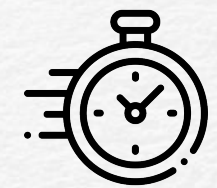


Profissão: Cientista de Dados

JORNADA DE APRENDIZAGEM



Carga horária total: 70 horas



**VAMOS
COMEÇAR!**

1

FASE 1

Python para Análise de Dados

- Variáveis & Tipos de Dados
- Estruturas de Dados
- Fluxo Condicional & Repetição
- Arquivos & Funções
- Programação Funcional
- Programação Orientada a Objetos
- Módulos e Pacotes
- Tratamento de Erros
- Scripting

Desenvolvimento em Ciência de Dados e Metodologia Crisp-DM

- Perspectivas de Carreira
- Metodologia
- Elementos básicos de Python e Numpy
- Introdução ao Pandas e carga de dados
- Limpeza e preparação de dados
- Descritiva I
- Árvores I
- Git/Github - Controle de versionamento

FASE 2

2

3

FASE 3

Desenvolvimento de aplicações com Python

- Agregações e operações em grupos
- Descritiva II
- Árvores de regressão
- Regressão I e II
- Scripting
- Streamlit I

BY
UIC

Profissão: Cientista de Dados

JORNADA DE APRENDIZAGEM

 Carga horária total: 70 horas

Desenvolvimento de Modelos com Pandas e Python

- Métodos de análise
- Árvores II
- Regressão Logística I
- Streamlit II
- Markdown/ Terminal

FASE 4

4

5

FASE 5

Gestão avançada de Dados e Desenvolvimento de aplicações

- PostgreSQL
- Descritiva III
- Combinação de modelos
- Cálculo
- Algelin
- PCA
- Streamlit III

Agrupamento e segmentação de dados

- K-means
- Hierárquicos/aglomerativos
- Streamlit IV

FASE 6

6



BY
UIC

Profissão: Cientista de Dados

JORNADA DE APRENDIZAGEM



Carga horária total: 70 horas

7

FASE 7

Modelagem estatística e Regressão com Python

- Probabilidade
- Inferência
- Regressão III e IV

Modelagem de dados categorizados

- Descritiva IV
- Regressão Logística II
- Streamlit V e Pycaret

FASE 8

8

9

FASE 9

SQL para análise de dados

- Base de dados e Linguagem SQL
- Tabelas em SQL
- Seleção e ordenação em SQL
- Filtragem e seleção condicional
- Agregações
- Múltiplas tabelas
- Técnicas avançadas



BY
UIC

Profissão: Cientista de Dados

JORNADA DE APRENDIZAGEM

Competência 4

Proficiência em análise avançada de dados com Pandas em Python, otimizando códigos e personalizando projetos. Desenvolvimento de modelos de classificação multinomial com árvores de decisão e validação cruzada. Aplicar regressão logística para análise e previsão. Criar relatórios interativos no Streamlit, integrando manipulação de arquivos e filtros. Manipular arquivos no Linux e criar documentos profissionais com Markdown e LaTeX no Jupyter Notebook.



Habilidade 1

Otimizar códigos reduzindo variáveis intermediárias. Utilizar for, list comprehension, .apply(), e tqdm para análise.

Habilidade 2

Aplicar classificação multinomial. Implementar cross-validation. Resolver problemas de classificação com dados de sensores.

Habilidade 3

Interpretar e aplicar regressão logística e utilizá-la para prever probabilidades. Avaliar a qualidade de modelos de regressão logística.

Habilidade 4

Utilizar Streamlit para otimizar leitura de arquivos. Manipular funções para gerenciar arquivos.

Habilidade 5

Aplicar comandos Linux para manipular arquivos. Criar documentos com Markdown e LaTeX no Jupyter Notebook.

FASE 4

