dados, incluindo probabilidade, inferência estatística, testes de hipóteses e análise exploratória de dados.
- Detalhar os conceitos de Business Intelligence, Big Data, Data Mining e Inteligência Artificial, e como eles se relacionam e contribuem para a Ciência de Dados.

# Data Science



- Ensinar como identificar, definir e estruturar problemas a serem resolvidos através de projetos de Ciência de Dados, focando na formulação de hipóteses e na definição de objetivos claros.
- Abordar técnicas de modelagem e estruturação de dados para análise, incluindo normalização, dimensionalidade e técnicas de redução de dados.
- Desenvolver habilidades práticas na preparação, limpeza e tratamento de dados para garantir a qualidade e precisão necessárias para análises confiáveis.

# Data Science

---

- Introduzir os fundamentos da linguagem R e sua aplicabilidade na análise estatística e modelagem de dados em projetos de Ciência de Dados.
- Ensinar os fundamentos da linguagem Python e o uso de bibliotecas específicas para Inteligência Artificial e Machine Learning, como Pandas, NumPy, Scikit-learn, TensorFlow e Keras.

# Data Science



---

## Conteúdo Programático

- 1 - Introdução à Ciência de Dados: Conceitos e Campo de Atuação
- 2 - Estatística Aplicada à Ciência de Dados
- 3 - Definição de Problemas e Modelagem de Dados em Ciência de Dados
- 4 - Preparação, Limpeza e Tratamento de Dados
- 5 - Python e Bibliotecas para Ciência de Dados

# Data Science



---

## Conteúdo Programático

6 - Bancos de Dados para Ciência de Dados: SQL, NoSQL, e Armazenamento Distribuído

7 - Aplicação de Machine Learning e Deep Learning em Ciência de Dados

8 - Introdução à Linguagem R e Aplicações em Ciência de Dados

# Avaliações e Materiais

# Data Science

---

## Formas de Acompanhamento

- Frequência nas aulas
- Participação em sala
- Entrega das atividades

# Data Science

---

## Material de Estudo:

- Slides no Classroom do Aluno
- GitHub
- Bibliografia da Disciplina
- Vídeos
- Trabalhos
- Estudos de Casos
- Atividades





# Data Science



---

## Metodologia de Avaliação

- **Primeiro Processo Avaliativo** – Entrega das atividades e do projeto 1.
- **Segundo Processo Avaliativo** - Entrega das atividades e do projeto 2.

# Conhecendo vocês

# Data Science

---

Agora é com vocês!

- Nome
- Idade
- Você está procurando emprego? Em qual área?
- Um fato aleatório sobre você

# Data Science

---

Fim das férias!

# Git & GitHub na nossa disciplina

# Data Science

---

**Git** é um sistema de controle de versão distribuído que permite aos desenvolvedores rastrear e gerenciar mudanças no código fonte ao longo do tempo. É uma ferramenta essencial para a colaboração em projetos de desenvolvimento de software.

# Data Science

---

**GitHub** é uma plataforma de hospedagem para repositórios Git que facilita a colaboração entre desenvolvedores. Além de armazenar código, o GitHub oferece ferramentas para revisão de código, gerenciamento de projetos e integração com outras ferramentas de desenvolvimento.

# Hora do Exercício



# Data Science

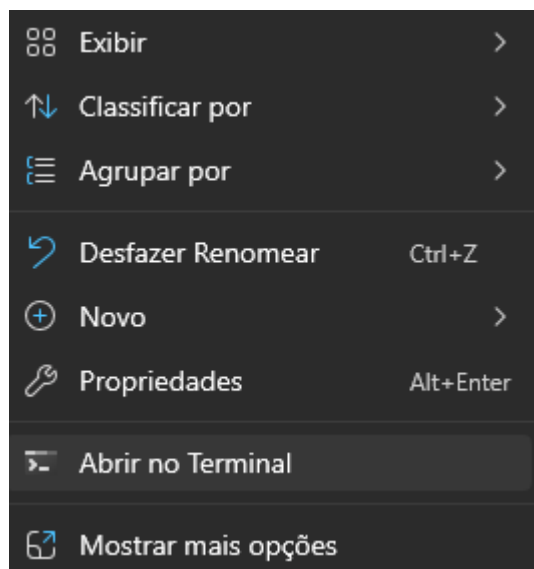
---

Exercício:

- 1 – Encontrar o usuário MatmJr no GitHub
- 2 – Encontrar o Repositório CursoDataScience
- 3 – Acessar o Material da Primeira Semana

# Data Science

Após criar a pasta chamada BigData, clicar com o botão do lado direito e selecionar a opção abrir no terminal



```
PS E:\Senac\Aulas\CursoDataScience> code .
```

# Dúvidas?

---



**Marco Mialaret, MSc**

**Telefone:**

**81 98160 7018**

**E-mail:**

**marcomialaret@gmail.com**

