



Curso: Pós-Graduação em Data Science

Duração: 162h

Área formativa: PG's & MBA's

Sobre o curso

A **Pós-Graduação em Data Science** destina-se a todos aqueles que queiram adquirir conhecimentos que lhes permitam tirar partido desta nova capacidade estratégica, dotando-os dos conhecimentos necessários para retirar o máximo valor dos dados, dando uma visão detalhada, teórica e prática, de conceitos e metodologias.

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 muitas 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.

O programa do ciclo de estudos de Pós-Graduação é constituído por um conjunto de 10 módulos curriculares, compostos por sessões de 3 horas cada.

Esta Pós-Graduação está ainda dividida em 2 ciclos de especialização – Data Science Manager e Data Science Researcher.

A especialização de Data Science Manager visa fornecer os principais conceitos e ferramentas usadas pelos Data Scientists, assim como os princípios de constituição e orientação de equipas de Data Science bem sucedidas.

A especialização de Data Science Researcher oferece uma forte base dos princípios necessários para o desenvolvimento de um projecto de Data Science, assim como exemplos de práticas de aplicação dos mesmos.

A **Pós-Graduação em Data Science** é desenvolvida ao abrigo de um protocolo de colaboração entre a **Atlântica - Escola Universitária de Ciências Empresariais, Saúde, Tecnologias e Engenharia** e a Rumos.

Destinatários

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

As candidaturas à Pós-graduação em Data Science estão abertas a:

- Todos os que tenham um grau académico de licenciatura ou superior nas áreas científicas;
- Todos os profissionais com ou sem grau académico, cuja experiência seja considerada adequada para que o aluno tenha sucesso no curso e as turmas resultem homogéneas.

A seleção será sempre feita mediante análise curricular pela Coordenação Científica do curso, que pode chamar o candidato a uma entrevista presencial. Em qualquer dos casos, a decisão será sempre fundamentada e apresentada por escrito ao candidato.

Adicionalmente, 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

A Pós-graduação em Data Science pode ser ministrada:

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

Programa

Ciclo de especialização Data Science Manager

- Fundamentos de Ciência de Dados (9 horas)
 - Docente: [Manuela Almeida](#), Data Scientist @ Talkdesk > Presencial (em Lisboa) / Online (outras localidades)
- Gestão de Equipas de Ciência de Dados (9 horas)
 - Docente: [Pedro Lopes](#), Big Data Scientist @ Aptoide
- Gestão do Processo de “Data Mining” (18 horas)
 - Docente: [Miguel Guedes](#), Diretor of Consulting na CGI
- Aplicação da Ciência de Dados (18 horas)
 - Docente: [Filipa Peleja](#), Data Scientist na Vodafone Portugal

Ciclo de especialização Data Science Researcher

- Programação em Python (27 horas)
 - Docente: [Luís Silva](#), AI Engineer @ OutSystems
- Metodologia de Recolha de Dados (18 horas)

- Docente: [Luís Silva](#), AI Engineer @ OutSystems
- Metodologia de Pré-Processamento de Dados (18 horas)
 - Docente: [Carlos Rodrigues](#), Data Scientist & Big Data Engineer @ Marionete
- Inferência Estatística (18 horas)
 - Docente: [Luís Baía](#), Data Scientist @ Farfetch
- Análise Exploratória de Dados (18 horas)
 - Docente: [Hugo Lopes](#), Data Scientist @ James
- Fundamentos Práticos de “Machine Learning” (18 horas)
 - Docente: [Daniel Ferrari](#), Data Scientist na Talkdesk

CICLO DE ESPECIALIZAÇÃO DATA SCIENCE MANAGER

- **Fundamentos de ciência de dados (9 horas)**
(Docente: Manuela Almeida, Data Scientist @ Morphing Portals)
 - 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 de equipas de ciência de dados (9 horas)**
(Docente: Pedro Lopes, Big Data Scientist @ Aptoide)
 - O papel do *Data Science Manager*
 - Perfil dos vários membros de uma equipa de *Data Science*
 - Métricas de sucesso de um projecto de Data Science
 - Entrevistas
- **Gestão do Processo de Data Mining (18 horas)**
(Docente: Miguel Guedes, Diretor of Consulting na CGI)
 - 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 (18 horas)**
(Docente: Filipa Peleja, Data Scientist na Vodafone Portugal)
 - 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

CICLO DE ESPECIALIZAÇÃO DATA SCIENCE RESEARCHER

• Programação com Python (27 horas)

(Docente: Luís Silva, Data Scientist @ Feedzai)

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

• Metodologias de recolha de dados (9 horas)

(Docente: Luís Silva, Data Scientist @ Feedzai)

- Processo ETL (extract, transform, load)
- Conceito de data governance
- Fontes de dados: Excel, XML, MySQL, Web, APIs

• Metodologias de pré-processamento de dados (18 horas)

(Docente: Carlos Rodrigues, Data Scientist & Big Data Engineer @ Marionete)

- Tipos de dados
- Limpeza de dados
- Transformação de dados

• Inferência estatística (18 horas)

(Docente: Luís Baía, Data Scientist @ Farfetch)

- 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

• Análise exploratória de dados (18 horas)

(Docente: Hugo Lopes, Data Scientist @ James)

- Seleção de Variáveis
- Sumário estatístico dos dados
- Redução de dimensão
- Visualização para exploração dos dados

• Fundamentos práticos de “Machine Learning” (18 horas)

(Docente: Daniel Ferrari, Data Scientist na Talkdesk)

- Predição, erros e validação cruzada
- Aprendizagem Supervisionada e não supervisionada
- Modelos de Regressão
- Modelos de Classificação
- Pacote Scikit-learn