# Fórum de Debate do Módulo 3

Olá alunos!

Sejam todos muito bem-vindos ao nosso Fórum Pontuado. Nessa roda de discussão que iremos formar, gostaria de discutir um pouco sobre mercado de trabalho e aplicação de tecnologias.

Vou levantar algumas questões abaixo e espero que vocês criem um texto englobando-as, sem necessariamente responder uma a uma separadamente.

**Vamos lá!**

Temos vivido uma verdadeira revolução no campo da computação. Três pilares estão atualmente sustentando essa revolução e inovação: Big Data, Internet das Coisas (IoT) e Machine Learning (Aprendizado de Máquina). São tantas tecnologias, metodologias e teorias, que acabamos perdidos sobre o que realmente vale a pena estudar e especializar.

Sei que essa é uma dúvida presente na cabeça de muitos de vocês. Diante dessa situação e com o foco voltado para a Ciência dos Dados, sabemos que é impossível dominar profundamente tudo isso. Sendo assim:

a) O que vale mais a pena? Aprofundar em uma única tecnologia e se tornar especialista nela ou aprender um pouco de cada coisa?

b) Quais tecnologias são mais relevantes nesse momento (linguagens, frameworks, técnicas, metodologias ou outras)?

c) E o futuro? O que nos reserva?

Um grande abraço e aguardo a participação de todos.

Prof. João Paulo

**Lembre-se que a sua postagem neste fórum será pontuada.**

Para que a experiência seja ainda mais enriquecedora, aconselhamos que você realize ao menos uma réplica, ou seja, deixe um comentário de resposta no depoimento do seu colega, revelando afinidades ou diferenças em relação à sua postagem.

Boa noite, pessoal!

Acredito que para essas questões levantadas, podem haver diversas opiniões e, em alguns momentos, divergentes umas das outras, pois é levado muito em consideração a área de formação dos indivíduos que estão adentrando e migrando para a ciência de dados, além de suas experiências profissionais no mercado de trabalho, pois certamente estão ou estiveram em contato com diferentes setores e tipos de negócios específicos. Além deste background anterior completo ser diferente para cada um, ainda tem o nível de conhecimento e experiência profissional aplicados na área da ciência de dados propriamente dita, pois as empresas que adotam práticas de análise de dados e produção de conhecimento para os seus negócios podem pensar e agir diferentemente de seus concorrentes ou parceiros. É curioso que a literatura também chega a ser levemente divergente com relação a diversas terminologias envolvidas com a área de dados e, dessa forma, tendo a acreditar que as diferentes perspectivas do mercado também acabam sendo espelhadas.

Na minha opinião, como um cientista da computação formado e como analista e desenvolvedor de sistemas, tenho a percepção de que tudo é válido para uma área tão promissora e atual. A necessidade do conhecimento do referencial teórico é muito importante, pois o cientista tem esse nome por explorar e conhecer a fundo o que está por trás dos fatos e conceitos relacionados a determinada área e, sendo um cientista de dados, este, certamente, terá que conhecer profundamente e verticalmente tudo que envolve a teoria por trás dos diferentes tipos de análise de dados. Além disso, é fundamental ter o conhecimento do arcabouço matemático e algorítmico que justifica essas análises. E, nesse ponto, pode ser levantada outra questão, pois as tecnologias e ferramentas podem ser trocadas e atualizadas constantemente, porém o conhecimento das técnicas e teorias são os mesmos e são esses os reais focos de estudo para um cientista.

Por outro lado, é importante para qualquer profissional da área de TI estar habituado com tecnologias e ferramentas ligadas à sua área de atuação, porém, apenas tendo um conhecimento superficial e horizontal, de forma que, ao ser incorporado a uma equipe de trabalho ou empresa específica possa aplicar o seu conhecimento técnico utilizando uma tecnologia ou ferramenta específica.

Além desses conhecimentos, há a necessidade de o cientista de dados conhecer o negócio que fará parte do seu dia-a-dia, porém isto será em um momento futuro, no qual, o profissional já estará adentrado numa equipe e envolvido com o *core* ou negócio da empresa.

Com o conhecimento da área de ciência de dados que eu já possuo, posso afirmar que as linguagens que estão mais em alta são ***Python***, ***R*** e ***Scala***. ***Python***, majoritariamente, está acima por ser uma linguagem de programação com diversas aplicações e que oferece uma simplicidade na forma de codificação, se comparada com outras linguagens de mais baixo nível. Além disso, ela já dispõe de bibliotecas matemáticas como ***scipy*** e ***numpy***, bibliotecas para análise de dados como ***pandas*** e, principalmente, para aprendizado de máquina como ***sklearn*** e ***tensorflow*** e para *Big Data* como ***pyspark***. Outros tipos de bibliotecas são de acesso a banco de dados *SQL* ou *NoSQL* e de coleta e obtenção de dados, pois também fazem parte do dia-a-dia do cientista de dados.

Em termos de projeções futuras, acredito que num período de cinco à dez anos teremos muitas profissões extintas no Brasil e no mundo, porém muitos profissionais também migrando para áreas da computação, sobretudo, relacionadas à dados, pois o mundo, de modo geral, estará mais tendencioso a um *mindset* *data-driven*, graças ao bum da tecnologia e geração desenfreada de dados atuais e à conscientização futura e natural de que não há mais espaço para suposições sem dados e de que os dados bem trabalhados podem ter uma precisão maior na tomada de decisão do que os próprios especialistas de cada área. Isso não significa que os especialistas não terão mais espaço, apenas terão suas atuações ligeiramente modificadas.