**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>

**Ìnicio das aulas:** 02 outubro

**Término das aulas:** 31 outubro

**Demoday:** 01 novembro

**Horário:** 9h às 17h

**Carga Horária:** 150 horas (132 horas síncronas e 18 horas assíncronas)

**Docente:** Hélio Bomfim de Macêdo Filho, Thais Salzer Procopio

**Monitores:** Bruno Garcia, Fábio Paim, Kalina Rabbani e Ricardo de Lima

**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:** |

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:** |

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 ℹ️, 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 ℹ️, 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 ℹ️;
* 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** | |
|  | Aspectos de Ciência de Dados 👨🏽‍🔬  - Conhecimentos e Habilidades do Cientista de Dados 👩‍💻  Introdução aos Ambientes de Desenvolvimento 🔬  - Notebooks 📒  Aspectos Nativos de Python 🐍(Básico)  - Tipos de Dados 🎲 Escalares 🪜  - Controle 🕹️ de Fluxo 🔄  Logo do Python |
| **2º Assunto** | |
|  | Aspectos Nativos de Python 🐍(Intermediário)  - Estruturas de Dados 🏛️ 🎲 e Sequências 🎰  - Funções🧑🏽‍🔧 (incluindo funções anônimas/lambda 🥷🏽)  - Geradores 🔂  - Erros 🚫 e tratamento de exceção ✅  Logo do Python |
| **3º Assunto** | |
|  | Fontes de Dados 🚰 🎲  - HTTP(S) 🕸️  - Sistema de Arquivos 🗂️  - Aspectos de Banco de Dados 🏦🎲  Logo do Python |
| **4º Assunto** | |
|  | Fontes de Dados 🚰 🎲  - Aspectos de Raspagem de Dados ✂️  - API📝  Logo do Python |
| **5º Assunto** | |
|  | Manipulação de Dados 🪡 🎲 em Alto Desempenho 🚀 🧞‍♂️  - Computação Vetorizada 🧮  NumPy logo 2020.svg |
| **6º Assunto** | |
|  | Análise de Dados 🕵🏽 🎲 em Alto Desempenho 🚀 🧞‍♂️  - Séries 🎞️  - DataFrame 📽️  - Descrição dos Dados ✍️ 🎲  Pandas logo.svg |
| **7º Assunto** | |
|  | Limpeza e Preparação dos Dados 🏗️ 🎲  - Filtro 🍝  - Seleção 🍕  Pandas logo.svg |
| **8º Assunto** | |
|  | Tratamento dos Dados 🪢 🎲  - Aspectos de Tratamento dos Dados 🪢 🎲  - Junção, Mescla e Concatenação 🥇  - Pivoteamento 🥈  Pandas logo.svg |
| **9º Assunto** | |
|  | Visualização dos Dados 🔍 🎲  - Aspectos de Visualização dos Dados 🔍 🎲  - Principais Gráficos  - Linhas 📈, Barras 📊, Histogramas 📶  - Densidade 🛎️ e Dispersão ⛬ |
| **10º Assunto** | |
|  | Agregação dos Dados 🎒 🎲  - Aspectos de Agrupamento  - Dividir ➗ para Conquistar ➕  Pandas logo.svg |
| **11º Assunto** | |
|  | Operações em Dados Agregados 🩺🧑‍⚕️🎲  - Aspectos de Amostragem Aleatória e Permutações 🔀  - Média Ponderada de Grupos e Correlação ☄️  - Aspectos de Regressão Linear 🔥  Pandas logo.svg |
| **12º Assunto** | |
|  | Desenvolvimento Colaborativo 🤝  Git-logo.svg  Desenvolvimento Web 🕸️ 💻  Streamlit documentation  Desenvolvimento Embarcado em Inteligência de Negócio 🧠  Pandas logo.svg |
| **Projeto Final** | |
|  | 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 dos docentes (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). |
|  | **Thais Salzer Procopio**  Bacharel em Ciências Econômicas (2010) pela Universidade Federal de Juiz de Fora (UFJF). Mestre em Economia Aplicada pela UFJF (2013). Atualmente é Auditora Federal de Finanças e Controle, em exercício na Secretaria de Reformas Econômicas (SRE). Atuou como Assessora Econômica na Secretaria do Tesouro Nacional (STN) (2015-2019) e Chefe de Divisão de Economia e Justiça na Secretaria de Política Econômica (SPE) (2019-2021). |