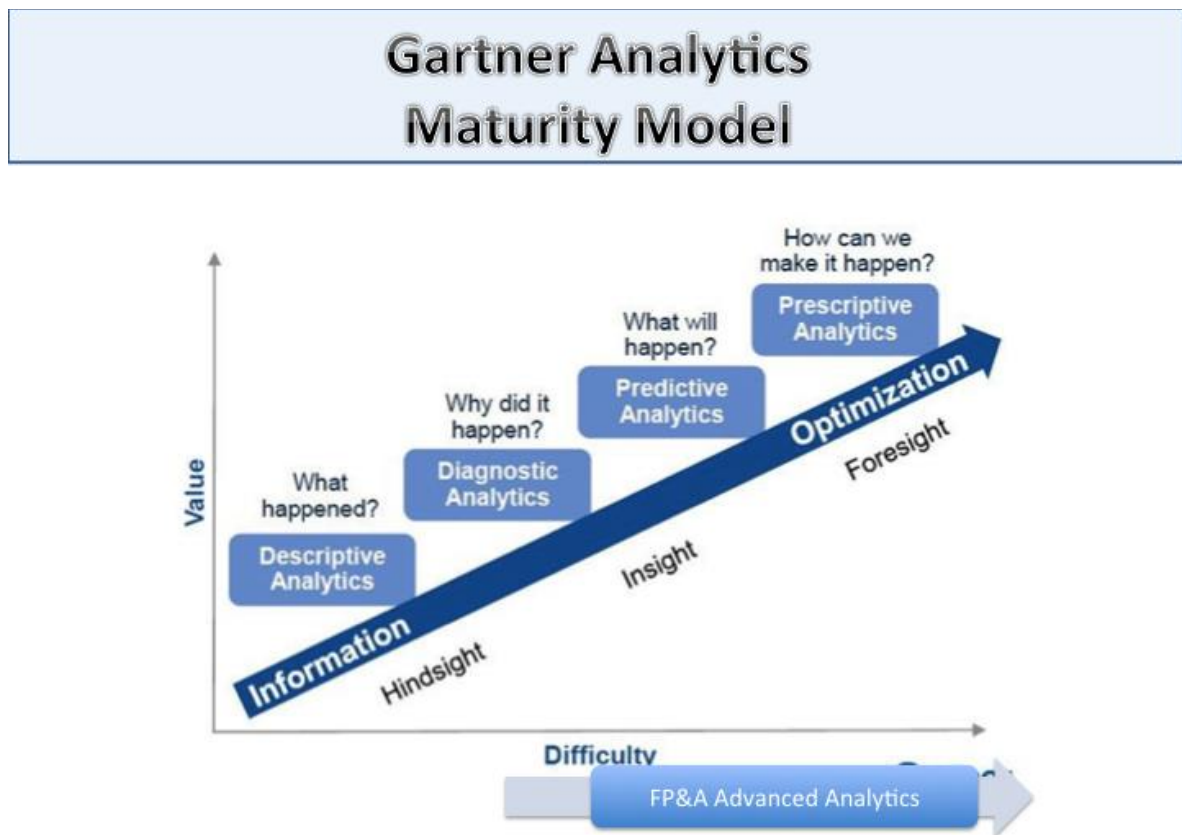


A OPERACIONALIZAÇÃO DA ANÁLISE PREDITIVA

A maioria das empresas dispõe hoje de massivas quantidades de dados, mas elas não sabem como extrair valor real destes dados. A análise preditiva é o uso de dados, algoritmos estatísticos e técnicas de aprendizado de máquina para identificar a probabilidade de resultados futuros com base em dados históricos.

Mas colocar Análise Preditiva em prática não é simples questão de instalar alguma nova tecnologia. As tecnologias impulsionadoras são fundamentais, mas é necessário também que a empresa adapte seus processos de negócio de modo a explorar os *insights* gerados. No gráfico a seguir observamos a evolução de uma empresa com o objetivo de adquirir a maturidade na Análise Preditiva.

O modelo da Gartner, mostrado abaixo, apresenta uma visão geral dos diferentes tipos de análises dentro de uma empresa.



Estima-se que empresas que atingem um alto grau de maturidade analítica e conseguem usar plenamente os seus dados poderão ver uma melhoria de 30% a 50% na eficiência global.

ESTÁGIOS DE MATURIDADE

As organizações têm diferentes pontos de partida e diferentes taxas de progressos ao longo da evolução apresentada no gráfico acima. Segundo o Prof. Thomas Davenport o estágio de maturidade analítica de uma empresa se enquadra em uma das cinco situações abaixo:

Estágio 1 (*Analytically Impaired*): a organização não tem um ou vários dos pré-requisitos para conduzir trabalho analítico sério, tais como dados, habilidades analíticas, ou interesse da gestão sênior;

Estágio 2 (*Localized Analytics*): há pequenos pacotes de atividade analítica na organização, mas eles não são coordenados ou focalizados em alvos estratégicos;

Estágio 3 (*Analytical Aspirations*): a organização visa um futuro mais analítico, tem capacidades analíticas estabelecidas, e tem umas poucas iniciativas significativas em curso, mas o progresso é lento- frequentemente porque algum fator crítico tem sido difícil de implementar;

Estágio 4 (*Analytical Companies*): a organização tem os recursos humanos e tecnológicos necessários, aplica análise preditiva regularmente, e conquista alguns benefícios. Mas o foco estratégico não está assentado em análise preditiva, e não a transformou numa vantagem competitiva;

Estágio 5 (*Analytical Competitor*): a organização rotineiramente usa análise preditiva como uma capacidade distintiva nos negócios. Ela enfoca toda a empresa, compromissou e envolveu lideranças, e tem atingido resultados em larga escala.

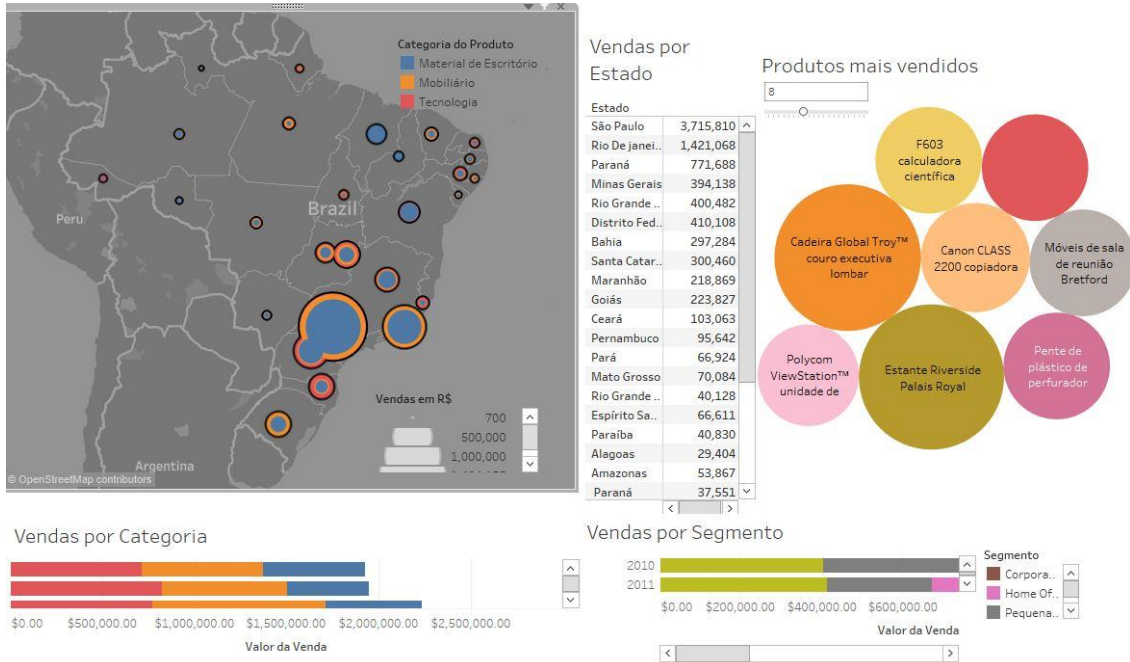
IMPLANTAÇÃO DA ANÁLISE PREDITIVA

Para avançar nos estágios acima temos três regras úteis:

Utilizar *dashboards*

A maioria das empresas não tem pessoal habilitado para encontrar *insights* em dados. *Dashboards* como o Tableau ou Qlik são ferramentas que os gerentes e supervisores de áreas podem usar e configurar sem ter conhecimento de programação.

Eles próprios podem gerar suas análises visuais em tempo real com painéis interativos que os ajudam a fazer descobertas e agir imediatamente.



Integrar os dados



Os dados não devem ser divididos por departamentos. Armazene os dados localmente ou na nuvem, sejam eles Big Data, um banco de dados SQL, uma planilha ou aplicativos na nuvem, como o Google Analytics, tornando mais fácil para os analistas e outros usuários finais extrair o valor desejado de uma grande e variada quantidade de informações sob demanda.

O Cientista de Dados

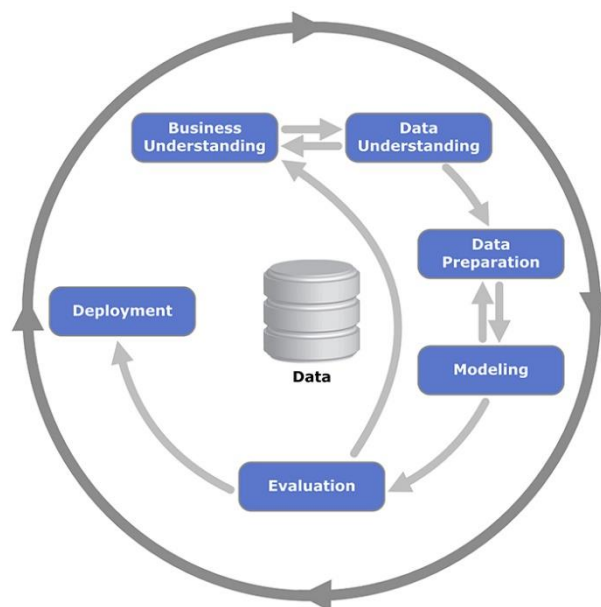
Finalmente é essencial dispor de *expertise*, com novas funções como o Cientista de Dados. O Cientista de Dados é um profissional multidisciplinar, com conhecimentos em ciência da computação, matemática, estatística e, principalmente, conhecimentos do negócio onde está inserido. Quem exatamente é esta figura ainda não está claro, mas

podemos dizer que o papel de um Cientista de Dados é gerir os dados da empresa, através de uma estratégia baseada em valor para o negócio.

FASES DE UM PROJETO

Entender o negócio:

Qual é o resultado que você obter e qual a variável de saída? Quais são as variáveis (preditores) a serem usadas para as previsões e que irão influenciar esse resultado? Se você construir um modelo e criar as probabilidades de resultados futuros, quais são as probabilidades aceitáveis? Os chefes de departamento ficarão convencidos pelo seu modelo analítico?



Enteder os dados

Que dados que são necessários para todos os preditores? Você quer dados para todas as regiões, linhas de produtos, etc.?

Preparar os dados

Os dados são suficientes para a análise? Os dados precisam ser limpos? Você tem todas as variáveis que você precisa? Os dados estão em formato que você pode importar facilmente para um pacote de software de análise preditiva?

Particionar os dados

Antes de construir o modelo precisamos dividir aleatoriamente o conjunto de dados em duas partes - dados de treinamento e dados para teste. Treinamos o modelo sobre a base de dados para treinamento (digamos 70% dos dados escolhidos aleatoriamente) e testamos o modelo com os 30% não usados para o treinamento. Os dados para o teste nunca são usados na construção do modelo.

Modelar

Seu modelo será supervisionado, não supervisionado ou misto? Você procura resultados conjuntos (*clusters*) ou individuais? Que algoritmos pretende usar? Existe bibliografia referente a aplicação para pesquisar?

Avaliar

Aplicamos o modelo gerado na base de dados separada para o teste. Baseado em resultados históricos, o modelo está funcionando satisfatoriamente? Se necessário, agora é a hora de tentar novos preditores e novas técnicas de modelagem.

Implantar

Como você implementará o modelo nos computadores da empresa? Como você treinará os operadores? Como medir os resultados ao longo do tempo. O modelo poderá variar ao longo do tempo? Deve haver uma tentativa de criar uma análise ROI para esta operação?

Michel Janos

7/2015