AWARI. - Intensivo de Data Science com mentoria individual

**02**

**Visão Geral**

**Tarefa: Responde ao Quiz**

**Quiz**

1. **Defina com suas palavras o que é Data Science.**

*É uma ciência que estuda os dados para transformar em informações importantes ao negócio, através de um ou mais processos de captura, transformação, geração e análise de dados.*

1. **Descreva resumidamente as principais etapas de um projeto de Data Science.**

*1. Aquisição de Dados*

*Definição das fontes de dados e forma de extração das fontes;*

*2. Limpeza de Dados*

*Processos de extração com qualidade dos dados, retirada de inconsistência, reformatar valores;*

*3. Análise Exploratória de Dados*

*Verificar os conjuntos de dados para resumir as principais características dessa forma definimos suposições e modelos para produzir resultados de análise.*

*4. Modelagem de Dados*

*São utilizadas técnicas de aprendizado de máquina aos dados para classificar e prever valores para identificar o melhor modelo que atenda às necessidades do negócio.*

*5. Interpretação dos Dados*

*Apresentação dos dados de forma a agregar valor aos negócios, respondendo perguntas realizadas para atender ao projeto gerando análises para identificar a probabilidade de resultados futuros.*

1. **Quais as habilidades essenciais de um cientista de dados? Você já possui alguma dessas habilidades?**

*Abaixo algumas habilidades do cientista de dados:*

*Business - Comunicação, Conhecimento do negócio, Gestão de projetos, Governança*

*Tecnologia - Gestão de dados Estruturados e não estruturados, Machine learning, Deep learning, Big data*

*Matemática/Modelagem - Algoritmos, estatística, Modelos gráficos*

*Programação - Conhecimento em linguagens voltadas a DS (R, Python, SQL)*

*Estatística - Data mining, Modelagem estatística*

*Dessas habilidades tenho bons conhecimento em*

*Comunicação, Gestão de dados estruturados, Gestão de projetos, Gestão de dados*

1. **Quais são as principais as ferramentas e linguagens de Data Science?**

*As principais ferramentas são:*

*Armazenamento e versionamento de códigos Git/github*

*Linguagem de programação R/Python*

*Gerenciamento de ambientes Virtualenv/Pipenv*

*Editor de código (notebooks) Jupyter/Colab*

*Bibliotecas matemáticas Matlab/Octave*

*Plataforma de desenvolvimento - Anaconda*

1. **Quais as diferenças entre Big Data, Business Intelligence e Data Science?**

***Data science****, é um processo que combina matemática, programação e resolução de problemas para extrair informações e encontrar padrões nos dados coletados.*

***Business Intelligence****, é um processo de coleta, estruturação dos dados, análise, divulgação e monitoramento de informações, que oferecem apoio a gestão de negócios.*

***Big data*** *é como um conjunto de ferramentas e práticas voltadas para o armazenamento e processamento de grandes volumes de dados, de forma mais eficiente e ágil.*

*As soluções de* ***Big Data / DS*** *coletam os dados de forma mais sistêmica e contínua, as ferramentas de BI analisam e classificam informações com objetivos específicos para tomada de decisão.*

1. **Quais as diferenças entre Inteligência Artificial, Machine Learning e Deep Learning?**

*Inteligência Artificial é a programação que simula a capacidade humana de raciocinar, prevendo o ambiente e a tomada de decisão.*

*Machine learning é o aprendizado da máquina avaliando os dados para descobrir um padrão para tomada de decisões e melhorando a atividade ao longo do tempo, utilizando a IA como base.*

*Deep learning é o aprendizado de máquina com técnicas mais especializadas inspiradas na estrutura e função do cérebro (redes neurais), por exemplo identificação de voz e imagens.*