

# Relatório 6 - O Que É Ciência De Dados

Lucas Scheffer Hundsdorfer

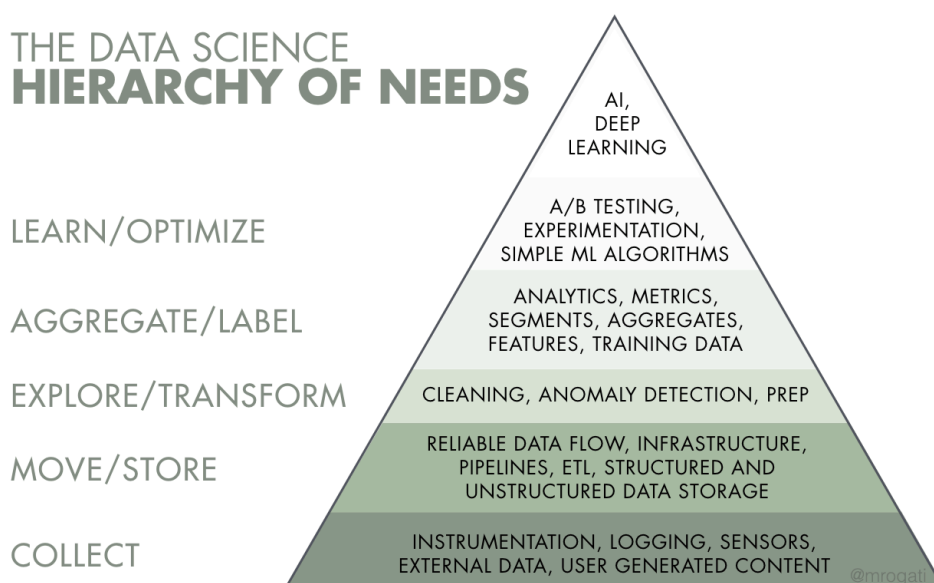
## Descrição da atividade

No começo do primeiro vídeo ele dá um exemplo bem simples de como um cientista de dados deveria pensar, analisando os 'dados' necessários e tomando a melhor decisão para melhorar o problema/situação. Muitas vezes essa tomada de decisão vem de 'insights' que não se parecem muito óbvios para todos, ciência de dados se trata de analisar todo o problema e prever possíveis cenários para a sua tomada de decisão, basicamente pensar em como usar os dados para saber mais do que eles realmente mostram, pensar fora da caixa.

Porém agora temos uma ferramenta que ajuda nessa área de ciência de dados, o Machine Learning, agora conseguimos passar dados para a máquina checar os dados que não são estruturados para a máquina conseguir analisar os padrões existentes. O grosso do trabalho de um cientista de dados pode ser também analisar se os dados passados para o aprendizado de máquina estão limpos, pois com dados ruins geram resultados errados e construindo modelos de aprendizado de máquina errados, o que não tem como corrigir.

Já no segundo vídeo ele traz uma análise do que é ser um cientista de dados mais perto do real mercado de trabalho, no começo do vídeo ele já deixa claro que ciência de dados é apenas trazer resoluções de problemas através de dados. Ele traz um pouco sobre a história da ciência de dados, que em 2001 William S. Cleveland queria elevar o Data Mining a outro nível, combinando ciência da computação com mineração de dados, tirando vantagem do poder do computador para estatísticas criando assim a ciência de dados.

Após isso ele traz um desenho de uma pirâmide, exemplo abaixo:



Essa pirâmide deixa claro o que é mais importante para a ciência de dados, na base dela obviamente vem a coleta de dados que sem ela não haveria ciência de dados, os dois próximos passos são apenas de armazenar e averiguar se esses dados são confiáveis, após isso vai para análise para ver se não há anomalias e atinge a otimização e o aprendizado profundo. Após isso ele traz uma análise mais detalhada do que o cientista de dados faz dependendo do tamanho da empresa, que é basicamente quanto maior a empresa, menos o cientista tem que se preocupar em áreas que não são do interesse dele e focar apenas no que ele é bom.

## Conclusões

Com esses 2 vídeos eu chego à conclusão que a ciência de dados é de suma importância para chegar em qualquer análise/'insight' mais preciso. Mostra também como os cientistas de dados têm que se tornar estrategistas e resolvedores de problemas.

## Referências

▶ What REALLY is Data Science? Told by a Data Scientist

▶ O que é ciência de dados | Nerdologia Tech