



COMO DISSEMINAR A LINGUAGEM DOS DADOS

| POR JOÃO LUIZ BECKER E EDUARDO HENRIQUE DINIZ

Todos podem aprender as habilidades de ler, escrever e compreender dados, desde que se mudem os paradigmas de ensino de *data literacy*.

Para interagir e conviver atualmente com outros indivíduos e com o mundo do trabalho, as pessoas cada vez mais necessitam ser “letradas” em dados. Pesquisa realizada com executivos pelo Grupo Gartner revela que 80% das organizações pretendem iniciar programas internos de *data literacy*, expressão que surgiu para identificar a habilidade de ler, escrever e compreender dados. O objetivo é cobrir deficiências de formação que impedem que se use o potencial dos dados para gerar valor para os negócios.

Precisamos, portanto, levar profissionais de administração a conhecerem, simultaneamente, ferramentas quantitativas, estatísticos, computacionais e de tecnologia da informação. A educação dos gestores dos setores privado e público deve envolver uma reflexão sobre como várias metodologias e técnicas podem ser usadas para resolver problemas práticos, capacitando tais gestores a solucionar questões complexas a partir da análise de grandes volumes de dados.

POTENCIAIS E DIFICULDADES

Considere-se, por exemplo, a crescente utilização de grandes volumes de dados para analisar a mobilidade humana. Com o rastreamento do uso de celular de milhões de usuários de telefonia monitorados on-line, é possível estudar o comportamento das pessoas em situações específicas. No caso da avaliação de riscos de uma catástrofe, torna-se viável compreender para onde e como as pessoas se deslocam e, com isso, propor modelos mais seguros para enfrentar momentos críticos. Em um plano micro, uma loja de departamentos pode avaliar o movimento dos consumidores em seu estabelecimento para propor-lhes melhores serviços.

Podemos também pensar em cruzamentos de bases de dados distintas com a finalidade de conhecer comportamentos que hoje só conseguimos analisar superficialmente. Um exemplo com muito potencial de estudo é o Cadastro Único, que contém dados das famílias mais pobres do Brasil e serve de base para mais de duas dezenas de políticas públicas do país, como o Programa Bolsa Família e o Programa de

O aprendizado centrado no participante e na prática cotidiana ajuda a valorizar as ferramentas analíticas de baixo para cima, em um processo de “aprender fazendo”.

Apoio à Conservação Ambiental. A dificuldade é avaliar o impacto desses programas na vida dos beneficiários de forma mais abrangente. Qual é a relação entre o Bolsa Família e o consumo de tecnologia das famílias beneficiárias, por exemplo? Como o Programa Nacional de Crédito Fundiário afeta a incidência de doenças no interior do país? Essas perguntas só podem ser respondidas com um olhar atento à lógica dos dados históricos existentes tanto na base do Cadastro Único quanto em bases da área de saúde e de sistemas de telefonia, por exemplo.

Há duas categorias de problemas para a materialização das análises desses casos que mencionamos anteriormente. A primeira é como acessar os dados de forma segura e sem ferir a privacidade dos indivíduos que serão “monitorados”. Essa primeira categoria de problemas é abordada, nesta edição, em dois artigos, um sobre a nova legislação de proteção de dados e o outro sobre projetos para o uso dos dados em benefício da população, como o da ferramenta OPAL (ou Open Algorithms), desenvolvida pela organização não governamental (ONG) Data-Pop Alliance.

Neste artigo, queremos explorar a segunda categoria de problemas, que está diretamente relacionada à deficiência de *data literacy*. A formação de cidadãos capazes de pensar em problemas que envolvem grandes massas de dados exige o domínio de certa “linguagem de dados”, o que não é elementar.

É necessário desenvolver capacidade crítica para o adequado uso dos dados, pois nem sempre as informações que vêm à tona são pertinentes ou relevantes. Tome-se, por exemplo, uma notícia divulgada no site de um município do interior do Brasil comparando índices de saúde da população antes e depois de determinada ação de suas autoridades. À primeira vista, passa-se a impressão de que a ação desencadeada promoveu a melhoria da saúde dos munícipes, aspecto ressaltado no comunicado. Mas um exame mais acurado, fazendo uma comparação com os demais municípios

da região, mostra que os índices melhoraram no entorno no mesmo ritmo, o que coloca em dúvida a relação entre a ação promovida pela prefeitura e a melhoria na saúde dos habitantes. O caso remete à postura crítica indispensável para quem utiliza dados massivamente, com permanente vigilância a respeito das premissas embutidas nas análises, de maneira especial sobre sua representatividade.

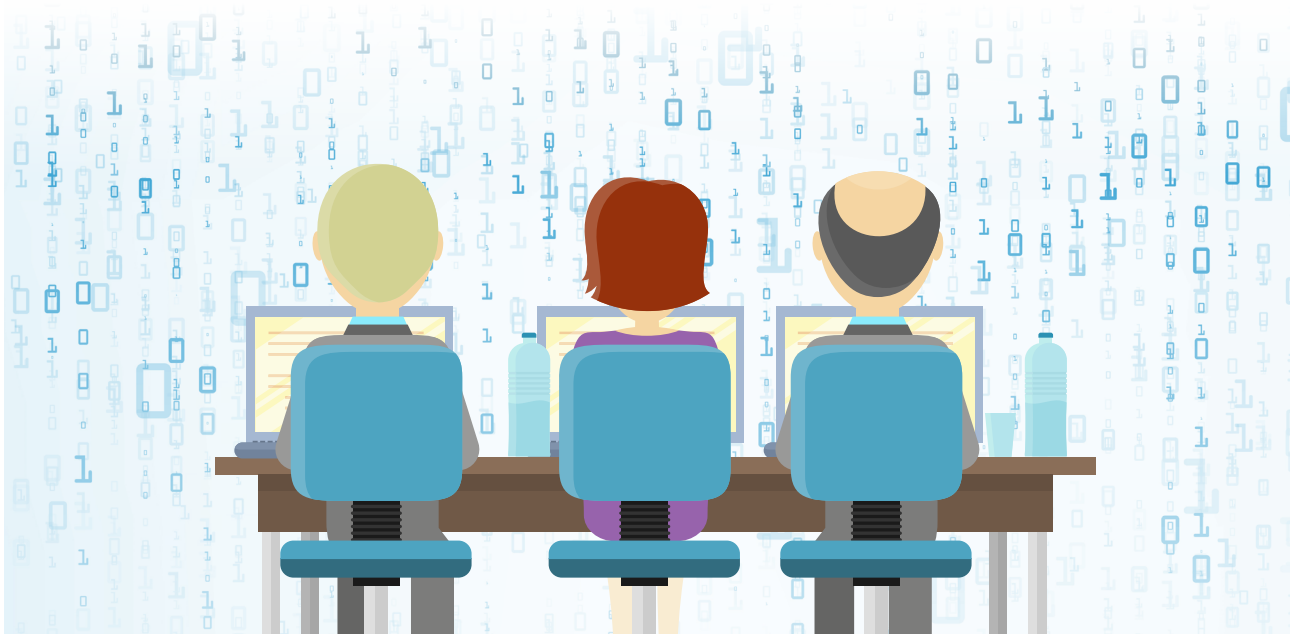
PROBLEMAS NO ENSINO

O que dificulta a disseminação de pensar sobre dados está ligado a processos arraigados de ensino de linguagens não naturais.

Tomemos como exemplo o ensino da matemática. Enquanto as crianças aprendem a linguagem oral naturalmente, apenas interagindo com um círculo relativamente pequeno de pessoas, só conseguimos aprender matemática em um processo sistemático, às vezes árduo, na escola.

Embora o ser humano tenha aprendido a contar logo após ter criado a linguagem oral, antes mesmo de ter inventado a escrita, a matemática exige nível alto de abstração. Para entender a ligação entre o concreto e o abstrato na matemática, considere a diferença entre a aritmética, que lida com números elementares e com algumas operações entre eles, e a álgebra, que estuda a manipulação formal de equações. Para as crianças pequenas, a matemática é a aritmética, a arte de contar, uma atividade ligada ao mundo que as cerca. A partir de certo momento no processo educacional, os estudantes são apresentados à álgebra, muito mais abstrata, momento em que a maioria deles passa a acreditar que não consegue aprender matemática, mesmo que tivessem sido bons em aritmética poucos anos antes.

O problema no ensino dessa linguagem formal está associado à falta de consciência dos instrutores de que a passagem da aritmética para a álgebra é um enorme salto em abstração para o aprendiz. Bastaria dizer que, enquanto a aritmética tem suas origens há mais de 30 mil anos, a álgebra



aparece “apenas” algumas dezenas de milhares de anos depois. Ou seja, foi necessário encadear conhecimentos ao longo de milênios para a humanidade fazer essa transição que exigimos que adolescentes percorram de um ano para o outro. Com isso, perdemos uma quantidade enorme de jovens que odiarão matemática para sempre.

Não bastasse a dificuldade dos instrutores de entender os interesses legítimos de seus pupilos em encontrar significado no ensino dessa linguagem formal, alguns fazem questão de tornar esse conhecimento mais hermético. Pitágoras (570 a.C. – 495 a.C.) criou toda uma simbologia em torno dos números e da geometria, considerados divinos, para ser apreciada apenas por iniciados. “Neopitagóricos” contemporâneos costumam dar mais importância às “revelações” dos números do que ao conteúdo prático que pode ser extraído deles.

DE BAIXO PARA CIMA

E o que o ensino da linguagem matemática tem a ver com a alfabetização para os dados? Muitas vezes, treinamentos oferecidos por provedores de ferramentas analíticas valorizam os processos quantitativos em detrimento da checagem das premissas envolvidas nas análises. Dissociada do contexto em que será utilizada, a linguagem de dados não ajuda no entendimento das mudanças que devem ser realizadas nos processos de negócios.

A competência numérica que devemos incentivar, ao contrário, deve promover a constante vigilância acerca dos pressupostos dos modelos quantitativos. Entendemos que o aprendizado centrado no participante ajuda a valorizar as

ferramentas analíticas de baixo para cima, em um processo de “aprender fazendo”, experimentando com a prática cotidiana. Adaptando os ensinamentos de Paulo Freire, a leitura dos dados deve ser precedida pela leitura de mundo que o aprendiz traz para o treinamento. É o diálogo dessas leituras que conecta as formas oral e escrita de interpretação, permitindo o entendimento e a transformação do mundo.

Em meio à abundância de dados em que vivemos, graças à ampla disseminação das tecnologias digitais, a tarefa que temos é formar jovens competentes para entender atividades de negócios, estruturas sociais e padrões de comportamento por meio da interpretação desses mesmos dados. Da mesma forma que desenvolvemos competências para criar a linguagem dos números ao longo dos séculos para lidar com problemas práticos, precisamos agora elaborar a linguagem dos dados com o objetivo de desenvolver mecanismos de representação da realidade que nos ajudem a transformar de modo positivo o nosso mundo.

Promover *data literacy* significa massificar o poder de manejar os dados que nos cercam. A disseminação dessa linguagem dos dados para todos, e não apenas para um grupo de iniciados que conhecem os “segredos dos números”, é essencial para a construção de uma sociedade mais justa. ●

PARA SABER MAIS:

- João Luiz Becker. *Estatística Elementar: transformando dados em informação*, 2015.
- Valerie Logan. *Fostering Data Literacy and Information as a Second Language: a Gartner trend insight report*, 2019.
- Paulo Freire. *Pedagogia da Autonomia: saberes necessários à prática educativa*, 2011.

JOÃO LUIZ BECKER > Professor da FGV EAESP > joao.becker@fgvbr
EDUARDO HENRIQUE DINIZ > Professor da FGV EAESP > eduardo.diniz@fgvbr