

# Curso

Arquiteturas e Sistemas  
para Utilização Intensiva  
de Dados em Data Science





## Curso

### Arquiteturas e Sistemas para Utilização Intensiva de Dados em Data Science

- » Modalidade: **online**
- » Duração: **6 semanas**
- » Certificação: **TECH Universidade Tecnológica**
- » Acreditação: **6 ECTS**
- » Horário: **ao seu próprio ritmo**
- » Exames: **online**

Acesso ao site: [www.techtute.com/pt/informatica/curso/arquiteturas-sistemas-utilizacao-intensiva-dados-data-science](http://www.techtute.com/pt/informatica/curso/arquiteturas-sistemas-utilizacao-intensiva-dados-data-science)

# Índice

01

Apresentação

---

*pág. 4*

02

Objetivos

---

*pág. 8*

03

Direção do curso

---

*pág. 12*

04

Estrutura e conteúdo

---

*pág. 18*

05

Metodologia

---

*pág. 22*

06

Certificação

---

*pág. 30*

# 01

# Apresentação

Nos anos 70, o modelo racional ganhou grande popularidade e ofereceu uma solução para a forma como a informação era representada nos sistemas de dados da altura. Desde então, foram desenvolvidas novas ferramentas e tecnologias para se adaptarem ao hardware e resolverem o problema do processamento de grandes quantidades de dados. Neste curso estudarás as tecnologias e ferramentas oferecidas pelo mercado e estabelecerá um critério de decisão para poder escolher a que melhor se adapta às necessidades da empresa.



“

*Combine processos técnicos e comerciais para utilizar as melhores ferramentas para o tratamento massivo de dados”*

O conteúdo deste Curso procurará mostrar aos alunos as ferramentas e tecnologias atualmente desenvolvidas para o processamento de dados em massa. Para o efeito, serão examinados os principais componentes de um sistema concebido para lidar com grandes volumes de dados.

No final do curso, o engenheiro informático interessado neste setor terá adquirido um critério que o ajudará a distinguir os programas que melhor se adaptam às necessidades da empresa. Os alunos serão também capazes de avaliar que aplicações amplamente utilizadas utilizam os fundamentos dos sistemas distribuídos para conceber os seus próprios programas, o que representa uma melhoria do seu perfil profissional.

Todo o curso é composto por uma série de casos práticos que apoiarão a aprendizagem dos alunos que procuram progredir nas suas carreiras e desafiar-se a si próprios para alcançar a excelência.

Este **Curso de Arquiteturas e Sistemas para Utilização Intensiva de Dados em Data Science** conta com o conteúdo educativo mais completo e atual do mercado. As suas principais características são:

- ◆ O desenvolvimento de estudos de caso apresentados por especialistas em engenharia centrada no ciclo integrado de dados
- ◆ O conteúdo gráfico, esquemático e eminentemente prático do livro fornece informações científicas e práticas sobre as disciplinas que são essenciais para a prática profissional
- ◆ Exercícios práticos onde o processo de autoavaliação pode ser levado a cabo para melhorar a aprendizagem
- ◆ A sua ênfase especial em metodologias inovadoras
- ◆ As lições teóricas, perguntas a especialistas, fóruns de discussão sobre questões controversas e atividades de reflexão individual
- ◆ A disponibilidade de acesso aos conteúdos a partir de qualquer dispositivo fixo ou portátil com ligação à Internet



*Analise a melhor forma de armazenar e recuperar informações em bases de dados maciças"*

“

*Avalie quais as aplicações amplamente utilizadas que utilizam fundamentos de sistemas distribuídos e conceba um sistema novo e inovador”*

O corpo docente do curso inclui profissionais do setor que trazem a sua experiência profissional para esta capacitação, para além de especialistas reconhecidos de sociedades de referência e universidades de prestígio.

Os seus conteúdos multimédia, desenvolvidos com a mais recente tecnologia educativa, permitirão ao profissional uma aprendizagem situada e contextual, ou seja, um ambiente simulado que proporcionará uma capacitação imersiva programada para praticar em situações reais.

A estrutura desta capacitação centra-se na Aprendizagem Baseada em Problemas, na qual o profissional deve tentar resolver as diferentes situações de prática profissional que surgem durante o Curso. Para tal, contará com a ajuda de um inovador sistema de vídeos interativos feitos por especialistas reconhecidos com vasta experiência em Arquiteturas e Sistemas para Utilização Intensiva de Dados em Data Science.

*Analisar as funcionalidades fundamentais para os sistemas distribuídos e a sua importância em diferentes tipos de sistemas.*

*Aprenda, através de um curso 100% online, a identificar formatos de codificação de dados.*



# 02

## Objetivos

O conteúdo deste curso ajudará os engenheiros informáticos a avaliar as ferramentas de análise de dados existentes no mercado para escolher uma que seja do interesse de uma empresa ou, em alternativa, desenvolver uma que funcione para atingir esse objetivo. Para o efeito, a TECH estabeleceu os seguintes objetivos gerais e específicos.





“

*Este curso não só o ajudará a melhorar o seu perfil académico, como também o ajudará a desenvolver o seu pensamento crítico”*



## Objetivos gerais

---

- ◆ Analisar os benefícios da aplicação de técnicas de análise de dados em cada departamento da empresa.
- ◆ Desenvolver as bases para a compreensão das necessidades e aplicações de cada departamento.
- ◆ Gerar conhecimentos especializados para selecionar a ferramenta adequada
- ◆ Propor técnicas e objetivos para serem tão produtivos quanto possível, de acordo com o departamento.





## Objetivos específicos

---

- ◆ Analisar como as bases de dados armazenam e recuperam informação
- ◆ Concretizar os diferentes modelos de replicação e os problemas associados
- ◆ Desenvolver as formas de partição e as transações distribuídas
- ◆ Determinar os sistemas por lotes e os sistemas (quase) em tempo real

“

*Faça deste curso a melhor oportunidade para desenvolver novas formas de pensar que irão gerar um novo sistema de utilização de dados”*

# 03

## Direção do curso

O Curso de Arquiteturas e Sistemas para Utilização Intensiva de Dados em Data Science reúne um grupo selecionado de profissionais com vários anos de experiência em análise de dados no setor empresarial. Desta forma, os conhecimentos transmitidos são-no por profissionais capazes de responder a todas as questões que os alunos possam ter e de lhes apresentar casos reais para melhor exemplificar os conteúdos do curso.



“

*Este Curso é apoiado pela vasta experiência profissional de um excelente corpo docente”*

## Diretor Internacional Convidado

O Doutor Tom Flowerdew é uma figura destacada internacionalmente no campo da ciência de dados. Assim, tem exercido o cargo de **Vice-Presidente de Ciência de Dados** na **MasterCard**, em **Londres**. Neste papel, tem sido responsável pela preparação, operação e estratégia de uma equipa consolidada nesta área, com a missão de apoiar um portfólio de **produtos inovadores em pagamentos, combater a lavagem de dinheiro (AML)** e analisar casos de uso de **criptomoedas**.

Além disso, foi **Diretor de Ciência de Dados em Soluções de Ciberinteligência**, também na **MasterCard**, onde liderou a integração de dados para apoiar produtos revolucionários baseados em **criptomoedas**. De fato, sua capacidade de lidar com **dados** complexos e desenvolver **soluções avançadas** tem sido fundamental para o sucesso de múltiplos projetos na área da **cibersegurança** e das **finanças**.

Igualmente, para a empresa **Featurespace**, ocupou vários papéis cruciais, incluindo o de **Chefe de Entrega de Produtos Estandarizados**, em **Cambridge**, liderando uma equipa e um projeto de transformação que reduziu o tempo e esforço de entrega em mais de 75%. Além disso, como **Diretor de Entrega**, na sede dos **Estados Unidos**, gerenciou todas as funções de entrega da empresa na **América do Norte**, melhorando significativamente a **eficiência operacional** e fortalecendo as relações com os **clientes**.

Adicionalmente, o Doutor Tom Flowerdew demonstrou sua habilidade em construir e liderar equipes de alto desempenho ao longo de sua carreira, destacando seu papel como **Cientista de Dados**, tanto em **Atlanta**, onde recrutou e gerenciou um grupo de especialistas na área, quanto em **Cambridge**. Assim, seu foco em **inovação** e **resolução de problemas** deixou uma marca indelével nas organizações onde trabalhou, consolidando-se como um **líder influente** no campo da **ciência de dados**.



## Dr. Flowerdew, Tom

---

- Vice-Presidente de Ciência de Dados na MasterCard, Londres, Reino Unido
- Diretor de Ciência de Dados em Soluções de Ciberinteligência na MasterCard, Londres
- Chefe de Entrega de Produtos Estandarizados na Featurespace, Cambridge
- Diretor de Entrega para os Estados Unidos na Featurespace, Cambridge
- Cientista de Dados na Featurespace, Atlanta, Geórgia, Estados Unidos
- Cientista de Dados na Featurespace, Cambridge
- Pesquisador em Estatística e Pesquisa Operacional na Universidade de Lancaster
- Doutorado em Pesquisa de Operações pela Universidade de Lancaster
- Graduado em Engenharia de Sistemas pela BAE Systems
- Licenciado em Matemáticas pela Universidade de York



*Graças à TECH, poderá aprender com os melhores profissionais do mundo”*

## Direção



### Doutor Arturo Peralta Martín-Palomino

- ♦ CEO e CTO na Prometheus Global Solutions
- ♦ CTO na Korporate Technologies
- ♦ CTO na AI Shephers GmbH
- ♦ Doutoramento em Engenharia Informática pela Universidade de Castilla la Mancha
- ♦ Doutoramento em Economia, Negócios e Finanças pela Universidade Camilo José Cela Prémio Extraordinário de Doutoramento
- ♦ Doutoramento em Psicologia pela Universidade de Castilla la Mancha
- ♦ Mestrado em Tecnologias Avançadas de Informação pela Universidade de Castilla la Mancha
- ♦ Mestrado MBA+E (Mestrado em Administração de Empresas e Engenharia Organizacional) pela Universidade de Castilla la Mancha
- ♦ Professor associado docente em cursos de licenciatura e mestrado em Engenharia Informática na Universidade de Castilla la Mancha
- ♦ Professor do Mestrado em Big Data e Data Science na Universidade Internacional de Valencia
- ♦ Professor do Mestrado em Indústria 4.0 e do Mestrado em Design Industrial e Desenvolvimento de Produto
- ♦ Membro do Grupo de Investigação SMILe da Universidade de Castilla la Mancha

## Professores

### Sr. Luis Javier Peris Morillo

- ◆ Technical Lead em Capitole Consulting. Lidera uma equipa na Inditex na unidade logística da sua plataforma aberta
- ◆ Senior Technical Lead e Delivery Lead Support em HCL
- ◆ Agile Coach e Director de Operações da Mirai Advisory
- ◆ Membro do Comité de direção em qualidade como Diretor de Operações
- ◆ Desenvolvedor, Team Lead, Scrum Master, Agile Coach, Product Manager em DocPath
- ◆ Engenharia Superior em Informática pela ESI de Ciudad Real (UCLM)
- ◆ Pós-graduação em Gestão de projetos pela CEOE - Confederação Espanhola de Organizações Empresariais
- ◆ +50 MOOCs concluídos, ensinados por universidades de renome como a Universidade de Stanford, Universidade de Michigan, Universidade de Yonsei, Universidade Politécnica de Madrid, etc.
- ◆ Várias certificações, algumas das mais notáveis ou recentes são Azure Fundamentals

# 04

## Estrutura e conteúdo

Os módulos deste curso têm como objetivo fornecer uma perspectiva teórica e prática que ajudará os alunos a conhecer as tecnologias e ferramentas existentes no mercado, examinando os diferentes componentes dos sistemas que foram concebidos para trabalhar com grandes volumes de dados. Esta será uma excelente oportunidade para avaliar a utilização destes programas e, em seguida, conceber o seu próprio. Desta forma, são cumpridos os objetivos do curso de capacitar engenheiros profissionais, íntegros e de grande prestígio.



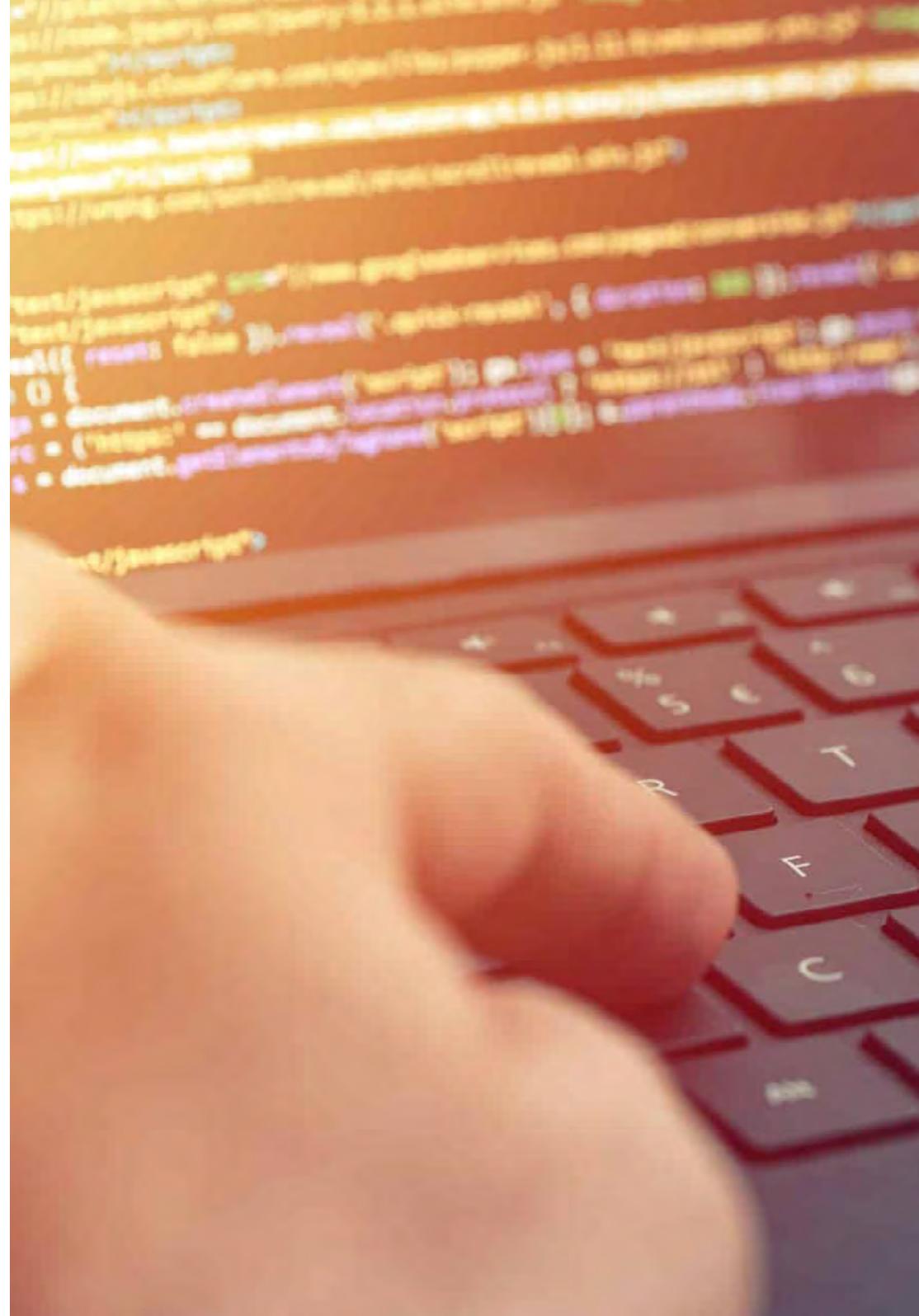


“

*Com este Curso, descobrirá uma vasta gama de possibilidades ao conceber o seu próprio sistema”*

## Módulo 1. Arquiteturas e sistemas para uso intensivo de dados

- 1.1. Requisitos não funcionais Pilares das aplicações de dados massivos
  - 1.1.1. Fiabilidade
  - 1.1.2. Adaptabilidade
  - 1.1.3. Capacidade de manutenção
- 1.2. Modelos de dados
  - 1.2.1. Modelo relacional
  - 1.2.2. Modelo documental
  - 1.2.3. Modelo de dados de rede
- 1.3. Bases de dados Gestão do Armazenamento e recuperação de dados
  - 1.3.1. Índices hash
  - 1.3.2. Armazenamento estruturado em log
  - 1.3.3. Árvores B
- 1.4. Formatos de codificação de dados
  - 1.4.1. Formatos específicos da linguagem
  - 1.4.2. Formatos estandardizados
  - 1.4.3. Formatos de codificação binários
  - 1.4.4. Fluxo de dados entre processos
- 1.5. Replicação
  - 1.5.1. Objetivos da Replicação
  - 1.5.2. Modelos de replicação
  - 1.5.3. Problemas com a Replicação
- 1.6. Transações distribuídas
  - 1.6.1. Transação
  - 1.6.2. Protocolos para transações distribuídas
  - 1.6.3. Transações serializáveis



- 1.7. Particionamento
  - 1.7.1. Formas de particionamento
  - 1.7.2. Interação de índice secundário e particionado
  - 1.7.3. Reequilíbrio de partições
- 1.8. Processamento de dados offline
  - 1.8.1. Processamento por lotes
  - 1.8.2. Sistemas de ficheiros distribuídos
  - 1.8.3. MapReduce
- 1.9. Processamento de dados em tempo real
  - 1.9.1. Tipos de Broker de mensagens
  - 1.9.2. Representação de bases de dados como fluxos de dados
  - 1.9.3. Processamento de fluxos de dados
- 1.10. Aplicações práticas na empresa
  - 1.10.1. Consistência nas leituras
  - 1.10.2. Abordagem holística dos dados
  - 1.10.3. Escala de um serviço distribuído

“

*Com este curso, atingirá todos os seus objetivos profissionais e pessoais num campo com grande procura internacional”*

# 05 Metodologia

Este programa de capacitação oferece uma forma diferente de aprendizagem. A nossa metodologia é desenvolvida através de um modo de aprendizagem cíclico: **o Relearning**. Este sistema de ensino é utilizado, por exemplo, nas escolas médicas mais prestigiadas do mundo e tem sido considerado um dos mais eficazes pelas principais publicações, tais como a ***New England Journal of Medicine***.



“

*Descubra o Relearning, um sistema que abandona a aprendizagem linear convencional para o levar através de sistemas de ensino cíclicos: uma forma de aprendizagem que provou ser extremamente eficaz, especialmente em disciplinas que requerem memorização”*

## Estudo de Caso para contextualizar todo o conteúdo

O nosso programa oferece um método revolucionário de desenvolvimento de competências e conhecimentos. O nosso objetivo é reforçar as competências num contexto de mudança, competitivo e altamente exigente.

“

*Com a TECH pode experimentar uma forma de aprendizagem que abala as fundações das universidades tradicionais de todo o mundo”*



*Terá acesso a um sistema de aprendizagem baseado na repetição, com ensino natural e progressivo ao longo de todo o programa de estudos.*



## Um método de aprendizagem inovador e diferente

Este programa da TECH é um programa de ensino intensivo, criado de raiz, que propõe os desafios e decisões mais exigentes neste campo, tanto a nível nacional como internacional. Graças a esta metodologia, o crescimento pessoal e profissional é impulsionado, dando um passo decisivo para o sucesso. O método do caso, a técnica que constitui a base deste conteúdo, assegura que a realidade económica, social e profissional mais atual é seguida.



*O nosso programa prepara-o para enfrentar novos desafios em ambientes incertos e alcançar o sucesso na sua carreira”*

*O estudante aprenderá, através de atividades de colaboração e casos reais, a resolução de situações complexas em ambientes empresariais reais.*

O método do caso tem sido o sistema de aprendizagem mais amplamente utilizado nas principais escolas de informática do mundo desde que existem. Desenvolvido em 1912 para que os estudantes de direito não só aprendessem o direito com base no conteúdo teórico, o método do caso consistia em apresentar-lhes situações verdadeiramente complexas, a fim de tomarem decisões informadas e valorizarem juízos sobre a forma de as resolver. Em 1924 foi estabelecido como um método de ensino padrão em Harvard.

Numa dada situação, o que deve fazer um profissional? Esta é a questão que enfrentamos no método do caso, um método de aprendizagem orientado para a ação. Ao longo do programa, os estudantes serão confrontados com múltiplos casos da vida real. Terão de integrar todo o seu conhecimento, investigar, argumentar e defender as suas ideias e decisões.

## Relearning Methodology

A TECH combina eficazmente a metodologia do Estudo de Caso com um sistema de aprendizagem 100% online baseado na repetição, que combina elementos didáticos diferentes em cada lição.

Melhoramos o Estudo de Caso com o melhor método de ensino 100% online: o Relearning.

*Em 2019 obtivemos os melhores resultados de aprendizagem de todas as universidades online do mundo.*

Na TECH aprende- com uma metodologia de vanguarda concebida para formar os gestores do futuro. Este método, na vanguarda da pedagogia mundial, chama-se Relearning.

A nossa universidade é a única universidade de língua espanhola licenciada para utilizar este método de sucesso. Em 2019, conseguimos melhorar os níveis globais de satisfação dos nossos estudantes (qualidade de ensino, qualidade dos materiais, estrutura dos cursos, objetivos...) no que diz respeito aos indicadores da melhor universidade online do mundo.



No nosso programa, a aprendizagem não é um processo linear, mas acontece numa espiral (aprender, desaprender, esquecer e reaprender). Portanto, cada um destes elementos é combinado de forma concêntrica. Esta metodologia formou mais de 650.000 licenciados com sucesso sem precedentes em áreas tão diversas como a bioquímica, genética, cirurgia, direito internacional, capacidades de gestão, ciência do desporto, filosofia, direito, engenharia, jornalismo, história, mercados e instrumentos financeiros. Tudo isto num ambiente altamente exigente, com um corpo estudantil universitário com um elevado perfil socioeconómico e uma idade média de 43,5 anos.

*O Relearning permitir-lhe-á aprender com menos esforço e mais desempenho, envolvendo-o mais na sua capacitação, desenvolvendo um espírito crítico, defendendo argumentos e opiniões contrastantes: uma equação direta ao sucesso.*

A partir das últimas provas científicas no campo da neurociência, não só sabemos como organizar informação, ideias, imagens e memórias, mas sabemos que o lugar e o contexto em que aprendemos algo é fundamental para a nossa capacidade de o recordar e armazenar no hipocampo, para o reter na nossa memória a longo prazo.

Desta forma, e no que se chama Neurocognitive context-dependent e-learning, os diferentes elementos do nosso programa estão ligados ao contexto em que o participante desenvolve a sua prática profissional.



Este programa oferece o melhor material educativo, cuidadosamente preparado para profissionais:



#### Material de estudo

Todos os conteúdos didáticos são criados pelos especialistas que irão ensinar o curso, especificamente para o curso, para que o desenvolvimento didático seja realmente específico e concreto.

Estes conteúdos são depois aplicados ao formato audiovisual, para criar o método de trabalho online da TECH. Tudo isto, com as mais recentes técnicas que oferecem peças de alta-qualidade em cada um dos materiais que são colocados à disposição do aluno.



#### Masterclasses

Existem provas científicas sobre a utilidade da observação por terceiros especializada.

O denominado Learning from an Expert constrói conhecimento e memória, e gera confiança em futuras decisões difíceis.



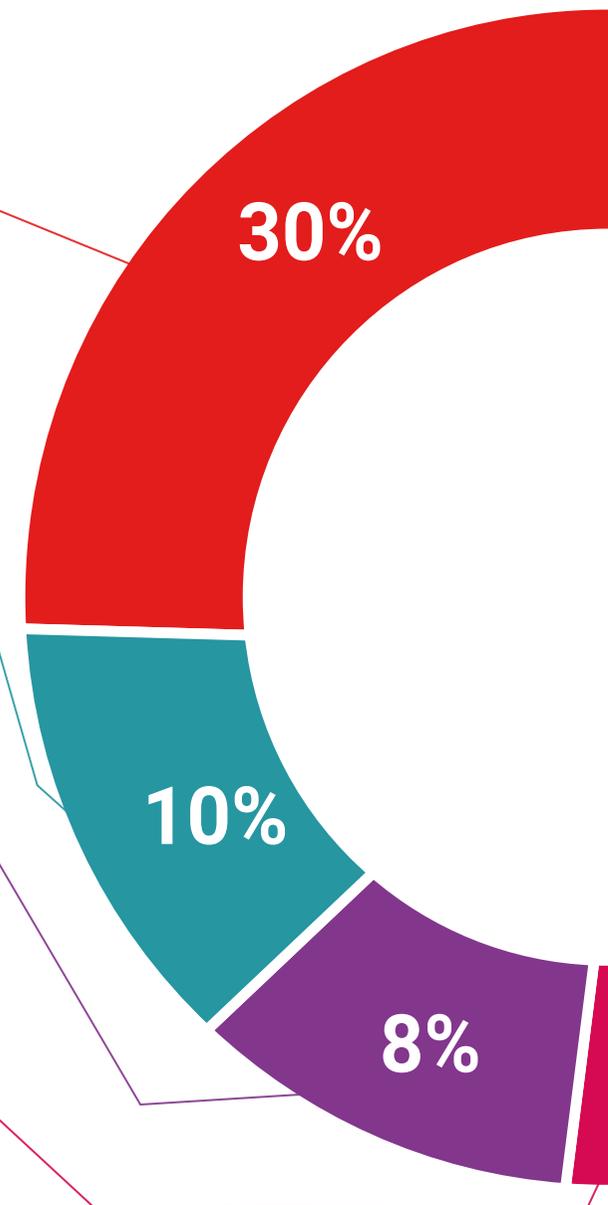
#### Práticas de aptidões e competências

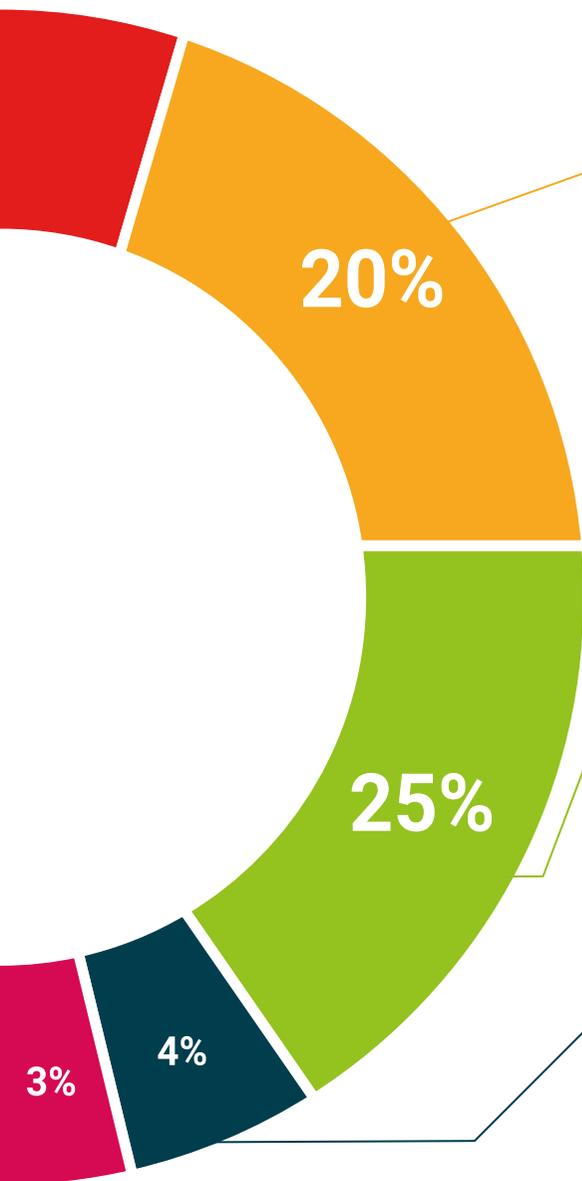
Realizarão atividades para desenvolver competências e aptidões específicas em cada área temática. Práticas e dinâmicas para adquirir e desenvolver as competências e capacidades que um especialista necessita de desenvolver no quadro da globalização em que vivemos.



#### Leituras complementares

Artigos recentes, documentos de consenso e diretrizes internacionais, entre outros. Na biblioteca virtual da TECH o aluno terá acesso a tudo o que necessita para completar a sua capacitação.





#### Case studies

Completarão uma seleção dos melhores estudos de casos escolhidos especificamente para esta situação. Casos apresentados, analisados e instruídos pelos melhores especialistas na cena internacional.



#### Resumos interativos

A equipa da TECH apresenta os conteúdos de uma forma atrativa e dinâmica em comprimidos multimédia que incluem áudios, vídeos, imagens, diagramas e mapas conceituais a fim de reforçar o conhecimento.

Este sistema educativo único para a apresentação de conteúdos multimédia foi premiado pela Microsoft como uma "História de Sucesso Europeu".



#### Testing & Retesting

Os conhecimentos do aluno são periodicamente avaliados e reavaliados ao longo de todo o programa, através de atividades e exercícios de avaliação e auto-avaliação, para que o aluno possa verificar como está a atingir os seus objetivos.



06

# Certificação

O Curso de Arquiteturas e Sistemas para Utilização Intensiva de Dados em Data Science garante, para além do conteúdo mais rigoroso e atualizado, o acesso a um certificado de Curso emitido pela TECH Universidade Tecnológica.



“

*Conclua este plano de estudos com sucesso e receba o seu certificado sem sair de casa e sem burocracias”*

Este **Curso de Arquiteturas e Sistemas para Utilização Intensiva de Dados em Data Science** conta com o conteúdo educacional mais completo e atualizado do mercado.

Uma vez aprovadas as avaliações, o aluno receberá por correio, com aviso de receção, o certificado\* correspondente ao título de **Curso** emitido pela **TECH Universidade Tecnológica**.

O certificado emitido pela TECH Universidade Tecnológica expressará a qualificação obtida no Mestrado Próprio, atendendo aos requisitos normalmente exigidos pelas bolsas de emprego, concursos públicos e avaliação de carreiras profissionais.

Certificação: **Curso de Arquiteturas e Sistemas para Utilização Intensiva de Dados em Data Science**

Modalidade: **online**

Duração: **6 semanas**

ECTS: **6**



\*Apostila de Haia: Caso o aluno solicite que o seu certificado seja apostilado, a TECH Universidade Tecnológica providenciará a obtenção do mesmo a um custo adicional.



## Curso

Arquiteturas e Sistemas  
para Utilização Intensiva  
de Dados em Data Science

- » Modalidade: online
- » Duração: 6 semanas
- » Certificação: TECH Universidade Tecnológica
- » Acreditação: 6 ECTS
- » Horário: ao seu próprio ritmo
- » Exames: online

# Curso

Arquiteturas e Sistemas  
para Utilização Intensiva  
de Dados em Data Science