

PLANO DE ENSINO

Programa de Formação em Análise de Dados

Formação completa em Dados com foco em Data Science

- **Aulas:**

1 aula (ao vivo) de 3.5h por semana 13h de conteúdo assíncrono, com 2 workshops em sextas feiras.

Aula das 19h às 22h30 - Terça-Feira

28 semanas - 100 horas de aula síncrono e total estimado 450 horas estimadas total

O Programa Formação em Análise de Dados é um programa de formação com foco em ciência de dados aplicado em análise. Ao final da formação o profissional estará capacitado a entender as várias frentes de trabalho na área de análise de dados, desenvolver análises exploratórias sobre os dados, exibição e apresentação de informações obtidas pelas análises, além de manutenção de armazenagem desses dados em bancos de dados.

Componentes curriculares

	Onboarding - 1 Semana(3,5 horas)
Bloco 01: Introdução à Área de Dados	<ul style="list-style-type: none">● Introdução à computação● Conceitos de dados, informação, conhecimento● Introdução a análise de dados● Subáreas da área de dados● Carreiras na área de dados
	Introdução a programação - Python I - 4 Semanas(14 horas)
	<ul style="list-style-type: none">● IDEs de desenvolvimento● Introdução a linguagem Python<ul style="list-style-type: none">○ Operadores○ Variáveis○ Tipos de Dados○ Condicionais○ Laços de repetição○ Funções○ Arquivos○ Exceções○ Bibliotecas
	Programação aplicada - Python II - 6 Semanas (21 horas)
	<ul style="list-style-type: none">● Introdução a manipulação de dados com Python e bibliotecas<ul style="list-style-type: none">○ Numpy<ul style="list-style-type: none">■ Array■ Métodos x Atributos■ Inicialização de Arrays Numpy■ Manipulação das dimensões de um array

- Slicing em Arrays
- Indexação
- Manipulação de Arrays
 - vstack/hstack
 - concatenate
 - append
- Funções Matemáticas
- Operações entre arrays
- Pandas
 - Series
 - Criação de uma series
 - Operações básicas com series
 - Filtros
 - Outras operações
 - Concatenação de series
 - Dataframes
 - Criação de um dataframe
 - Leitura de dados de csv e excel
 - Atributos de um dataframe
 - Acesso aos dados
 - Ordenação e descrição
 - Transformação de dados
 - Filtros
 - Mudança condicional
 - Export de dataframes
 - Limpeza e Manipulação de dados
 - Investigação e definição de tipos de dados
 - Análise de inconsistências
 - Dados faltantes
 - Análise de duplicatas
 - Remoção de colunas constantes
- Visualização de Dados
 - Matplotlib
 - Seaborn

Banco de Dados Relacional - 4 Semanas (14 Horas)

- Conceitos
 - Bancos de dados.
 - SGBD.
 - Banco de dados relacional
 - Propriedades ACID (Atomicidade, Consistência, Isolamento, Durabilidade).
- Linguagem SQL
 - SQL ANSI
 - DDL
 - Create
 - Alter
 - Drop
 - DML e DQL
 - Select

	<ul style="list-style-type: none"> ■ Update ■ Insert ■ Delete ○ Joins ○ Comandos avançadas <ul style="list-style-type: none"> ■ CASE ○ Views ○ Índices
Bloco 02 - Análise de Dados	Exploração de dados na internet e API - 2 Semanas (7 Horas)
	<ul style="list-style-type: none"> ● Conceito, e entendimento de plataformas online <ul style="list-style-type: none"> ○ Considerações éticas na coleta e uso de dados online. ○ Regulamentações de privacidade. ○ Fontes de dados online. ○ Coleta e análise de dados de redes sociais. ○ Integração com APIs de plataformas sociais. ○ JSON. ○ XML. ○ API <ul style="list-style-type: none"> ■ Conceitos básicos. ■ Uso de APIs para coleta de dados. ■ Exemplos práticos de integração com APIs populares.
	Projeto Prático I - 2 Semanas - (7 horas)
	<p>O projeto 1 é sobre a extração, limpeza e manipulação de dados no INEP sobre o Censo da Educação Superior.</p>
	<p>Vamos disponibilizar os dados e os alunos precisarão imputar os dados de arquivos csv para um banco de dados sql. Após isso, vão precisar realizar diversas queries de SQL para responder questões ligadas ao negócio</p>
	Workshop I - Uma introdução a IA - Foco em Machine Learning
	<p>Conteúdo extra, para trazer um overview sobre um assunto relevante da área.</p>
Matemática e Estatística Aplicadas - 4 Semanas (14 horas)	
<ul style="list-style-type: none"> ● Estatística Descritiva <ul style="list-style-type: none"> ○ Tipos de Variáveis ○ Medidas "Resumo" <ul style="list-style-type: none"> ■ Medidas de Tendência Central ■ Medidas de Dispersão ■ Quantis, quartis e percentis ○ Análise Bidimensional <ul style="list-style-type: none"> ■ Associação entre variáveis qualitativas ■ Associação entre variáveis quantitativas ■ Correlação <ul style="list-style-type: none"> ● Correlação de Spearman ● Correlação de Pearson ■ Associação entre variáveis quali e quantitativas <ul style="list-style-type: none"> ● Correlação Point-Biserial ● R2 	

- Distribuição de Probabilidade
 - Distribuições Discretas
 - Distribuição de Bernoulli
 - Distribuição Binomial
 - Distribuições Contínuas
 - Distribuição Normal
 - Distribuição Normal Padrão
- Estatística Inferencial
 - O que é inferência
 - Amostragem
 - Lei dos Grandes Números
 - Teorema do Limite Central
 - Intervalo de Confiança
 - Teste de Hipótese
 - O que é um teste de hipótese
 - Exemplo simples
 - Metodologia Genérica
 - P-valor
 - Exemplos Práticos
 - Teste para a média
 - Teste de normalidade
- Análise Exploratória de Dados (EDA)
- Regressão Linear

Visualização de Dados e Storytelling - 1 Semana (3,5 horas)

- Técnicas de apresentação de dados para análises explanatórias.
- Comunicação: história através dos dados, visualizações e escolha para cada demanda.
- Detalhes e os erros mais comuns na visualização de dados.
- Noção básica sobre acessibilidade na visualização de dados (pessoas daltônicas).

Workshop II - Uso de bibliotecas python para plotagem de dash

Foco em plotly

Introdução a visualização de dados com uso de ferramentas de BI - 2 Semanas (14 horas)

- Princípios de Design de Visualização:
 - Escolha de gráficos apropriados para diferentes tipos de dados.
- Business Intelligence (BI):
 - Definição e objetivos do BI.
 - Como o BI suporta a tomada de decisões nas organizações.
- Power BI: Introdução e Demonstração Prática:
 - Visão geral do Power BI e sua integração com o BI.
 - Demonstração prática: Criação de relatórios simples no Power BI.
 - Conectando e Importando Dados no Power BI:
- Importação de dados de diferentes fontes.
 - Transformação e limpeza de dados no Power BI.
 - Criando Dashboards Interativos:

Projeto Prático - 2 Semanas (7 horas)

Projeto onde o aluno irá usar as mesmas bases de dados do projeto 1, porém aqui o foco vai ser em visualização de dados e geração de insights.

	Parte 1: Construir um dashboard em Power BI para se fazer uma análise do ensino superior no Brasil
--	--

	Parte 2: Usar o python para gerar insights e estatísticas e apresentar (preferencialmente em um ppt) as conclusões/insights para a turma.
--	---