



## Programa do curso Bootcamp em Análise de Dados (turma exclusiva para mulheres) – 2023

**Entidade Promotora:** Escola Nacional de Administração Pública - Enap  
**Endereço:** Asa Sul SPO área especial 2-A CEP: 70.610-900 Brasília - DF  
**CNPJ:** nº 00.627.612/0001-09

**Contato:** [bootcamp@enap.gov.br](mailto:bootcamp@enap.gov.br)  
**site:** <https://www.enap.gov.br/pt/cursos/coding-bootcamp>

**Ínicio das aulas:** 02 outubro  
**Término das aulas:** 31 outubro  
**Demoday:** 01 de novembro

**Horário:** 9h às 17h  
**Carga Horária:** 150 horas (132 horas síncronas e 18 horas assíncronas)

**Docente:** Hélio Bomfim de Macêdo Filho  
**Monitores:** Bruno Garcia, Fernanda Teixeira, Thais Salzer, Kalina Rabbani

**Nível de formação:** Certificação Avançada em Análise de Dados

### **Público-alvo:**

Servidoras públicas federais que trabalham com planilhas eletrônicas ou base de dados e desejam ser analistas de dados ou pretendem iniciar a trajetória como cientistas de dados ou engenheiras de *machine learning*.

### **Pré-requisitos:**

Conhecimento avançado na linguagem em programação Python, conhecimento introdutório nas bibliotecas numpy e pandas, bem como domínio do inglês instrumental.

### **Processo Seletivo:**

Aplicação de teste de habilidade técnica para prospecção do perfil da turma.

### **Objetivo / Competência: “o que vc vai aprender neste curso”**

A participante experimentará o mundo da programação de computadores para ciência de dados no contexto da Administração Pública e será capaz de:



- Reconhecer o escopo de problemas da Administração Pública que podem ser endereçados por programação de computadores e análise de dados;
- Esquematizar um arcabouço de linguagem de programação e suas bibliotecas para a produção de conhecimento a partir de dados brutos;
- Utilizar o pensamento computacional e a capacidade analítica para aumentar o desempenho das servidoras públicas nas suas atividades laborais; e
- Desenhar soluções de alto desempenho e com economia de escala na Administração Pública.

### **Na Prática: “substituir o Ciencia de Dados”**

A servidora pública federal será capacitada para trabalhar, individualmente ou em equipe, em todas as etapas que envolvem a extração de dados a partir de diversas fontes; a transformação dos dados a partir dos seus cruzamentos; bem como a produção de gráficos úteis para serem disponibilizados em aplicativos web. Mais detalhadamente, teremos:

- Inserção no campo do conhecimento que mescla dados, computação e métodos para a extração de informação útil, bem como aquisição de vocabulário e fluência nos fundamentos técnicos mais utilizados na análise de dados;
- Habilidade para extrair dados das mais diversas fontes, incluindo consulta a banco de dados, realização de download de arquivo de página na *web*, consulta a interface de programação de aplicações (API) e carregamento de arquivos armazenados localmente no computador todos por meio de linguagem de programação;
- Fluência nas ferramentas projetadas para trabalhar com grandes volumes de dados e com maior expressividade computacional;
- Capacidade em limpar e preparar os dados de eventuais ruídos, mesclar dados pulverizados em locais distintos, bem como dispor os dados em formatos mais fáceis de analisar;
- Criatividade para construir visualizações informativas, tanto para explorar o universo dos dados sendo trabalhados, quanto para produzir informações com significado;



- Prática em ferramentas para trabalho em equipe, bem como ferramentas que requerem pouca ou nenhuma experiência em desenvolvimento *web* para criar aplicativos na Internet com suas visualizações informativas;
- Cobertura de todo o ciclo de vida de um projeto de análise de dados da Administração Pública Federal, bem como a consolidação de material de referência para consulta futura.

## Ementa:

1. Aspectos de Ciência de Dados
2. Introdução aos Ambientes de Desenvolvimento
3. Aspectos Nativos de Python
4. Fontes de Dados
5. Manipulação de Dados em Alto Desempenho
6. Análise de Dados em Alto Desempenho
7. Limpeza e Preparação dos Dados
8. Tratamento dos Dados
9. Visualização dos Dados
10. Agregação dos Dados
11. Operações em Dados Agregados
12. Desenvolvimento Web
13. Desenvolvimento Colaborativo
14. Desenvolvimento Embarcado na Inteligência de Negócio

## Metodologia de Ensino:

Os encontros são síncronos e, diariamente, 50% da aula é expositiva com conceitos teóricos mesclados com aplicações práticas e os 50% restantes são separados para estimular as participantes a resolver desafios. As seguintes diretrizes são chave para a condução da aprendizagem:

- **Prática:** as aulas são conduzidas em Python e com o uso de bases de dados reais como instrumentos didáticos para consolidar o domínio das técnicas apresentadas e fomentar a autonomia da participante na sua vida profissional;
- **Assessorada:** o desenvolvimento da solução dos desafios pelas participantes é assessorado por monitores qualificados para apoiar as participantes em todos os aspectos dos desafios;



- **Colaborativa:** além da assessoria dos monitores, as participantes são estimuladas a aprender continuamente com os facilitadores e especialmente com as outras participantes;
- **Escalonada:** os desafios são desenvolvidos em ordem crescente de complexidade, proporcionando um ambiente estimulante e incremental de aprendizagem para alcançar, com equidade, a todos as participantes; e
- **Continuada:** o curso apresenta o conteúdo respeitando uma ordem lógica e as aulas são interligadas por meio de revisão em cada início de aula.

## Avaliação da Aprendizagem:

O curso é composto de exposição a conceitos teóricos e desafios diários, seguidos de atividades práticas supervisionadas em que as participantes são estimuladas a aplicar os conhecimentos do curso na sua esfera de atuação profissional. A avaliação da aprendizagem será composta pelo trabalho desenvolvido nas atividades práticas supervisionadas.

### PLANO DE AULA:

#### 1º Assunto

	<p>Aspectos de Ciência de Dados</p> <ul style="list-style-type: none"> <li>- Conhecimentos e Habilidades do Cientista de Dados</li> </ul> <p>Introdução aos Ambientes de Desenvolvimento</p> <ul style="list-style-type: none"> <li>- Notebooks</li> </ul> <p>Aspectos Nativos de Python (Básico)</p> <ul style="list-style-type: none"> <li>- Tipos de Dados Escalares</li> <li>- Controle de Fluxo</li> </ul>	
--	---	--







#### 2º Assunto

	<p>Aspectos Nativos de Python (Intermediário)</p> <ul style="list-style-type: none"> <li>- Estruturas de Dados e Sequências</li> <li>- Funções (incluindo funções anônimas/lambda )</li> <li>- Geradores</li> <li>- Erros e tratamento de exceção</li> </ul>	
--	--	--















#### 3º Assunto

	<p>Fontes de Dados</p> <ul style="list-style-type: none"> <li>- HTTP(S)</li> </ul>	
--	--	--



	<ul style="list-style-type: none"> <li>- Sistema de Arquivos</li> <li>- Aspectos de Banco de Dados</li> </ul>	 
<b>4º Assunto</b>		
	<p>Fontes de Dados</p> <ul style="list-style-type: none"> <li>- Aspectos de Raspagem de Dados</li> <li>- API</li> </ul>	 
<b>5º Assunto</b>		
	<p>Manipulação de Dados em Alto Desempenho</p> <ul style="list-style-type: none"> <li>- Computação Vetorizada</li> </ul>	 
<b>6º Assunto</b>		
	<p>Análise de Dados em Alto Desempenho</p> <ul style="list-style-type: none"> <li>- Séries</li> <li>- DataFrame</li> <li>- Descrição dos Dados</li> </ul>	 
<b>7º Assunto</b>		
	<p>Limpeza e Preparação dos Dados</p> <ul style="list-style-type: none"> <li>- Filtro</li> <li>- Seleção</li> </ul>	 
<b>8º Assunto</b>		
	<p>Tratamento dos Dados</p> <ul style="list-style-type: none"> <li>- Aspectos de Tratamento dos Dados</li> <li>- Junção, Mescla e Concatenação</li> <li>- Pivoteamento</li> </ul>	 
<b>9º Assunto</b>		
	<p>Visualização dos Dados</p> <ul style="list-style-type: none"> <li>- Aspectos de Visualização dos Dados</li> </ul>	



	<ul style="list-style-type: none"> <li>- Principais Gráficos</li> <li>- Linhas, Barras, Histogramas</li> <li>- Densidade e Dispersão</li> </ul>	  
<b>10º Assunto</b>		
	<p>Agregação dos Dados</p> <ul style="list-style-type: none"> <li>- Aspectos de Agrupamento</li> <li>- Dividir para Conquistar</li> </ul>	 
<b>11º Assunto</b>		
	<p>Operações em Dados Agregados</p> <ul style="list-style-type: none"> <li>- Aspectos de Amostragem Aleatória e Permutações</li> <li>- Média Ponderada de Grupos e Correlação</li> <li>- Aspectos de Regressão Linear</li> </ul>	
<b>12º Assunto</b>		
	<p>Desenvolvimento Colaborativo</p>	 
	<p>Desenvolvimento Web</p>	 
	<p>Desenvolvimento Embarcado em Inteligência de Negócio</p>	
		  
<b>Projeto Final</b>		
	Atividade Prática Supervisionada	
	Apresentação dos Trabalhos	

## Bibliografia Básica:

McKinney W (2022). Python for data analysis, Third Edition. O'Reilly Media, Inc.



Mohammad Khorasani, Mohamed Abdou, Javier Hernández Fernández (2022). Web Application Development with Streamlit: Develop and Deploy Secure and Scalable Web Applications to the Cloud Using a Pure Python Framework. Apress.

---

**Bibliografia Complementar:**

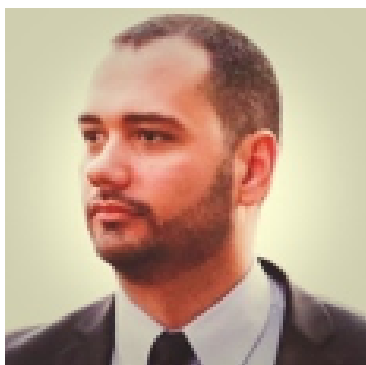
---

John V. Guttag (2021). Introduction to Computation and Programming Using Python: With Application to Computational Modeling and Understanding Data. The MIT Press

---

**Currículo resumido do docente (com foto):**

---

**Hélio Bomfim de Macêdo Filho**

Bacharel em Ciência da Computação (2007) pela Universidade Federal do Ceará (UFC). Mestre e Doutor em Engenharia de Sistemas e Computação (2010 e 2014) pela Universidade Federal do Rio de Janeiro (UFRJ). Atualmente, é servidor público na Anvisa e atuou como Analista Administrativo (2015 - 2016), Coordenador de Segurança Digital (2016 - 2019) e Assessor da Gerência-Geral de Tecnologia de Produtos para Saúde (2019 - Atual).