



Especialização Data Science Foundations

Tecnologias de Informação - Data & Analytics

- **Nível:** Intermédio
 - **Duração:** 90h
-

Sobre o curso

A quantidade de dados gerados pelas organizações tem vindo a crescer e conseqüentemente a necessidade de pessoas especializadas capazes de retirar valor desses dados. Quando devidamente explorados, permitem suportar decisões estratégicas para o negócio, criação de novos produtos, oferta de serviços mais personalizados entre muito outras potencialidades. **Data Science** é por isso uma área transversal a todos o tipos de organizações e a todos os sectores, sendo o seu valor reconhecido pelas vantagens competitivas que potencia.

A **Especialização Data Science Foundations** fornece aos alunos os principais conceitos e ferramentas usadas por Data Scientists e Data Analysts.

Destinatários

Destina-se a todos aqueles que queiram adquirir conhecimentos que lhes permitam tirar partido desta nova capacidade estratégica, dotando-os dos conhecimentos necessários para conseguirem retirar o máximo valor dos dados e com isso aportarem valor competitivo para as organizações, nomeadamente profissionais que desempenhem funções diretas de tomada de decisão ou que intervenham em processos de análise de dados para a tomada de decisão sobre os negócios da organização.

Pré-requisitos

Os alunos deverão ter:

- Conhecimentos básicos de Matemática e Estatística;
- Conhecimentos básicos de utilização de uma qualquer linguagem de programação;
- Bom nível de capacidade de leitura em língua inglesa.

Metodologia

Esta Especialização pode ser ministrada:

- em modelo híbrido (**Presencial e Live Training** – online em tempo real)
- em 100% **Live Training**.

A metodologia pedagógica está focada no saber fazer, pelo que os conceitos e teoria de base são fortemente explorados em treino orientado para a colocação dos conhecimentos em prática.

Os alunos devem ter em conta que para além da exigência das aulas há a exigência de estudo adicional e de desenvolvimento de trabalhos práticos que permitam um cabal desenvolvimento do saber fazer nesta área de conhecimento.

Programa

- **Fundamentos de Ciência de Dados (9 horas)**
 - Presencial (em Lisboa) / Online (outras localidades)
 - por: [Manuela Almeida](#), Data Scientist na Talkdesk
- **Gestão do Processo de “Data Mining” (18 horas)**
 - Presencial (em Lisboa) / Online (outras localidades)
 - por: [Adelaide Leitão](#), Business Unit Diretor na Closer Consulting
- **Aplicação da Ciência de Dados (18 horas)**
 - Presencial (em Lisboa) / Online (outras localidades)
 - por: [Rodrigo Veríssimo](#), Senior Data Scientist na Farfetch
- **Inferência Estatística (18 horas)**
 - Presencial (no Porto) / Online (outras localidades)
 - por: Inês Rosete, Research Scientist na Nimble Portal
- **Programação em Python (27 horas)**
 - Presencial (em Lisboa) / Online (outras localidades)
 - por: [Luís Silva](#), AI Engineer na OutSystems
- **Fundamentos de ciência de dados**
 - O que é Data Science (DS)
 - O papel da Data Science nos diferentes tipos de organizações
 - Estruturação de um projeto de Data Science
 - Resultados de um projeto de Data Science
 - As ferramentas básicas de um Data Scientist

- **Gestão do Processo de Data Mining**

- Etapas do processo de *data mining*
- Tipos de questões e características de boas questões, expectativas e objetivos
- Conceito de recolha de dados
- Conceito de Exploratory Data Analysis
- Conceito de inferência estatística
- Conceito de modelos preditivos
- Critério de paragem
- Comunicação de resultados

- **Aplicação da ciência de dados**

- Comparação do cenário ideal versus o cenário real
- Qualidade dos dados fonte
- Factores que afectam os resultados
- Inferência estatística versus predição
- Dimensão dos dados
- Interpretação de resultados
- Escalabilidade
- Reprodutibilidade
- Casualidade versus confusão
- A/B Testing
- Manutenção dos modelos

- **Inferência estatística**

- Funções massa e densidade em probabilidade
- Probabilidades condicionadas e teorema de Bayes
- Valores expectáveis
- Média, desvio padrão e variância
- Distribuições binomial, normal e de Poisson
- Teorema Limite Central
- Testes de hipóteses
- Significância estatística e valor P
- Pacotes NumPy e statsModels

- **Programação com Python**

- Instalação do python
- Básicos da programação com python
- Pacotes Pandas
- Jupyter notebooks