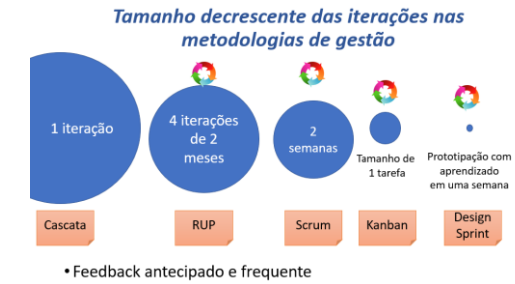


Introdução a Ciência de Dados

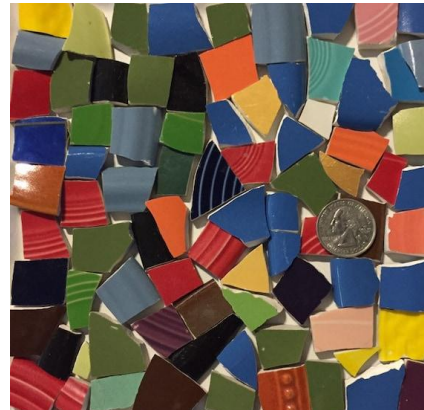


Professor: Alex Pereira

Iteração (Aula 11)



- Transforma incrementos do Desconhecido Desconhecido
 - Em Desconhecido Conhecido
- Analogia sobre desenvolvimento de software
 - e a certeza da iteração, aprendizado e mudança
- Criar um mosaico quadrado com 2 tipos de peças
 - Um tipo você constroi e o outro você recebe (depois da que construiu)
 - ✓ A peça que você vai construir depende do que você vai receber



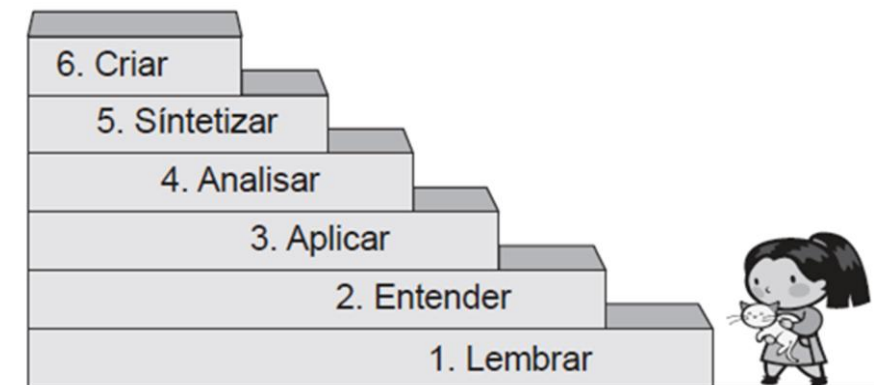
Como se manter útil num contexto de alto nível de automação

Soft skills vs Hard skills

Hard skills serão automatizados mais cedo

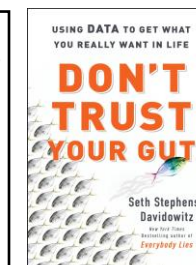
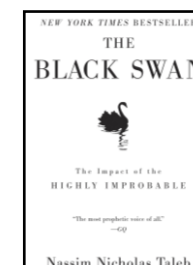
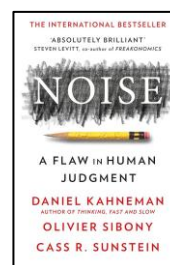
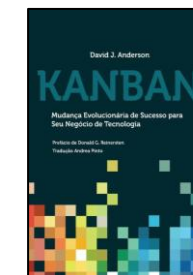
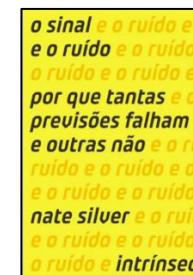
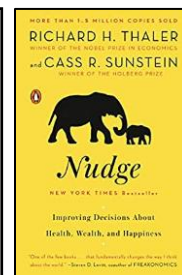
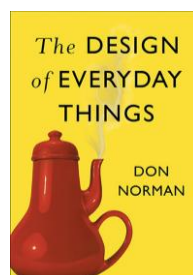
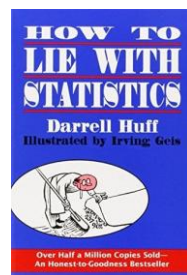
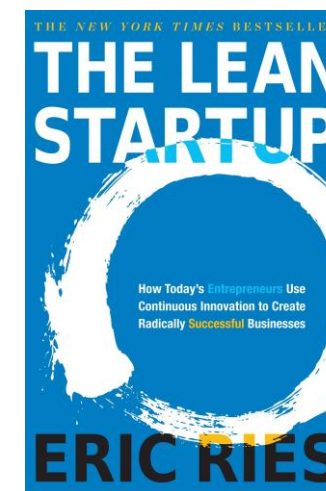
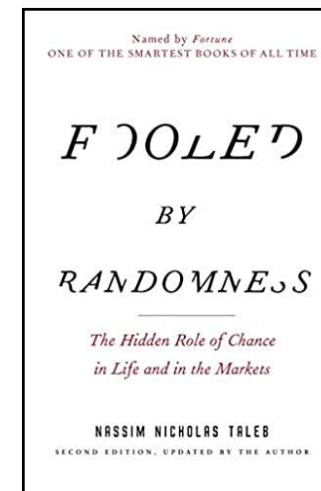
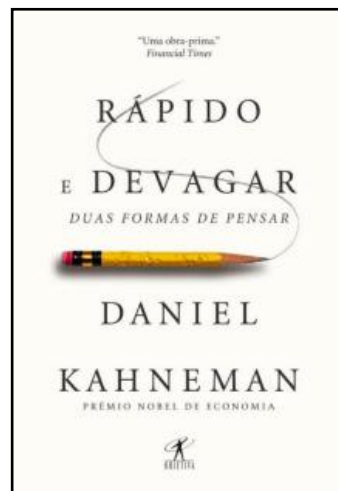
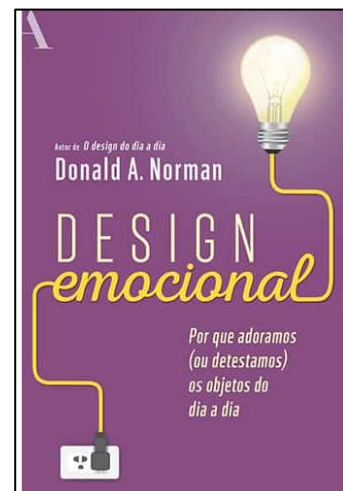
Habilidades e Habitos a serem adquiridos no contexto do avanço da IA

- Questionar o óbvio
 - Por que a multiplicação de dois números positivos resulta num número par e a multiplicação de dois números negativos também?
- Aprofundamento de entendimento/pesquisa
 - O porque em vários níveis
- **Taste (gosto) / bom senso (arquitetura, estética, design) - Aula**
- Dar nota/avaliar o trabalho de outro/IA
- Empatia (fonte: [Po-Shen Loh](#))

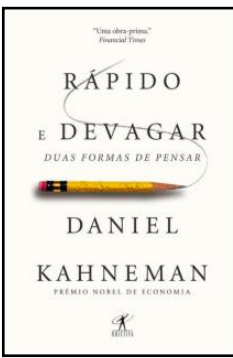


Livros afetos a ciência de dados e escopo da aula

Aula

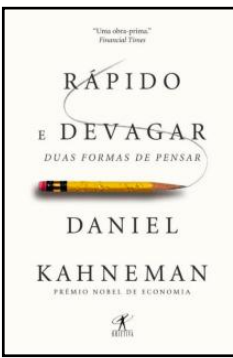


Atalhos Mentais (Heurísticas)



- São regras práticas (heurísticas)
 - que o cérebro usa para decidir rápido com pouco esforço
 - ✓ são estratégias rápidas que simplificam decisões sob incerteza e limitação de atenção
- Na maioria dos contextos cotidianos, isso aumenta eficiência;
 - em contextos novos, ambíguos ou estatisticamente complexos,
 - ✓ pode distorcer julgamentos.
- Exemplos comuns
 - Disponibilidade
 - ✓ estimar frequência/probabilidade pelo que vem fácil à mente, como notícias recentes ou vívidas.
 - Representatividade
 - ✓ julgar pela “semelhança” com um protótipo, ignorando base rate.
 - Status quo e inércia
 - ✓ preferência por manter a opção padrão ou não agir.
 - Reconhecimento
 - ✓ escolher o que é conhecido quando falta informação comparativa.

Vieses Cognitivos



- Padrões sistemáticos de erro no julgamento e na tomada de decisão
 - que surgem de **atalhos mentais**, emoções e limitações de atenção e memória.
- Eles aceleram escolhas no dia a dia,
 - mas podem distorcer a percepção da realidade e levar a conclusões imprecisas.
- São tendências automáticas do pensamento
 - que simplificam o processamento de informações,
 - porém frequentemente desviam da lógica estatística
 - ✓ ou da racionalidade normativa

Gatilhos Mentais existem à revelia da sua vontade

Linda tem 31 anos, é solteira, franca e muito inteligente. É formada em filosofia. Quando era estudante, preocupava-se profundamente com questões de discriminação e justiça social, e também participava de manifestações antinucleares.

- **Qual a alternativa mais provável?**
 - A) Linda é uma caixa de banco.
 - B) Linda é uma caixa de banco ou astronauta
 - C) Linda é uma caixa de banco e é ativa no movimento feminista.

Análise dos Problemas apresentados anteriormente

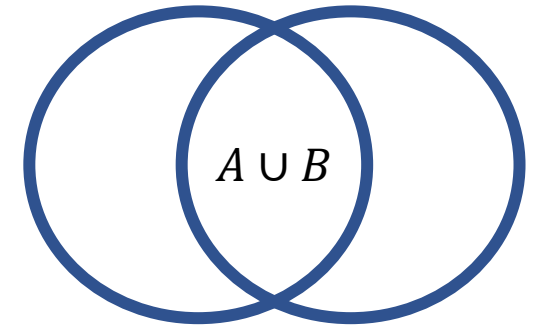
- Linda

- Teoria de conjuntos

- ✓ $P(A \cup B) \geq P(A) \geq P(A \cap B)$

- Respondemos intuitivamente outra pergunta

- ✓ Qual descrição melhor se encaixa no perfil de Linda



- O enunciado te desviou o foco da parte importante do problema?

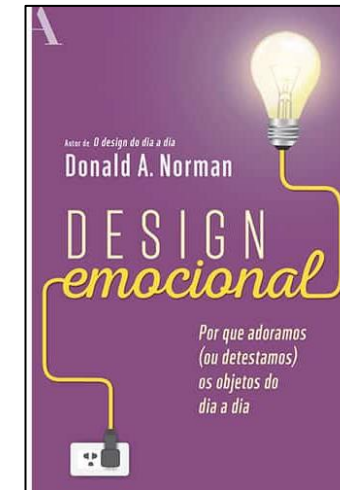
- O quanto comum é haver um “enunciado” complexo na vida real?

- Sob uma pergunta difícil para a qual não conhecemos a resposta

- Mas conhecemos a resposta de outra pergunta mais simples

- ✓ Intuitivamente, damos como resposta esta outra

Design Emocional – Don Norman



Três níveis complementares

de como pessoas percebem e se relacionam com produtos

Viceral

- Primeira impressão sensorial

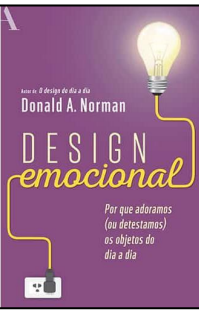
Comportamental

- Qualidade do uso no dia a dia

Reflexivo

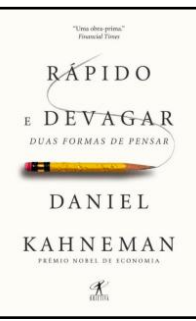
- Significado/identidade atribuídos ao produto

Design Emocional – Don Norman



- Visceral
 - respostas imediatas e subconscientes aos estímulos sensoriais do produto, como forma, cor, textura, som e movimento;
 - é a primeira impressão **estética** e de “**segurança**” ou “**atração**”.
- Comportamental
 - qualidade da experiência durante o uso;
 - envolve usabilidade, feedback, fluidez de tarefas, sensação de controle e eficiência na execução.
- Reflexivo
 - interpretação consciente e de longo prazo;
 - inclui memórias, valores, autoimagem, status percebido, confiança na marca e o que o produto comunica socialmente.

Cognitive Ease (fluência cognitiva)

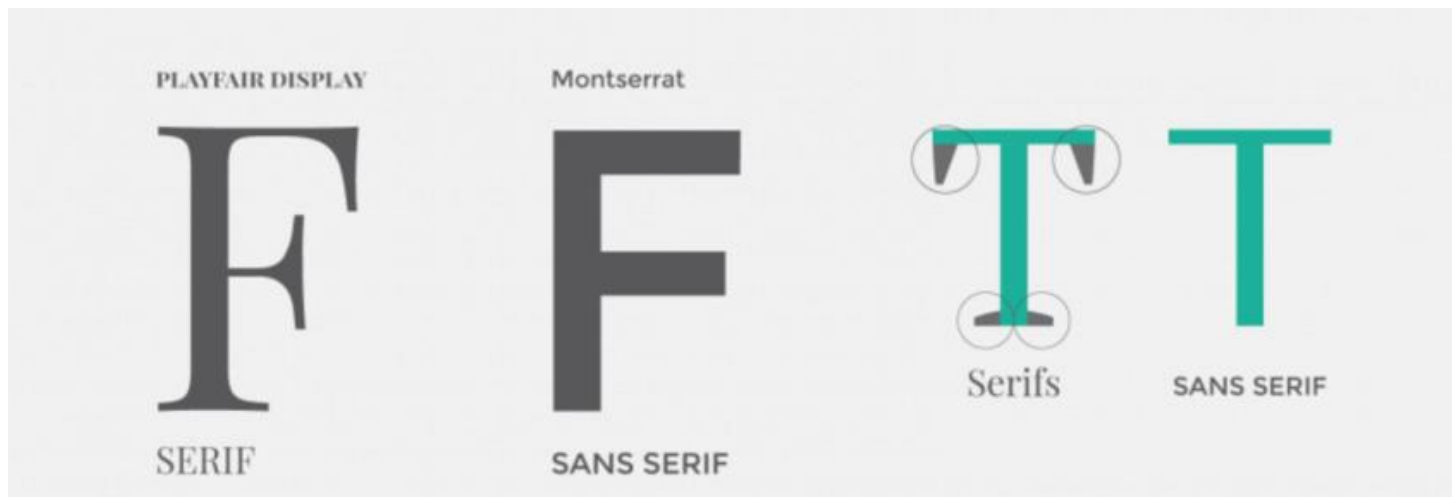


- É o estado em que o cérebro processa informações com pouco esforço,
 - levando a respostas rápidas, sensação de familiaridade e maior aceitação
- Há processamento fluente graças a
 - clareza, familiaridade, boa legibilidade, instruções simples, padrões conhecidos e bom humor, o que aumenta confiança e velocidade de resposta.
- Gatilhos de ease
 - **tipografia legível**, alto contraste, linguagem simples, chunking de informação, layouts previsíveis, repetição/priming e feedbacks claros.

Cognitive Strain (esforço cognitivo)

- É o estado de esforço atencional que torna o pensamento mais lento, analítico e vigilante.
 - É o estado que você vai querer estar ao assinar um contrato importante ou comprar um carro / casa
 - ✓ Aciona o Sistema 2 (Deliberado, ordenado e demanda esforço)
- Há processamento custoso diante de
 - novidade, ambiguidade, baixa legibilidade, regras incomuns, multitarefa ou pressão,
 - ✓ o que aciona atenção focada, checagem e ceticismo.
- Gatilhos de strain
 - jargões, instruções longas, baixa previsibilidade, muitos passos, escolhas similares demais, **fontes pequenas, baixos contrastes** e interrupções.

Categorias de Fontes (Com serifa e sem)



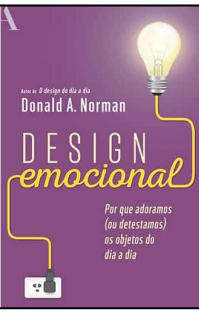
Bom uso da tipografia (fontes)

- Contraste(W3C)
 - **4.5:1** para texto padrão e
 - **3:1** para texto grande
- Depende do objetivo / ocasião
 - Ex.: Outdoor promocional vs livro técnico vs dashboards
- Outdoor promocional
 - exige leitura instantânea a distância com tipografia enorme, alto contraste, poucas palavras e caixa mista para maximizar cognitive ease e legibilidade em movimento.
- Livro técnico
 - privilegiar leitura contínua com largura de linha controlada, altura de linha confortável, hifenização bem configurada

Bom uso da tipografia (fontes) - Dashboards

- Dashboards digitais pedem tipografia para leitura de relance
 - Usar tamanhos maiores em KPI,
 - hierarquia consistente e
 - contraste suficiente para reduzir esforço cognitivo e acelerar decisões.
- Trate números como elemento de primeira classe
 - Alinhar colunas numéricas à direita/decimal,
 - Alinhar dígitos em números tabulares,
- Rotular diretamente no gráfico
- Priorizar fontes Sans Serif (sem serifa)
- Limitar a variedade de estilos e tamanhos
 - para manter escaneabilidade

Nível Visceral nos Dashboards

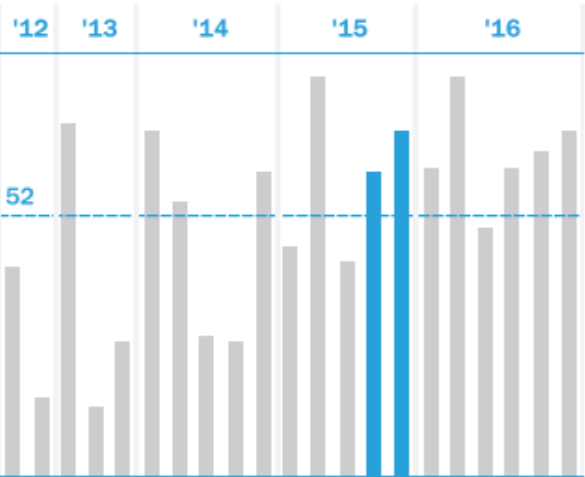


- O design visceral
 - é a reação intestinal e imediata à aparência de algo.
 - É o "amor à primeira vista".
- Nos dashboards,
 - se conecta aos princípios que governam a harmonia visual
- Princípios estudados na aula 1
 - Ex.: Use a cor com propósito e Evite a desordem (clutter)
- Reação visceral positiva
 - Paleta de cores harmoniosa, tipografia legível e um layout limpo com bom uso de espaço em branco

Bom Exemplo

Course Metrics

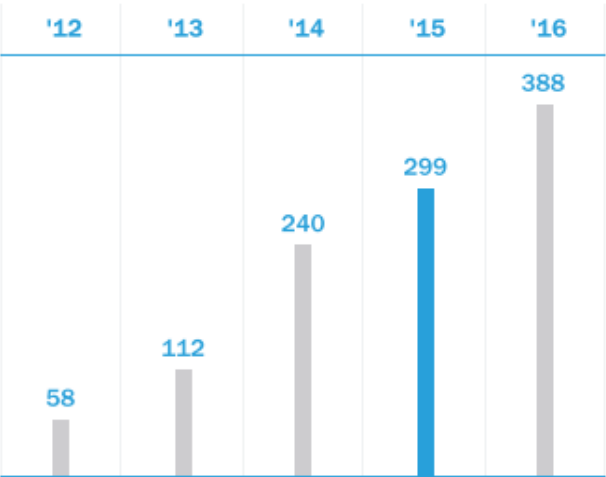
Students



1097

Total students in five years

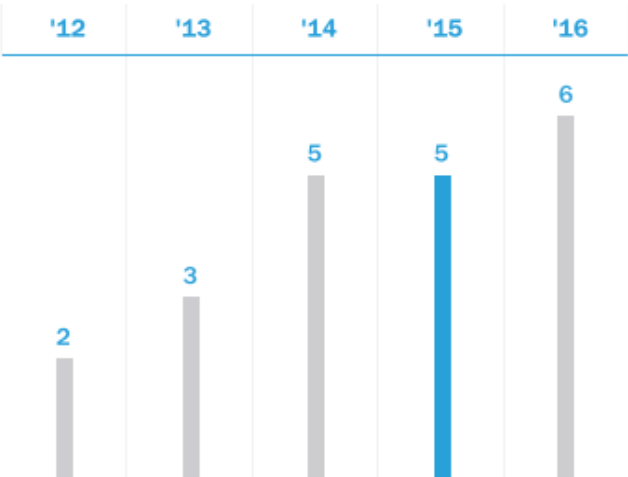
Enrollments



687

Total students in 2015-2016

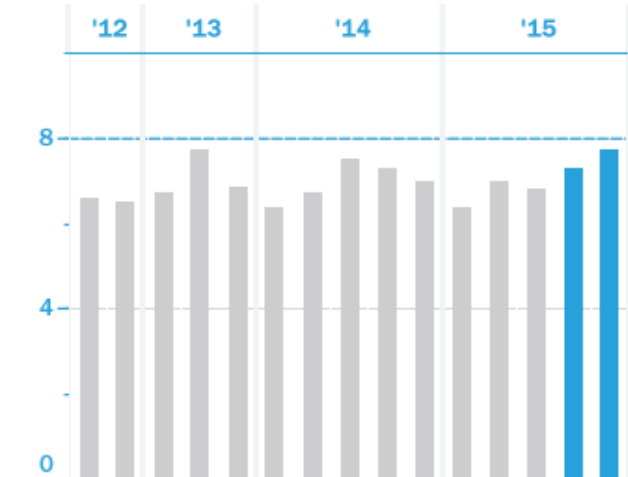
Classes



21

Total classes in five years

Ratings



7.7 of 8

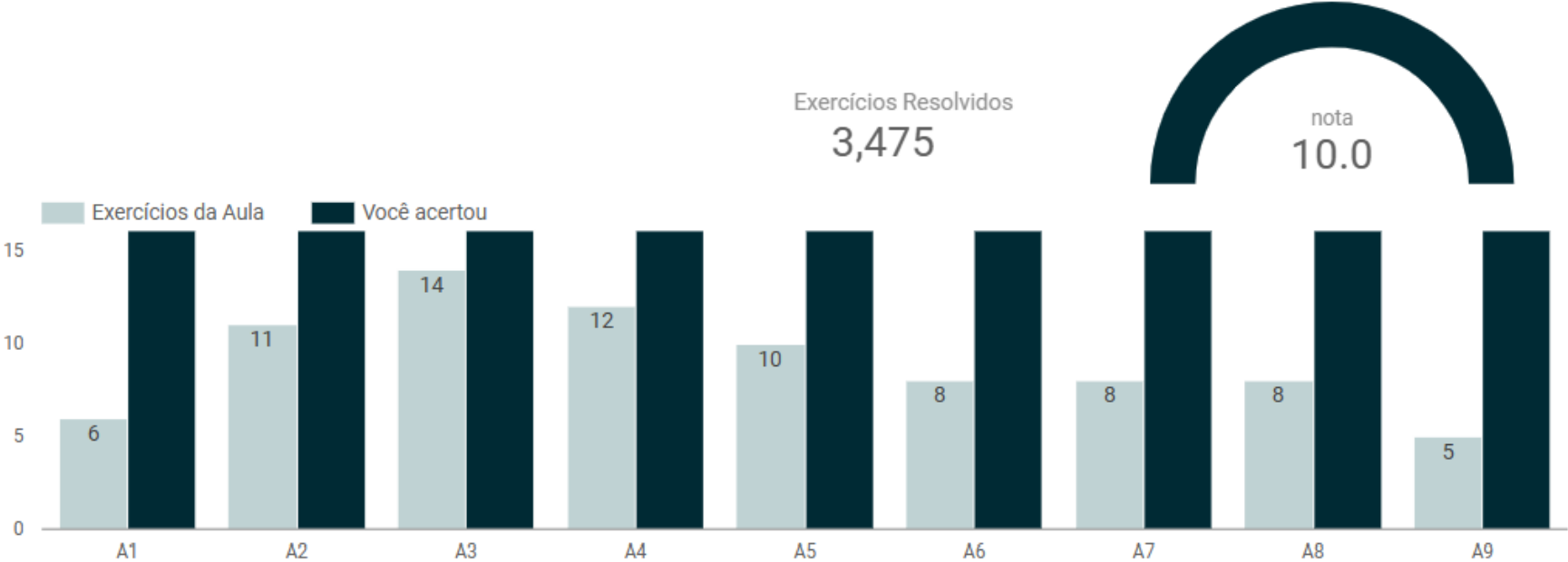
Most Recent Instructor Rating (out of 8.0)

Bom Exemplo



Introdução a Programação em Python

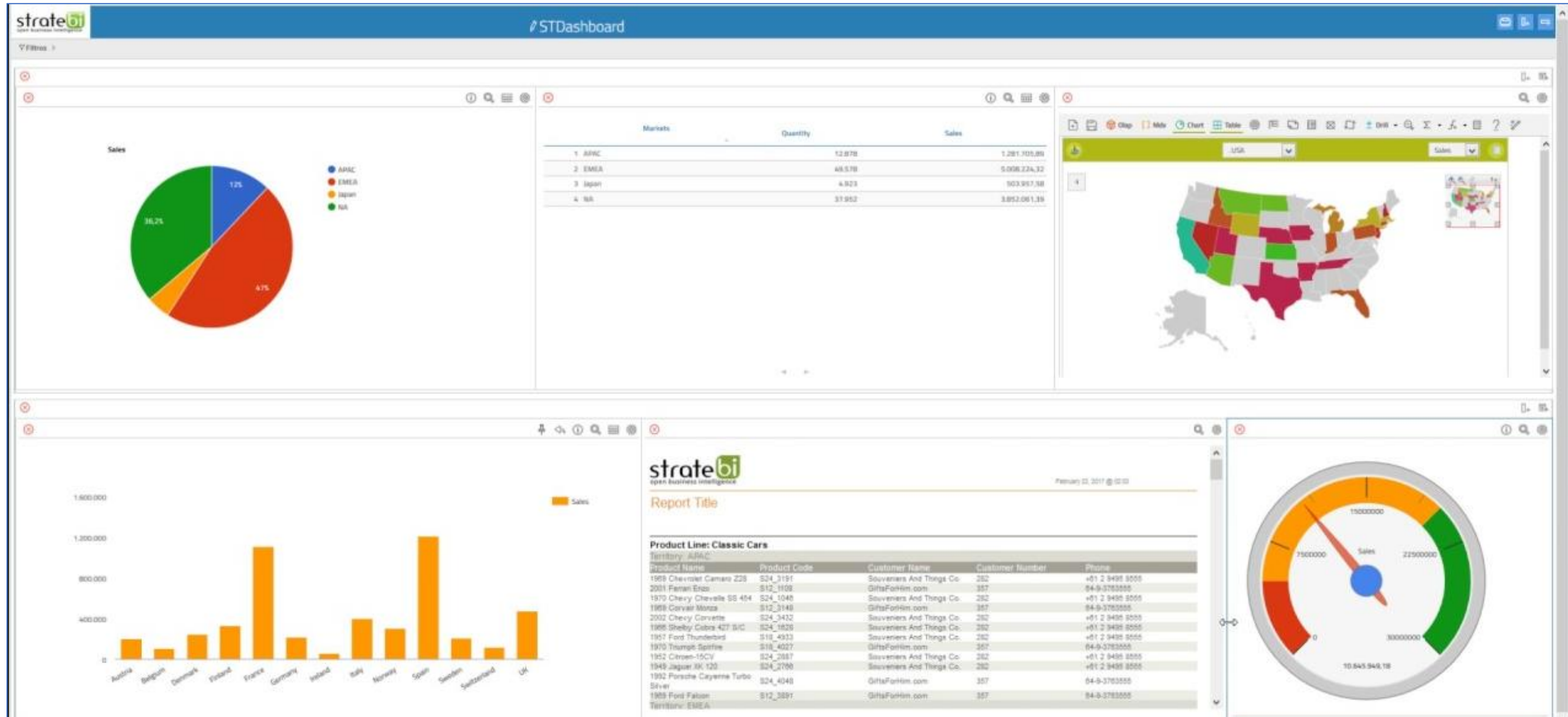
Olá Alex, confira o seu desempenho



Aula ▾

Aula 1 ▾	Exercicio 2 ▾	Acertou	Observações	Deadline	Hora do Envio
A9	R1	True	The student's reflection demonstrates a clear depth ...	2025-02-05T09:00:00	2025-02-03T14:56:57
A9	R1	True	A reflexão demonstra um bom nível de profundidad...	2025-02-05T09:00:00	2025-02-03T23:53:17
A9	R1	True	A reflexão demonstra uma boa profundidade de ent...	2025-02-05T09:00:00	2025-02-05T17:28:28
A9	R1	True	The reflection demonstrates some understanding of...	2025-02-05T09:00:00	2025-01-06T20:25:09

O que pode melhorar?



O que pode melhorar?



Exemplos de Contraste não apropriados (WCAG)

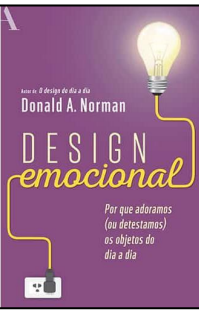
Preço Médio no Período (U\$)

112.7K

Food and Drug Administration

um anti-inflamatório não esteroidal (AINE)
izado para aliviar dores e febre, sendo eficaz

Nível Comportamental nos Dashboards



- O design comportamental foca
 - na usabilidade,
 - na função e no sentimento de controle e eficácia
 - ✓ que um usuário experimenta ao usar um produto.
 - É sobre a alegria de um produto
 - ✓ que "simplesmente funciona".
- Princípios estudados na aula 1
 - Escolha o gráfico certo para a mensagem e Crie uma hierarquia visual clara
- Emoção comportamental positiva
 - Usar o gráfico correto para os dados (ex: um gráfico de linhas para uma tendência temporal) e
 - organizar a informação de forma lógica,
 - ✓ o usuário entende a informação sem esforço.
 - Essa fluidez gera um sentimento de domínio e competência

Bom Exemplo



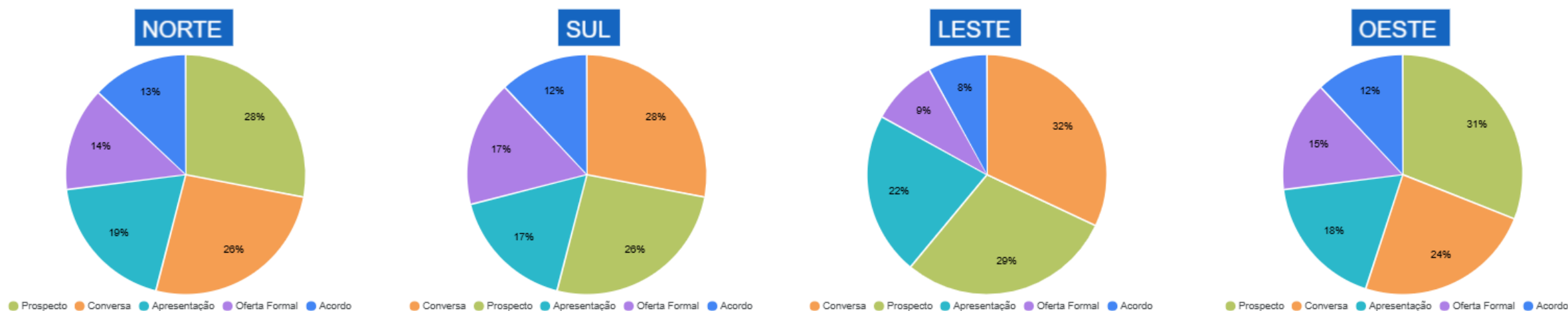
Atividade 12.1 – Makeover (refazer)

- Dados e Dashboard

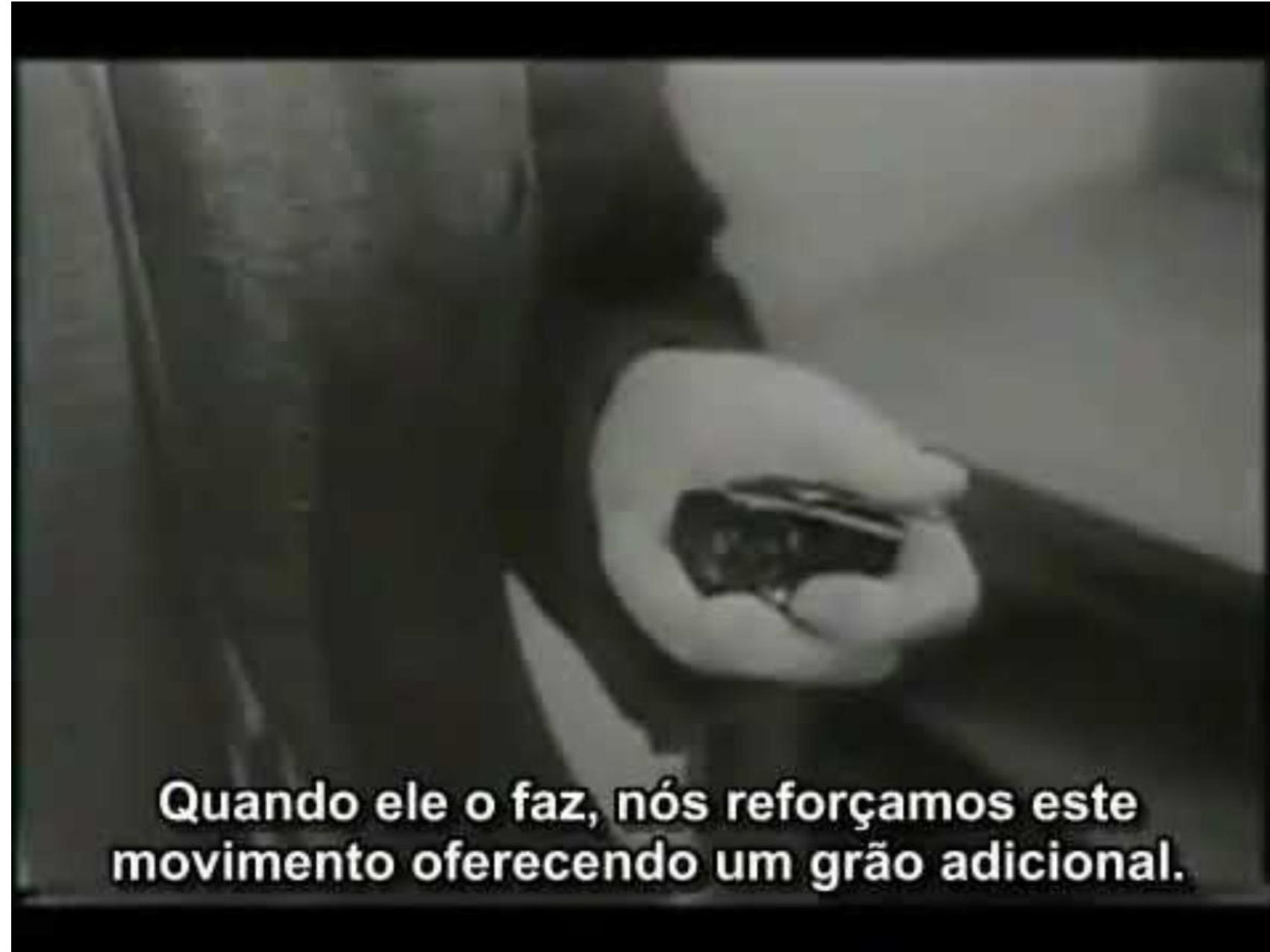
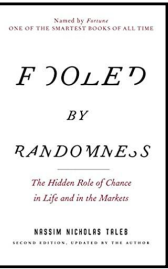
- Seu cliente é uma empresa que tem o pipeline de vendas abaixo
 - ✓ Ele te contratou para entender melhor suas vendas
 - e eventualmente melhora-las
- Melhore o gráfico a fim de torna-lo mais útil (acionável)

Análise do Pipeline de Vendas por Região

Fases: Prospecto, Conversa, Apresentação, Oferta Formal e Acordo

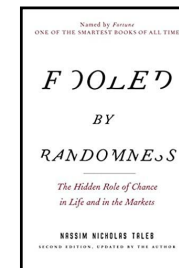


Skinner Pigeon Experiment – Enxergar sinal onde só tem ruído



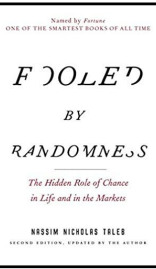
Quando ele o faz, nós reforçamos este movimento oferecendo um grão adicional.

Uma rara doença afetou 3 crianças de uma mesma igreja nos EUA



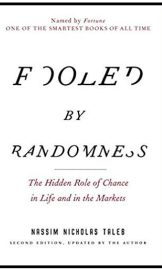
- Narrativa falaciosa
 - Essa igreja/comunidade deve estar causando/transmitindo esta doença
- Esse evento passa a não ser tão improvável
 - quando se contabiliza todos os clusters (igrejas) dos EUA
 - ✓ Prob. 3 ou mais casos ≈ 0.000155
 - Em 50mil igrejas haverá pelo menos 8 com 3 casos
 - $50,000 \times 0.000155 \approx 8$

Viés Cognitivo: Ilusão de Competência



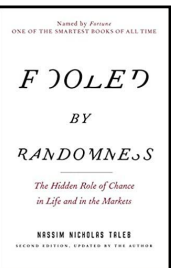
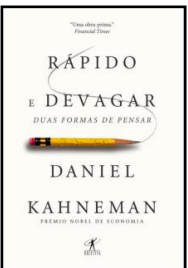
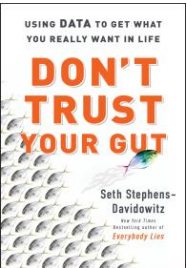
- Superestimar seu nível de conhecimento ou habilidade em um assunto,
 - levando a acreditar que entende algo quando, na verdade, não entende
- Um material familiar ou de fácil compreensão cria uma sensação enganosa de domínio,
 - o que pode dificultar o aprendizado genuíno e causar excesso de confiança,
 - ✓ torna pessoas menos propensas a se envolver em estratégias ativas de aprendizagem que solidificariam sua compreensão.

Viés Cognitivo: Ilusão de Competência



- Trading de ações
 - Skinner “aciona o botão” ou outra fonte de ruído
 - ✓ Ao mesmo tempo em que um trader lucra com ações -> Eu dei causa (competência)
 - Nassim Taleb: há momentos num mercado em crescimento que quase todos os papeis da Bolsa estão em ascensão
 - Está atribuindo para si um evento de causa externa
- Sucesso no empreendedorismo, segundo Kahneman
 - Sucesso = talento + sorte; E
 - ✓ Grande sucesso = um pouco mais de talento + muita sorte
 - Narrativas de sucesso são contam histórias coerentes no retrospecto,
 - mas ignoram “não-eventos” e caminhos onde um pequeno acaso teria mudado tudo, gerando a ilusão de causalidade

Base Rates



THE SEXY PATH TO QUICK FAILURE

Field	Median Time a Business in the Field Lasts (For Comparison, the Average Dentist's Business Lasts 19.5 Years)
Record stores	2.5 years
Amusement arcades	3.0
Hobby, toy, and game stores	3.25
Bookstores	3.75
Clothing stores	3.75
Cosmetics and beauty supply stores	4.0

Source: Luo and Stark (2014)

TOP 5 BUSINESSES WITH THE GREATEST NUMBER OF MILLIONAIRES (NOTE: THIS IS A MISLEADING CHART FOR PICKING A GREAT BUSINESS)

Field	Owners in the Top 0.1 %
Lessors of real estate	12,573
Activities related to real estate	10,911
Automobile dealers	5,236
Offices of physicians	4,711
Restaurants	4,471

Source: online appendix of Smith et al. (2019); this only includes data for S Corporations.

Ignorando Base Rates

Um indivíduo foi descrito por outro como segue:

- Steve é muito tímido e retraído, invariavelmente prestativo, mas com
- pouco interesse nas pessoas ou no mundo real.
- De índole dócil e organizada, tem necessidade de ordem e estrutura, e
- uma paixão pelo detalhe.
- Há maior probabilidade de Steve ser
 - um bibliotecário ou um fazendeiro?

Análise dos Problemas apresentados anteriormente

- Steve
 - Ignoramos os base rates (as taxas de recorrência a priori)
 - ✓ Existem muito mais fazendeiros do que bibliotecários
 - Respondemos intuitivamente outra pergunta
 - ✓ Qual descrição melhor se encaixa no perfil de Steve \neq Qual a probabilidade
- O enunciado te desviou o foco da parte importante do problema?
- Sob uma pergunta difícil para a qual não conhecemos a resposta
 - Mas conhecemos a resposta de outra pergunta mais simples
 - ✓ Intuitivamente, damos como resposta esta outra

Exercício 12.1

- Faça um makeover de um dashboard de um colega seu
 - Do exercício 7.1 ou do exercício 8.1
 - ✓ Um colega pode emprestar o seu dashboard para mais de um colega
 - Tire uma cópia do dashboard antes de compartilhá-lo com permissão de edição
 - ✓ Compartilhe o data source também
- Submeta [aqui](#) no campo “Link para o seu dashboard” a URL do dashboard melhorado
 - No campo “Sua Análise” a URL do dashboard original