



Certified Data Science Practitioner (CDSP™)

CertNexus

- **Nível:** Intermédio
- **Duração:** 30h

Sobre o curso

Os dados são um dos ativos mais valiosos que as organizações possuem. Quando devidamente tratados, o conhecimento obtido através da sua análise pode reduzir custos, aumentar a produtividade e orientar os objetivos futuros.

Para que um negócio prospere no mundo atual, impulsionado por dados, é essencial tratá-los como um dos seus ativos mais importantes. Os dados são fundamentais para compreender onde a organização está e para onde se dirige. Para além de revelar insights, os dados também orientam decisões e influenciam as operações do dia a dia.

Isto exige uma equipa robusta de profissionais capazes de analisar, interpretar, manipular e apresentar dados dentro de um quadro de processos eficaz. Este curso irá capacitá-lo a gerar valor para a organização, aplicando conceitos de ciência de dados na prática.

Exame

- **Número de questões:** 90, das quais 75 contam para a pontuação final.
- **Pontuação mínima para aprovação:** 72% ou 69%, dependendo da versão do exame (as versões são estatisticamente equivalentes).
- **Duração:** 120 minutos (*Nota: O tempo publicado inclui os 10 minutos destinados à leitura e assinatura do Acordo do Candidato e à revisão das instruções do exame*).
- **Opções de exame:** Online através da Pearson OnVUE ou presencialmente nos centros de exames Pearson VUE.
- **Formato das questões:** Escolha múltipla / Resposta única.

Destinatários

Este curso foi desenvolvido para profissionais que utilizam dados para resolver desafios empresariais.

O formando deverá ter vários anos de experiência em tecnologias, incluindo alguma aptidão para programação.

No entanto, este curso não se destina a um único perfil. Os participantes podem ser:

- Programadores que procuram expandir os seus conhecimentos sobre como tomar decisões empresariais através da recolha, organização, análise e manipulação de dados com código.
 - Analistas de dados com formação em matemática aplicada e estatística que pretendam elevar as suas competências para um nível mais avançado.
 - Outros profissionais que trabalhem com dados e queiram extrair insights de forma mais eficaz, aplicando esse conhecimento para resolver problemas de negócio e gerar maior valor para a organização.
 - Além disso, este curso foi concebido para ajudar os formandos a prepararem-se para a certificação **CertNexus® Certified Data Science Practitioner™ (CDSP) (Exame DSP-210)**.
-

Objetivos

Neste curso, irá implementar técnicas de ciência de dados para alcançar objetivos organizacionais:

- Aplicar princípios de ciência de dados para resolver desafios de negócio.
 - Utilizar o processo de Extract, Transform, and Load (ETL) para preparar conjuntos de dados.
 - Aplicar diversas técnicas para analisar dados e extrair insights valiosos.
 - Desenvolver uma abordagem de Machine Learning para responder a desafios empresariais.
 - Treinar, ajustar e avaliar modelos de classificação.
 - Treinar, ajustar e avaliar modelos de regressão e previsão.
 - Treinar, ajustar e avaliar modelos de clustering.
 - Concluir um projeto de ciência de dados, apresentando modelos a uma audiência, colocando-os em produção e monitorizando o seu desempenho.
-

Pré-requisitos

- Compreensão geral dos conceitos fundamentais de ciência de dados, incluindo, entre outros: tipos de dados, funções dentro da área de Data Science, o ciclo de vida dos dados e os benefícios e desafios associados.
 - É recomendável ter experiência com linguagens de programação de alto nível, como Python. Estar confortável com bibliotecas fundamentais de Data Science em Python, como NumPy e pandas, será uma mais valia.
 - É importante ter experiência com bases de dados, incluindo a utilização de linguagens de consulta como SQL
-

Programa

- Abordar Desafios Empresariais com Data Science
- Extração, Transformação e Carregamento de Dados (ETL)

- Análise de Dados
- Desenvolvimento de uma Abordagem de Machine Learning
- Desenvolvimento de Modelos de Classificação
- Desenvolvimento de Modelos de Regressão
- Desenvolvimento de Modelos de Clustering
- Conclusão de um Projeto de Data Science

Abordar Desafios Empresariais com Data Science

- Iniciar um projeto de Data Science
- Democratizar o acesso aos dados
- Formular um problema de Data Science

Extração, Transformação e Carregamento de Dados (ETL)

- Extrair dados
- Transformar dados
- Carregar dados

Análise de Dados

- Examinar dados
- Explorar a distribuição subjacente dos dados
- Utilizar visualizações para analisar dados
- Pré-processar dados

Desenvolvimento de uma Abordagem de Machine Learning

- Identificar conceitos de Machine Learning
- Identificar conceitos de Deep Learning baseados em Transformers
- Testar uma hipótese

Desenvolvimento de Modelos de Classificação

- Treinar e ajustar modelos de classificação
- Avaliar modelos de classificação

Desenvolvimento de Modelos de Regressão

- Treinar e ajustar modelos de regressão
- Avaliar modelos de regressão

Desenvolvimento de Modelos de Clustering

- Treinar e ajustar modelos de clustering
- Avaliar modelos de clustering

Conclusão de um Projeto de Data Science

- Comunicar resultados a stakeholders

- Demonstrar modelos numa aplicação web
- Implementar e testar pipelines de produção

Apêndice A: Mapeamento do conteúdo do curso para a certificação CertNexus® Certified Data Science Practitioner™ (CDSP) (Exame DSP-210)