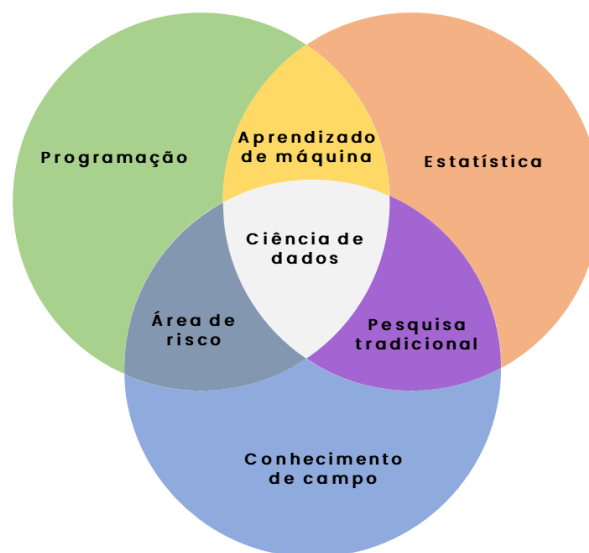


## Relatório 6 - Vídeo: O Que É Ciência De Dados (I)

Beatriz Almeida Felício

### Descrição da atividade

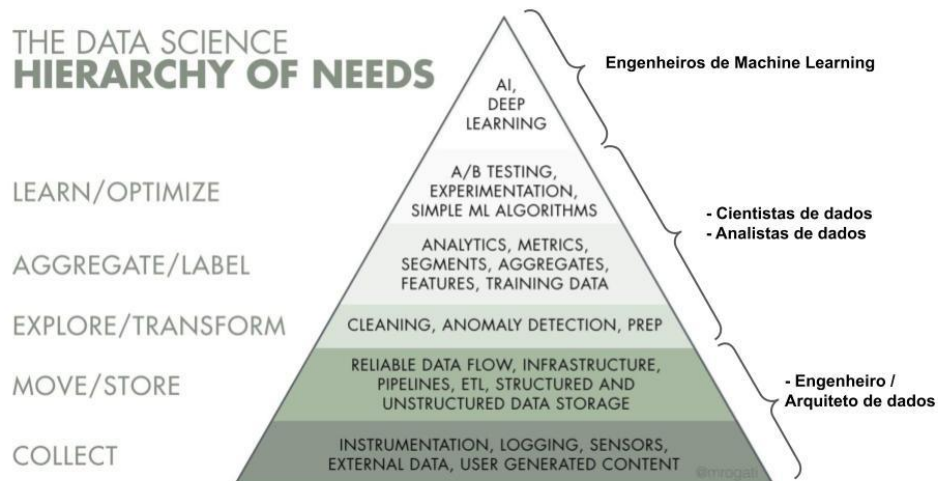
A ciência de dados é uma disciplina que utiliza dados para extrair informações e criar valor estratégico, indo além do que os números aparentam mostrar. O avanço no volume de dados disponíveis atualmente nos permite desenvolver modelos preditivos capazes de antecipar situações futuras, oferecendo às organizações uma vantagem competitiva. Nesse contexto, a estatística desempenha um papel fundamental, sendo a área da matemática que nos ajuda a entender, categorizar e explicar fenômenos, transformando observações em padrões quantificáveis.



Um dos pilares da ciência de dados é o uso de *machine learning* (aprendizado de máquina), uma abordagem que analisa os dados disponíveis e constrói automaticamente regras para prever resultados. No entanto, a qualidade dos dados é um aspecto crítico: dados ruins inevitavelmente geram modelos imprecisos e ineficazes. É importante observar que muitas das informações disponíveis hoje não são estruturadas, exigindo técnicas avançadas de pré-processamento e organização para torná-las úteis.

Para maximizar o impacto da ciência de dados em uma empresa, seja por meio de *insights* acionáveis, produtos orientados por dados (*data products*), ou sistemas de recomendação de produtos, é essencial compreender a hierarquia de necessidades em ciência de dados. Esse processo começa pela coleta de dados, que pode vir de dispositivos variados, como sensores, transações digitais ou redes sociais. Após a coleta, os dados são armazenados em sistemas apropriados, como *Data Warehouses*, bancos de dados tradicionais ou arquivos em formatos como .csv. Em seguida, aplica-se o processo de ETL (*Extract, Transform, Load*), que é crucial para garantir que os dados estejam prontos para análise.

A etapa de análise é onde os dados processados são utilizados para guiar decisões estratégicas, como identificar métricas de desempenho, avaliar a viabilidade de produtos ou definir ações que podem melhorar os resultados da empresa. Por fim, a melhoria contínua dos modelos analíticos ocorre por meio de testes A/B, experimentação controlada e a inclusão de novos conjuntos de dados, garantindo que os sistemas se adaptem às mudanças e mantenham sua eficácia.



## Conclusões

A ciência de dados é uma disciplina multidisciplinar e estratégica que combina estatística, aprendizado de máquina e gestão de dados para transformar informações brutas em soluções práticas e preditivas. Seu impacto vai além da análise descritiva, proporcionando às empresas a capacidade de antecipar tendências, otimizar operações e criar valor em um cenário competitivo e orientado por dados. Contudo, o sucesso dessa abordagem depende fortemente da qualidade dos dados, da aplicação de técnicas robustas de análise e de uma visão estratégica voltada para a melhoria contínua. Assim, investir em ciência de dados não é apenas uma decisão tecnológica, mas também uma aposta no futuro da inovação e da competitividade organizacional.

## Referências

**YouTube.** Disponível em: <<https://youtu.be/ykSILAQQu6o?si=k-ThBJK4ZILmx6nP>>. Acesso em: 14 dez. 2024.

**YouTube.** Disponível em: <[https://youtu.be/xC-c7E5PK0Y?si=CrZPNpVOtAeLUX0\\_>](https://youtu.be/xC-c7E5PK0Y?si=CrZPNpVOtAeLUX0_>). Acesso em: 15 dez. 2024.