



Photo by [rawpixel](#) on [Unsplash](#)

Sistemas de Informação

Para Smart Factories

Sistemas de Informação para Smart Factories
Prof. Pedro Nascimento de Lima

THE SMARTEST WAY TO

Learn Data Science Online

The skills people and businesses need to succeed are changing. No matter where you are in your career or what field you work in, you will need to understand the language of data. With DataCamp, you learn data science today and apply it tomorrow.

[Start Learning For Free](#)

Create Your Free Account

LinkedIn

Facebook

Google

or



Email address



Password

[Create Free Account](#)

By continuing you accept the Terms of Use and Privacy Policy, that your data will be stored outside of the EU, and that you are 16 years or older.



DESAFIO DATA SCIENCE

Datacamp

- Plataforma (em inglês) para aprender Data Science (R, python, SQL, etc.) com mais de 200 cursos interativos;
- Custo ~ R\$ 120,00 / mês;
- Organiza os cursos em “carreiras”;
- Nível de qualidade alto e global.

Nada Mal

The screenshot shows three career paths on DataCamp:

- R Programmer**: 40 hours, 10 courses. Base salary: \$118,000
- Data Scientist with R**: 40 hours, 10 courses. Base salary: \$118,000
- Data Analyst with R**: 36 hours, 10 courses. Base salary: \$62,000

A blue arrow points from the "Nada Mal" text box to the "Data Scientist with R" card.

A Python Developer uses their ~~ammer~~ programming skills to wrangle data and build tools for data analysis.

Learn More >

03/09/2019

Minha Proposta

- Para quem se interessar:
 - Escolha um curso e eu “pago” ele para você;
 - A cada curso que terminar, eu libero outro que você escolher para você realizar;
 - Sugestão: Use o R se não tiver um bom motivo para escolher o python;
 - Você ganha um ponto extra se fizer o curso;
 - Tudo isso é em **paralelo** às disciplinas e aos outros compromissos;

03/09/2019

“Não quero ser programador, porquê isso é útil?”

- R – bibliotecas free para diversos tipos de tarefas:
 - Previsão de Demanda: forecast, prophet;
 - Simulação por Eventos Discretos: simmer, Arena2r;
 - Simulação de Dinâmica de Sistemas: deSolve;
 - Análise de Causalidade: CausalImpact;
 - Algoritmos Genéticos: GA
 - Teoria de Filas: queueing
 - Controle de Qualidade: qicharts
 - Programação Linear: lpSolve
 - Análise de Confiabilidade: Weibull-R
 - e mais de 13 mil outros pacotes.
- Conclusão: Muitos problemas da empresa podem ser resolvidos com dois elementos:
 - Software livre;
 - Alguém que saiba usá-lo!

03/09/2019

Links

- Cursos Disponíveis:
<https://www.datacamp.com/courses>
- Link para Inscrição (até 11/03; apenas os primeiros 70 inscritos receberão acesso):
<https://goo.gl/forms/NXIhRocS2UXCDWpI2>
- Link para Acesso (após 11/03):
<https://www.datacamp.com/enterprise/desafio-data-science-2019-1>

03/09/2019

Regras (para a pontuação extra)

- Para poder participar o aluno deve:
 - Criar uma conta no site www.datacamp.com;
 - Navegar nos cursos e escolher um curso de seu interesse (qualquer curso que ainda não tenha feito);
 - Até a segunda semana de aula, realizar a inscrição pelo google forms;
 - Acessar o link: <http://bit.ly/bonus20191>
 - **Durante seis meses após o início das aulas, você tem acesso a qualquer curso.**
 - Se concluir mais de 4 horas de conteúdo até a prova do grau A, recebe 1 ponto;
 - Se concluir mais de 6 horas de conteúdo até a prova do GB, recebe 1 ponto no GB;
 - Pontos extra só valem para uma disciplina.

03/09/2019



INTRODUÇÃO À DISCIPLINA

Sistemas de Informação para Smart Factories
Prof. Pedro Nascimento de Lima

Aula Introdutória

- Quem Somos?
- Porque estamos aqui?
- Competências e Conteúdos desta Disciplina;
- Plano de Aula e Contrato Pedagógico;
- O que você vai levar desta disciplina.
- Introdução à Análise de Decisão.

03/09/2019

Quem Somos?

- Nome;
- Cidade;
- Curso / Semestre;
- Empresa e setor onde trabalha (se aplicável);

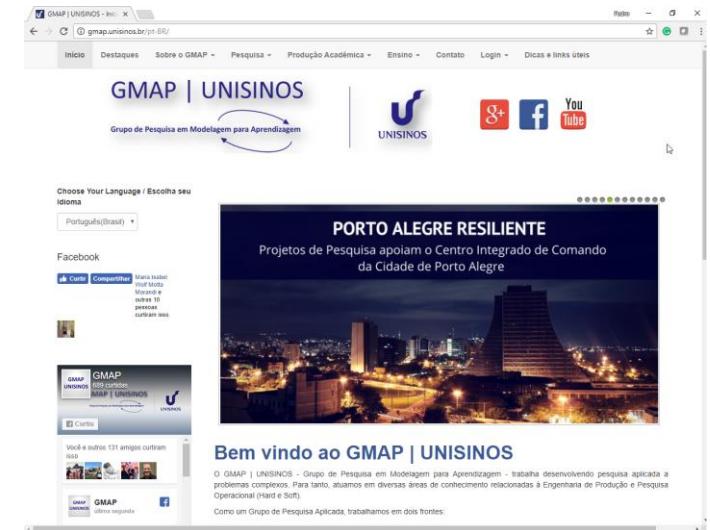
03/09/2019

Sobre o GMAP | UNISINOS

- Grupo de Pesquisa em **Modelagem para Aprendizagem.**



<http://youtube.com/gmapunisinos>



<http://gmap.unisinos.br>

03/09/2019

“Correr com as máquinas vs Correr contra as máquinas”



Erik Brynjolfsson: The key to growth? Race with the machines

99,807 views



SHARE

...

<https://www.youtube.com/watch?v=sod-eJBf9Y0>

03/09/2019

Para pensar...

- Uma máquina pode substituir o seu trabalho hoje ou no futuro?
- Qual é o papel do Engenheiro de Produção frente a este Cenário?
- Que competências você precisa desenvolver para ser competitivo:
 - No Passado;
 - Hoje;
 - No Futuro.
- O que fazer?

03/09/2019

Competências

- Utilizar **tecnologias que promovam a troca de informações** dentro e entre empresas, a fim de melhorar o gerenciamento dos sistemas produtivos;
- Buscar aprendizagem constante, **identificando os avanços tecnológicos** e organizacionais, a fim de **aplica-los no dia-a-dia** da sua prática profissional, contribuindo para a melhoria da produtividade dos sistemas produtivos;
- Utilizar ferramentas computacionais para desempenhar suas atividades;
- **Projetar, gerir e avaliar o fluxo de informações** nas empresas;
- Projetar para a transformação organizacional, estando atento à evolução dos cenários produtivos;
- Associar e aplicar **integradamente os conceitos de sistema de informação, sistema de produção e estratégia de produção**;
- Compreender as diferentes categorias de sistema de informação analisando sua aplicação em sistemas de produção;
- Estabelecer a relação da **gestão de informação empresarial** com novas tecnologias e tecnologias baseadas em conhecimento.

03/09/2019

Como esta Disciplina irá contribuir

Grau A:

- **Análise de Decisão:**
 - Análise de Decisão: Como organizar informações para tomar decisão em situações de **risco e incerteza**;
 - Análise Multicritério: Como organizar informações para decisões considerando **múltiplos critérios de decisão**;
- **Mineração de Dados / Machine Learning:**
 - Introdução, Modelos Lineares;
 - Modelos de Classificação;
- Atividade em Duplas: Lista de Exercícios – Análise de Decisão.
- Atividade Extra: Desafio Data Science

03/09/2019

Como esta Disciplina irá contribuir

Grau B:

- Sistemas de Informação – Conceitos;
- Sistemas de Informação Industriais: Quais são os principais sistemas de informação industriais e qual é o seu propósito.
- Conduzindo um Projeto de Melhoria de Sistemas de Informação:
 - Arquitetura de Sistemas de Informação - Quais são os componentes genéricos dos Sistemas de Informação.
 - Engenharia de Requisitos: Como definir requisitos para um Sistema de Informação.
 - Bases de Dados SQL – Como Projetar;
 - Bases de Dados SQL – Como Consultar.
- Atividade em Grupo: Projetando / Melhorando um Sistema de Informação Empresarial. Será necessária a entrega de um trabalho escrito.

03/09/2019



ANÁLISE DE DECISÃO

Decision Analysis

- Campo do Conhecimento fundado por Ronald Howard, prof. em Stanford, em 1969.
- Técnicas de Análise de Decisão são aplicadas em diversos setores, e industrias, frequentemente as que são de capital intensivo.



Ronald Howard

03/09/2019

Aplicações da Análise de Decisão

- Principais Áreas de Aplicação:
 - Energia;
 - Manufatura e Serviços;
 - Área Médica;
 - Aplicações Militares.

KEEFER, D. L.; KIRKWOOD, C. W.; CORNER, J. L. Perspective on Decision Analysis Applications, 1990–2001. **Decision Analysis**, v. 1, n. 1, p. 4–22, 2004. Disponível em: <<http://pubsonline.informs.org/doi/abs/10.1287/deca.1030.0004>>..

Table 3 Number of Applications Articles by Application Area, with Trends

	Number of articles	
	1970–1989	1990–2001
ENERGY	24	26
Bidding (and pricing)	3	3
Environmental risk	NA*	5
Product and project selection	4	7
Regulation	5	NA
Site selection	8	NA
Strategy	NA	3
Technology choice	4	5
Miscellaneous	NA	3
MANUFACTURING AND SERVICES	16	23
Budget allocation	3	NA
Finance	NA	2
Product planning	4	5
R&D project selection	NA	8
Strategy	5	7
Miscellaneous	4	1
MEDICAL	16	5
MILITARY	NA	13
PUBLIC POLICY	20	13
Standard setting	8	NA
Miscellaneous	12	13
GENERAL	9	6

*Not applicable.

03/09/2019

Exemplo: PETROBRAS: Seu petróleo é nosso

- Você desconfia que existe petróleo em um terreno de sua empresa.
- Sua empresa pode decidir se irá perfurar o terreno ou se irá vendê-lo para a PETROBRAS.
- Você é o Gestor da empresa e, o dono lhe pergunta: “O que eu faço: Vendo este terreno ou vamos perfurar o terreno para tentar encontrar petróleo?”
- Qual seria a sua proposta para abordar este problema?



03/09/2019

Exemplo: PETROBRAS: Seu petróleo é nosso

- Um geólogo lhe informou que ele acredita que a chance de encontrar petróleo neste terreno seja de 25 %.
- Ao mesmo tempo, a PETROBRAS lhe ofereceu 90 k reais pelo terreno, para que ela mesma possa perfurar o terreno.
- O custo de perfuração é de 100 k. Se for encontrado petróleo, a receita esperada seria de 800k, logo o lucro esperado da empresa seria de 700k.
- Qual seria a decisão que você recomendaria?
- (Anote esta decisão).

03/09/2019

Exemplo - PETROBRAS

- Estruturando a Decisão - Tabela de Payoffs:

Alternativa	Prêmio	
	Tem Petróleo	Não Tem Petróleo
Estado da Natureza		
Perfurar para procurar petróleo	700k	-100k
Vender o Terreno	90k	90k
Probabilidade a priori da Condição do Terreno	1 em 4	3 em 4

03/09/2019

Como abordar este problema se...

- A probabilidade de encontrar petróleo for desconhecida?
- A probabilidade de encontrar petróleo não puder ser estimada com precisão?
- Os stakeholders não concordarem sobre a probabilidade de encontrar petróleo?

03/09/2019

Tipos de Decisão

- **Decisão sob Certeza:** Pressupõe-se que apenas um estado da natureza é possível.
- **Decisão Sob Risco:** Diversos estados da natureza são possíveis, e há uma estimativa a respeito da probabilidade de cada um ocorrer.
- **Decisão sob Incerteza:** Diversos estados da natureza são possíveis.

03/09/2019

Estrutura Geral da Análise de Decisão

- O tomador de decisão escolhe uma dentre as possíveis alternativas de decisão.
- A “natureza” escolherá um dentre os possíveis resultados possíveis.
- Cada combinação de uma alternativa de decisão e um estado resulta em um prêmio para o decisor, que é estimada na tabela de payoff.
- A tabela de payoff é utilizada para encontrar uma estratégia para o decisor, de acordo com algum critério de decisão estabelecido.
- O tomador de decisão pode ter informações a respeito da probabilidade relativa de cada um dos estados da natureza possíveis.

03/09/2019

Uso da Probabilidade em Análise de Decisão

- Premissa da Análise de Decisão: Os estados da Natureza **são uma variável aleatória**. (A Natureza geralmente não é “malevolente”).
- O campo do conhecimento que lida com oponentes malevolentes é a Teoria dos Jogos.
- A probabilidade estimada a priori de cada estado acontecer é considerada uma probabilidade subjetiva (subjetiva, pois é inferida a partir de conhecimentos de especialistas).

03/09/2019

Tipos de Análise de Decisão

- **Com Experimentação:** A empresa pode realizar algum tipo de experimento para reduzir a incerteza relacionada ao problema.
- **Sem experimentação:** Não é possível realizar um experimento antes de comprometer recursos da empresa.

03/09/2019

Critérios para Tomada de Decisão

- Em situações de incerteza, é possível utilizar diferentes critérios de decisão para escolher uma alternativa:
 - **Critério MaxiMax:** *Agressivo.* “Escolha a alternativa que tenha o maior prêmio possível.”
 - **Critério MaxiMin:** *Conservador.* “Escolha a alternativa que reduza a sua possível perda.”
 - **Critério de Hurwicz:** *Procure um Equilíbrio.* “Escolha um coeficiente de otimismo, e faça uma média entre o critério MaxiMax e o Maximin”.
 - **Critério de Laplace:** *Faça a Média.* “Se você não sabe o que irá acontecer, parta do pressuposto que todos os estados são equiprováveis.”

03/09/2019

Critério de Decisão MaxiMax

- Para cada uma das possíveis alternativas de decisão, encontre o prêmio máximo ao longo de todos os estados de natureza possíveis.
- Em seguida, encontre o máximo desses prêmios máximos. Escolha a alternativa cujo prêmio mínimo fornece esse máximo.

	Tem petr.	N tem petr.
Perfurar	700	-100
Vender Terreno	90	90

03/09/2019

Critério de Decisão MaxiMax

- Este critério adota um ponto de vista otimista, buscando o melhor resultado possível.

	Tem petr.	N tem petr.	Payoff Máximo
Perfurar	700	-100	700
Vender Terreno	90	90	90
Critério MaxiMax (Máximo dos Máximos)			700

1

2

03/09/2019

Critério de Decisão MaxiMin

- Para cada uma das possíveis alternativas de decisão, encontre o prêmio mínimo ao longo de todos os estados de natureza possíveis.
- Em seguida, encontre o máximo desses prêmios mínimos. Escolha a alternativa cujo prêmio mínimo fornece esse máximo.

	Tem petr.	N tem petr.
Perfurar	700	-100
Vender Terreno	90	90

03/09/2019

Critério de Decisão MaxiMin

- Este critério adota um ponto de vista pessimista, porém fornece a melhor garantia de payoff que pode ser obtido.
- O raciocínio é válido quando se está jogando contra um player inteligente e malevolente.

1

	Tem petr.	N tem petr.	Payoff Mínimo da Alternativa
Perfurar	700	-100	-100
Vender Terreno	90	90	90
Critério MaxiMin (Máximo dos Mínimos)			90

2

03/09/2019

Critério de Laplace

- Faça a média dos payoffs e escolha a alternativa com a melhor média.
- Este critério equivale a dizer que ambos os cenários são igualmente prováveis.

	Tem petr.	N tem petr.	Média dos Payoffs
Perfurar	700	-100	300
Vender Terreno	90	90	90
Critério de Laplace (Máximo das Médias)			300

1

2

03/09/2019

Critério de Hurwicz

- Escolha a alternativa com o melhor payoff médio ponderado, considerando um coeficiente α de otimismo.
- Payoff Ponderado = $\alpha * \text{MelhorPayoff} + (1 - \alpha) * \text{PiorPayoff}$

1

	Tem petr.	N tem petr.	Payoff Ponderado (alfa = 0,7)
Perfurar	700	-100	460
Vender Terreno	90	90	90
Critério de Hurwicz (Máxima Média Pond.)			460

2

03/09/2019

Critério MiniMax Regret (Savage)

- Calcule o Custo de Oportunidade (Regret) de Cada Alternativa, por cenário:
$$CO = \text{Máx Payoff do Cenário} - \text{Payoff da Decisão no Cenário}$$
- Analise o máximo Custo de Oportunidade de cada Alternativa;
- Escolha a alternativa com o menor custo de oportunidade máximo.

	Tem petr.	N tem petr.
Perfurar	700	-100
Vender Terreno	90	90

	Tem petr.	N tem petr.	Máximo Custo de Oportunid.
1	0	190	190
2	610	0	610

03/09/2019

2

Decisão sob Risco

- Normalmente há uma estimativa sobre a probabilidade de ocorrência dos diferentes estados.
- Nestas condições, é possível utilizar ainda os critérios de decisão da probabilidade máxima ou o critério de Bayes.
- O método mais comum de tomada de decisão sobre o risco é o critério de Bayes (ou Valor Monetário Esperado / Máxima Utilidade Esperada).

03/09/2019

Critério de Decisão da Probabilidade Máxima

- Encontre o estado da natureza mais provável. Para este estado, encontre a alternativa de decisão que gere o maior prêmio possível.

	Tem petr.	N tem petr.	Payoff Mínimo da Alternativa
Perfurar	700	-100	-100
Vender Terreno	90	90	90
Probabilidade	0.25	0.75	
Critério Prob. Máxima - Vender Terreno			

03/09/2019

Critério de Decisão de Bayes

- Calcule o Valor Monetário Esperado para cada uma das alternativas, utilizando a probabilidade de ocorrência dos estados da natureza como fatores de ponderação. Encontre a alternativa de decisão com o maior utilidade esperada.
- **No longo prazo**, a regra de Bayes tende a gerar resultados melhores do que os demais critérios.

	Tem petr.	N tem petr.	Valor Monetário Esperado
Perfurar	700	-100	100
Vender Terreno	90	90	90
Probabilidade	0.25	0.75	
Critério de Bayes - Perfurar			

03/09/2019

Exemplo

- Sua empresa está avaliando ampliar a capacidade do setor de prensas (setor que representa a restrição atual da empresa). Após conduzir um estudo de simulação deste investimento, sua empresa calculou a seguinte tabela de payoffs.
- Pressuponha que esta decisão será tomada sob incerteza. Informe qual decisão a empresa deve tomar de acordo com os critérios apresentados anteriormente. (alfa de Hurwicz = 0,3)

Alternativa	Payoff (Valor Presente Líquido da Decisão)	
	Estado	Estado
	Baixa Demanda	Alta Demanda
Manter a Capacidade Atual	100.000,00	120.000,00
Expandir em 25 %	-70.000,00	180.000,00
Prob.	30%	70%

03/09/2019

Problema do Bourbon

Hillier e Lieberman 15.2.3

- Joana Silva é a gerente do Bourbon. Ela precisa reabastecer seu estoque de morangos orgânicos. Seu fornecedor regular é capaz de fornecer quantas caixas ela quiser. Entretanto, pelo fato de esses morangos já estarem muito maduros (e serem orgânicos!), ela precisará vendê-los no dia seguinte e, depois, jogar fora o que não for vendido. Joana estima que será capaz de vender 12, 13, 14 ou 15 caixas amanhã. Ela pode comprar os morangos a US\$ 7 por caixa e vendê-los a US\$ 18 por caixa. Joana agora precisa decidir quantas caixas deve comprar.
- Joana consultou registros anteriores da loja referentes a vendas diárias de morangos. Com base nisso, ela estima que as probabilidades prévias sejam 0,1, 0,3, 0,4, e 0,2 para vendas de 12, 13, 14 e 15 caixas de morangos amanhã.
- Estruture as decisões e estados da natureza e payoffs em uma tabela de Payoffs;
- Quantas Caixas Joana deve Comprar se usar cada um dos critérios de decisão sob incerteza e risco aprendidos em aula?

03/09/2019

Análise de Sensibilidade com a Regra de Bayes

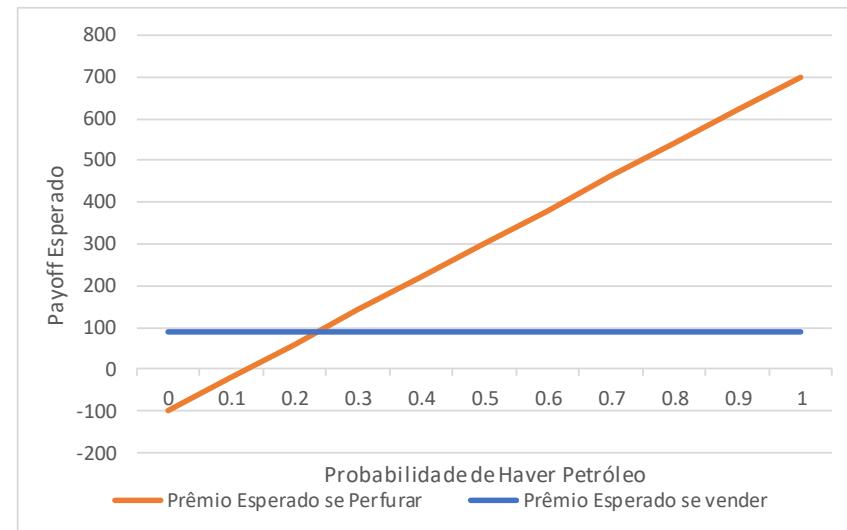
- E se o geólogo estiver errado?
- E se as probabilidades estimadas forem imprecisas?
- O quanto as probabilidades podem estar erradas, e ainda assim faz sentido perfurar para tentar encontrar petróleo?

03/09/2019

Análise de Sensibilidade com a Regra de Bayes

- $E[\text{Prêmio}(\text{Perfuração})] = 700p - 100(1-p) = 800p - 100$
- Ou seja, o Valor esperado da perfuração, em função da probabilidade prévia p de encontrar petróleo ou não é igual a **800 p – 100**

p (Prob. de Haver Petroleo)	Prêmio Esperado se Perfurar	Prêmio Esperado se vender
0	-100	90
0.1	-20	90
0.2	60	90
0.3	140	90
0.4	220	90
0.5	300	90
0.6	380	90
0.7	460	90
0.8	540	90
0.9	620	90
1	700	90



03/09/2019

Encontrando o Ponto de Cruzamento

- Qual deve ser a probabilidade prévia mínima de encontrar petróleo para que a decisão de perfurar o poço seja justificada pela regra de Bayes?
- $E[\text{Prêmio(Perfuração)}] = E[\text{Prêmio(Vender)}]$
- $800 p - 100 = 90$
- $p = 190 / 800$
- **$p = 0,2375$**
- R: A probabilidade de encontrar petróleo deve ser maior do que 23,75 % para que a decisão de perfurar terreno seja justificável pela regra de Bayes.

03/09/2019

Problema da Empresa InovaSINOS

Hillier e Lieberman 15.2.2

- A empresa InovaSINOS desenvolveu um novo produto que lhe permitirá vender um novo componente para a indústria automotiva.
- A empresa também tem a alternativa de vender os direitos de comercialização do produto por 15 milhões.
- Se a empresa optar por fabricar o dispositivo, a lucratividade do negócio dependerá de quantas unidades a empresa conseguirá vender no primeiro ano.
- Atualmente, a empresa estima que tem capacidade de vender até 10 mil unidades do seu produto. Apesar disso, se o produto for bem aceito, acredita-se que é possível vender 100 mil unidades em um ano.
- Para fins de análise, estes são os dois cenários que a empresa deseja avaliar, mas não está claro qual é a probabilidade de cada um deles ocorrer.
- Se a empresa decidir produzir e comercializar o novo produto, ela terá de arcar com o custo de implantação de 6 milhões. O ganho unitário de cada produto é 600 reais.

03/09/2019

Problema da Empresa InovaSINOS

Hillier e Lieberman 15.2.2

- a) Identifique as alternativas de decisão, estados da natureza e tabela de payoffs para este problema.
- b) Qual deve ser a decisão de acordo com os Critérios MaxiMin, MaxiMax, Hurwicz ($\alpha=0,2$) e Laplace?
- c) Qual seria a decisão recomendada por você para a empresa? Porquê?
- d) Utilizando a regra de Bayes, qual deve ser a decisão da empresa se a probabilidade dela vender 10 mil computadores for de 95 %? Qual seria o valor esperado desta decisão?
- e) Esboce um gráfico que represente o prêmio esperado para cada uma das alternativas de decisão versus a probabilidade de vender 10 mil computadores.
- f) Qual deve ser a probabilidade de vender 10 mil computadores para que seja vantajoso produzir o produto internamente segundo a regra de Bayes?

03/09/2019

Problema da Empresa InovaSINOS

Hillier e Lieberman 15.2.2

- a) Identifique as alternativas de decisão, estados da natureza e tabela de payoffs para este problema.

		Unidades Vendidas	
		10000	100000
Alternativa	Vender os Direitos	15.00	15.00
	Fabricar o Dispositivo	0.00	54.00

03/09/2019

Problema da Empresa InovaSINOS

Hillier e Lieberman 15.2.2

b) Qual deve ser a decisão de acordo com os Critérios MaxiMin, MaxiMax, Hurwicz ($\alpha=0,2$) e Laplace?

	Unidades Vendidas		Decisões Segundo Cada Critério			
	10000	100000	MaxiMin	MaxiMax	Laplace	Hurwicz ($\alpha = 0,2$)
Vender os Direitos	15.00	15.00	15.00	15.00	15.00	15.00
Fabricar o Dispositivo	0.00	54.00	0.00	54.00	27.00	10.80

03/09/2019

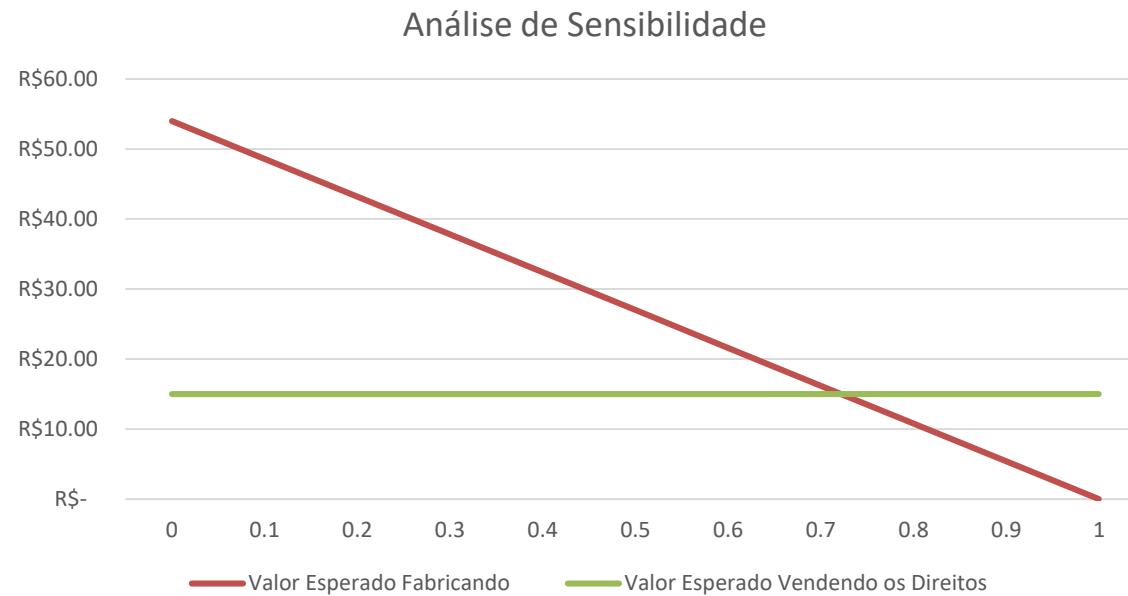
Problema da Empresa InovaSINOS

Hillier e Lieberman 15.2.2

c) Esboce um gráfico que represente o prêmio esperado para cada uma das alternativas de decisão versus a probabilidade de vender 10 mil computadores.

$$E[\text{Prêmio}(\text{Fabricar})] = 54 * 10^6 (1 - p), \quad \text{logo:}$$

$$E(p) = 54 - 54p$$



03/09/2019

Problema da Empresa InovaSINOS

Hillier e Lieberman 15.2.2

Resolução:

b) Equação para o Prêmio Esperado de Fabricar em Função da Probabilidade de vender 10 mil:

$$E[\text{Prêmio}(\text{Fabricar})] = p * 0 + (1 - p) * 54 * 10^6$$

$$E[\text{Prêmio}(\text{Fabricar})] = 54 * 10^6 (1 - p)$$

c) Qual deve ser a decisão da empresa se a probabilidade dela vender 10 mil computadores for de 95 %? Qual seria o valor esperado desta decisão?

$$E[\text{Prêmio}(\text{Fabricar})] = (1 - 0,95) * 54 *= 2,7 * 10^6$$

$$E[\text{Prêmio}(\text{Vender Direitos})] = 15 * 10^6$$

Logo a Decisão deve ser Vender os Direitos, com um prêmio esperado de $15 * 10^6$

d) Qual deve ser a probabilidade de vender 10 mil computadores para que seja vantajoso produzir o produto internamente segundo a regra de Bayes?

$E[\text{Prêmio}(\text{Fabricar})] > E[\text{Prêmio}(\text{Vender Direitos})]$, logo:

$$54 * 10^6 (1 - p) > 15 * 10^6$$

$$(1 - p) > \frac{15 * 10^6}{54 * 10^6}$$

$$1 - p > 0,277$$

$$-p > 0,277 - 1$$

$$-p > -0,722$$

$$p < 0,722$$

03/09/2019

Análise de Decisão com Experimentação

- Podemos utilizar testes adicionais para melhorar as estimativas preliminares dos estados da natureza;
- Os experimentos podem se tornar fontes de informação relevantes para reduzir a incerteza residual sobre o problema;
- Deste modo, diversas decisões podem ser tomadas considerando uma **sequência específica**.

03/09/2019

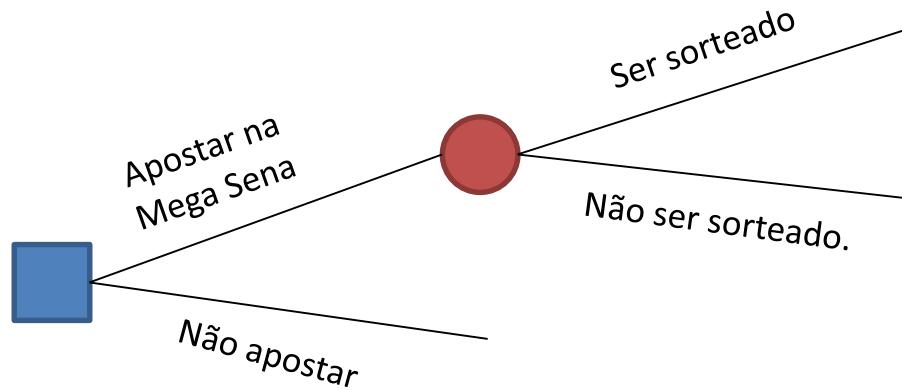
Construindo uma Árvore de Decisão

- Uma Árvore de Decisão possui dois tipos de nós:
 - **Nó de Decisão** : indica que uma decisão deve ser tomada naquele momento.
 - **Nó de Evento**: Indica um evento aleatório, que ocorre naquele momento.

03/09/2019

Construindo uma Árvore de Decisão

- Uma Árvore de Decisão possui dois tipos de nós:
 - – **Nó de Decisão** : indica que uma decisão deve ser tomada naquele momento.
 - – **Nó de Evento**: Indica um evento aleatório, que ocorre naquele momento.



03/09/2019

Passos para Utilizar uma Árvore de Decisão

- **Passo 1:** Organize a sequência de decisões a tomar de modo visual, utilizando nós de decisão e nós de incerteza.
- **Passo 2:** Resolva a Árvore de Decisão da **direita para a esquerda**, uma coluna por vez. Para cada coluna, utilize o procedimento **a** ou **b**, de acordo com o tipo de nó de decisão.
 - Para Nós de Evento:** Calcule o prêmio esperado, multiplicando o prêmio de cada ramificação pela probabilidade de cada ramificação;
 - Para Nós de Decisão:** Compare os prêmios esperados de cada ramificação e escolha a alternativa com maior prêmio esperado.

03/09/2019

Problema da Empresa InovaSINOS

Hillier e Lieberman 15.2.2 –

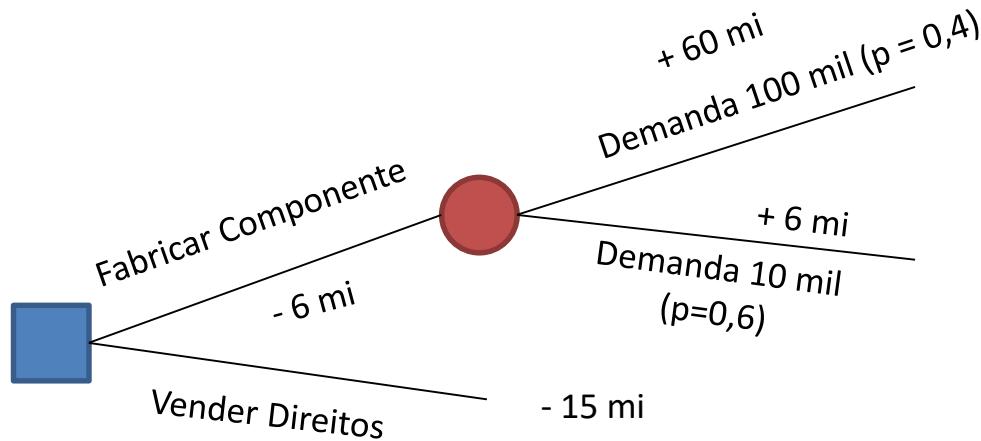
Resolva o Mesmo Problema com uma Árvore de Decisão

- A empresa InovaSINOS desenvolveu um novo produto que lhe permitirá vender um novo componente para a indústria automotiva.
- A empresa também tem a alternativa de vender os direitos de comercialização do produto por **15 milhões**.
- Se a empresa optar por fabricar o dispositivo, a lucratividade do negócio dependerá de quantas unidades a empresa conseguirá vender no primeiro ano.
- Atualmente, a empresa estima que tem capacidade de vender até 10 mil unidades do seu produto. Apesar disso, se o produto for bem aceito, acredita-se que é possível vender 100 mil unidades em um ano.
- Para fins de análise, estes são os dois cenários que a empresa deseja avaliar, mas não está claro qual é a probabilidade de cada um deles ocorrer.
- Se a empresa decidir produzir e comercializar o novo produto, ela terá de arcar com o custo de implantação de **6 milhões**. O ganho unitário de cada produto é **600 reais**.
- **Pressuponha que a Probabilidade Estimada de Vender 100 mil unidades é de 60%.**

03/09/2019

Exemplo Realizado em Aula

Passo 1 – Construir a Árvore (->>)

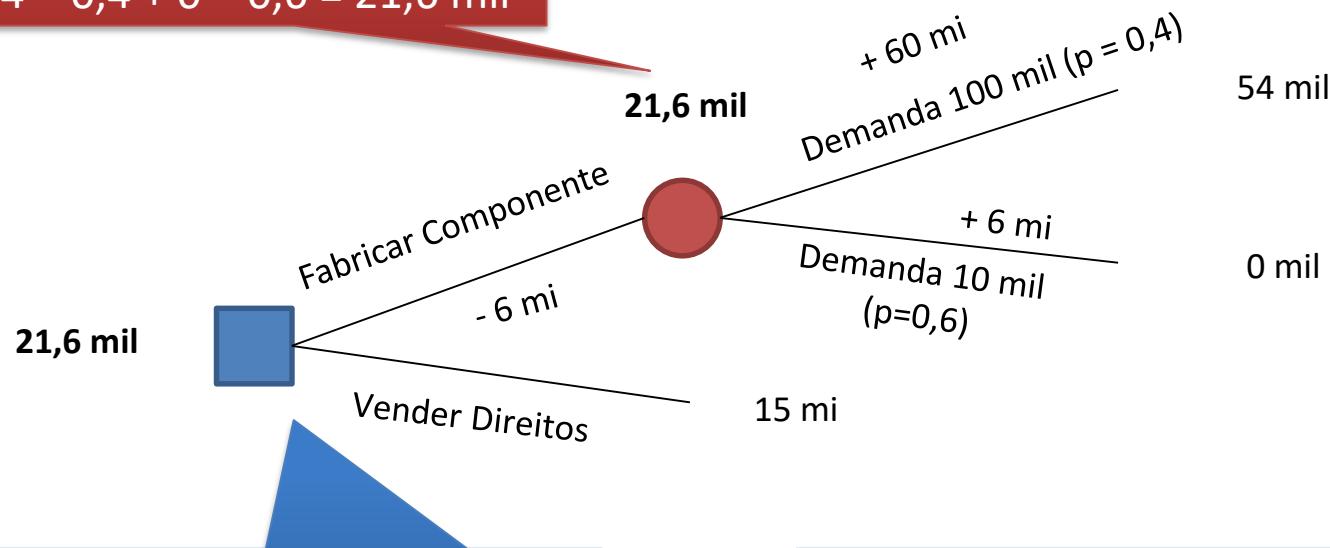


03/09/2019

Exemplo Realizado em Aula

Passo 2 – Resolver a Árvore (<<-)

Nó de Evento: Calcular Valor Esperado:
 $54 * 0,4 + 0 * 0,6 = 21,6 \text{ mil}$



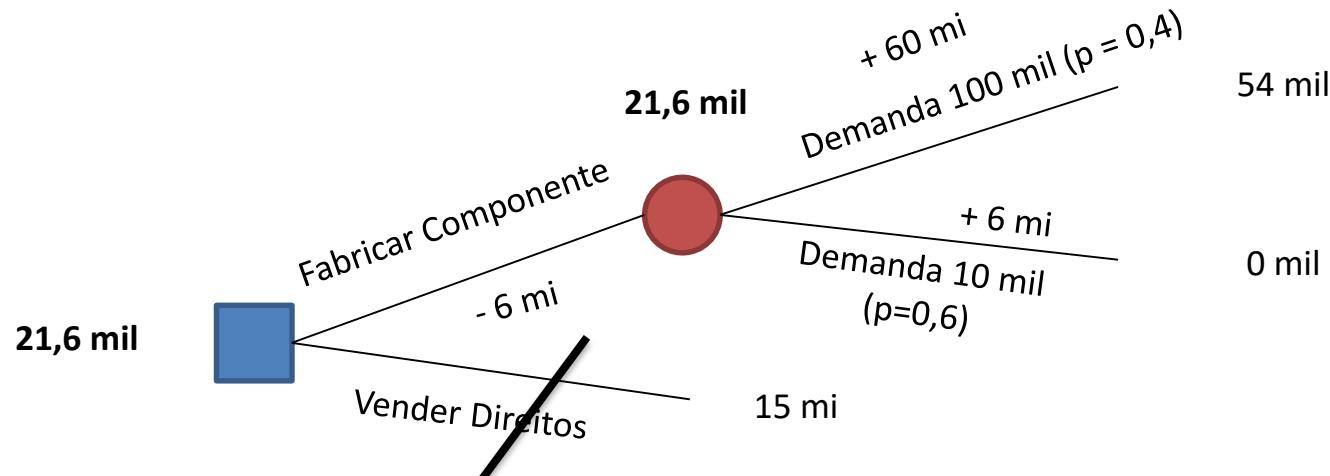
Nó de Decisão: Escolher o Maior Valor Esperado (Regra de Bayes!)
Escolho Fabricar Componentes

Decisão: Fabrique o Componente com o valor esperado de 21,6 mil (e assuma as consequências!)

03/09/2019

Exemplo Realizado em Aula

Passo 2 – Resolver a Árvore (<<-)



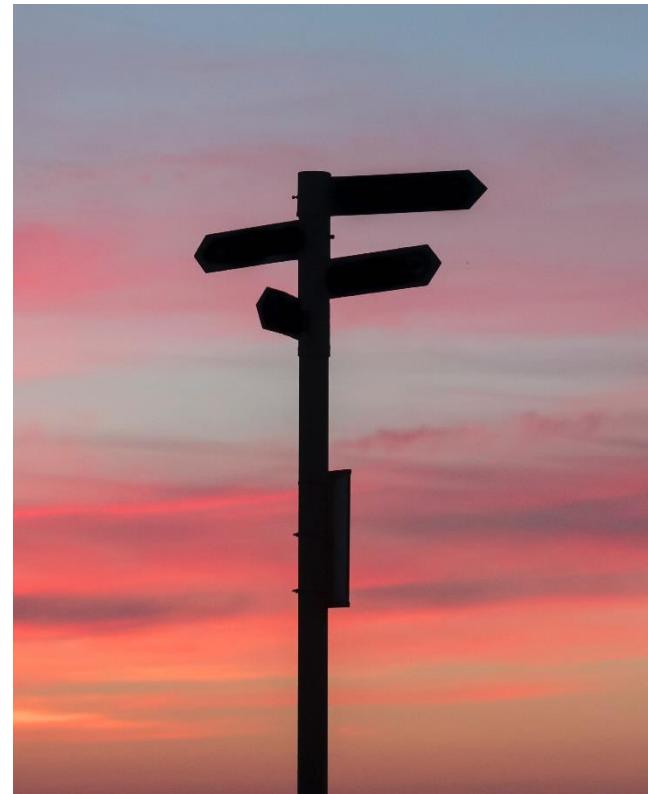
03/09/2019



VALOR DA INFORMAÇÃO

Quando Há Incerteza, Qual é o Valor da Informação?

- Quanto Vale saber...
 - Se o seu fornecedor irá ou não aumentar o preço...
 - Se um novo produto terá bom desempenho no mercado ou não...
 - Se há ou não petróleo em um poço...
 - Se o dólar irá disparar ou não?



03/09/2019

Decidindo entre realizar ou não experimentos

- Quando há incerteza em relação ao mercado, uma decisão frequente é se deve-se ou não realizar **experimentos** em relação às variáveis incertas;
 - Exemplo da PETROBRAS: Há indícios da existência de petróleo?
 - Novos Produtos: Qual será a aceitação do mercado em relação ao novo produto?
- Nestas condições a Análise de Decisão oferece a análise do **Valor da Informação** como ferramenta para suportar estas decisões.

03/09/2019

Quando há Incerteza, qual é o valor da informação?

- Considere novamente o exemplo da INOVASinos, considerando estas probabilidades:

Unidades Vendidas		
	10000	100000
Vender os Direitos	15.00	15.00
Fabricar o Dispositivo	0.00	54.00
Probabilidade	40%	60%

- Se fosse possível realizar uma pesquisa de mercado que indique com precisão qual será o cenário que irá acontecer, você estaria disposto a pagar que quantia para ter a informação perfeita?
 - 5 milhões
 - 10 milhões
 - 15 milhões

03/09/2019

O Valor da Informação

- **Valor Esperado da Informação Perfeita**
 - Quando é possível, de alguma forma, reduzir a incerteza obtendo informação a respeito de qual cenário irá se concretizar, é útil saber até quanto se deve pagar pela informação;
 - O cálculo do VEIP oferece um **limite superior** do valor potencial do experimento.
 - Se o custo do experimento for maior do que o Valor Esperado da Informação Perfeita, então não faz sentido realizar o experimento.

03/09/2019

Valor Esperado da Informação Perfeita (VEIP)

- O Valor Esperado da Informação Perfeita (VEIP) pode ser calculado por esta equação:
- $VEIP = PEIP - VME$
 - Onde:
 - PEIP – Prêmio Esperado com a Informação Perfeita.
 - VME – Valor Monetário Esperado (calculado com a Regra de Bayes).

03/09/2019

Calculando o VME

	Unidades Vendidas	
	10000	100000
Vender os Direitos	15.00	15.00
Fabricar o Dispositivo	0.00	54.00
Probabilidade	40%	60%

- Neste Caso o VME (Valor Monetário Esperado), sem experimentação será:
 - $E[\text{Prêmio(Vender)}] = 15 * 0,4 + 15 * 0,6 = 15$
 - $E[\text{Prêmio(Fabricar)}] = 0 * 0,4 + 54 * 0,6 = 32,4$
 - Logo, o Valor Monetário Esperado é **32,4**, Fabricando o Dispositivo.

03/09/2019

Calculando o PEIP

Unidades Vendidas		
	10000	100000
Vender os Direitos	15.00	15.00
Fabricar o Dispositivo	0.00	54.00
Probabilidade	40%	60%

- O Prêmio Esperado com a Informação Perfeita é calculado considerando que sabe-se qual será o cenário antes de se decidir, logo:
 - $PEIP = 0,4 * 15 + 0,6 * 54 = 38.4$

03/09/2019

Calculando o VEIP

- $VEIP = PEIP - VME$
- $VEIP = 38,4 - 32,4 = 6$ milhões
- Ou seja, não faz sentido gastar mais do que 6 milhões para “testar” este mercado.

03/09/2019

Exemplo 2 – Valor da Informação

- Considerando esta tabela de Payoffs (em milhares de Reais), responda:

Alternativa	Cenário 1	Cenário 2	Cenário 3
A1	6	1	1
A2	1	3	0
A3	4	1	2
Prob. Prévia	30 %	40 %	30 %

- Qual decisão deve ser tomada de acordo com o critério de Bayes?
- Qual seria o Prêmio Esperado se o decisor soubesse qual o cenário iria se concretizar antes de tomar a decisão?
- Qual é o VEIP?
- Se alguém lhe oferecesse a informação perfeita por R\$ 1200,00, você deveria ou não comprar esta informação segundo o critério do VEIP?

03/09/2019

Exemplo 2 – Valor da Informação

a) Qual decisão deve ser tomada de acordo com o critério de Bayes?

- $E[\text{Prêmio}(A1)] = 0,3 * 6 + 0,4 * 1 + 0,2 * 1 = 2,5$
- $E[\text{Prêmio}(A2)] = 0,3 * 1 + 0,4 * 3 + 0,2 * 0 = 1,5$
- $E[\text{Prêmio}(A3)] = 0,3 * 4 + 0,4 * 1 + 0,2 * 2 = 2,2$
- R: Escolha a alternativa 1, pois tem o maior valor esperado ($VME = 2,5$).

03/09/2019

Exemplo 2 – Valor da Informação

- a) b) Qual seria o Prêmio Esperado se o decisor soubesse qual o cenário iria se concretizar antes de tomar a decisão?

Alternativa	Cenário 1	Cenário 2	Cenário 3
A1	6	1	1
A2	1	3	0
A3	4	1	2
Máximo	6	3	2
Prob. Prévia	30 %	40 %	30 %

$$PEIP = 6 * 0,3 + 3 * 0,4 + 2 * 0,3 = 3,6$$

03/09/2019

Exemplo 2 – Valor da Informação

c) Qual é o VEIP?

$$\text{VEIP} = \text{PEIP} - \text{VME}$$

$$\text{VEIP} = 3,6 - 2,5 = 1,1$$

d) Não, pois $\text{VEIP} < 1,2$

03/09/2019



TEORIA DA UTILIDADE

Testando a Teoria da Utilidade

1. O que você prefere:

- a) Uma chance de 50 % de ganhar R\$ 100.000,00
- b) Ganhar R\$ 40.000,00 com certeza.

A maioria responde que prefere ganhar R\$ 40 mil com certeza, embora o valor esperado da primeira opção seja R\$ 50 mil.

03/09/2019

Testando a Teoria da Utilidade

1. O que você prefere (seguro do carro):

- a) Pagar R\$ 2000,00 em um seguro para um carro de R\$ 30.000,00, considerando um risco de 3 % de sofrer um sinistro.
- b) Não pagar um seguro.

Muitas pessoas preferem pagar um seguro, embora o valor esperado de não ter um seguro seja maior do que o de ter o seguro.

Valor esperado com o seguro: $- 2000 + 0,03 * 30k = - 1,1$ mil

Valor esperado sem o seguro: $- 0,03 * 30 k = - 0,9$ mil

03/09/2019

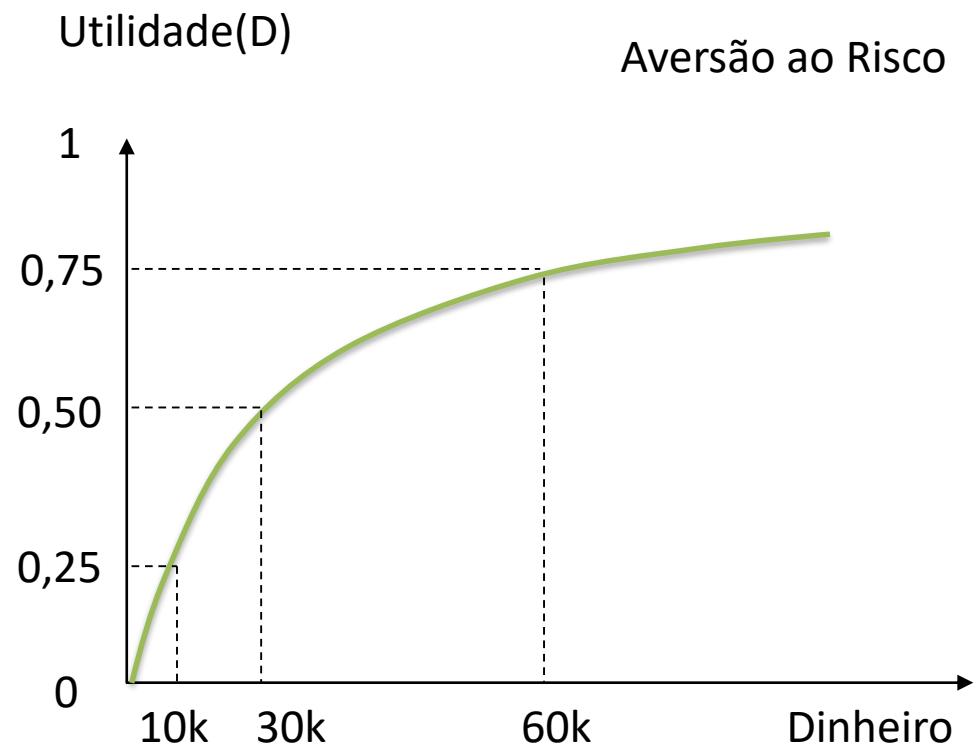
Teoria da Utilidade

- Quando utilizamos a Regra de Bayes, partimos do pressuposto que o prêmio esperado seja a medida adequada para avaliar cada uma das alternativas;
- Nos exemplos anteriores, utilizando a Regra de Bayes, a decisão correta seria sempre apostar no jogo, e não fazer o seguro.
- Estas diferenças estão relacionadas às preferências dos decisores em **relação ao risco**.
- A Teoria da Utilidade fornece uma maneira de considerar as nossas preferências em relação ao risco, por meio da função de **Utilidade Monetária**.

03/09/2019

Função de Utilidade Monetária

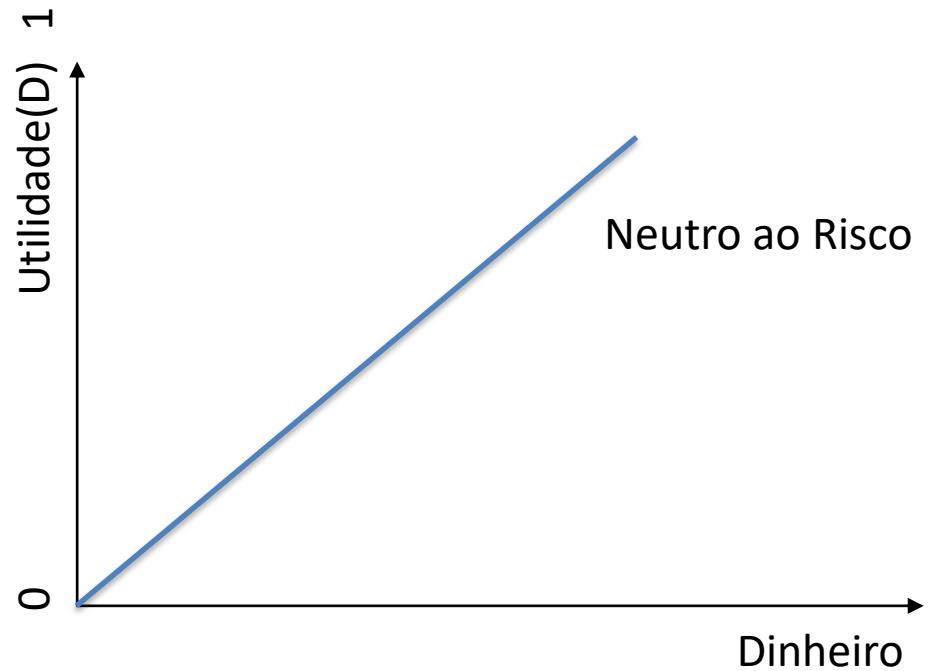
- Alguém com esta função de utilidade prefere ter 30 k “o dobro do” que ter 10 k.
- Isto ocorre pois as necessidades básicas desta pessoa são atendidas por R\$ 10 k.
- Este indivíduo possui utilidade marginal decrescente para o dinheiro.



03/09/2019

Funções de Utilidade Monetária

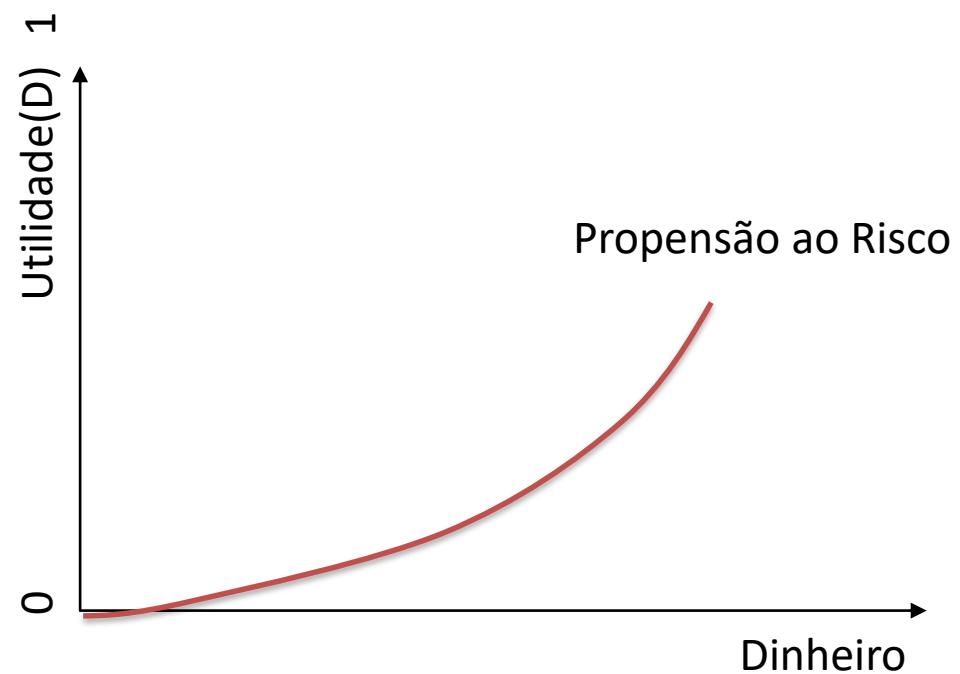
- Um indivíduo neutro ao risco possui esta função de utilidade monetária:



03/09/2019

Funções de Utilidade Monetária

- Este indivíduo possui propensão ao risco. Para ele, pouco dinheiro possui pouca utilidade.
- Este indivíduo possui utilidade marginal crescente para o dinheiro.



03/09/2019

Caracterizando a Função de Utilidade

Método da Loteria Equivalente

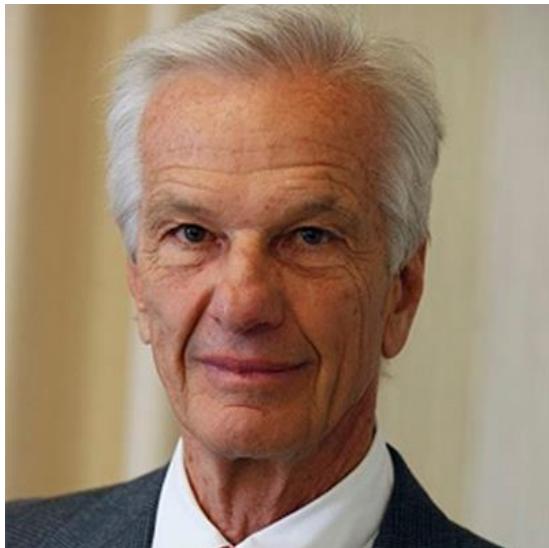
1. Encontre o maior prêmio potencial, e atribua a ele a maior utilidade (1);
2. Encontre o menor prêmio potencial, e atribua a ele a menor utilidade (0);
3. Para determinar a utilidade de outro prêmio potencial, oferecer ao tomador de decisão duas alternativas hipotéticas:
 1. Obter o prêmio máximo com probabilidade p , e obter um prêmio mínimo com probabilidade $1-p$.
 2. Obter definitivamente um prêmio M.
4. Perguntar: Qual valor de p o deixaria indiferente em escolher qualquer uma das alternativas?

Obs: A utilidade do prêmio M é: $U(M) = p * 1 + 1-p$

03/09/2019

Exemplo 1 – Teoria da Utilidade

- Dois voluntários....



Jorge Paulo Lehman
Patrimônio: USD 24 bi



Eike Batista
Patrimônio: USD 116 mi
(Caiu de USD 32 bi)

03/09/2019

O quê você prefere....

50 milhões com
certeza

500 milhões com
probabilidade p

Qual probabilidade “ p ” te
tornaria indiferente?

03/09/2019

O quê você prefere....

100 milhões com certeza

500 milhões com probabilidade p

Qual probabilidade “ p ” te tornaria indiferente?

03/09/2019

O quê você prefere....

200 milhões com certeza

500 milhões com probabilidade p

Qual probabilidade “ p ” te tornaria indiferente?

03/09/2019

O quê você prefere....

400 milhões com certeza

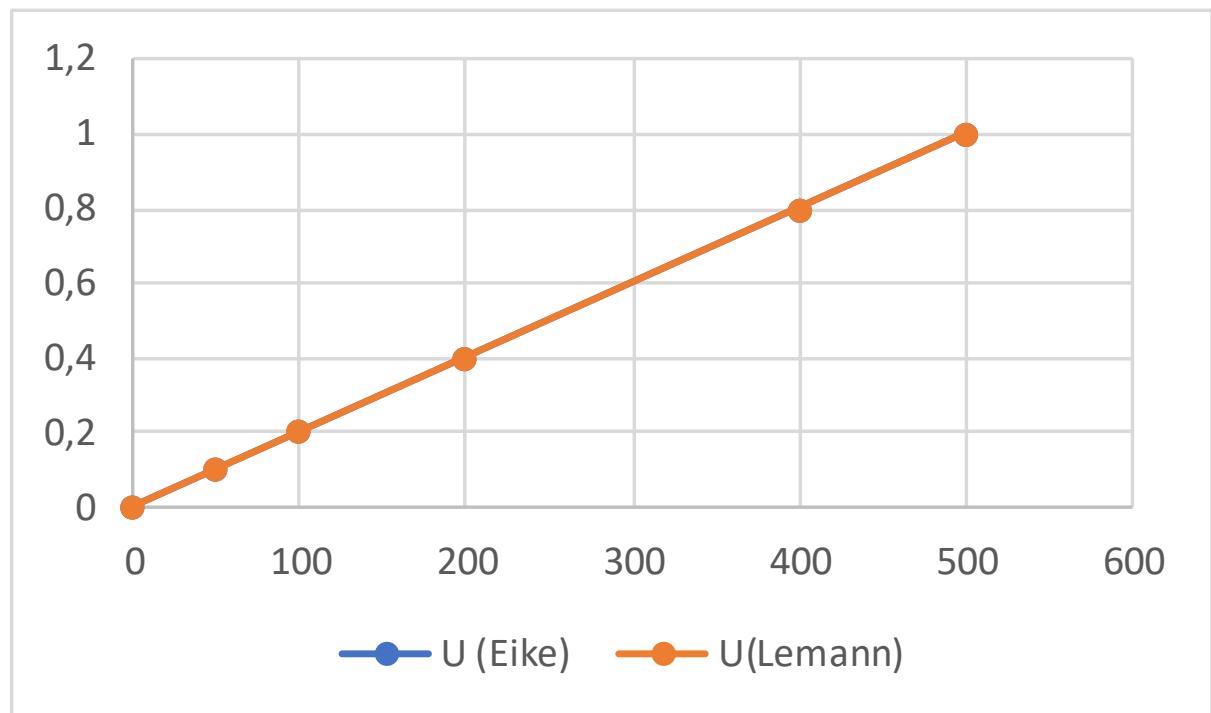
500 milhões com probabilidade p

Qual probabilidade “ p ” te tornaria indiferente?

03/09/2019

Gráfico da Utilidade se os dois fossem decisores “Racionais”

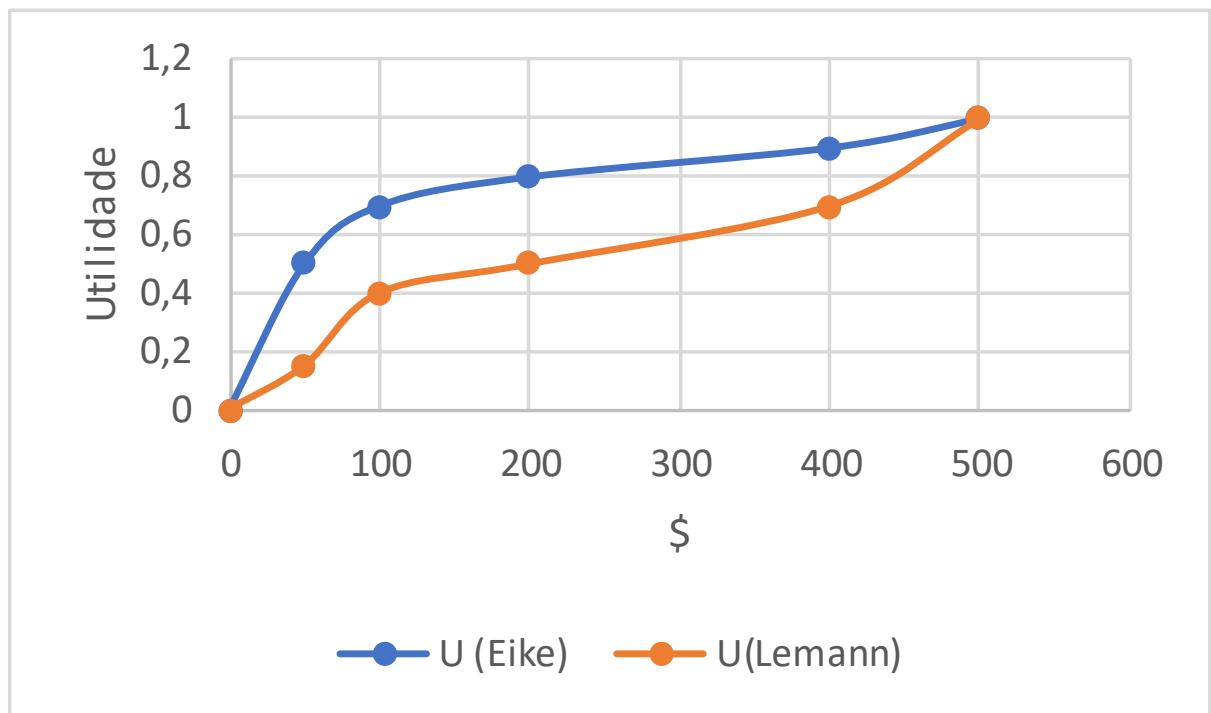
\$	U (Eike)	U(Lemann)
0	0	0
50	0,10	0,10
100	0,20	0,20
200	0,40	0,40
400	0,80	0,80
500	1	1



03/09/2019

Criando uma Função de utilidade

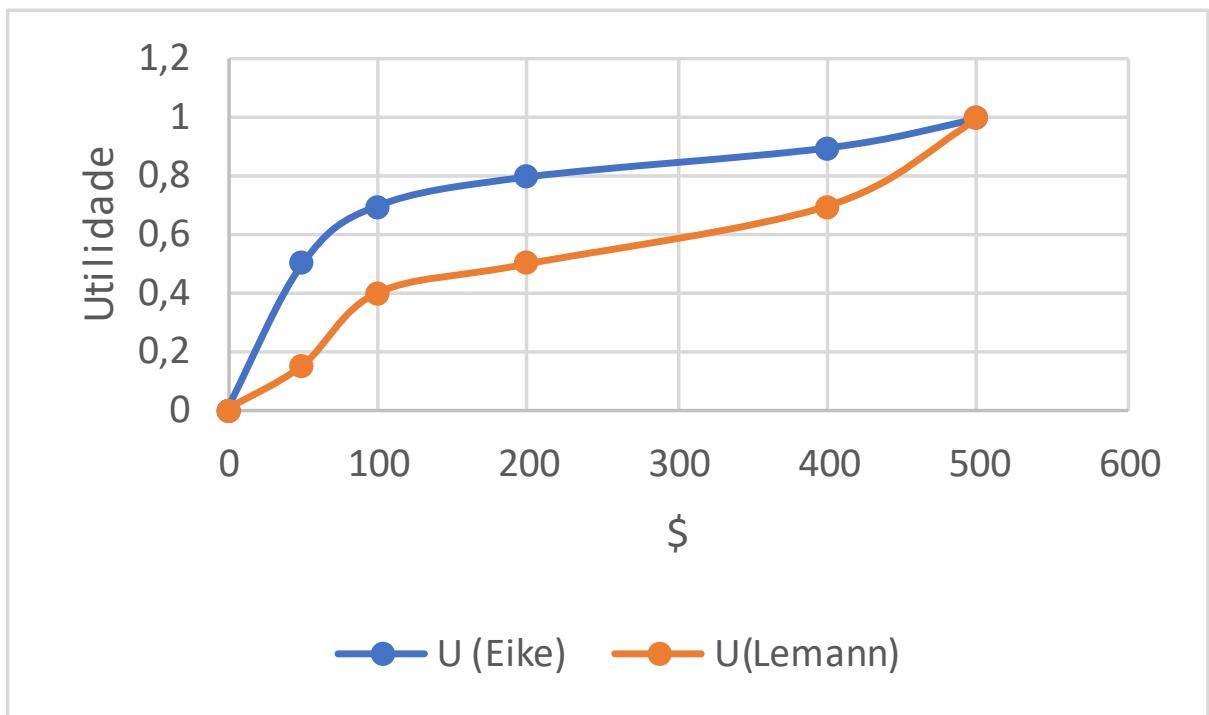
\$	U (Eike)	U(Lemann)
0	0	0
50	0,50	0,15
100	0,7	0,4
200	0,8	0,5
400	0,9	0,7
500	1	1



03/09/2019

Criando uma Função de utilidade

\$	U (Eike)	U(Lemann)
0	0	0
50	0,50	0,15
100	0,7	0,4
200	0,8	0,5
400	0,9	0,7
500	1	1



03/09/2019

Exemplo 2 – Teoria da Utilidade

- Imagine que você deseja traduzir o problema da INOVASinos para considerar a função de utilidade do dono desta empresa.
- Após utilizar o método da loteria equivalente, você calculou as seguintes utilidades, conforme esta tabela de payoff.
- Indique qual seria a decisão de sua empresa, aplicando a regra de Bayes sobre a Utilidade.

Unidades Vendidas		
	10000	100000
Vender os Direitos	0,7	0,7
Fabricar o Dispositivo	0,00	1
Probabilidades	50 %	50 %

03/09/2019

Aplicando a Teoria da Utilidade

- Para aplicar a teoria da utilidade, é necessário:
 - identificar a utilidade do decisor para cada um dos prêmios possíveis;
 - aplicar a regra de bayes considerando o valor de utilidade, ao invés do valor monetário.
 - escolher a decisão com a maior **utilidade esperada**.

03/09/2019

Aplicando a Teoria da Utilidade

- $U(\text{Vender}) = 0,7 * 0,5 + 0,7 * 0,5 = 0,7$
- $U(\text{Fabricar}) = 0 * 0,5 + 1 * 0,5 = 0,5$
- Logo, a decisão para Maximizar a **Utilidade Esperada** seria vender os Direitos.

03/09/2019

Exemplo para Revisão

Adaptado de Hillier e Lieberman 15.2-7

- Você foi contratado como gerente de uma empresa agrícola que possui 1000 acres de terra cultivável. Esta fazenda apenas planta uma cultura por estação.
 - A produtividade (sacas / acre) desta fazenda, bem como a margem líquida por saca dependem do clima da próxima estação.
 - O agricultor precisa decidir qual cultura escolher para a próxima estação, que possui um clima incerto. Uma empresa de consultoria afirma que pode predizer o clima com maior precisão, ao custo de 10.000 reais.
 - A tabela a seguir apresenta a produtividade esperada e margem líquida para cada uma das quatro culturas .
- a) Defina as alternativas de decisão, estados da natureza e tabela de prêmios para este problema.
b) Defina a melhor decisão de acordo com a Regra de Bayes.
c) Indique se o agricultor deve contratar a consultoria ou não com base no VEIP.

Clima	Produtividade Esperada (sacas / acre)					Prob. Prévia
	Cultura 1	Cultura 2	Cultura 3	Cultura 4		
Seco	30	25	40	60	-	0.5
Moderado	50	30	35	60	-	0.3
Úmido	60	40	35	60	-	0.2
Margem / saca	R\$ 3.00	R\$ 4.50	R\$ 3.00	R\$ 1.50	-	

03/09/2019

Exemplos para Revisão

Adaptado de Hillier e Lieberman 15.2-7

- a) Defina as alternativas de decisão, estados da natureza e tabela de prêmios para este problema.

Clima	Payoffs							
	Cultura 1	Cultura 2		Cultura 3		Cultura 4		Prob. Prévia
Seco	R\$ 90,000	R\$ 112,500	R\$ 120,000	R\$ 90,000	0.5			
Moderado	R\$ 150,000	R\$ 135,000	R\$ 105,000	R\$ 90,000	0.3			
Úmido	R\$ 180,000	R\$ 180,000	R\$ 105,000	R\$ 90,000	0.2			

Atenção! Nem sempre a tabela apresentada é a tabela de payoffs.
Neste caso: $90.000 = 30 * 3 * 1000$

03/09/2019

Exemplos para Revisão

Adaptado de Hillier e Lieberman 15.2-7

b) Utilize a Regra de Bayes para definir qual cultura deve ser plantada.

Clima	Payoffs								Prob. Prévia
	Cultura 1	Cultura 2		Cultura 3		Cultura 4			
Seco	R\$ 90,000	R\$ 112,500	R\$ 120,000	R\$ 90,000	0.5				
Moderado	R\$ 150,000	R\$ 135,000	R\$ 105,000	R\$ 90,000	0.3				
Úmido	R\$ 180,000	R\$ 180,000	R\$ 105,000	R\$ 90,000	0.2				
Valor Esperado	R\$ 126,000	R\$ 132,750	R\$ 112,500	R\$ 90,000	-				

03/09/2019

Exemplos para Revisão

Adaptado de Hillier e Lieberman 15.2-7

Clima	Payoffs								Prob. Prévia
	Cultura 1		Cultura 2		Cultura 3		Cultura 4		
Seco	R\$ 90,000	R\$ 112,500	R\$ 120,000	R\$ 90,000	0.5				
Moderado	R\$ 150,000	R\$ 135,000	R\$ 105,000	R\$ 90,000	0.3				
Úmido	R\$ 180,000	R\$ 180,000	R\$ 105,000	R\$ 90,000	0.2				
Valor Esperado	R\$ 126,000	R\$ 132,750	R\$ 112,500	R\$ 90,000	-				

c) Encontre o VEIP

$$\text{VEIP} = \text{PEIP} - \text{VME}$$

$$\text{PEIP} = 120 * 0,5 + 0,3 * 135 + 0,2 * 180 = 141.000$$

$$\text{VEIP} = 141 - 132,75 = 8,25 \text{ mil reais},$$

ou seja, o agricultor não deveria gastar mais do que 8,25 mil reais para obter mais informações sobre o cenário mais provável.

03/09/2019

Exemplos para Revisão

Adaptado de Hillier e Lieberman 15.5-7

- A sua empresa está estudando a introdução de um novo produto do qual se acredita ter uma chance de 60 % de ser bem sucedido.
- Se o produto for bem sucedido, o lucro da empresa será de 40 milhões. Se o produto não for bem sucedido, a empresa terá uma perda de 15 milhões.
 - a) Identifique as alternativas de decisão, estados da natureza e tabela de prêmios.
 - b) Aplique os critérios de Maximin, Maximax, Laplace e Hurwicz ($\alpha = 0,8$), e defina as suas respectivas decisões, pressupondo que esta decisão será tomada sem considerar as probabilidades prévias.
 - c) Aplique a regra de decisão de Bayes e determine a decisão ótima com base neste critério;
 - d) Encontre o EVPI (Valor Esperado da Informação Perfeita).
 - e) Realize uma análise de sensibilidade da probabilidade de sucesso do produto, e determine a partir de qual probabilidade faz sentido implantar o produto no mercado.

03/09/2019

Exemplos para Revisão

Adaptado de Hillier e Lieberman 15.5-7

- Identifique as alternativas de decisão, estados da natureza e tabela de prêmios.
- Aplique os critérios de Maximin, Maximax, Laplace e Hurwicz ($\alpha = 0,8$), e defina as suas respectivas decisões, pressupondo que esta decisão será tomada sem considerar as probabilidades prévias.

Alternativas	Bem Sucedido	Mal Sucedido	Maximin	MaxiMax	Laplace	Hurwicz ($\alpha = 0,8$)
Introduzir Novo Produto	40	-15	-15	40	12.5	29
Não Introduzir	0	0	0	0	0	0

03/09/2019

Exemplos para Revisão

Adaptado de Hillier e Lieberman 15.5-7

- c) Aplique a regra de decisão de Bayes e determine a decisão ótima com base neste critério;
- d) Encontre o VEIP (Valor Esperado da Informação Perfeita).

Alternativas	Bem Sucedido	Mal Sucedido	VME (Regra de Bayes)
Introduzir Novo Produto	40	-15	18
Não Introduzir	0	0	0
Prob. Prévia	60%	40%-	

$$VEcIP = 40 * 0,6 + 0 * 0,4 = 24$$

$$VEIP = VEcIP - VME = 24 - 18 = 6 \text{ milhões}$$

03/09/2019

Exemplos para Revisão

Adaptado de Hillier e Lieberman 15.5-7

- e) Determine a partir de qual probabilidade p de sucesso do produto faz sentido implantar o produto no mercado. Represente o problema graficamente.

$$E[\text{Introduzir}] = 40 p - 15 * (1-p)$$

$$E[\text{Introduzir}] = 40 p - 15 + 15 p$$

$$E[\text{Introduzir}] = 55 p - 15$$

$$E[\tilde{N} \text{ Introduzir}] = 0$$

Faz sentido introduzir o produto se o valor esperado de Produzir for maior do que o valor esperado de não introduzir:

$E[\text{Introduzir}] > E[\tilde{N} \text{ Introduzir}]$, ou seja:

$$55 p - 15 > 0$$

$$p > 15 / 55$$

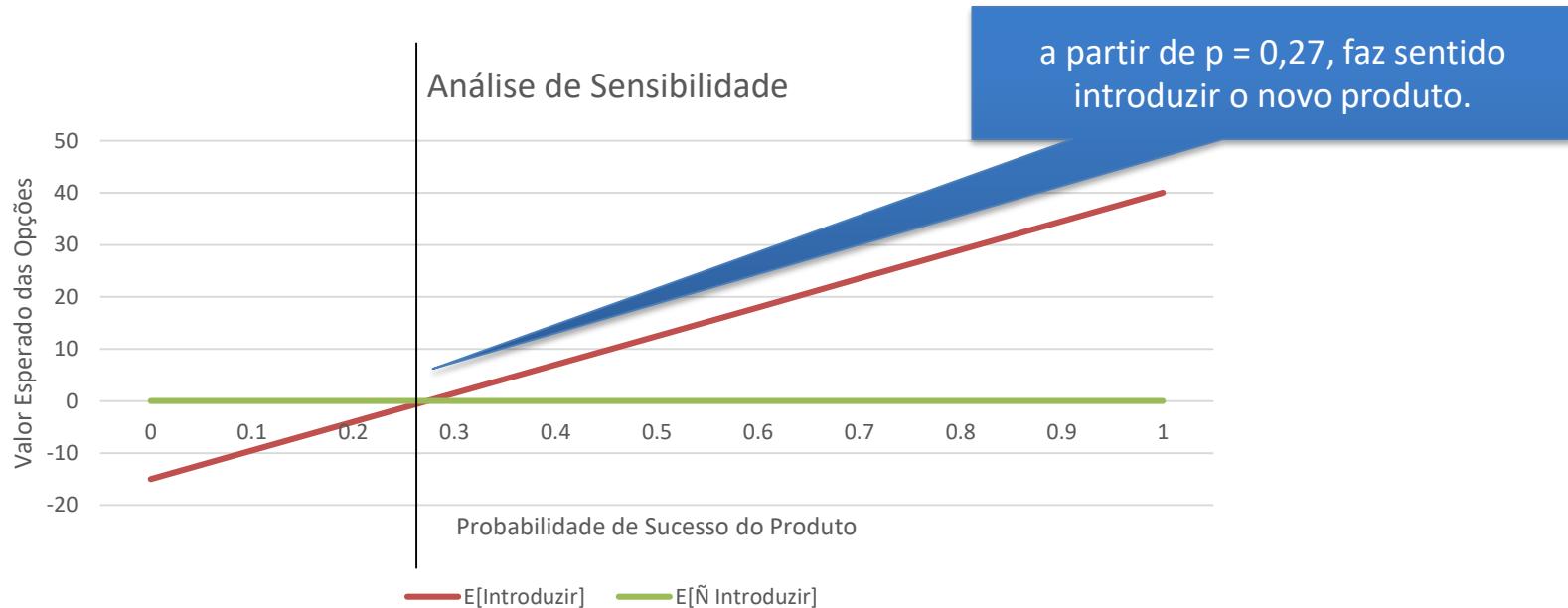
$$p > 27,27 \%$$

03/09/2019

Exemplos para Revisão

Adaptado de Hillier e Lieberman 15.5-7

- e) Determine a partir de qual probabilidade p de sucesso do produto faz sentido implantar o produto no mercado. Represente o problema graficamente.



03/09/2019

Exemplo para Revisão

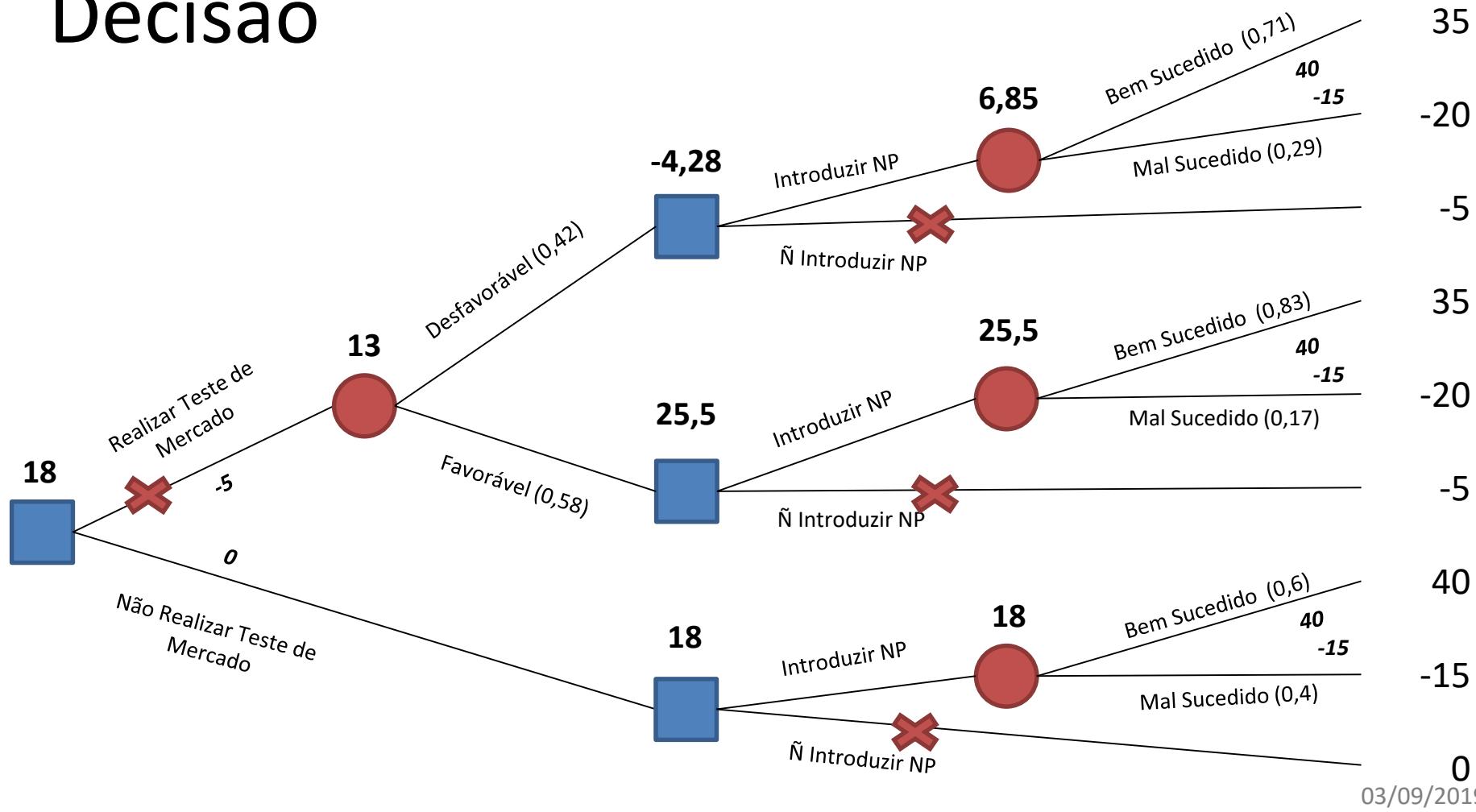
Adaptado de Hillier e Lieberman 15.5-7

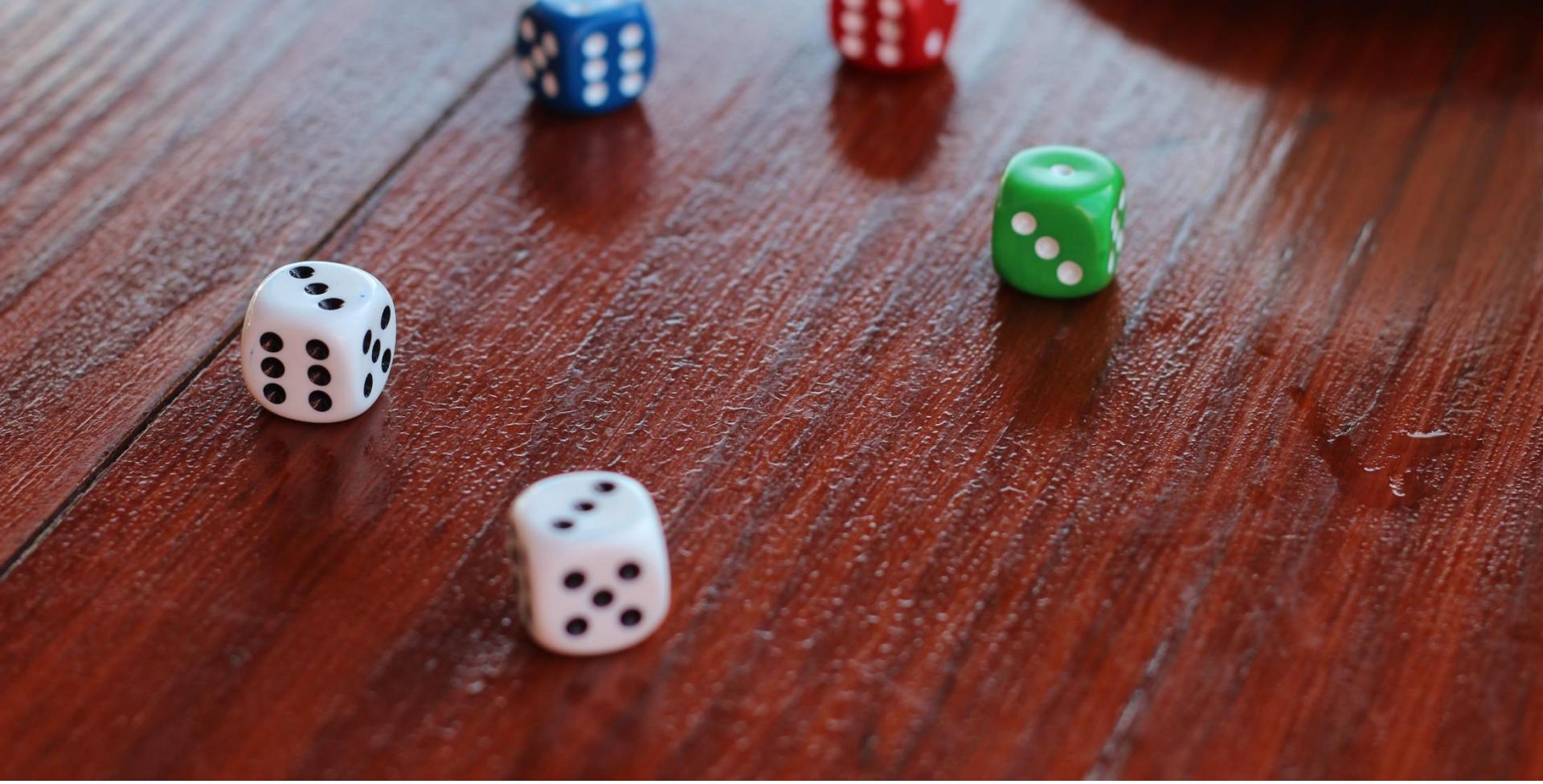
- Agora considere que uma opção é experimentar o produto no mercado, ao custo de 5 milhões. Uma empresa de consultoria em testes de mercado calculou as seguintes probabilidades para o teste da empresa:
 - A probabilidade do Teste ser favorável é de 58 %.
 - Se o teste no mercado for favorável, a probabilidade de sucesso do produto é de 83 %;
 - Se o teste no mercado for desfavorável, a probabilidade de sucesso do produto é de apenas 29 %.
- a) Construa uma árvore de decisão para o problema, agora considerando esta nova opção.
- b) Resolva a Árvore de Decisão e indique qual deve ser a estratégia ótima da empresa para este problema.

03/09/2019

Resolução da Árvore de Decisão

Prêmio





ANÁLISE DE DECISÃO SOB RISCO – SIMULAÇÃO DE MONTE CARLO

Simulação de Monte Carlo – Exemplo Simplificado

- Exemplo:
- Desejo saber quando serei capaz de comprar um imóvel, porém estou em dúvida sobre quando irei ter capacidade de adquirir este bem. Minha esposa me perguntou se iremos comemorar ou não o natal em nossa casa nova.
- Com base nos dados históricos apresentados, utilize a Simulação de Monte Carlo para responder: **Qual é a probabilidade** de você passar o natal em sua casa nova?

03/09/2019

Simulação de Monte Carlo – Exemplo Simplificado

A data na qual posso adquirir o meu bem depende de diversos fatores, até certo ponto imprevisíveis:

- a) Quanto dinheiro irei receber nos próximos meses;
 - b) Quanto dinheiro irei gastar nos próximos meses;
 - c) Quanto o meu dinheiro poupado irá render nos próximos meses;
 - d) O quanto o valor do bem que quero comprar irá subir nos próximos meses.
-
- Apesar disto, possuo informação histórica suficiente para **estimar** estas informações, e informar qual é a data esperada de compra do meu imóvel. Também posso saber qual é a probabilidade desta data ser menor do que a data do natal.

03/09/2019

Passos da Simulação de Monte Carlo

1. Modelar o Problema;
2. Coletar Dados sobre o problema e estimar os parâmetros necessários para a simulação;
3. Simular “n” rodadas do problema;
4. Estimar o valor esperado da variável de resposta e outras medidas de interesse;
5. Interpretar os Resultados e prover suporte à decisão.

03/09/2019

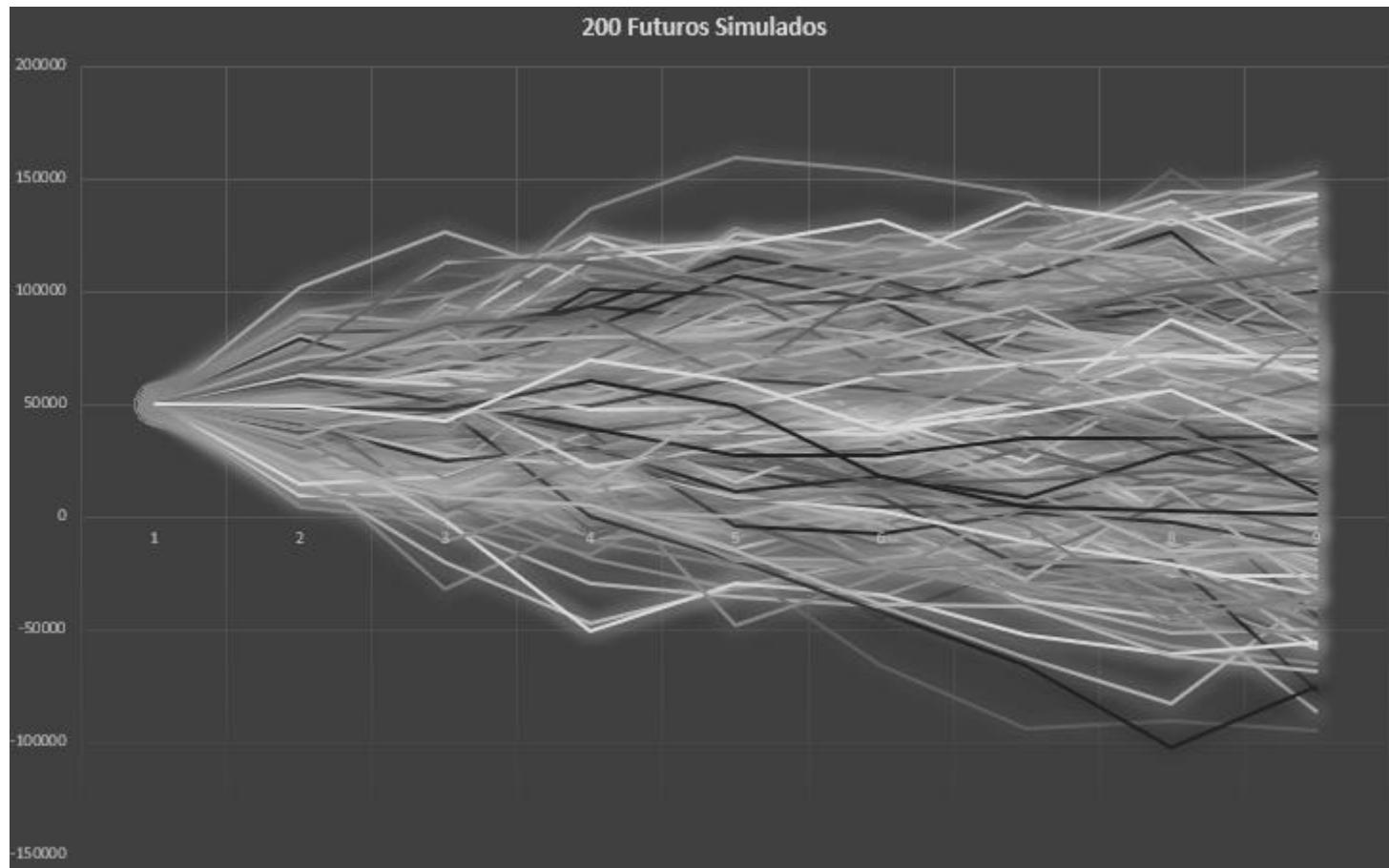
Passos da Simulação de Monte Carlo

The screenshot shows a Microsoft Excel spreadsheet titled "Exemplo-Monte-Carlo-Financas.xlsx". The ribbon menu is visible at the top, showing "Arquivo", "Página Inicial", and other tabs like "Inserir", "Layout da Página", etc. The "Fonte" (Font) tab is selected in the ribbon. The formula bar displays the formula =INV.NORM.N(ALEATÓRIO();0.00058;0.001). The main table has columns A through F. Column A contains "Quanto Ganhei" values, column B contains "Quanto Gastei" values, column C contains "Rendimento Mensal dos Investimentos" percentages, column D contains "Rendimentos do Mês Anterior" values, column E contains "Saldo Final" values, and column F is empty. Row 10 is highlighted with a green border. The bottom navigation bar shows tabs for "Historico", "Parametros", "Simulacoes", and a new button. The status bar at the bottom left says "Pronto".

	A	B	C	D	E	F
1	Quanto Ganhei	Quanto Gastei	Rendimento Mensal dos Investimentos	Rendimentos do Mês Anterior	Saldo Final	
2	1527.05	1147.41	0.057%	0.00	379.64	
3	1704.53	1898.26	0.096%	0.36	185.91	
4	1299.75	1221.78	0.054%	0.10	263.87	
5	1950.50	1009.87	0.107%	0.28	1204.50	
6	1735.58	913.75	0.169%	2.03	2026.33	
7	1394.89	1277.13	0.151%	3.07	2144.09	
8	1614.74	1198.58	0.013%	0.28	2560.25	
9	1021.76	1382.14	0.199%	5.09	2199.87	
10	1352.94	1350.49	0.129%	2.84	2202.33	
11	1529.87	1759.67	0.039%	0.85	1972.52	
12	1565.25	801.78	0.068%	1.34	2735.99	
13	1847.26	1144.66	0.086%	2.26	3428.50	

03/09/2019

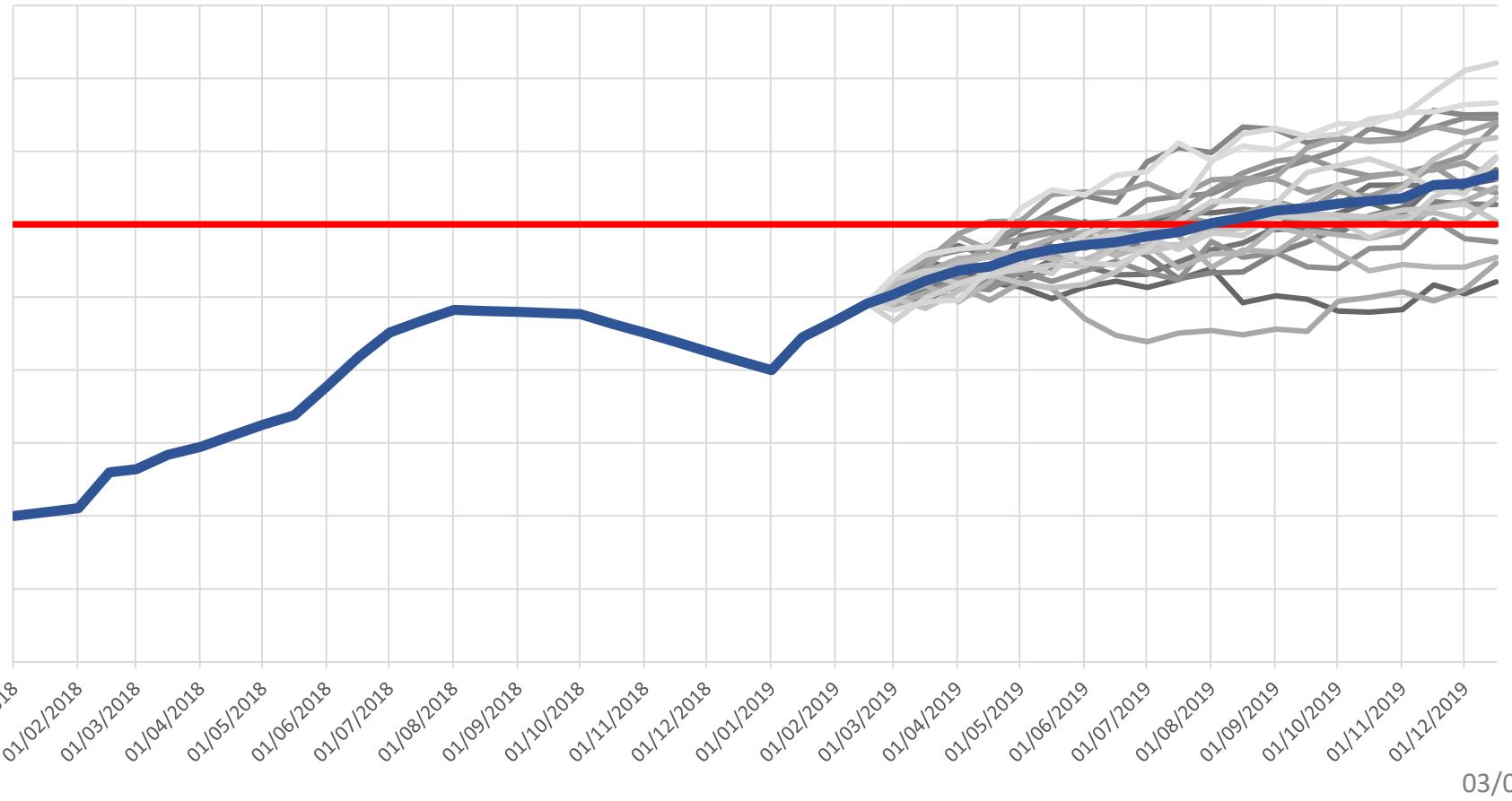
Passos da Simulação de Monte Carlo



03/09/2019

Exemplo Real

Quando Será Possível Comprar?

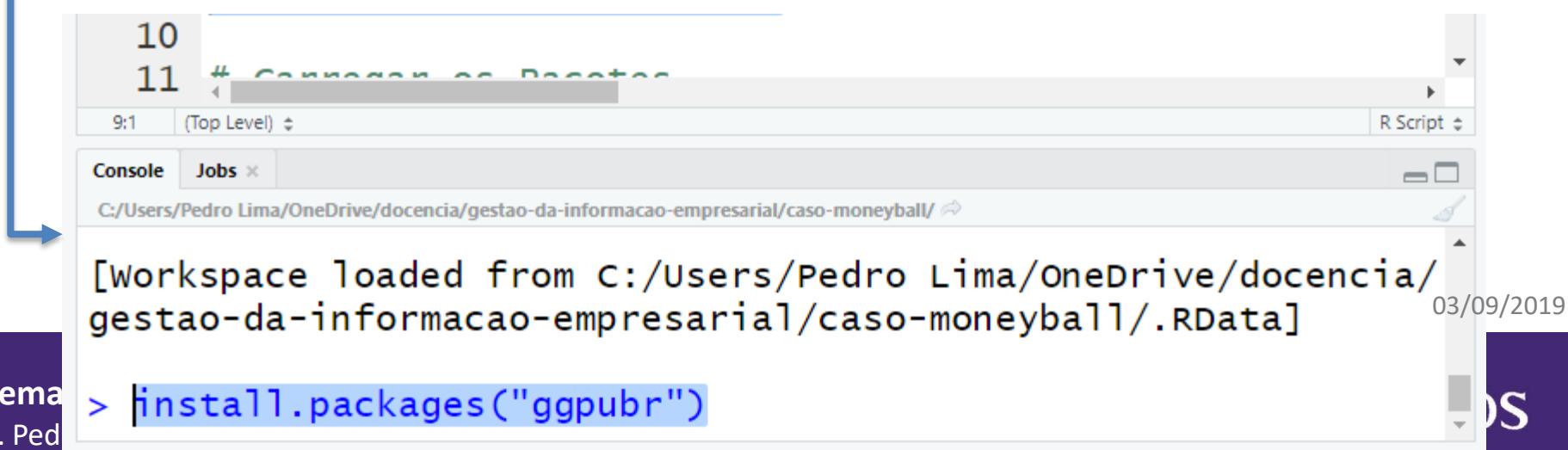




MACHINE LEARNING

Antes de Começar a Aula..

1. Baixar o arquivo baseball.csv, salvar em uma pasta;
2. Procurar Abrir o Programa RStudio;
3. Criar um arquivo “analise-moneyball.R” nesta pasta;
4. Fechar o Rstudio;
5. Abrir a pasta novamente e dar dois cliques no arquivo “analise-moneyball.R” (isso vai facilitar as coisas).
6. Ir no “console” do R, digitar install.packages("ggpubr") e dar ENTER.
7. Diversas telas abrirão.



The screenshot shows the RStudio interface. At the top, there's a code editor window with two lines of R code: line 10 contains the number '10' and line 11 contains the command '# Carregando os Pacotes'. Below the code editor is a toolbar with tabs for 'Console' and 'Jobs'. The 'Console' tab is selected. In the console area, the path 'C:/Users/Pedro Lima/OneDrive/docencia/gestao-da-informacao-empresarial/caso-moneyball/' is shown. The message '[Workspace loaded from C:/Users/Pedro Lima/OneDrive/docencia/gestao-da-informacao-empresarial/caso-moneyball/.RData]' is displayed. At the bottom of the console, the command '> install.packages("ggpubr")' is visible. A blue arrow points from the left towards the RStudio interface. The bottom left corner of the slide has the text 'Sistema Prof. Ped' and the bottom right corner has 'OS'.

```
10
11 # Carregando os Pacotes
9:1 (Top Level) R Script
Console Jobs
C:/Users/Pedro Lima/OneDrive/docencia/gestao-da-informacao-empresarial/caso-moneyball/
[Workspace loaded from C:/Users/Pedro Lima/OneDrive/docencia/gestao-da-informacao-empresarial/caso-moneyball/.RData]
> install.packages("ggpubr")
```

INTRODUÇÃO

Nestas duas aulas nós iremos...

- Conhecer o que é Machine Learning;
- Entender quais são as principais técnicas de machine learning;
- Aprender como ML funciona com dois casos utilizando o R:
 - **Um Exemplo com Regressão - Moneyball** – Entendendo modelagem e regressão linear múltipla com o caso do Moneyball;
 - **Um Exemplo com Classificação - ML para Serviços** - Melhorando operações de um callcenter de um banco com modelos de classificação.

03/09/2019

Porquê Machine Learning é Possível

- **A realidade possui padrões:**
 - Gremistas vestem azul, Colorados vestem vermelho;
 - Gaúchos comem mais carne do que pessoas em outros estados;
 - Clientes nos estados X,Y,Z, compram mais produtos a,b,c na época do ano w.
- **Conhecer** estes padrões pode salvar vidas, ou **gerar valor esperado**.
- Com um volume de dados suficientemente grande, é possível “aprender” padrões a partir destes dados computacionalmente, e codificar estes padrões em um **modelo**.
- Este **modelo** pode ser usado por alguém para fazer **previsões**, levando à aumento de receita (lembre do valor da informação).

03/09/2019

Qual é o valor de saber...

- para onde é provável que você vá na quarta-feira às 18h?
- que produtos você mais compra na sexta-feira?
- quais são seus hábitos alimentares, e que doenças você tem tido?
- se um funcionário tem alta probabilidade de colocar a empresa na justiça?
- quais produtos um cliente é mais propenso a comprar?

03/09/2019

O que é Machine Learning (ML)

- Machine learning: Construir e usar algoritmos que “aprendem” a partir de dados passados;
- Machine Learning é comumente usado para:
 - Classificar;
 - Predizer Valores (Regressão);
 - Agrupar (Clustering).
- Objetivo comum: **Construir um modelo** capaz de **predizer alguma variável** com base em **informações prévias** que leve algum sistema a tomar melhores decisões.

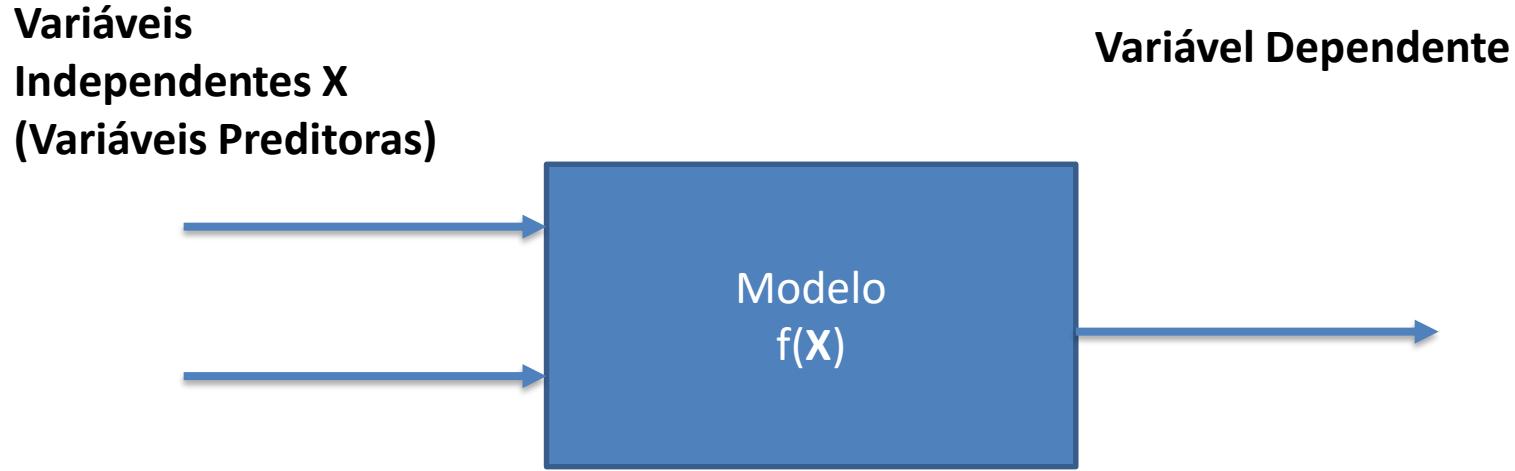
03/09/2019

Aplicações Comuns

- Aplicações Comuns em Negócios:
 - “Shopping Basket Analysis”: Que tipos de produtos as pessoas normalmente compram junto?
 - Sistemas de Recomendação: O que recomendar para uma pessoa que está comprando protetor solar?
 - Carros Autônomos: A placa logo à frente é uma placa de pare ou não? O sinal está vermelho ou não? Existe uma pessoa no meio da rua?

03/09/2019

O que é um Modelo?



- Cada Técnica de Machine Learning possui um formato “diferente” de Modelo.

03/09/2019

Problemas Típicos Abordados pela ML

- **Classificação:**
 - Objetivo: Predizer a categoria de uma nova observação;
 - **Variável Dependente é categórica e conhecida;**
 - Pode ser uma transformação de uma variável contínua (ex.: Reprovou ou não).
- **Régressão:**
 - Objetivo: Predizer o valor de uma variável de uma nova observação;
 - Variável Dependente é numérica e conhecida;
 - Ex.: Qual é o salário estimado de um aluno dada a sua nota?
- **Clusterização:**
 - Objetivo: Agrupar objetos similares em clusters;
 - Não é necessário saber as classes a priori.

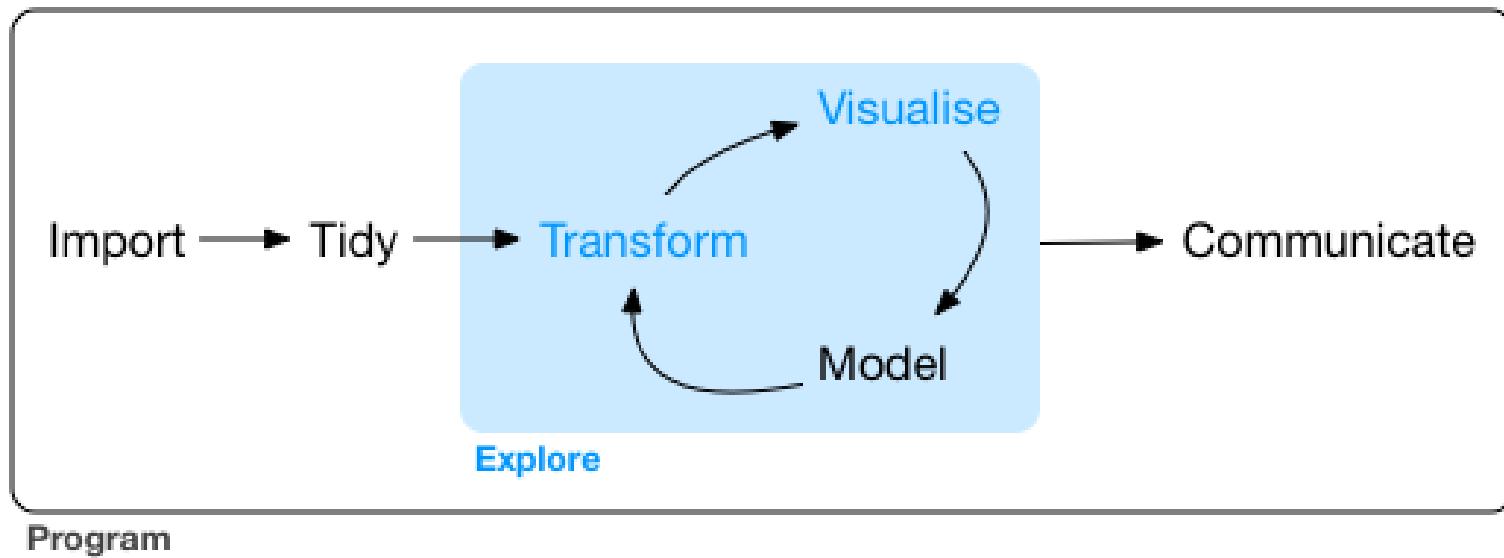
03/09/2019

Classificação das Técnicas de ML

Classe	Tipo de Problema	Variável Dependente a priori	Tipo da Variável Dependente	Técnicas Conhecidas
Aprendizagem Supervisionada	Classificação	Sim	Categórica	Random Forests Decision Trees; Boosted Machines; Regressão Logística
	Regressão	Sim	Numérica	Regressão Linear Redes Neurais “Deep Learning”
Aprendizagem Não-Supervisionada	Clusterização	Não	“Categórica – Clusters”	k-means

03/09/2019

Principais Tarefas em Data Science



<https://r4ds.had.co.nz/explore-intro.html>

03/09/2019

TÉCNICA 1 – REGRESSÃO LINEAR MÚLTIPLA

Questão....

Qual é a maneira **mais óbvia** de estabelecer /estudar a relação entre...

- O quanto alunos tiram no Enem e o quanto eles ganham?
- O quanto as empresas investem em P&D e o quanto elas lucram?
- O quanto você vai à academia e o quanto você emagrece?
- O quanto o professor lê os slides o quanto os alunos aprendem?

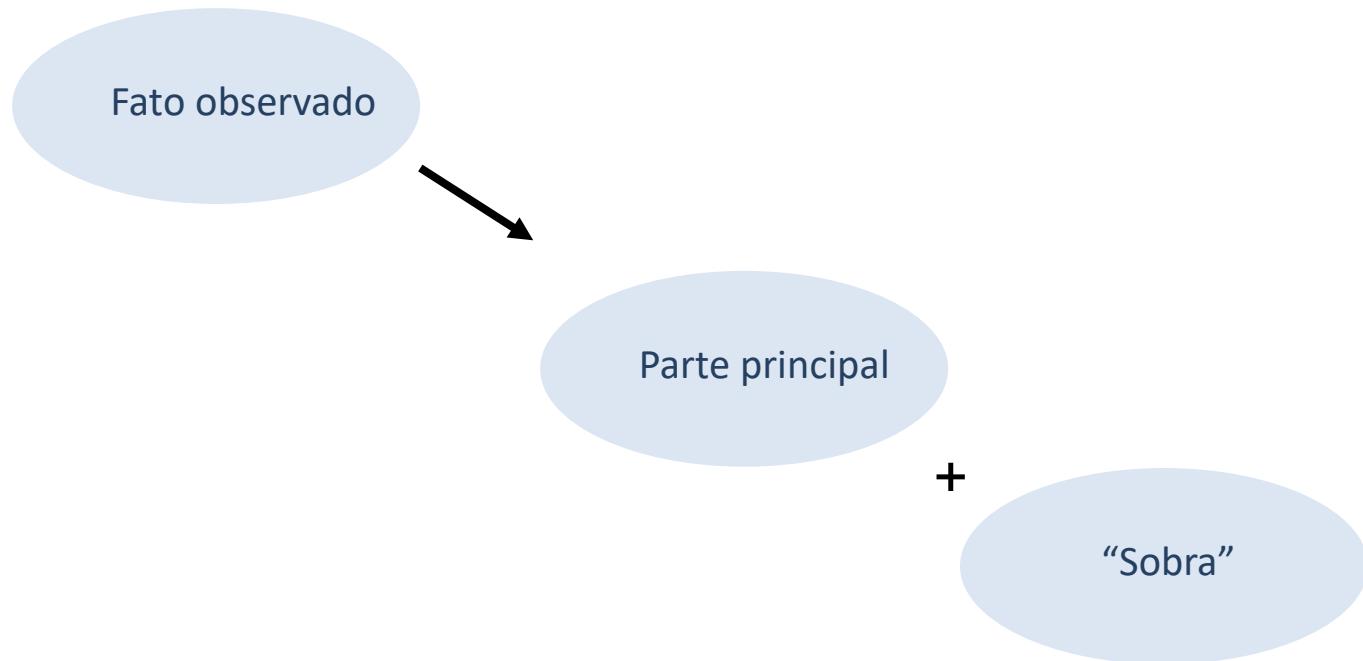
03/09/2019

Conteúdo desta Seção

- Nesta seção abordaremos:
 - O que é a Análise de Regressão Linear (e porquê aprendê-la antes de aprender outros algoritmos de machine learning);
 - Diferentes usos da Regressão: Estudo do Sistema e Predição;
 - Como a regressão pode ser usada em um caso de análise de performance esportiva (Moneyball)?
- Nesta seção não abordaremos:
 - Avaliação de Pressupostos da Regressão Linear;
 - Outras formas de regressão.

03/09/2019

“Quebrar” e “Remontar”



O.P.S.

Observado, Principal, Sobra

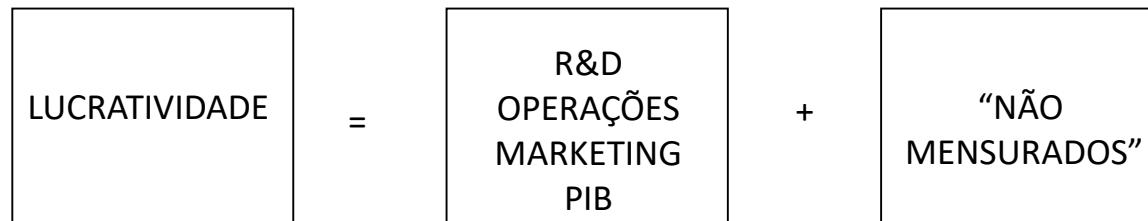
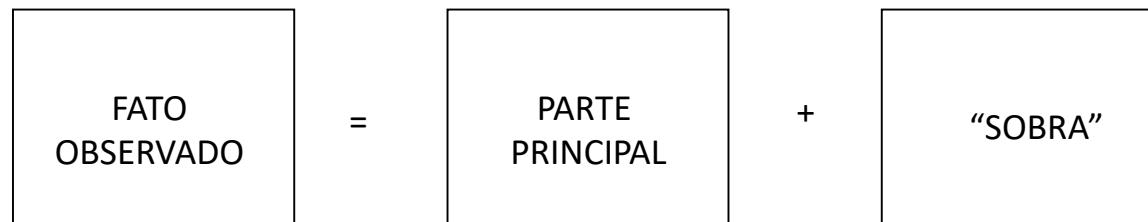


POR QUE FAZEMOS ISSO?

03/09/2019

O.P.S.

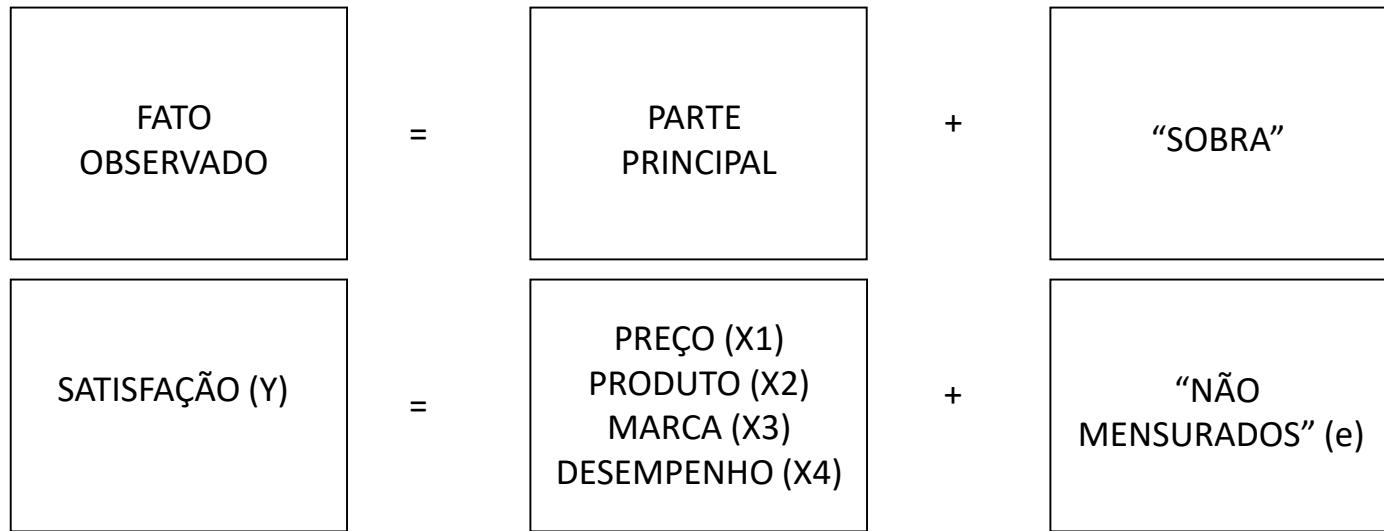
Observado, Principal, Sobra



03/09/2019

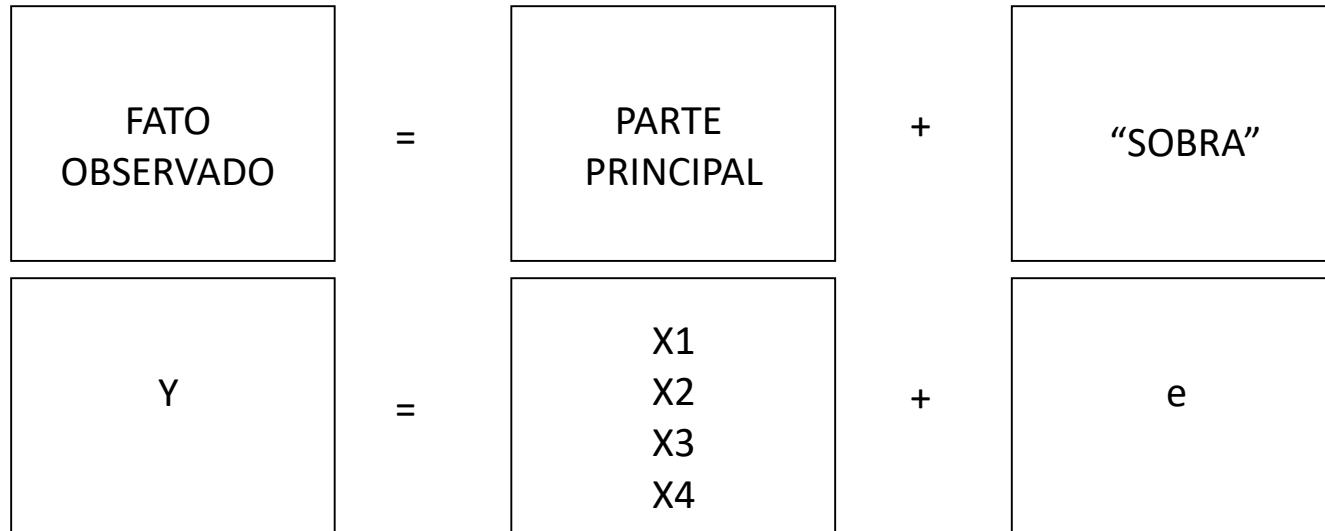
O.P.S.

Observado, Principal, Sobra



03/09/2019

Análise de regressão linear



Analisa as relações entre um FATO OBSERVADO e as PARTES PRINCIPAIS que a compõe/afetam.

03/09/2019

O Modelo de Regressão Linear

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + e$$

**Entender como as PARTES PRINCIPAIS
afetam/relacionam-se com o
FATO OBSERVADO**

03/09/2019

O Modelo de Regressão Linear

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + e$$

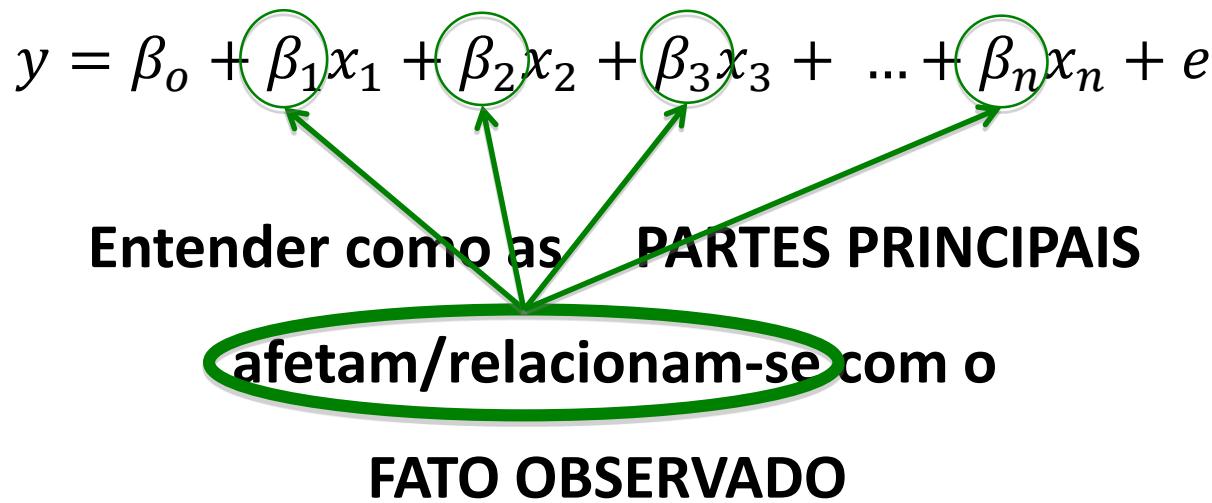
Entender como as **PARTES PRINCIPAIS**

afetam/relacionam-se com o

FATO OBSERVADO

03/09/2019

O Modelo de Regressão Linear



03/09/2019

O Modelo de Regressão Linear

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + e$$

Entender como as PARTES PRINCIPAIS

afetam/relacionam-se com o

FATO OBSERVADO

03/09/2019

O Modelo de Regressão Linear

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + e$$

**Entender como as PARTES PRINCIPAIS
afetam/relacionam-se com o
FATO OBSERVADO**

03/09/2019

O Modelo de Regressão Linear

Variável dependente (VD)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + e$$

Variáveis independentes (VI)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + e$$

Coeficientes (ou parâmetros)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + e$$

Erro

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + e$$

03/09/2019

O Modelo de Regressão Linear

Variável dependente (VD)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + e$$

A variável dependente é a VARIÁVEL DE INTERESSE do pesquisador

Ex.: Lucratividade, satisfação, desempenho de operações, etc.

Ela é “afetada” por variações nas variáveis independentes

A VD é medida pelo pesquisador

A VD DEVE ser uma variável contínua

A VD normalmente é denominada por “y”

03/09/2019

O Modelo de Regressão Linear

Variáveis independentes (VI)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + e$$

São as variáveis que afetam a variação na VD

São as VARIÁVEIS EXPLANATÓRIAS que “explicam” o quanto a VD varia

Ex.: Investimento em R&D, satisfação dos clientes, melhorias nos processos

As VI's são aquelas hipotetizadas teoricamente e que afetam
nossa variável de interesse

As VI's são medidas pelo pesquisador

A VI's são normalmente denominadas por “x”

03/09/2019

O Modelo de Regressão Linear

Coeficientes (ou parâmetros)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + e$$

São os efeitos que as VI's têm sobre a DV

Estabelecem uma relação entre VI e DV, positiva, negativa ou neutra

Servem para testar a existência de uma relação

Ex.: Investimento em R&D afeta lucratividade em XX

Ex.: Satisfação dos clientes afeta lucratividade em XX

Os coeficientes do modelo são calculados utilizando-se a estatística

Os coeficientes são normalmente denominados por “b” (beta)

03/09/2019

O Modelo de Regressão Linear

Erro

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + e$$

É o quanto “erramos” ao tentar estabelecer a relação entre VI’s e VD

É a parte do modelo que inclui todas as outras VI’s não incluídas (pensadas) no modelo e que também afetam a VD

É a diferença entre o valor medido (observado) da VD e o valor que a análise de regressão prediz

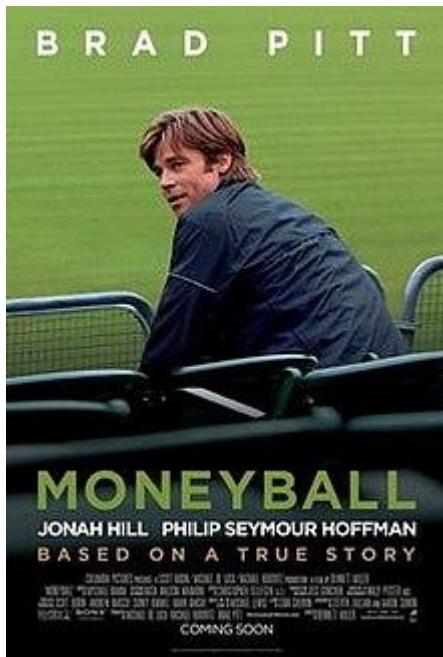
O erro é normalmente denominado por “e”

03/09/2019

O Modelo de Regressão Linear

- O objetivo da análise de regressão é verificar o quanto a VD varia em função da variação nas VI's ;
- A análise de regressão calcula os coeficientes populacionais para estabelecer a relação de cada VI's com a VD;
- A análise de regressão calcula (“prediz”) valores para a VD que mais se aproximam com os valores medidos (observados) ;
- A análise de regressão é baseada no método dos mínimos quadrados que será ilustrado posteriormente.

03/09/2019



REGRESSÃO LINEAR

CASO 1 – MONEYBALL



O homem que mudou o jogo | Trailer legendado | Em exibição nos cinemas

<https://www.youtube.com/watch?v=BumI-Yh0P1M&t=2s>

Exemplo - MoneyBall

Este exemplo usa dados relacionados ao filme “Moneyball” para apresentar a técnica de regressão linear com o R.

Este exercício e a ideia de usar o exemplo do Moneyball é baseada em uma aula do MIT, da plataforma Edx:

https://courses.edx.org/courses/course-v1:MITx+15.071x_3+1T2016/

03/09/2019

Moneyball e o Oakland A's

- Moneyball é o livro que conta a história sobre como o Data Analytics mudou a indústria do baseball;
- Oakland A's: Um dos times mais pobres do baseball. Foi vendido e teve seu orçamento cortado;
- Em 2002 o time perdeu três jogadores principais (é desse ponto que o filme começa);

03/09/2019

SaberMetrics – Estatística do Baseball

- Para quê isto é utilizado?
 - Comparar performance de jogadores;
 - **Predizer performance de jogadores ou de times no futuro;**
 - Determinar a contribuição de um jogador para o time, e portanto, definir o quanto cada jogador deveria ganhar.
- **Os times de baseball americanos provavelmente usam mais estatística do que a sua empresa.**
- Mais Informações:
- <https://en.wikipedia.org/wiki/Sabermetrics>

03/09/2019

Qual é a meta de um time de Baseball?

- Ir para as Playoffs!
- Quantos jogos um time precisa ganhar para chegar às playoffs?
- Paul DePodesta calculou que um time precisa de 95 vitórias para chegar às Playoffs.

03/09/2019

Número de Vitórias Necessário

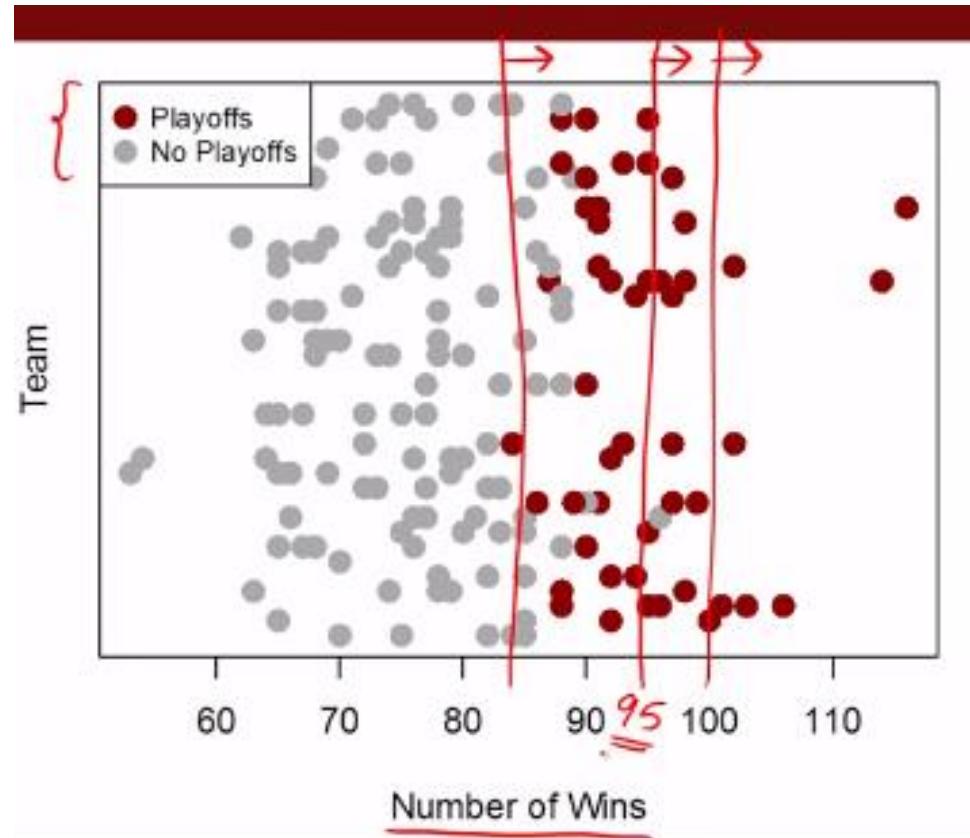


Imagen Playoffs

03/09/2019

Como se vence 95 jogos?

- Fazendo mais “Runs” do que o oponente.
- Quantos “Runs” a mais?
- Eles calcularam que precisariam fazer **135 Runs a mais do receberam** para ganhar 95 jogos.
- **Como calcular isso?**

03/09/2019

Lendo Dados em CSV com `read.csv()`

- Normalmente lemos dados no formato .csv no R para realizar as análises. É possível também ler dados em outros formatos.
- *# Definindo o Working Directory (pasta base na qual estaremos trabalhando).*
`setwd("D:/dev/sigepro-intro-R/moneyball")`
Lendo Dados em CSV
`baseball <- read.csv("baseball.csv")`

03/09/2019

Conhecendo os Dados com str()

- Antes de rodar qualquer análise precisamos conhecer a estrutura dos dados. Os dados contém **uma linha para cada time e ano de 1962 a 2012** para todas as temporadas. Dados incluem **Runs Scored (RS)**, **Runs Allowed (RA)** e **Vitórias (W)**.
- *# Podemos fazer isso usando a função str() (que mostra a estrutura)*
`str(baseball)`
- ```
'data.frame': 1232 obs. of 15 variables:
 $ Team : Factor w/ 39 levels
 "ANA", "ARI", "ATL", ...: 2 3 4 5 7 8 9 10 11 12 ...
 $ League : Factor w/ 2 levels "AL", "NL": 2 2 1 1
 2 1 2 1 2 1 ...
 $ Year : int 2012 2012 2012 2012 2012 2012 2012
 2012 2012 2012 2012 : num 0.415 0.378 0.403 0.428 0.424 0.405 0.39 0.43 0.47
 0.402 ...
```

03/09/2019

# Resumindo com o summary()

- Também podemos ter uma ideia dos dados usando o summary. Ele nos retorna médias, quartis, valores mínimos e máximos.
- `summary(baseball)`

|          | Team   | League       | Year           | RS       | RA |
|----------|--------|--------------|----------------|----------|----|
| BAL : 47 | AL:616 | Min. :1962   | Min. : 463.0   | Min. :   |    |
| 472.0    |        |              |                |          |    |
| BOS : 47 | NL:616 | 1st Qu.:1977 | 1st Qu.: 652.0 | 1st Qu.: |    |
| 649.8    |        |              |                |          |    |
| CHC : 47 |        | Median :1989 | Median : 711.0 | Median : |    |
| 709.0    |        |              |                |          |    |
| CHW : 47 |        | Mean :1989   | Mean : 715.1   | Mean :   |    |
| 715.1    |        |              |                |          |    |
| CIN : 47 |        | 3rd Qu.:2002 | 3rd Qu.: 775.0 | 3rd Qu.: |    |
| 774.2    |        |              |                |          |    |
| CLE : 47 |        | Max. :2012   | Max. :1009.0   | Max.     |    |
| :1103.0  |        |              |                |          |    |

03/09/2019

# Acessando variáveis específicas de um DataFrame

- Podemos acessar variáveis específicas de um DataFrame usando algumas notações possíveis.

`baseball$Year`

- [1] 2012 2012 2012 2012 2012 2012 2012 2012 2012  
2012 2012 2012 2012 2012  
[14] 2012 2012 2012 2012 2012 2012 2012 2012  
2012 2012 2012 2012 2012  
[27] 2012 2012 2012 2012 2011 2011 2011  
2011 2011 2011 2011 2011

03/09/2019

# Selecionando Linhas Específicas do DF

- Vamos selecionar apenas dados até o ano de 2002 (que foram os dados que eles possuíam em 2002)
- *# Considerando apenas anos exibidos pelo moneyball*  
`moneyball = subset(baseball, Year < 2002)`  
`str(moneyball)`
- `'data.frame': 902 obs. of 15 variables:`  
`$ Team : Factor w/ 39 levels`  
`"ANA", "ARI", "ATL", ... : 1 2 3 4 5 7 8 9 10 11`  
`...`

03/09/2019

# Relembrando nosso objetivo

- **Objetivo:** Saber como chegar às playoffs.
- **Em outras Palavras:** Saber quantos Runs um time deve fazer a mais do que “leva” para ter mais do que 95 vitórias.
- **Como:** Usando uma **Ressão Linear para predizer Vitórias em função do Run Differences.**

03/09/2019

# Calculando a RUN Difference

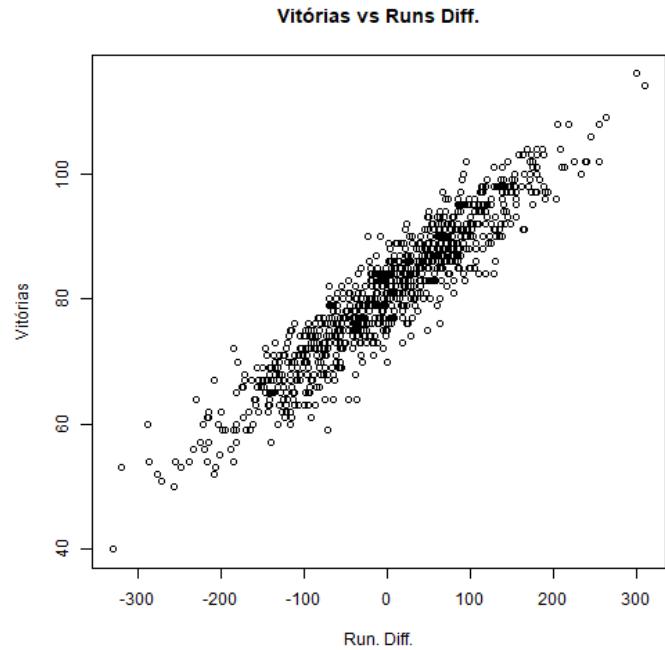
- Criando uma nova variável para calcular a “Run Difference”
- *# Calculando a Run Difference*  
moneyball\$RD = moneyball\$RS - moneyball\$RA  
**str**(moneyball)
- 'data.frame': 902 obs. of 16 variables:  
  \$ Team : Factor w/ 39 levels  
  "ANA", "ARI", "ATL", ... : 1 2 3 4 5 7 8 9 10 11  
  ...  
  ...

03/09/2019

# Existe uma Relação entre Run Difference e Vitórias?

Só faz sentido usar uma regressão linear se é plausível a existência de uma relação linear entre as variáveis.

```
plot(moneyball$RD, moneyball$W,
main = "Vitórias vs Runs Diff.",
xlab = "Run. Diff.", ylab =
"Vitórias")
```



03/09/2019

# Um teste de Correlação

Existe uma correlação alta entre estas duas variáveis?

```
cor(x = moneyball$RD, y = moneyball$W, method=c("pearson",
"kendall", "spearman"))
[1] 0.938515
```

03/09/2019

# Como predizer o número de vitórias com base em Run Differences?

Usando uma Regressão Linear!

$$y = \beta_0 + \beta_1 x_1 + e$$

Ou, em outras palavras...

$$\begin{aligned} Wins &= \beta_0 + \beta_1 RunDiff + e \\ W &= \beta_0 + \beta_1 RD + e \end{aligned}$$

## Como fazer isso no R?

```
modelo_vitorias = lm(W ~ RD, data=moneyball)
```

03/09/2019

# Analizando o Modelo para Predizer Vitórias

Podemos predizer o número de vitórias que um time terá a partir de um número de Home Runs.

```
summary (modelo_vitorias)
```

Call:

```
lm(formula = W ~ RD, data = moneyball)
```

Residuals:

| Min      | 1Q      | Median | 3Q     | Max     |
|----------|---------|--------|--------|---------|
| -14.2662 | -2.6509 | 0.1234 | 2.9364 | 11.0370 |

B0 = 80.881

Coefficients:

|             | Estimate  | Std. Error | t value | Pr(> t ) |
|-------------|-----------|------------|---------|----------|
| (Intercept) | 80.881375 | 0.131157   | 616.67  | <2e-16   |
| RD          | 0.105766  | 0.001297   | 81.55   | <2e-16   |

B1 = 0.1057

R<sup>2</sup> Ajust. = 0.88

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 .. 0.1 . 1

Residual standard error: 3.939 on 900 degrees of freedom

Multiple R-squared: 0.8808, Adjusted R-squared: 0.8807

F-statistic: 6651 on 1 and 900 DF, p-value: < 2.2e-16

03/09/2019

# Será que é são necessários 135 Runs a mais para chegar à Playoffs?

A partir da regressão linear nós sabemos que

$$W = 80.88 + 0.1057 RD$$

E também sabemos que W deve ser maior do que 95, logo quanto RUNS precisamos para ir às playoffs?

$$W \geq 95$$

$$80.88 + 0.1057 RD \geq 95$$

$$RD \geq \frac{95 - 80.88}{0.1058} \geq 133.446$$

Ou seja, sabemos que se um time quer ir para as playoffs ele precisa fazer **133,4** Runs a mais do que seus oponentes.

03/09/2019

# O que temos até Agora

- Para ir para as playoffs o time precisa de 95 vitórias ou mais.
- Para ter 95 vitórias, o time precisa de 133 ~ 135 Runs a mais do que os oponentes.
- Para isso, o time precisa:
  - Fazer mais Runs.
  - Levar menos Runs.

03/09/2019

# Como Avaliar um Jogador?

- Percentual de Rebatidas? (Batting Average - BA)
- Percentual de tempo que o Jogador passa na Base? (incluindo walks) (On-Base Percentage);
- Slugging Percentage (SLG). O quanto longe um jogador chega na sua vez de rebater;
- Quais destas estatísticas são mais importantes para considerar quando é necessário **comprar um jogador?**

03/09/2019

# Recorrendo à Regressão Linear Novamente!

- Na nossa base de dados estas estatísticas estão indicadas nas variáveis RS (Runs Scored), On-Base Percentage (OBP), Slugging Percentage (SLG) e Batting Average (BA).

```
str(moneyball)
'data.frame': 902 obs. of 16 variables:
 $ Team : Factor w/ 39 levels "ANA", "ARI", "ATL", ... : 1 2 3 4 5 7 8 9 10
11 ...
 $ League : Factor w/ 2 levels "AL", "NL": 1 2 2 1 1 2 1 2 1 2 ...
 $ Year : int 2001 2001 2001 2001 2001 2001 2001 2001 2001 ...
 $ RS : int 691 818 729 687 772 777 798 735 897 923 ...
 $ RA : int 730 677 643 829 745 701 795 850 821 906 ...
 $ W : int 75 92 88 63 82 88 83 66 91 73 ...
 $ OBP : num 0.327 0.341 0.324 0.319 0.334 0.336 0.334 0.324 0.35
0.354 ...
 $ SLG : num 0.405 0.442 0.412 0.38 0.439 0.43 0.451 0.419 0.458
0.483 ...
 $ BA : num 0.261 0.267 0.26 0.248 0.266 0.261 0.268 0.262 0.278
0.292 ...
```

03/09/2019

# Predizendo o Número de Runs - Sem Batting Average

- O modelo mais simples tem menos variáveis e ainda tem um R ao quadrado alto:

$$\text{Runs} = -804.3 + 27377.77 * \text{OBP} + 1584.91 * \text{SLG}$$

```
modeloruns_sBA = lm(formula = RS ~ OBP + SLG, data=moneyball)
summary(modeloruns_sBA)
```

Call:

```
lm(formula = RS ~ OBP + SLG, data = moneyball)
```

Residuals:

| Min     | 1Q      | Median | 3Q     | Max    |
|---------|---------|--------|--------|--------|
| -70.838 | -17.174 | -1.108 | 16.770 | 90.036 |

Coefficients:

|             | Estimate | Std. Error | t value | Pr(> t )   |
|-------------|----------|------------|---------|------------|
| (Intercept) | -804.63  | 18.92      | -42.53  | <2e-16 *** |
| OBP         | 27377.77 | 90.68      | 30.19   | <2e-16 *** |
| SLG         | 1584.91  | 42.16      | 37.60   | <2e-16 *** |

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 24.79 on 899 degrees of freedom

Multiple R-squared: 0.9296, Adjusted R-squared: 0.9294

F-statistic: 5934 on 2 and 899 DF, p-value: < 2.2e-16

03/09/2019

# Predizendo o Número de Runs Allowed

Com essa regressão, podemos estimar as **Runs Permitidas** com a equação:

$$\text{RunsAllowed} = -837 + 2913.6 * \text{OOBP} + 1514.29 * \text{OSLG}$$

```
modelorunsallowed = lm(formula = RA ~ OOBP + OSLG, data=moneyball)
summary(modelorunsallowed)
```

Call:

```
lm(formula = RA ~ OOBP + OSLG, data = moneyball)
```

Residuals:

| Min     | 1Q      | Median | 3Q     | Max    |
|---------|---------|--------|--------|--------|
| -82.397 | -15.178 | -0.129 | 17.679 | 60.955 |

Coefficients:

|                | Estimate | Std. Error | t value | Pr(> t )     |
|----------------|----------|------------|---------|--------------|
| (Intercept)    | -837.38  | 60.26      | -13.897 | < 2e-16 ***  |
| OOBP           | 2913.60  | 291.97     | 9.979   | 4.46e-16 *** |
| OSLG           | 1514.29  | 175.43     | 8.632   | 2.55e-13 *** |
| ---            |          |            |         |              |
| Signif. codes: | 0        | '***'      | 0.001   | '**'         |
|                | 0.01     | '*'        | 0.05    | '..'         |
|                | 0.1      | '.'        | 1       |              |

Residual standard error: 25.67 on 87 degrees of freedom

(812 observations deleted due to missingness)

Multiple R-squared: 0.9073, Adjusted R-squared: 0.9052

F-statistic: 425.8 on 2 and 87 DF, p-value: < 2.2e-16

03/09/2019

# Agora vem a parte legal

- Com os nossos modelos, agora é possível tentar predizer quantos jogos o Oakland A's vai ganhar em um determinado ano.
- Estamos tentando predizer quantos jogos o time vai ganhar antes da temporada começar, com o objetivo de suportar a decisão sobre quais jogadores queremos comprar.
- **Pressuposto # 1:** Performance passada dos jogadores do time que estamos montando tem correlação com a performance futura.
- **Pressuposto # 2:** A análise assume que haverão poucas lesões.
- **Pressuposto # 3:** Podemos estimar estatísticas para 2002 usando estatísticas dos jogadores coletadas em 2001.

03/09/2019

# Estimando Runs para 2002

- Com base na temporada de 2001, com o grupo que tivemos sabemos que a média do OBP é 0.339, e do SLG é 0.430.
- Nossa Regressão para Runs foi :  
 $\text{Runs} = -804.3 + 27377.77 * \text{OBP} + 1584.91 * \text{SLG}$
- Então a Estimativa de Runs é..

Runs = -804.3 + 2737.77 \* 0.339 + 1584.91 \*

0.430

Runs

[1] 805.3153

03/09/2019

# Estimando Runs Allowed para 2002

- Com base na temporada de 2001, com o grupo que tivemos sabemos que a média do OOBP é 0.307, e do OSLG é 0.373.
- Nossa Regressão para Runs Permitida foi

$$\text{RunsAllowed} = -837 + 2913.6 * \text{OOBP} + 1514.29 * \text{OSLG}$$

- Podemos fazer o mesmo para Runs Allowed

```
RunsAllowed = -837 + 2913.6 * 0.307 + 1514.29 *
0.373
```

```
RunsAllowed
[1] 622.3054
```

03/09/2019

# Quantos Jogos Esperamos Ganhar com esse Time?

- Nosso modelo de vitórias diz que

$$\text{Vitorias} = 80.88 + 0.1057 \text{ Run_Diffs}$$

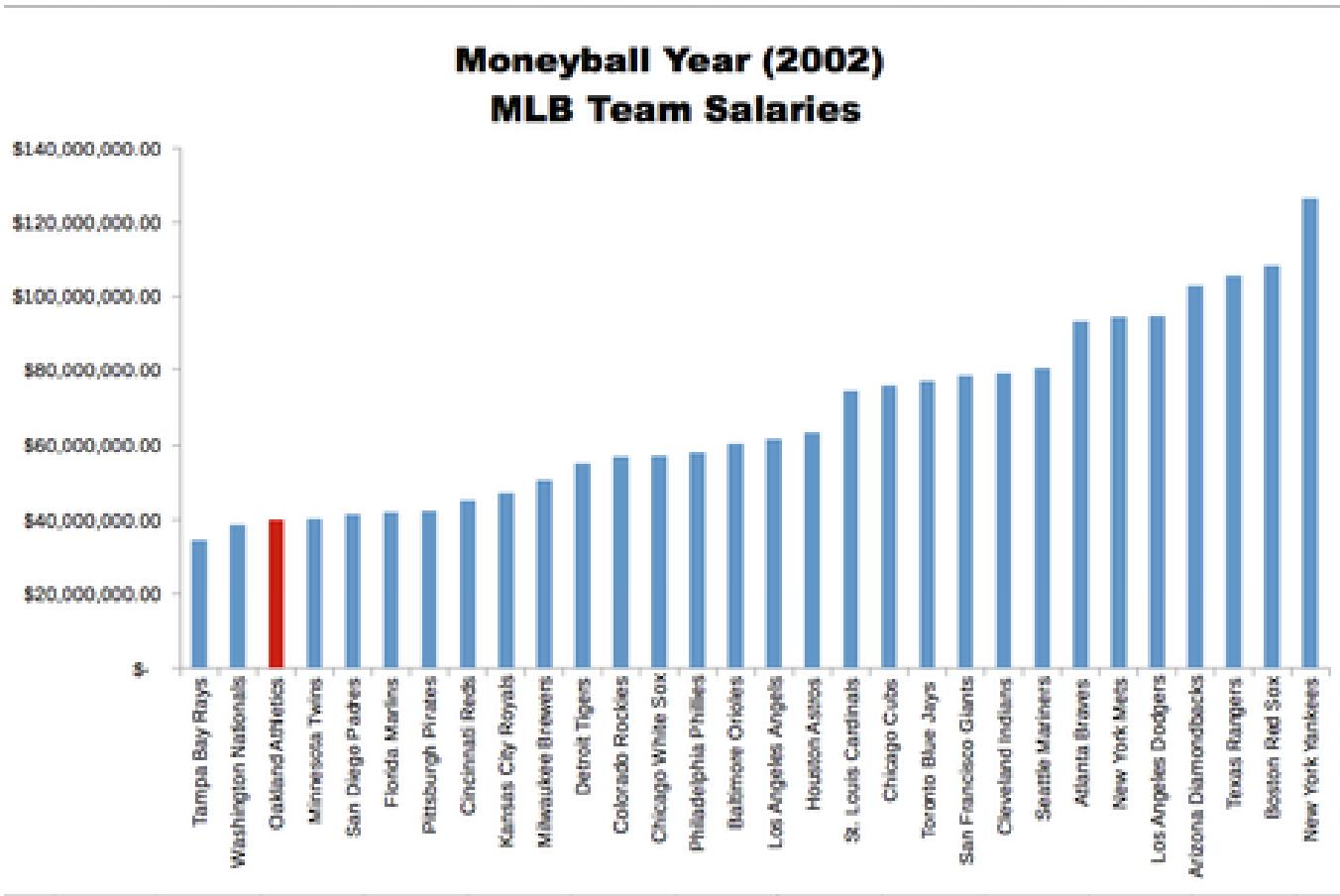
Então..

```
Vitorias = 80.88 + 0.1057 * (Runs - RunsAllowed)
if (Vitorias >= 95) {
 paste("Esse time deve chegar nas Playoffs com ",
Vitorias, " vitórias.")
} else {
 paste("Compre outros Jogadores, este time não
chega nas playoffs com apenas ", Vitorias, "
vitórias!")
}
[1] "Esse time deve chegar nas Playoffs com
100.224152772 vitórias."
```

- A abordagem de Paul foi parecida com essa.

03/09/2019

# A hora da verdade....



03/09/2019

# A hora da verdade

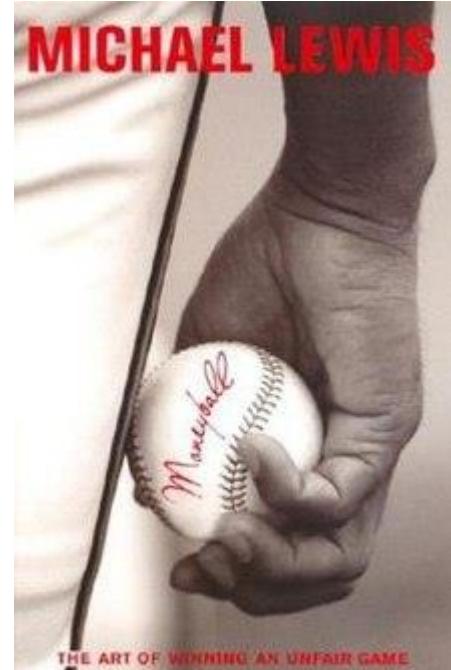
– Nosso modelo serve para alguma coisa?

| Variável     | Nosso Modelo | Modelo do Paul | Realizado |
|--------------|--------------|----------------|-----------|
| Runs         | 805          | 800 - 820      | 800       |
| Runs Allowed | 622          | 650 - 670      | 653       |
| Vitórias     | 100          | 93 - 97        | 103       |

- O Oakland A's ganhou 20 jogos em sequência nesse ano, mas não ganhou o campeonato;
- O Oakland A's conseguiu ir para as Playoffs mais uma vez!

03/09/2019

# Conclusão...



- Atualmente, análises estatísticas são mainstream na indústria do baseball ;
- Outros times adotaram as análises estatísticas, reduzindo a vantagem competitiva do Oakland A's;
- Paul Depodesta e Billy Beane continuam com carreiras

03/09/2019

# **CLASSIFICAÇÃO COM ÁRVORES DE DECISÃO E RANDOM FORESTS**

# Antes de Começar...

1. Baixar os arquivos credit.zip e caso-banco.zip;
2. Descompactar;
3. Deixar todos os arquivos na mesma pasta;
4. Abrir os arquivos que terminam com .R
5. Rodar os Comandos:
  1. install.packages("caret")
  2. install.packages("randomForest")
  3. install.packages("rpart.plot")
  4. install.packages("e1071")

03/09/2019

# Técnicas de Classificação

- Técnicas de Classificação Disponíveis para Aprendizagem Supervisionadas:
  - k-Nearest Neighbors (kNN)
  - Naive Bayes
  - Regressão Logística;
  - **Métodos Baseados em Árvores:**
    - Classification Trees;
    - Random Forests;
    - Gradient Boosted Machines (GBM);

03/09/2019

# Aprendizagem Supervisionada

- Encontre uma função  $f(x)$  que seja capaz de estimar a **categoria** ou **valor** de uma nova observação baseada em um conjunto de dados.
- Dado: um Conjunto de dados que contenha a variável de resposta.

03/09/2019

# Exemplo – Clientes de Um Banco Alemão

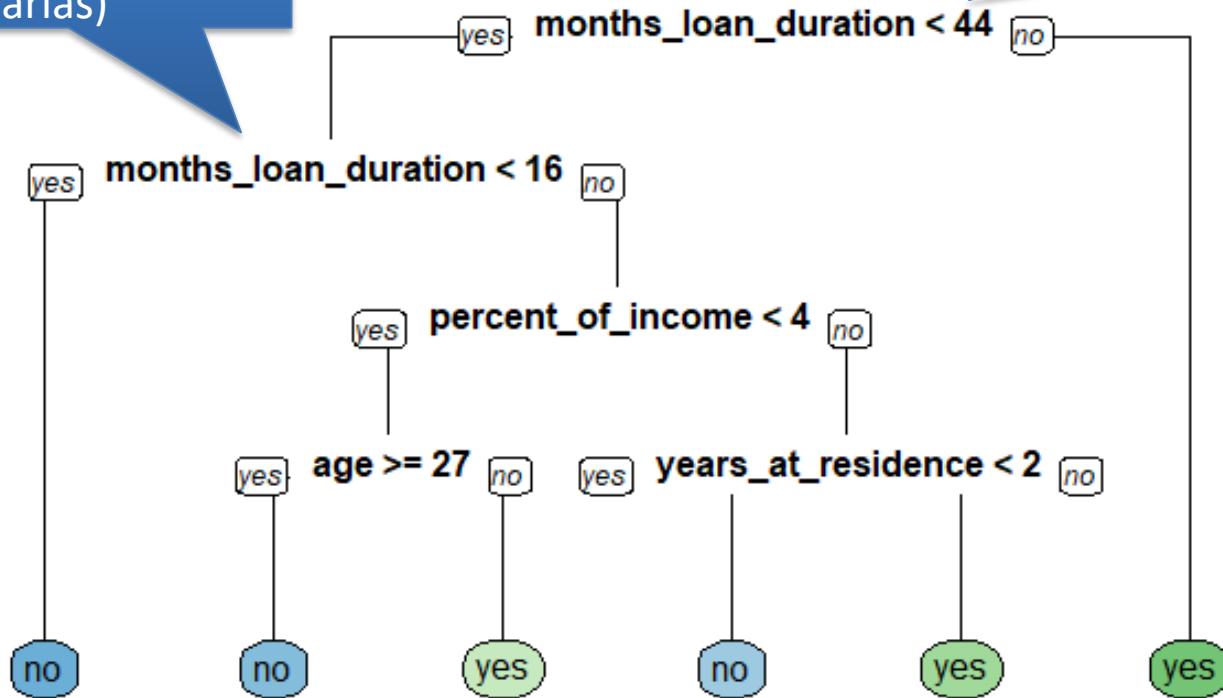
- <https://archive.ics.uci.edu/ml/datasets/Statlog%28German+Credit+Data%29>
- Cada um dos clientes do banco é representado por uma linha do dataset;
- O objetivo da análise é criar um modelo preditivo para indicar se os clientes irão ou não pagar o empréstimo.

03/09/2019

# O que é uma Árvore de Decisão

Nó de Decisão (sempre resultam em decisões binárias)

Nó Raiz



Resultado da Previsão: Neste caso o “Yes” quer dizer que o cliente não pagou

03/09/2019

# Processo para uso de modelos de Classificação baseados em árvores

1. Obter Dataset de dados;
2. Dividir o Dataset em um conjunto para **treinamento** ( ex.: 80% para criar a árvore) e outro para **testes** (ex.: 20% para avaliar a qualidade da árvore);
3. Treinar o modelo de classificação **com a base de treinamento**;
4. Avaliar a qualidade do modelo **com a base de testes**;
5. Usar o Modelo para predição.

03/09/2019

# Avaliando a Qualidade do Modelo

- Critérios para Avaliação da Qualidade do Modelo:
  - Acurácia: O quanto o Modelo Acerta;
  - Tempo de computação: O tempo necessário para gerar o modelo.
  - Interpretabilidade: Qual é a facilidade de interpretar o modelo.

03/09/2019

# Qualidade para Qualificação

- Acurácia e Erro.
- Fração de previsões que o modelo de classificação acertou.

$$Acurácia = \frac{Classificações Corretas}{Número Total de Classificações}$$

- Erro:  $1 - Acurácia$ .

03/09/2019

# Acurácia não é tudo

- Um médico olha um exame e indica se o paciente tem câncer ou não.
- O classificador indica que nenhum paciente tem câncer, e portanto não indica tratamento.
- Se 99 pacientes sobrevivem e um morre, o médico teve 99 % de acurácia!
- Você gostaria de ir nesse médico?

03/09/2019

# Matriz de Confusão (Confusion Matrix)

- Linhas e Colunas possuem todas as categorias.
- As linhas correspondem ao valor verdadeiro e as colunas às previsões.
- Exemplo do Médico:

|           |            | Previsto   |       |
|-----------|------------|------------|-------|
|           |            | Tem Câncer | Ñ Tem |
| Realidade | Tem Câncer | 0          | 1     |
|           | Ñ Tem      | 0          | 99    |

03/09/2019

# Exemplo 1

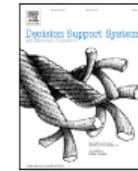
- Baixar Exemplo credit;
- Executar os passos da classificação;
- Observar a qualidade do modelo;
- Observar e interpretar a matriz de confusão.

03/09/2019

# Tópicos Não Abordados

- O que é Bagging, Bootstraping e Boosting, e como cada um dos métodos funciona;
- Otimização dos hyperparâmetros;
- Outras Técnicas de Classificação:
  - GBMs, etc;
- Outros Métodos de Avaliação da Qualidade do Modelo:
  - Curva ROC;
  - AUC.
- Avaliação de Importância das Variáveis.

03/09/2019



## A data-driven approach to predict the success of bank telemarketing



Sérgio Moro <sup>a,\*</sup>, Paulo Cortez <sup>b</sup>, Paulo Rita <sup>a</sup>

<sup>a</sup> ISCTE-IUL, Business Research Unit (BRU-IUL), Lisboa, Portugal

<sup>b</sup> ALGORITMI Research Centre, Univ. of Minho, 4800-058 Guimarães, Portugal

### ARTICLE INFO

#### Article history:

Received 1 November 2013

Received in revised form 28 February 2014

Accepted 4 March 2014

Available online 13 March 2014

#### Keywords:

Bank deposits

Telemarketing

Savings

Classification

Neural networks

### ABSTRACT

We propose a data mining (DM) approach to predict the success of telemarketing calls for selling bank long-term deposits. A Portuguese retail bank was addressed, with data collected from 2008 to 2013, thus including the effects of the recent financial crisis. We analyzed a large set of 150 features related with bank client, product and social-economic attributes. A semi-automatic feature selection was explored in the modeling phase, performed with the data prior to July 2012 and that allowed to select a reduced set of 22 features. We also compared four DM models: logistic regression, decision trees (DTs), neural network (NN) and support vector machine. Using two metrics, area of the receiver operating characteristic curve (AUC) and area of the LIFT cumulative curve (ALIFT), the four models were tested on an evaluation set, using the most recent data (after July 2012) and a rolling window scheme. The NN presented the best results ( $AUC = 0.8$  and  $ALIFT = 0.7$ ), allowing to reach 79% of the subscribers by selecting the half better classified clients. Also, two knowledge extraction methods, a sensitivity analysis and a DT, were applied to the NN model and revealed several key attributes (e.g., Euribor

# CASO 2 – CLASSIFICANDO CLIENTES DE ALTO POTENCIAL

# Caso 2 - Classificação

- Você trabalha em um grande banco e é responsável pela **prospecção de novos clientes para uma linha investimentos utilizando um call center**. Uma equipe de 10 pessoas trabalha sob a sua supervisão realizando ligações para os clientes. A equipe é capaz de realizar **100 ligações por dia**.
- Cada cliente que adquire o produto financeiro leva o banco a ganhar, em média, **800 reais**.
- Operar o Call Center Custa **3000** reais por dia.
- Tradicionalmente, as pessoas da equipe ligam para os clientes do banco aleatoriamente (não é possível ligar para todos os clientes!) oferecendo a linha de investimentos, porém a taxa de aceitação não é significativa (aprox. 11%).
- Após aprender que é possível **identificar padrões e utilizar modelos de classificação**, você reconheceu que a equipe poderia utilizar um modelo de classificação para melhor selecionar clientes.
- O modelo será usado para indicar quais clientes deveriam receber ligações, e assim a equipe iria se concentrar nestes clientes.

03/09/2019

# Caso 2 - Classificação

- Após aprender que é possível **identificar padrões e utilizar modelos de classificação**, você reconheceu que a equipe poderia utilizar um modelo de classificação para melhor selecionar clientes.
- O modelo será usado para indicar quais clientes deveriam receber ligações, e assim a equipe iria se concentrar nestes clientes.
- Você Deve:
  - Criar um Modelo de Classificação;
  - Avaliar a Qualidade deste Modelo;
  - Identificar qual é o percentual de clientes que aceitariam o investimento que é corretamente identificado pelo modelo;
  - Identificar qual é o Valor da Informação que o Modelo Agrega ao processo de Decisão.

03/09/2019

# Dados Disponíveis

## # Dados dos Clientes

- 1 - age (numeric) - idade
- 2 - job : tipo do emprego do cliente(categorical:  
"admin.", "unknown", "unemployed", "management", "housemaid", "entrepreneur", "student",  
"blue-collar", "self-employed", "retired", "technician", "services")
- 3 - marital : status conjugal (categorical: "married", "divorced", "single"; note: "divorced" means divorced or widowed)
- 4 - education : nível de educação (categorical: "unknown", "secondary", "primary", "tertiary")
- 5 - default: está devendo ao banco? (binary: "yes", "no")
- 6 - balance: saldo médio no banco(numeric)
- 7 - housing: tem empréstimo habitacional? (binary: "yes", "no")
- 8 - loan: tem empréstimo pessoal? (binary: "yes", "no")

## # Relacionado ao ultimo contato da campanha de marketing:

- 9 - contact: tipo de contato (categorical: "unknown", "telephone", "cellular")
- 10 - day: número dia do mês no qual foi feito contato (numeric)
- 11 - month: ultimo mês no qual foi feito contato (categorical: "jan", "feb", "mar", ..., "nov", "dec")
- 12 - duration: duração do ultimo contato em segundos (numeric)

## # Outros Atributos:

- 13 - campaign: número de contatos realizados durante esta campanha para este cliente (numeric, includes last contact)
- 14 - pdays: número de dias que passaram após ultimo contato da última campanha (numeric, -1 means client was not previously contacted)
- 15 - previous: número de contatos realizados antes desta campanha para este cliente (numeric)
- 16 - poutcome: resultado da última campanha de marketing (categorical: "unknown", "other", "failure", "success")

## Variável de Output:

- 17 - y – O cliente executou o investimento? (binary: "yes", "no")

03/09/2019

# Conclusões do Exercício

- O Modelo Tem uma Acurácia de 90 %, para todos os casos;
- Dentre os clientes que o Modelo sugere que irá comprar o produto, o modelo acerta 64%, ao invés de 11%.
- Considerando as premissas adotadas, o modelo agrega **42 mil reais por dia** ao processo do call-center.

03/09/2019



# TRABALHO EM GRUPO - GB

# Informações - Trabalho do Grau B

- **Objetivo do Trabalho:** Apresentar o **Projeto de um Sistema de Informação** para a execução de melhorias na empresa de **um dos integrantes do grupo.**
- Imagine que você tem a necessidade de liderar uma melhoria em um sistema de informação em sua empresa. Você recebeu a tarefa de projetar as mudanças neste sistema e gerenciar todo o projeto.
- Antes de o projeto iniciar, você precisa definir a especificação do projeto e submeter o projeto à aprovação da diretoria. É necessário escrever um documento contendo o **escopo** do projeto, e apresentar o projeto à seu chefe.
- Usaremos este trabalho para fazer exatamente isto.
- Entrega uma semana antes da prova do grau B.

03/09/2019

# Estrutura do Trabalho Escrito e Apresentação:

## 1. Definição do Problema (1 página):

1. Definição da Situação – Problema: Qual é a situação e qual é o problema que o sistema ou a melhoria no sistema pode evitar. Sugestão: Utilize um problema real para facilitar o seu trabalho.
2. Estimativa de Custos a serem evitados com o sistema, ou receitas a serem geradas.

## 2. Interfaces com Processos Organizacionais e outros Sistemas Existentes (0,5 páginas):

1. Discutir quais são os processos organizacionais da empresa relacionados ao sistema e como o sistema de informação deve suportar tais processos.

## 3. Exploração de Soluções Alternativas (0,5 página):

1. Explorar 3 diferentes alternativas para a abordagem da solução (ex.: adquirir e implantar um sistema pronto, desenvolver internamente, ou desenvolver com suporte externo), discutindo vantagens e desvantagens das alternativas.

03/09/2019

# Estrutura do Trabalho Escrito e Apresentação:

## 1. Análise de Requisitos do Sistema (6 a 8 páginas):

1. Definição de Requisitos Funcionais;
2. Definição dos Requisitos Não-Funcionais;
3. Definição de Casos de Uso;
4. Protótipo de Telas / gráficos (utilizando o Tableau)
5. Protótipo da Estrutura de Dados:
  1. Diagrama de Entidade e Relacionamento utilizando o dbdiagram.io
  2. Print das tabelas de teste criadas no excel.

## 2. Conclusões (1 página):

1. Determinar se o grupo **recomendaria ou não** o projeto à empresa do aluno, considerando o custo projetado do sistema e os benefícios potenciais.
2. Estimativa de Tempo para a Execução do Projeto;
3. Estimativa de Custos do Projeto.

03/09/2019

# Entregas do Trabalho

1. Relatório escrito (de acordo com os tópicos indicados anteriormente);

## 2. Critérios de Avaliação do Trabalho:

a. 30 % : Participação e Construção do Trabalho ao longo do Grau B. (Presença do Grupo nos momentos de condução do trabalho).

b. 10 % : Qualidade da Apresentação.

c. 60 % : Qualidade do Trabalho Escrito.

03/09/2019



# SISTEMAS DE INFORMAÇÃO - CONCEITOS

# O que é um Sistema de Informação?

O que é um Sistema de Informação?

*“Um Sistema de Informação é um sistema especializado no processamento e na comunicação de dados (máquinas) ou de informações (organismos vivos).”*

(Mattos 2010, p.5)  
09/09/2019

# Imagine-se nesta situação...

- Você precisa propor melhorias nos Sistemas de Informação ERP de Sua Empresa.
- Por onde você começaria? O que deve levar em consideração?



03/09/2019

# Conceitos: Dados, Informação, Conhecimento

Fatos básicos, cifras, observações e medições, sem contexto ou classificação

Dados processados: os dados foram organizados e interpretados

Um entendimento (ou modelo) sobre pessoas, objetos ou eventos, derivados de informações sobre eles



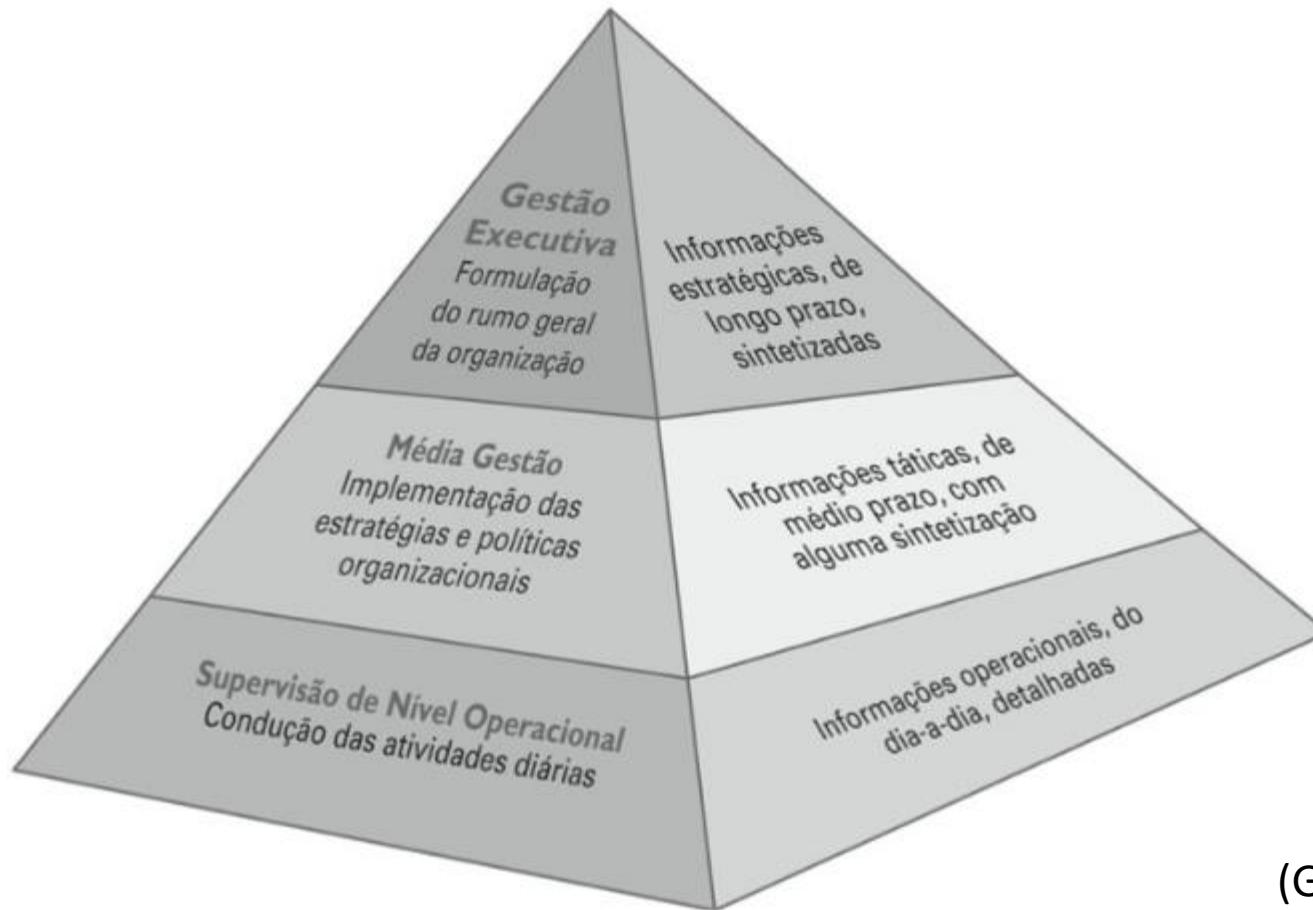
|      |     |      |
|------|-----|------|
| 12,5 | 3,7 | 1,75 |
| 45,1 | 3,8 | 2,22 |
| 19,8 | 3,9 | 7,81 |

|                      |    |
|----------------------|----|
| Relatório de Estoque |    |
| Número # Disponível  |    |
| 105                  | 39 |
| 106                  | 12 |

As vendas aumentaram 10% sobre o mesmo período do ano anterior com maior intensidade em...

(Gordon 2006, p.5)  
03/09/2019

# Níveis Organizacionais e Demandas por Informação



(Gordon 2006, p.13)  
03/09/2019

# Níveis de Decisões Gerenciais

| Características                   | Estratégico           | Tático                 | Operacional |
|-----------------------------------|-----------------------|------------------------|-------------|
| Objetivo                          | Aquisição de Recursos | Utilização de Recursos | Execução    |
| Horizonte Temporal                | Longo                 | Médio                  | Curto       |
| Nível de Gerência                 | Alta                  | Média                  | Baixa       |
| Escopo                            | Amplo                 | Médio                  | Focalizado  |
| Fonte de Informações              | Ext. e Intern.        | Ext. e Intern.         | Interno     |
| Nível de detalhes das informações | Altamente Agregado    | Moderadamente Agregado | Desagregado |
| Grau de Incerteza                 | Alto                  | Moderado               | Baixo       |
| Grau de Risco                     | Alto                  | Moderado               | Baixo       |

Bradley, S., Hax, A., & Magnanti, T. (1977). Applied mathematical programming. (p. 177)

03/09/2019

# Tipos de Decisões Gerenciais

| Tipo de Decisão                                        | Métodos Convencionais                     | Métodos Modernos          |
|--------------------------------------------------------|-------------------------------------------|---------------------------|
| Decisões Estruturadas<br>(Repetitivas,<br>Programadas) | Procedimentos<br>Organizacionais, Hábitos | Modelos Matemáticos       |
| Decisões Não-<br>Estruturadas<br>(Únicas, Complexas)   | Políticas, Intuição,<br>Regras            | Teoria da Decisão<br>PSMs |

Bradley, S., Hax, A., & Magnanti, T. (1977). Applied mathematical programming. (p. 179)

03/09/2019

# Classificações de Sistemas de Informação

Sistemas de Informação

(O'BRIEN, MARAKAS, 2012, p. 11)

Sistemas de Apoio às Operações

Sistemas de Processamento de Transações

Sistemas de Controle de Processos

Sistemas de Colaboração Empresarial

Sistemas de Processamento Especializado

Sistemas de Suporte Gerencial

Sistemas de Informações Gerenciais

Sistemas de Apoio à Decisão

Sistemas de Informação Executiva

Sistemas de Processamento Especializado

Sistemas Especialistas

Sistemas de Gestão do Conhecimento

Sistemas de Informação Estratégicos

Sistemas Funcionais de Negócios

03/09/2019



JESUÍTAS BRASIL

UNISINOS

# Conhecimentos Necessários Sobre Sistemas de Informação



(O'BRIEN, MARAKAS, 2012, p. 5)  
03/09/2019

# Sistemas de Suporte às Operações

- **Sistemas de Processamento de Transações:** Processam os dados resultantes das transações de negócios, atualizam bancos de dados operacionais e produzem documentos de negócios. Exemplos: sistemas de processamento de vendas e estoque e de contabilidade.
- **Sistemas de Controle de Processos:** Monitoram e controlam o processo industrial. Exemplos. Sistemas de controle de produção em indústria discreta (MES), ou sistemas supervisórios em indústria contínua (SCADA).
- **Sistemas de Colaboração Empresarial:** Oferecem suporte à comunicação e colaboração às equipes, grupos de trabalho e empresa. Exemplos: E-mail, Sistemas de controle de processos gerenciais.

(O'BRIEN, MARAKAS, 2012, p. 12)  
03/09/2019

# Sistemas de Suporte Gerencial

- **Sistemas de Informação Gerencial (MIS – Management Information Systems)** Fornecem informação em forma de relatórios para suporte à decisão de negócios. Exemplo: Sistemas de relatório e análise de vendas, desempenho da produção, etc.
- **Sistemas de Suporte à Decisão (DSS – Decision Support Systems):** Dão suporte ad hoc aos processos de tomada de decisão. Exemplos: Sistemas de Análise de risco, sistemas de otimização do mix. Sistemas de Sequenciamento.
- **Sistemas de Informação Executiva (EIS – Executive Information Systems):** Fornecem informações destinadas a executivos a partir de diversas fontes internas e externas de dados. Exemplos: Sistemas para análise do desempenho global do negócio, sistemas que analisam a performance de concorrentes para suporte ao planejamento estratégico.

(O'BRIEN, MARAKAS, 2012, p. 12)  
03/09/2019

# Outras Categorias de Sistemas de Informação

- **Sistemas Especialistas (Expert Systems):** Atuam como consultores técnicos para suporte à decisão de seus usuários. Exemplos: IBM Watson.
- **Sistemas de Gestão de Conhecimento:** Baseiam-se em conhecimento e dão suporte à criação, organização e disseminação do conhecimento dos negócios dentro da empresa. Exemplos: Sistemas de consultas às melhores práticas de processos.
- **Sistemas Funcionais do Negócio:** Dão suporte às aplicações operacionais e gerenciais das funções básicas de uma empresa. Exemplos: Sistemas de suporte aos setores de Operações, Marketing e Vendas, Finanças e Recursos Humanos.

(O'BRIEN, MARAKAS, 2012, p. 14)  
03/09/2019

# Processo de Desenvolvimento de Sistemas de Informação



(O'BRIEN, MARAKAS, 2012, p. 17)  
03/09/2019

# Componentes dos Sistemas de Informação



(Gordon 2006, p.8)  
03/09/2019

# Sistemas de Informação: Dimensões

| Dimensão       | Componente    |
|----------------|---------------|
| Tecnológica    | Hardware      |
|                | Software      |
|                | Dados         |
| Organizacional | Procedimentos |
| Humana         | Pessoas       |

(Audy et al. 2011, p. 116)  
03/09/2019

# Tecnologia da Informação: Para quê?

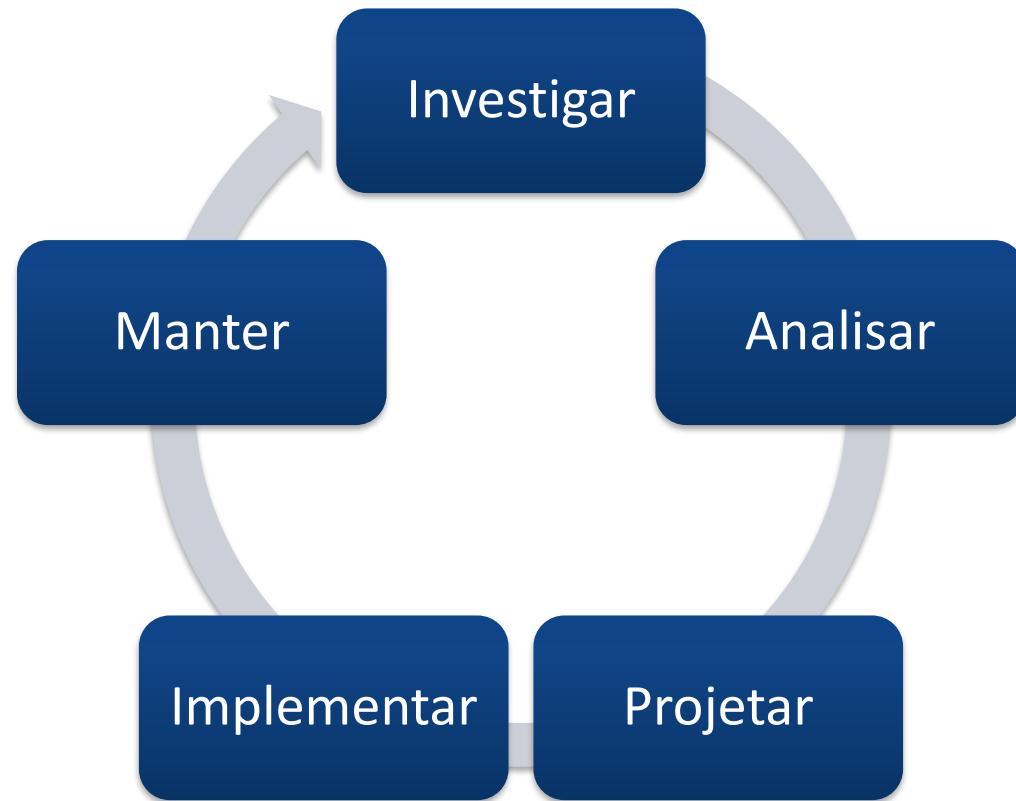
- Sistemas de Informação deveriam ser criados ou modificados para sustentar ou ampliar a vantagem competitiva das organizações. Sendo assim, elas podem ser utilizadas para...
- **Reducir Custos:**
  - Utilizar a TI para redução de custos empresariais, para clientes ou fornecedores.
- **Diferenciação:**
  - Desenvolver novos atributos utilizando TI para diferenciação de produtos e processos;
  - Desenvolver novos nichos de mercado com o auxílio da TI;
- **Inovação:**
  - Criar novos produto e serviços utilizando componentes da TI;
  - Desenvolver novos e
- **Promover e Sustentar o Crescimento:**
  - Utilizar a TI para suportar a expansão da empresa.
- **Desenvolver Alianças:**
  - Utilizar a TI para criar parcerias virtuais;
  - Utilizar a TI para criar sistemas de informação. (Ex.: Uber)

(O'BRIEN, MARAKAS, 2012, p. 47)



# PROJETANDO SISTEMAS DE INFORMAÇÃO

# O Ciclo de Desenvolvimento de Sistemas de Informação



O ciclo de desenvolvimento de Sistemas (O'BRIEN, MARAKAS, 2012, p. 17)  
03/09/2019

# Processos de Desenvolvimento de Sistemas de Informação

A literatura existente em **Engenharia de Software** apresenta diferentes processos de desenvolvimento de software. [@Wazlawick2013]

- O processo de desenvolvimento “Waterfall” é o processo mais tradicional e parte de certas premissas [@Wazlawick2013]:
- É possível definir os requisitos do software em um único ciclo;
- Uma vez definidos os requisitos do software, basta desenvolver o melhor sistema possível que atenda a estes requisitos.

03/09/2019

# O Processo de Desenvolvimento Padrão - Waterfall

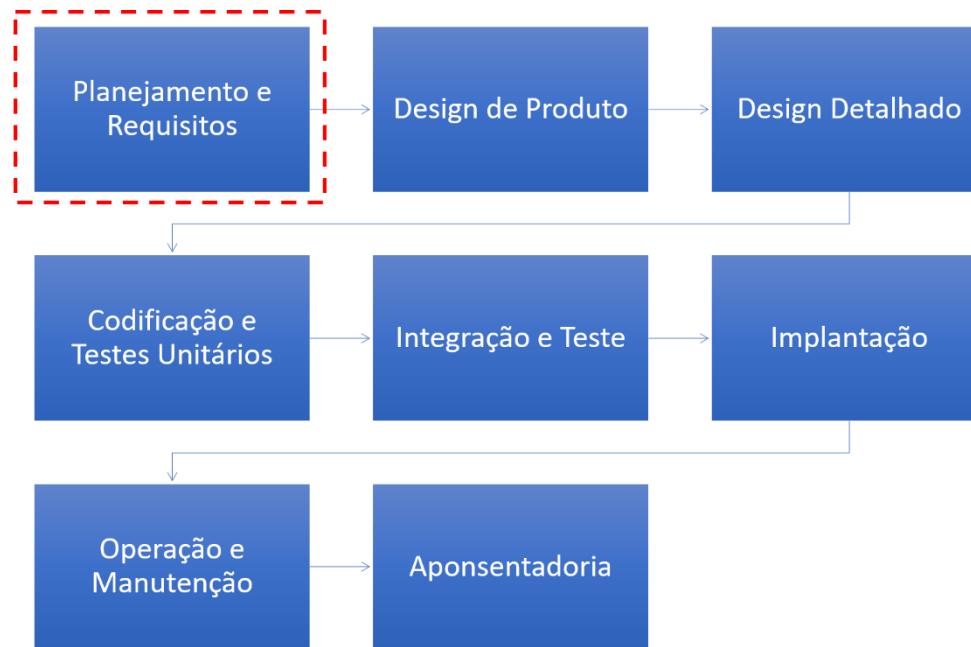
Em um processo de desenvolvimento de sistemas tradicional, o processo é organizado, em termos gerais, nesta ordem:

- Definir o Escopo do Sistema;
- Analisar os Requisitos do Sistema;
- Projetar o Sistema;
- Construir o Sistema;
- Validar o Sistema;
- Entregar o Sistema.

03/09/2019

# O Processo “Waterfall”

O processo Waterfall é composto por uma série de etapas sequenciadas [@Wazlawick2013]:



03/09/2019

# Modelos de Desenvolvimento Ágeis

Modelos de desenvolvimento ágeis possuem menos ênfase em definir o que deve ser feito em cada uma das atividades de desenvolvimento, e mais ênfase em como conduzir o processo de modo a **flexibilizar** o processo de desenvolvimento [@Wazlawick2013].

Exemplos de Processos Ágeis:

- FDD - Feature Driven Development;
- SCRUM;
- XP - Extreme Programming.

03/09/2019

# O Manifesto Ágil

"Nós estamos descobrindo formas melhores de desenvolver software fazendo e ajudando os outros a fazer. Através deste trabalho, chegamos aos seguintes valores:

- **Indivíduos e a interação entre eles** mais do que **processos e ferramentas**.
- **Software em Funcionamento** mais do que **documentação abrangente**.
- **Colaboração com o Cliente** mais do que **negociação de contratos**.
- **Responder à Mudanças** mais do que **seguir um plano**.

Mesmo havendo valor nos itens à direita, valorizamos mais os itens à esquerda". <http://agilemanifesto.org/iso/ptbr>

03/09/2019

# Os 12 princípios do Manifesto Ágil

- Qualquer semelhança com o Lean thinking **não** é mera coincidência.

Nós seguimos estes princípios:

1. Nossa maior prioridade é satisfazer o cliente através da entrega contínua e adiantada de software com **valor agregado**.
2. **Mudanças nos requisitos** são bem-vindas, mesmo tardivamente no desenvolvimento. Processos ágeis tiram vantagem das mudanças visando **vantagem competitiva** para o cliente.
3. **Entregar frequentemente** software funcionando, de poucas semanas a poucos meses, com preferência à menor escala de tempo.

03/09/2019

# Os 12 princípios do Manifesto Ágil

1. **Pessoas de negócio** e desenvolvedores devem trabalhar diariamente em conjunto por todo o projeto.
2. Construa projetos em torno de **indivíduos motivados**. Dê a eles o ambiente e o suporte necessário e confie neles para fazer o trabalho.
3. O método mais eficiente e eficaz de transmitir informações para e entre uma equipe de desenvolvimento é através de **conversa face a face**.

03/09/2019

# Os 12 princípios do Manifesto Ágil

1. Software funcionando é a **medida primária de progresso**.
2. Os processos ágeis promovem **desenvolvimento sustentável**. Os patrocinadores, desenvolvedores e usuários devem ser capazes de manter um **ritmo constante** indefinidamente.
3. Contínua atenção à **excelência técnica** e bom design (leia-se, projeto) aumenta a agilidade.

03/09/2019

# Os 12 princípios do Manifesto Ágil

1. **Simplicidade**—a arte de **maximizar a quantidade de trabalho não realizado** —é essencial.
2. As melhores arquiteturas, requisitos e designs emergem de **equipes auto-organizáveis**.
3. Em **intervalos regulares**, a equipe reflete sobre como se tornar mais eficaz e então refina e ajusta seu comportamento de acordo.

<http://agilemanifesto.org/iso/ptbr/principles.html>

03/09/2019

# Sobre o SCRUM

- O SCRUM é um Método de Gestão de Projetos, criado a partir das necessidades da indústria de software [@Cohn2011];
- O método foi criado após sucessivos fracassos da aplicação dos métodos tradicionais de gerenciamento de projetos à projetos de software.
- Atualmente é um dos padrões para o desenvolvimento de software.

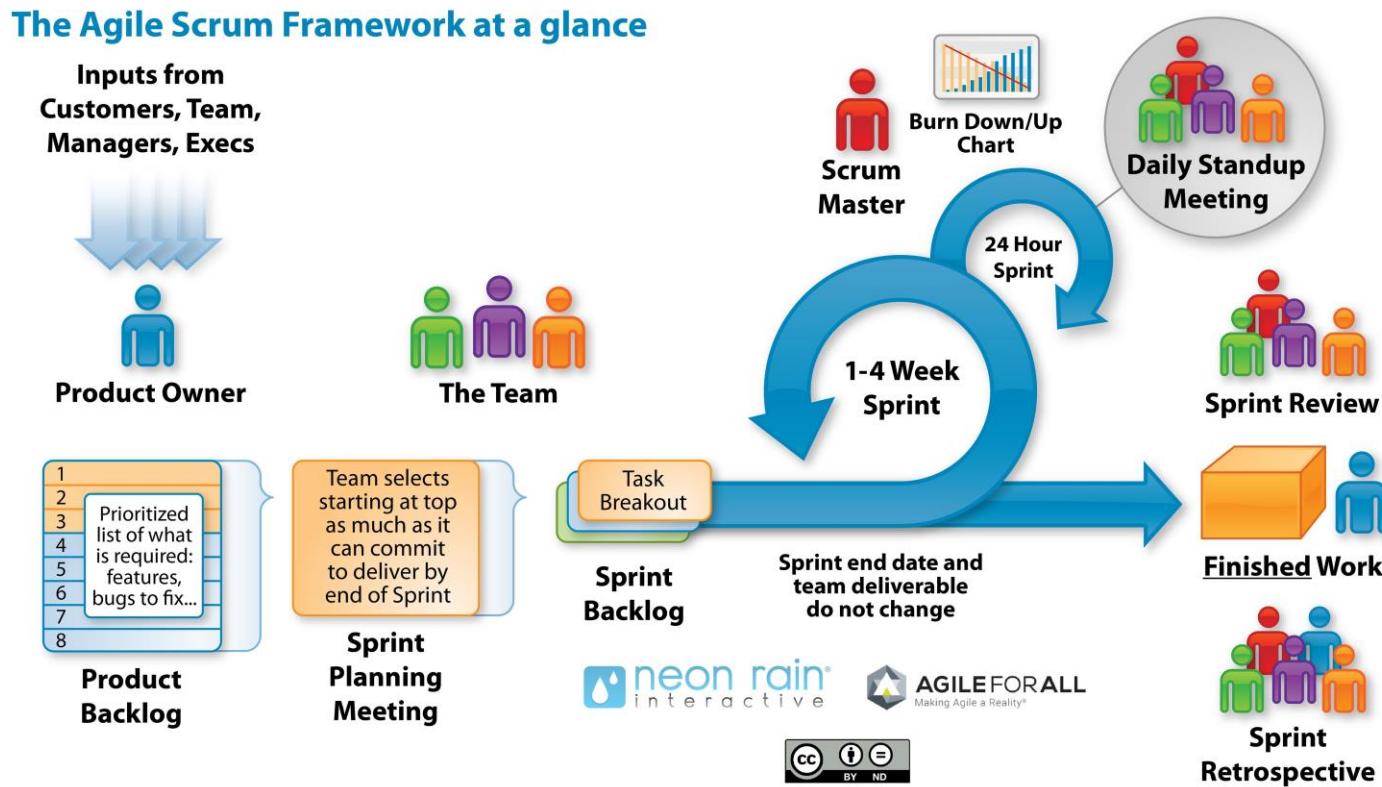
03/09/2019

# Aviso Prévio: O SCRUM não é uma panacéia

- O SCRUM foi criado especificamente para projetos de software que precisam de alto grau de flexibilidade;
- A contribuição desta abordagem é permitir a **flexibilização** do escopo, e a rápida **resposta** à mudanças de escopo.
- O método pode contribuir para projetos de escopo aberto;
- Exemplo: que ganhos de produtividade a sua empresa pode ter a partir da aplicação de técnicas de inteligência articial? Você é capaz de definir detalhadamente os requisitos deste projeto?

03/09/2019

# O Ciclo de Desenvolvimento com o SCRUM



## O Processo Scrum

03/09/2019

# Papéis no Scrum

- **Product Owner (Dono do Produto - Provavelmente serão vocês!):** Prioriza o escopo, e decide quais serão as prioridades.
- **Time:** Executa o projeto, e estima o esforço necessário para execução das atividades. O time define o que pode ser entregue em cada sprint.
- **Scrum Master:** Garante que o processo Scrum é seguido.
- Observação: Estes são os papéis, e não há um “gerente do projeto”. Geralmente, todos na equipe do projeto “trabalham”, e alguém tem a responsabilidade adicional de ser o Scrum Master.

03/09/2019

# Início do Ciclo Scrum - Planejamento do Sprint

Ao início de cada “Sprint” (ex. 2 semanas), a equipe se reúne para definir:

- Quais são os requisitos a serem priorizados naquele sprint (**definidos pelo PO**);
- Quais destes requisitos “cabem” no sprint;
- Como quebrar e dividir estas tarefas;
- Meta: Ao final do Sprint há um produto “potencialmente vendável”.

03/09/2019

# Planejamento do Sprint - Poker Planning

- Para estimar a duração das atividades, é possível utilizar a técnica **poker planning**. - Nesta técnica cada um dos integrantes da equipe escolhe uma carta com a estimativa em horas de esforço para completar o requisito.
- Se todos concordarem com a estimativa, ela é indicada;
- Se houverem estimativas discrepantes, quem estimou o maior tempo explica porquê acredita que o tempo é maior, e quem estimou o menor tempo explica porquê estimou pouco tempo.

03/09/2019

# Gerenciamento dentro do Sprint - Reuniões Diárias



Reunião Diária

03/09/2019

# Gerenciamento dentro do Sprint - Reuniões Diárias

No início de cada dia, 15 minutos são gastos para que cada integrante do grupo responda:

- O que ele **concluiu** no dia anterior;
- O que ele irá **concluir** hoje;
- Se há ou houveram **impedimentos** para realizar suas atividades;
- Atualizar a **estimativa de tempo restante** das atividades que ele está realizando.

03/09/2019

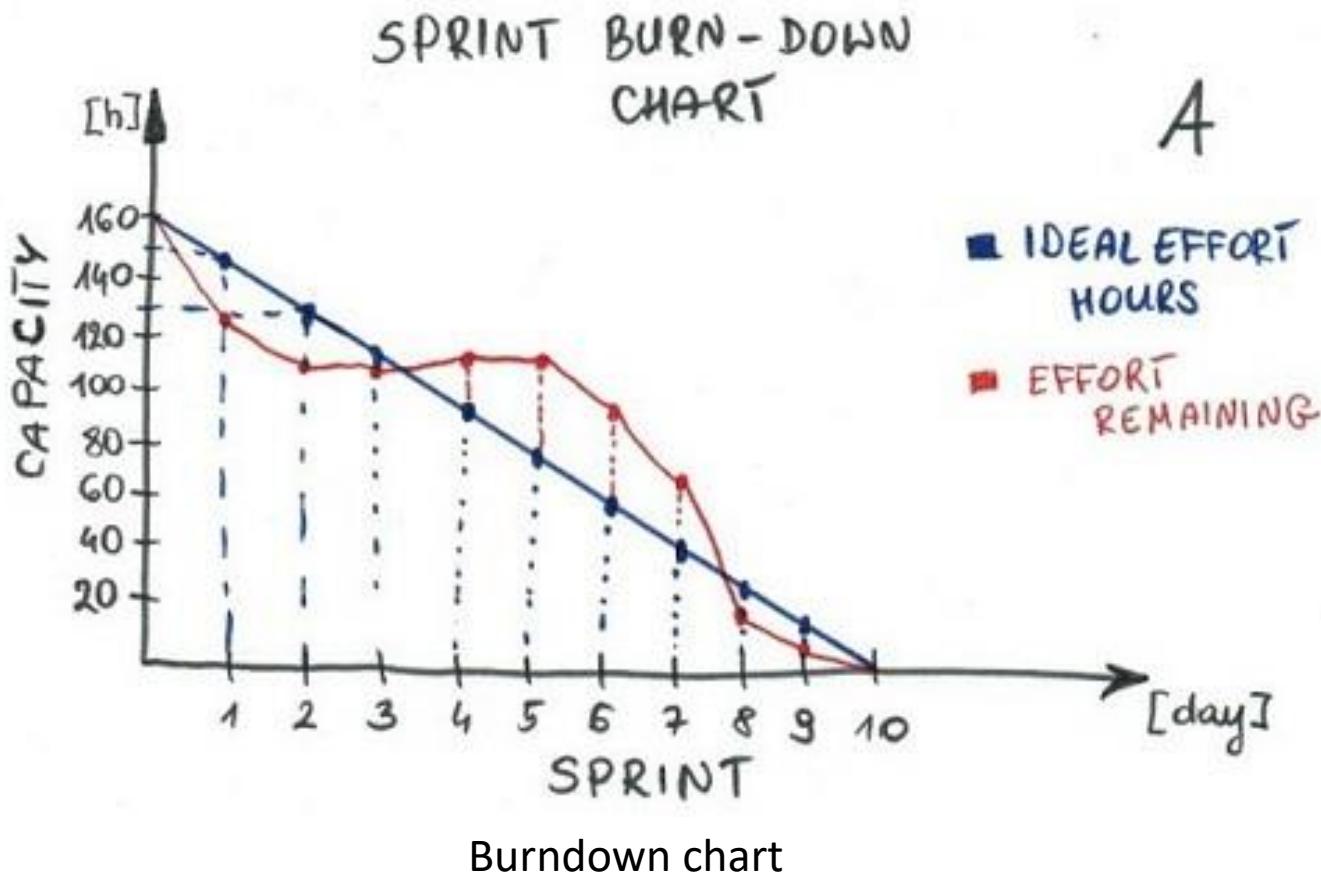
# Gerenciamento dentro do Sprint - Quadro SCRUM



Quadro Scrum

03/09/2019

# Gerenciamento dentro do Sprint - Gráfico Burndown



03/09/2019

# Finalizando o Sprint - Sprint Review

- Ao final do Sprint, é realizada uma revisão do Sprint passado, junto ao **Dono do Produto**, e possíveis clientes;
- A equipe mostra o resultado do trabalho realizado, e as funcionalidades do software são demonstradas;
- O escopo geral do projeto é re-discutido e revisado;
- Novas funcionalidades podem ser sugeridas neste momento.

03/09/2019

# Revisando o SCRUM

- O Scrum é uma abordagem criada especificamente para situações onde se deseja **flexibilidade**;
- Confia-se que o time não fará “corpo mole”. Em muitos casos, se recomenda que o time trabalhe próximo aos usuários fisicamente;
- Engenharia de Software não é só SCRUM. Existem outras técnicas para desenvolvimento (programação em pares, testes unitários, etc...).

03/09/2019

# **ANÁLISE DE REQUISITOS DE SISTEMAS DE INFORMAÇÃO**

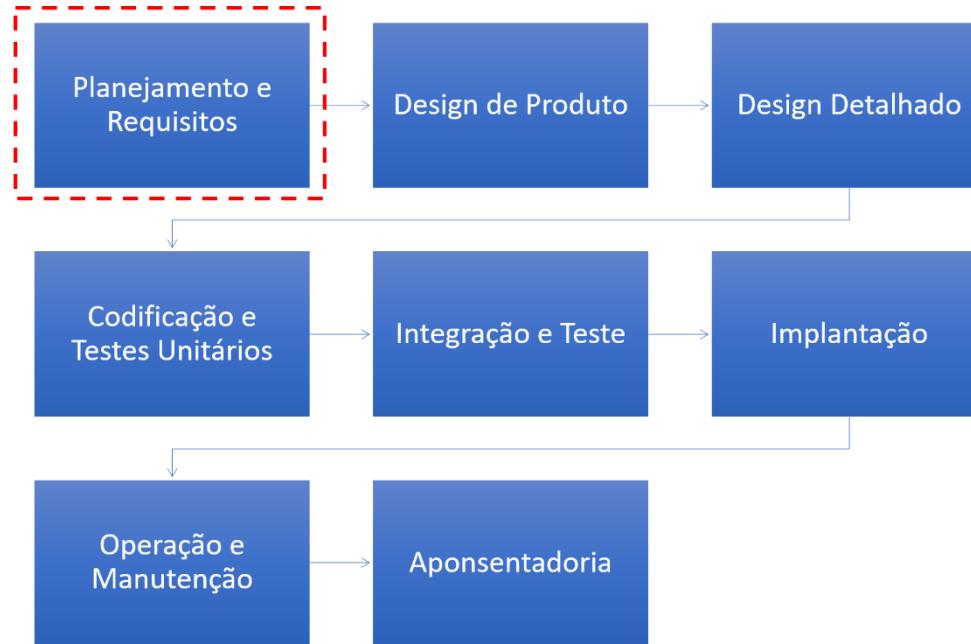
# Análise de Requisitos

- **Requisitos** em um projeto de software são as funcionalidades que o sistema deverá possuir, e as restrições sob as quais o sistema irá operar;
- Seja em um método Ágil, ou tradicional, a análise de requisitos é responsável pela definição e especificação do **escopo** de um projeto de software.
- **Engenharia de Requisitos** é o ramo da Engenharia de Software que trata da análise e gerenciamento de requisitos de software [@Sommerville2011].

03/09/2019

# Localização da Análise de Requisitos

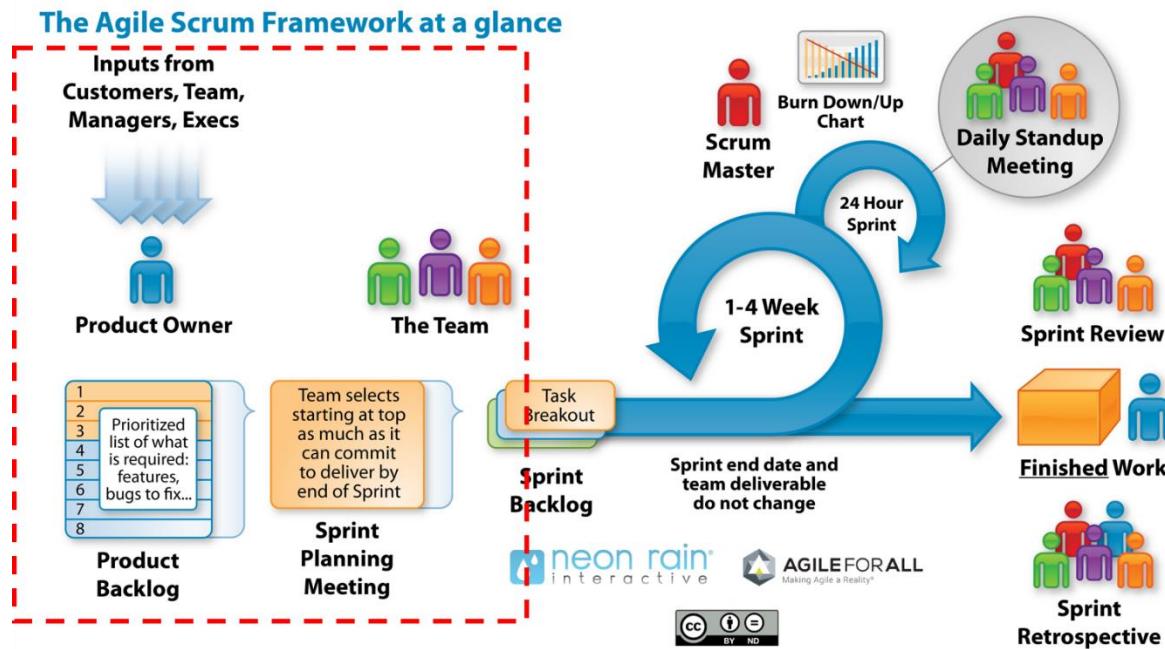
Em um ambiente que utiliza uma abordagem “waterfall”, a análise de requisitos é o primeiro passo do processo:



03/09/2019

# Localização da Análise de Requisitos

Em um ambiente que utiliza uma abordagem como o SCRUM, a análise de requisitos acontece semanalmente, por meio da priorização do Sprint:



03/09/2019

# Exemplo de Projeto com Cronograma, mas sem Escopo:

Se sua empresa tem um projeto de software, mas você **não sabe quais são os requisitos funcionais e não funcionais**, o projeto não tem escopo e tende ao infinito.

- Exemplo Típico de Projeto de Software sem Escopo definido:
- Analisar as necessidades do setor de PCP;
- Propor um software para melhorar o sequenciamento;
- Entregar o software para a fábrica.
- Pergunta: **Este projeto tem um escopo definido?**

03/09/2019

# Classificação de Requisitos de Software

Requisitos podem ser classificados em dois grupos[@Souza2011; @Sommerville2011]:

- **Requisitos Funcionais (RF)**: São as funcionalidades que o sistema deve oferecer a seus usuários, e podem ser descritos em linguagem natural ou diagramas. O foco destes requisitos é o cliente.
- Exemplos: Cadastro de clientes; Relatório de Vendas; Pesquisa de Produtos, etc.
- **Requisitos Não-Funcionais (RNF)**: Tratam-se de restrições técnicas sob as quais o sistema deve operar, e podem tratar do sistema como um todo.
- Exemplos: O sistema deve atender às normas da ANVISA de proteção à informação. O sistema não deve demorar mais do que 15 minutos para gerar o mix de produção ótimo. O sistema deve suportar 1 milhão de acessos simultâneos.

03/09/2019

# Usuários de um Documento de Requisitos de Software

Um documento de requisitos será utilizado por diversos usuários [adaptado de @Sommerville2011]:

## Cientes do Sistema

- Especificam e Leem os requisitos para verificar se o sistema atende às suas necessidades, e especificam alterações nos requisitos.

## Gerentes

- Usam o documento para planejar uma proposta para o sistema e o projeto de desenvolvimento do sistema.

## Engenheiros de Sistemas / Programadores

- Entendem o sistema que será desenvolvido.

## Engenheiros de Teste

- Desenvolvem testes de validação do sistema.

## Engenheiros de Manutenção / Programadores

- Entendem o sistema e suas demais partes.

03/09/2019

# Estrutura de um Documento de Requisitos

Um documento de requisitos pode possuir as seguintes seções[adaptado de @Sommerville2011]:

- Seções Introdutórias: Prefácio; Introdução; Glossário;
- Definição de Requisitos Funcionais\*;
- Arquitetura do Sistema\*: Uma visão geral sobre os módulos do sistema e a distribuição de funções entre estes módulos.
- Definição de Requisitos Não-Funcionais\*;
- Modelos do Sistema\*: Apresentação de modelos visuais que representam o sistema e seus componentes internos ou o sistema e o ambiente. Podem incluir modelos de objetos, modelos de fluxo de dados, modelos de transição, modelos de entidade e relacionamento, etc.;
- \* Tópicos obrigatórios para o trabalho do Grau B.

03/09/2019

# Exemplos - Quadro de Requisitos Funcionais

| Requisito Funcional                             | Descrição                                                                                                                                                                                                                                                                       |
|-------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RF-01 - Login por Rede Social                   | Permite que o usuário acesse o sistema utilizando sua conta do facebook.                                                                                                                                                                                                        |
| RF-02 - Login por Acesso Anônimo                | Permite que o usuário acesse o sistema sem informar dados pessoais. Ainda assim, permitirá a rastreabilidade do usuário.                                                                                                                                                        |
| RF-03 - Busca de Locais / Instituições Próximas | Permitirá que o usuário localize as instituições que estão mais próximas da sua localização atual. Será possível visualizar as instituições em um mapa ou em uma lista.                                                                                                         |
| RF-04 – Tela de Resumo da Instituição           | Exibirá informações de identificação da instituição pública (Nome, Endereço, Índice de Qualidade Calculado, Posição no Ranking Estadual).                                                                                                                                       |
| RF-05 - Tela de Detalhes da Instituição         | Esta tela conterá informações mais detalhadas sobre a instituição sendo avaliada. Será possível consultar o seu índice de qualidade, a evolução deste indicador ao longo dos últimos 6 meses, bem como a sua posição nos rankings nacionais, regionais, estaduais e municipais. |

Quadro de Requisitos Funcionais

03/09/2019

# Exemplos - Quadro de Requisitos Não-Funcionais

| Requisito Não Funcional                       | Descrição                                                                                                                                                                                                                                                                                                                            | Necessário Para RF... |
|-----------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| RNF-01 – Integração com o Google Places       | Para que a lista de locais de avaliação seja disponibilizada ao usuário, será necessário implementar uma integração da base de dados do sistema á um serviço que disponibilize dados sobre a localização das instituições de serviço público. Além disto, esta base de dados deve ser atualizada periodicamente.                     | RF-03; RF-04          |
| RNF-02 – Tratamento de Dados de Questionários | O sistema deve conter barreiras que diminuam a probabilidade de um usuário tentar burlar os dados do índice de qualidade do serviço público. Para isto, o sistema identificará usuários com alta frequência de avaliação e definirá um limite de avaliações de um mesmo usuário para uma mesma instituição em um intervalo de tempo. | RF-08                 |

Quadro de Requisitos Funcionais

03/09/2019

# Exemplos - Sketch de Telas

**03 – Busca de Instituições - Lista**

Funcionalidade de busca de instituições por nome ou proximidade geográfica. Exibe uma lista ordenada por itens de menor para maior distância do usuário.

The sketch shows a mobile application interface. At the top, a blue header bar displays the title "Instituições próximas" and a search icon. Below the header, there are two tabs: "LISTA" (selected) and "MAPA". The main content area lists six institutions with their names, addresses, and distances from the user:

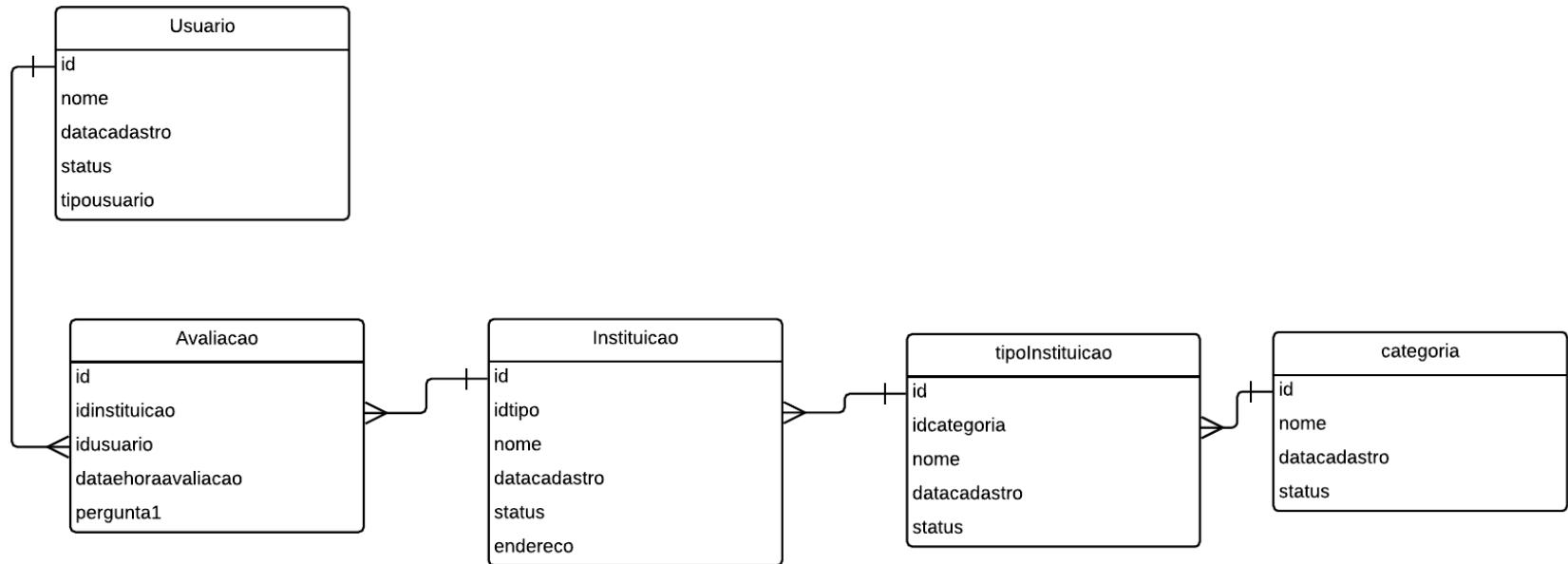
| Instituição                  | Endereço                     | Distância |
|------------------------------|------------------------------|-----------|
| UPA Zona Norte               | Praça Ernest Ludwig Herman   | 110m      |
| Linha T1 - Carris            | Porto Alegre, RS             | 345m      |
| Secretaria Estadual da Saúde | Av. Borges de Medeiros, 1501 | 380m      |
| Hospital Cristo Redentor     | R. Domingos Rubbo, 20        | 699m      |
| Centro de Saúde IAPI         | R. Três de Abril, 90         | 1.3km     |
| Hospital Conceição           | Av. Francisco Trein, 596     | 1.4km     |

At the bottom of the screen, there is a black navigation bar with three icons: a triangle pointing left, a circle, and a square.

Sketch de Telas

03/09/2019

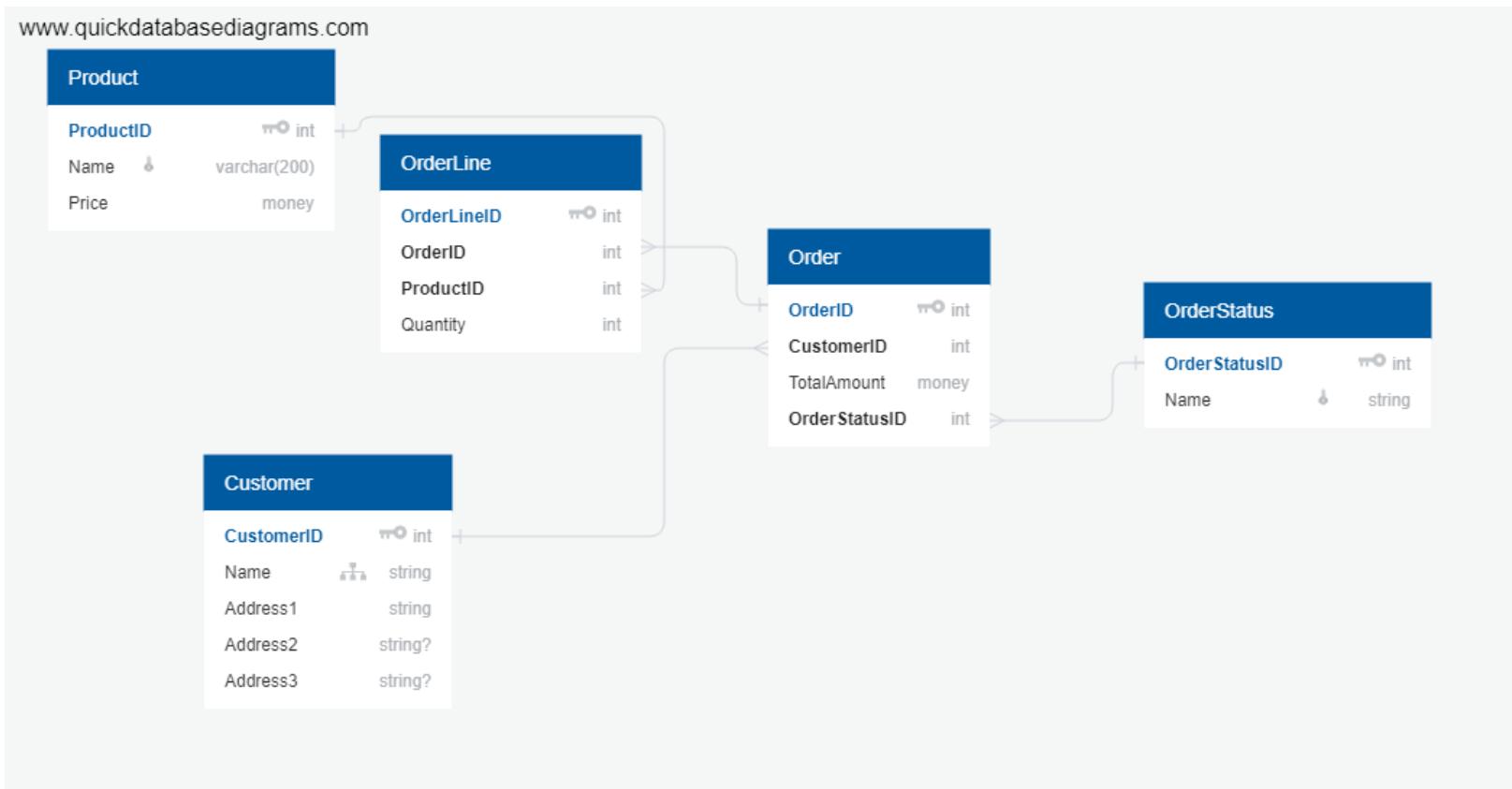
# Exemplos - Modelo da Estrutura dos Dados



DER Inicial - Avalia Brasil

03/09/2019

# Exemplo com o Quick Database Diagram:



DER - Exemplo do App Quick Database

03/09/2019

# O Processo é Iterativo

O processo de engenharia de requisitos é altamente iterativo [adaptado de @Sommerville2011]:



03/09/2019

# Técnicas para Definição / Elicitação de Requisitos

Requisitos podem ser definidos utilizando diversas técnicas, incluindo:

- **Entrevistas;**
- **Definição de Cenários:** Exemplos concretos de uso do sistema;
- **Definição de Casos de Uso:** Técnica de descoberta de requisitos presente na UML (Unified Modeling Language).

03/09/2019

# Detalhando os Requisitos com Casos de Uso

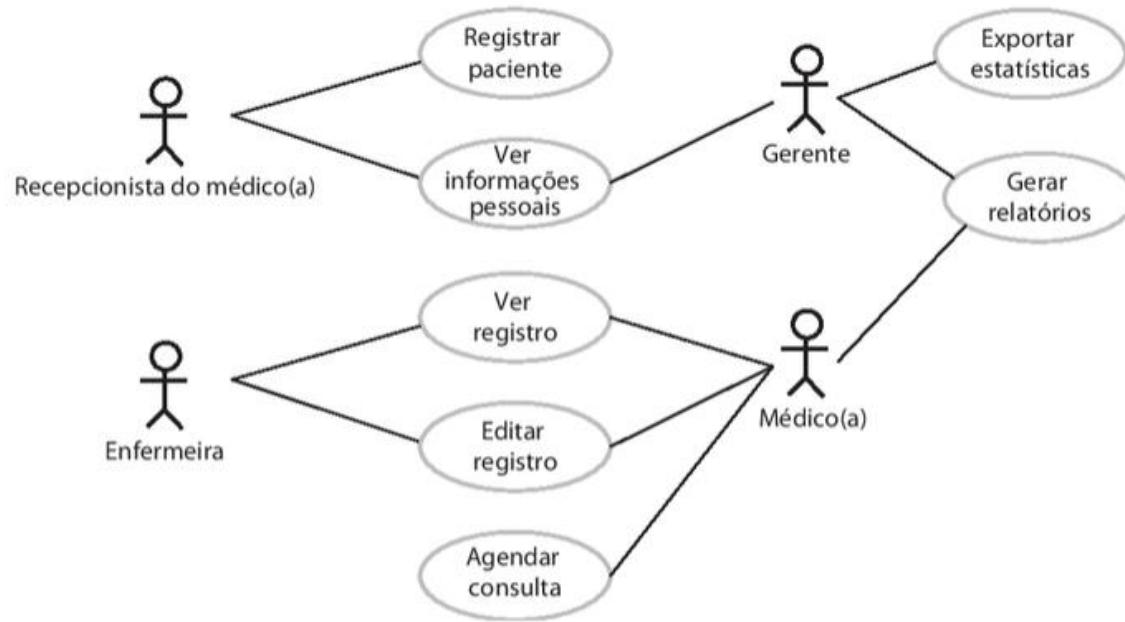
Em sua forma mais simples, um caso de uso identifica [@Sommerville2011]:

- **Atores:** Quais serão os atores envolvidos em interações com o sistema, e;
- **Interações:** Qual é o conjunto de interações que ocorrerá no sistema.
- Estes dois elementos são descritos em um diagrama de caso de Uso.

03/09/2019

# Diagrama de Casos de Uso

Exemplo diagrama de Caso de Uso  
[@Sommerville2011, p. 75]:



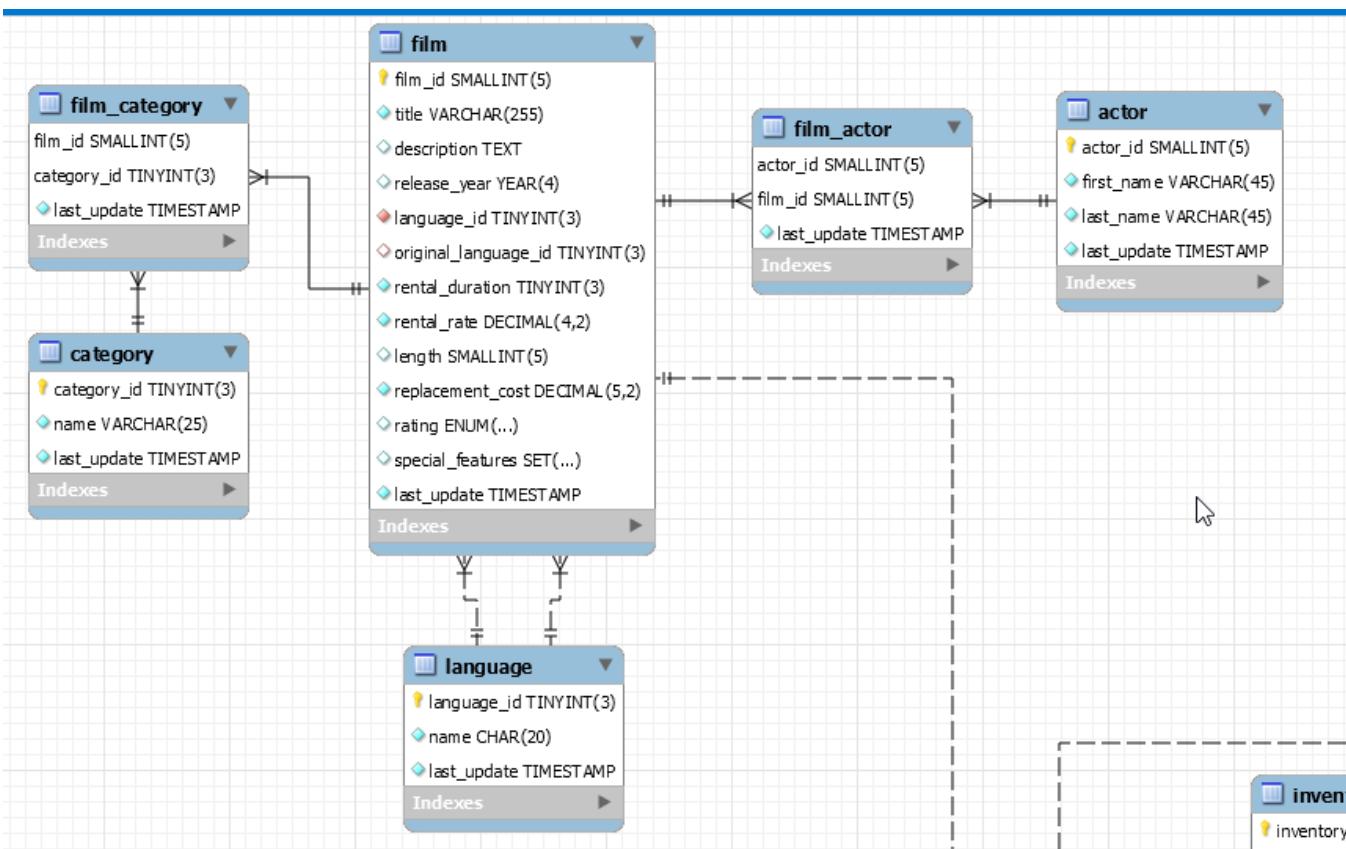
03/09/2019

# Descrição do Caso de Uso

Para cada um dos casos de uso identificados, deve ser elaborada uma **descrição das interações** entre os atores. Exemplo:

- Usuário insere e-mail e senha na tela de login;
- Sistema valida informações do usuário;
- Usuário acessar tela de cadastro de novo evento;
- ...

03/09/2019



# SISTEMAS DE GESTÃO DE BANCO DE DADOS

# SGBDs - Importância

Os SGDBs estão relacionados a...

- Onde ficam, as informações dos seus clientes, pedidos, ordens de produção, etc.;
- Como estes dados são **organizados e estruturados**;
- Se a empresa quiser modificar seus sistemas atuais para permitir novos tipos de informação, o que será necessário modificar;
- Se você quiser criar um novo sistema melhor do que o atual, a sua atual estrutura de dados permite, ou não?
- Até que ponto estes dados são **íntegros e confiáveis?**

03/09/2019

# SGBDs - Importância

## Exemplo Ilustrativo:

A empresa UrnasSINOS contratou um novo sistema MES. O Sistema, prometendo levar a empresa a novos patamares de produtividade, foi comprado e instalado por R\$ 467.000,00, sob a sua autorização. **Você é o responsável** pelo seu sucesso ou seu fracasso. O sistema foi **customizado** para atender as necessidades da empresa, e você supervisionou todo o processo.

Nos primeiros meses de instalação, o sistema funcionou bem, porém alguns sintomas estranhos começaram a aparecer....

03/09/2019

# SGBDs - Importância

Exemplo Ilustrativo:

**1. Eficiências maiores do que 100 %:** Em alguns momentos, o sistema indica que a eficiência de produção foi maior do que 100%. Todos da equipe, evidentemente, ficaram surpresos por tamanho **façanha** do sistema. Ninguém poderia acreditar que o sistema apresentasse tanto êxito. Para resolver o problema, **você** pediu que a equipe de TI simplesmente fizesse com que, caso o calculo chegasse a um valor maior do que 100 %, o sistema indicasse em seu relatório o valor 100 %.

Em um primeiro momento, isto resolveu o problema, e salvou a sua reputação dentro da empresa, porém alguns meses depois, outro problema ocorreu...

03/09/2019

# SGBDs - Importância

Exemplo Ilustrativo:

**1. Eficiências flutuantes:** Após alguns meses de operação, as eficiências calculadas pelo sistema simplesmente mudaram abruptamente para menos (uma mudança na ordem de 20 %), porém a fábrica está trabalhando como nunca!

A partir deste momento, o sistema entrou em descrédito, e todos pararam de acreditar nos resultados do sistema...

03/09/2019

# Banco de Dados - Conceitos

- Um **banco de dados** pode ser definido como uma coleção de dados relacionados. [@Amadeu2015];
- Um **sistema de Gerenciamento de Banco de Dados** (SGDB, correspondente ao inglês Data Base Management System - DBMS) é um software responsável pelo gerenciamento dos dados de um sistema.
- Os Sistemas SGBDs permitem **definir, construir, manipular e compartilhar** bancos de dados de modo estruturado;
- A maior parte dos sistemas de informação industriais utiliza um ou mais SGDBs para armazenar os seus dados. É comum que as empresas possuam dezenas de bases de dados controlando seus processos.

03/09/2019

# Principais SGBDs

SGDBs mais populares:

| Rank        |             |             | DBMS                 | Database Model    | Score       |             |             |
|-------------|-------------|-------------|----------------------|-------------------|-------------|-------------|-------------|
| Oct<br>2018 | Sep<br>2018 | Oct<br>2017 |                      |                   | Oct<br>2018 | Sep<br>2018 | Oct<br>2017 |
| 1.          | 1.          | 1.          | Oracle               | Relational DBMS   | 1319.27     | +10.15      | -29.54      |
| 2.          | 2.          | 2.          | MySQL                | Relational DBMS   | 1178.12     | -2.36       | -120.71     |
| 3.          | 3.          | 3.          | Microsoft SQL Server | Relational DBMS   | 1058.33     | +7.05       | -151.99     |
| 4.          | 4.          | 4.          | PostgreSQL           | Relational DBMS   | 419.39      | +12.97      | +46.12      |
| 5.          | 5.          | 5.          | MongoDB              | Document store    | 363.19      | +4.39       | +33.79      |
| 6.          | 6.          | 6.          | DB2                  | Relational DBMS   | 179.69      | -1.38       | -14.90      |
| 7.          | ↑ 8.        | ↑ 9.        | Redis                | Key-value store   | 145.29      | +4.35       | +23.24      |
| 8.          | ↓ 7.        | ↑ 10.       | Elasticsearch        | Search engine     | 142.33      | -0.28       | +22.09      |
| 9.          | 9.          | ↓ 7.        | Microsoft Access     | Relational DBMS   | 136.80      | +3.41       | +7.35       |
| 10.         | 10.         | ↓ 8.        | Cassandra            | Wide column store | 123.39      | +3.83       | -1.40       |

[www.db-engines.com/en/ranking](http://www.db-engines.com/en/ranking)

03/09/2019

# Como a Informação é organizada em um Banco de Dados?

- Para que seja útil, é importante estruturar a informação em um banco de dados de modo coerente;
- É necessário definir quais são as relações entre as diferentes “entidades” em uma empresa;
- Para tanto, é comum a utilização de **modelos do banco de dados** durante as fases de projeto do banco de dados;
- No contexto desta disciplina, focalizaremos nossa atenção para a **compreensão dos principais conceitos** presentes na modelagem e uso de bancos de dados.

03/09/2019

# Níveis de Modelagem do Banco de Dados

Existem três níveis de modelos do banco de dados:

- **Modelo Conceitual:** Representa um primeiro nível de abstração sobre as informações a serem retratadas no banco de dados. É uma definição de alto nível, sem possuir muitos detalhes sobre o modelo de dados.
- **Modelo Lógico:** Apresenta mais detalhes sobre as entidades (incluindo seus atributos e relações de modo detalhadado). Idealmente, apenas é criado após a definição do modelo lógico.
- **Modelo Físico:** Descreve detalhadamente as estruturas físicas de armazenamento dos dados incluindo tipos e tamanhos dos campos, índices, domínios, campos obrigatórios, etc.

03/09/2019

# Modelos de Bases de Dados - O Modelo Entidade-Relacionamento

A modelagem de entidade e relacionamento foi proposta originalmente por Peter Chen, e possui os seguintes elementos básicos:

- **Entidade:** Representa um objeto ou conceito do mundo real (exemplo: turma, funcionário, cliente, ordem de produção). Em um banco de dados real, cada entidade terá a sua própria tabela.
- **Atributo:** Representa propriedades de interesse de uma entidade (exemplo: nome, salário, data de nascimento de um funcionário). Em um banco de dados real, cada atributo será uma coluna de uma tabela da entidade.
- **Relacionamento:** Representa uma relação entre duas ou mais entidades (exemplo: um aluno pode participar em diversas turmas, e uma turma possui diversos alunos). Em um banco de dados real, as relações serão expressas em tabelas que relacionando entidades, bem como por suas chaves estrangeiras).

03/09/2019

# Características dos Relacionamentos:

Os relacionamentos entre as entidades podem ser caracterizados segundo:

- Sua **Cardinalidade**: Descreve como as entidades se relacionam entre si. Exemplo - Um pai tem “n” filhos (relação 1 para n). Existem restrições de cardinalidade mínima (exemplo, uma pessoa tem no mínimo 0 filhos, porém uma casa tem no mínimo um proprietário) e máxima (uma pessoa tem no máximo dois pais biológicos, porém pode ter “n” filhos).
- O número de **Graus** de um Relacionamento: Binário, ternário e e-nário. Refere-se à existência de duas, três ou mais entidades envolvidas no fato que o relacionamento representa.

03/09/2019

# Cardinalidade - Classificações Comuns

- **Um-para-Um (1:1)**: Cada ocorrência de cada entidade se associa a uma ocorrência da outra Entidade. (Exemplo - “Um Departamento é gerenciado por um e somente um Funcionário. Cada funcionário gerencia um e somente um Departamento”). Este tipo de relação é raro.
- **Um-para-Muitos (1:n)**: Uma entidade pode se relacionar a “n” outras instâncias de outra entidade. Exemplo: Uma máquina possui “n” ordens de serviço. Um pai possui “n” filhos. Um aluno está vinculado a “n” provas individuais.
- **Muitos-para-Muitos (n:n)**: Neste caso, existe uma relação mútua de muitos para muitos. Exemplos: Uma nota fiscal possui “n” produtos, e um produto pode estar em “n” notas fiscais.

03/09/2019

# Tipos de Atributos

- **Atributos Identificadores (Chaves)**: São utilizados únicamente para identificar a ocorrência de uma entidade. Em uma base relacionanal, normalmente corresponde à chave primária (ex.: O Seu ID na Unisinos). Usualmente, as chaves serão números inteiros, gerados sequencialmente, e automaticamente pela base de dados.
- **Atributos Descritores**: Todos os outros que servem para descrever uma entidade. Estes atributos podem ser de vários tipos (Caracteres, Números, Datas, etc.).

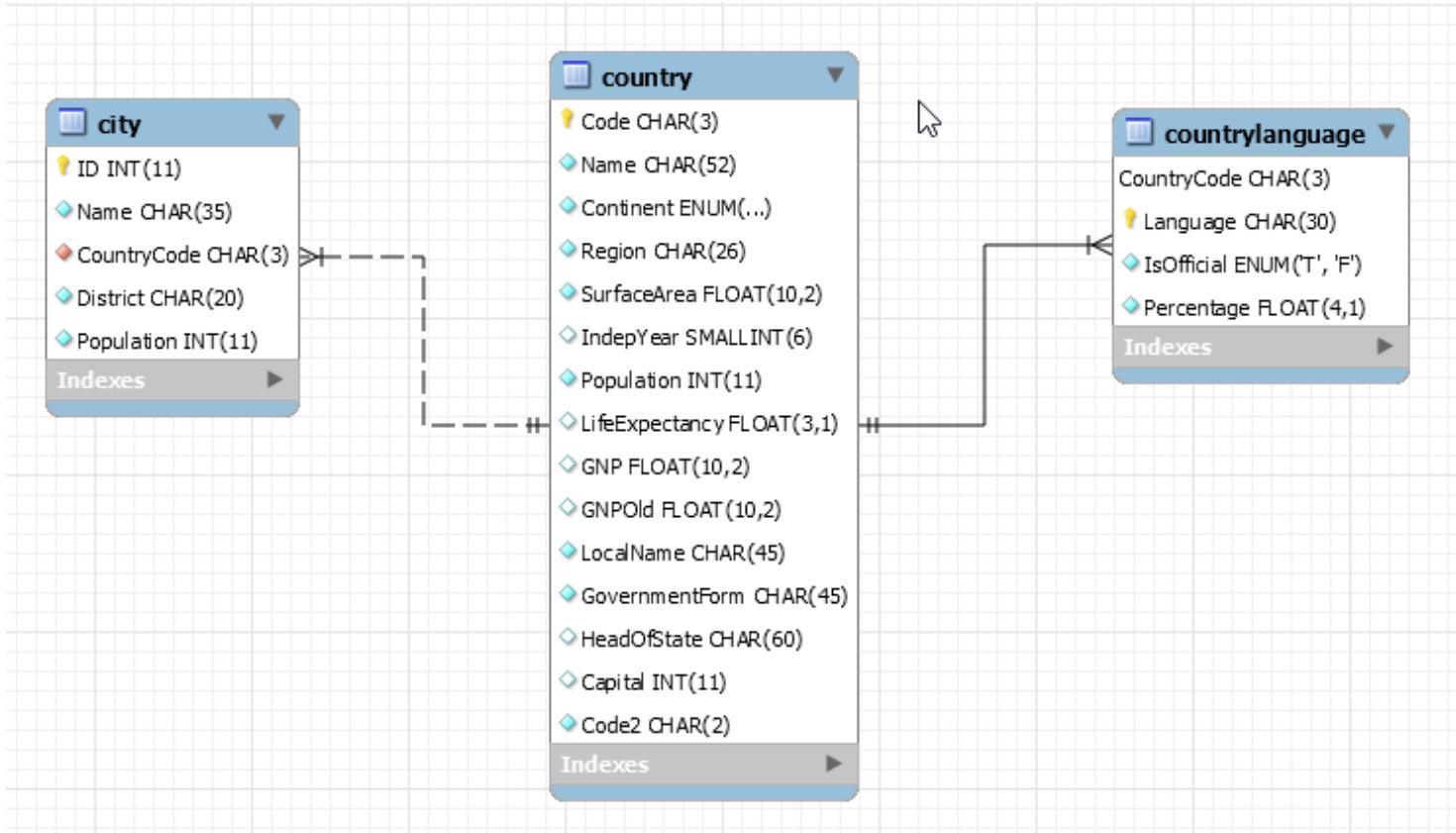
03/09/2019

# Notações para Diagramas ER:

- A notação de **Chen** (original), contém os conceitos apresentados anteriormente, porém não é suportada pela maioria dos softwares de projeto de bases de dados. É uma notação propícia para a modelagem conceitual dos dados;
- A notação de **James Martin** (Ou notação da “Engenharia de Informação”, “Crow’s Foot”, ou “Pé de Galinha”) é utilizada pela maioria dos softwares de projeto de bases de dados, e permite a definição do modelo lógico;
- Existem ainda outras notações (Bachman, Barkers, EXPRESS, IDEF1X, UML, etc.).

03/09/2019

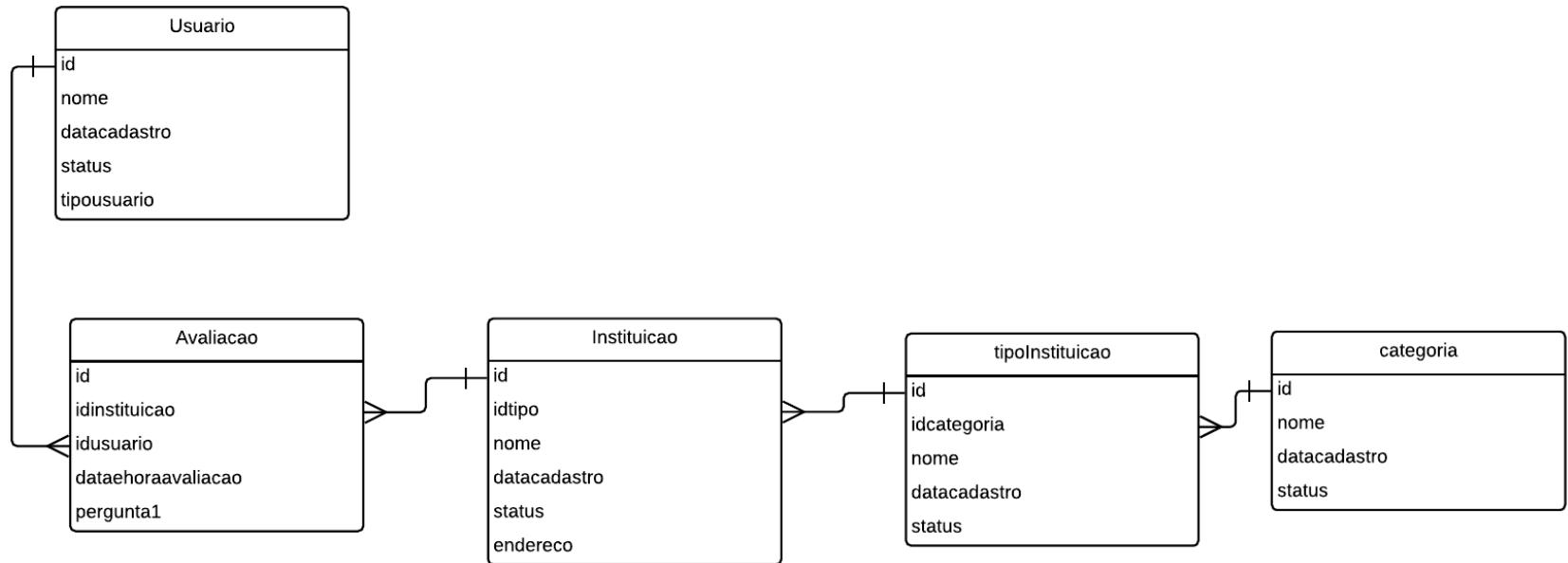
# Exemplo 1 - Modelo de James Martin



Exemplo 1 - Modelo de Martin

03/09/2019

# Exemplo 2 - Modelo de James Martin



DER Inicial - Avalia Brasil

03/09/2019

# Da Teoria para a Prática: Esboce um DER

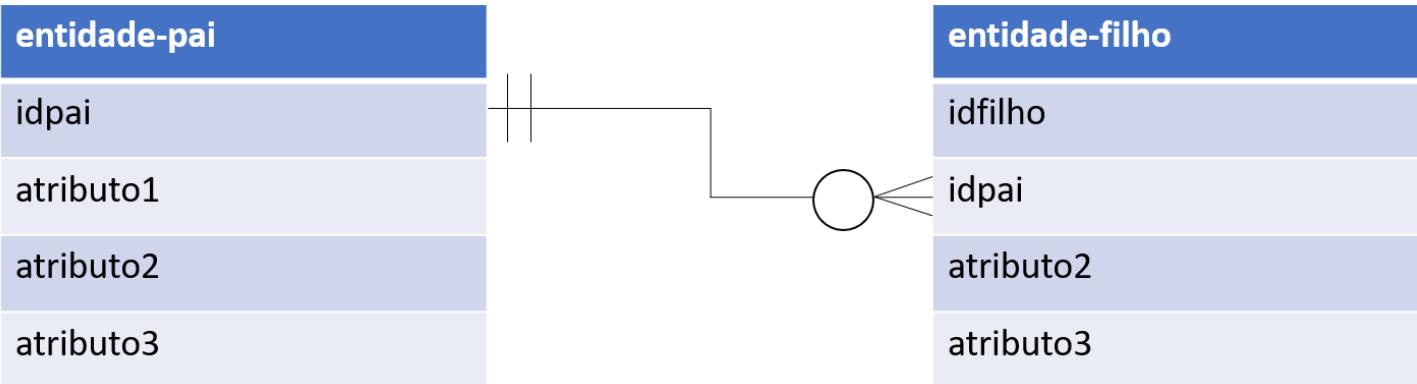
Imagine que você deseja estruturar as informações necessárias em uma base de dados de uma locadora de filmes. Apenas se preocupe com o módulo de cadastro de filmes (mantenha fora do seu escopo todos os aspectos operacionais - aluguel de filmes, pagamentos, controle de inventário, etc).

**Esboce o melhor DER que conseguir em 15 minutos,** representando as entidades e relações. Caso sobre tempo, inclua outras entidades que julgue necessário.

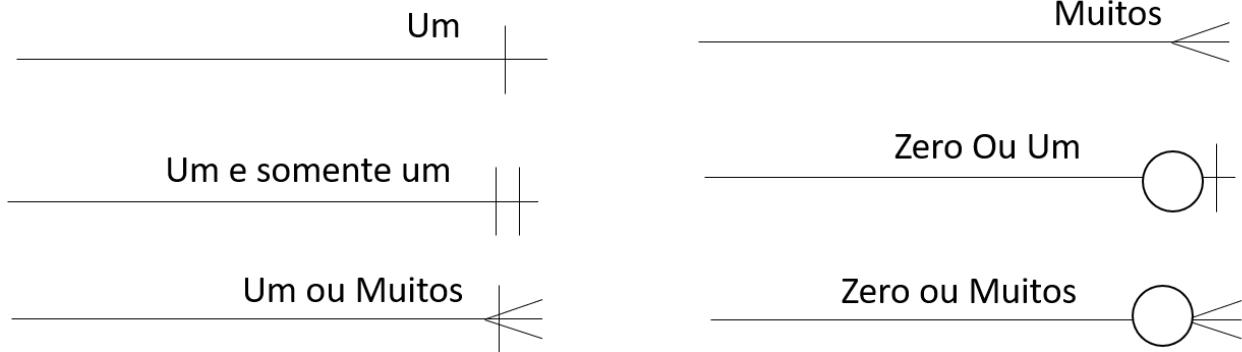
03/09/2019

# Cola - Notação – James Martin (Crow's foot)

Exemplo de Entidades e Relação entre elas



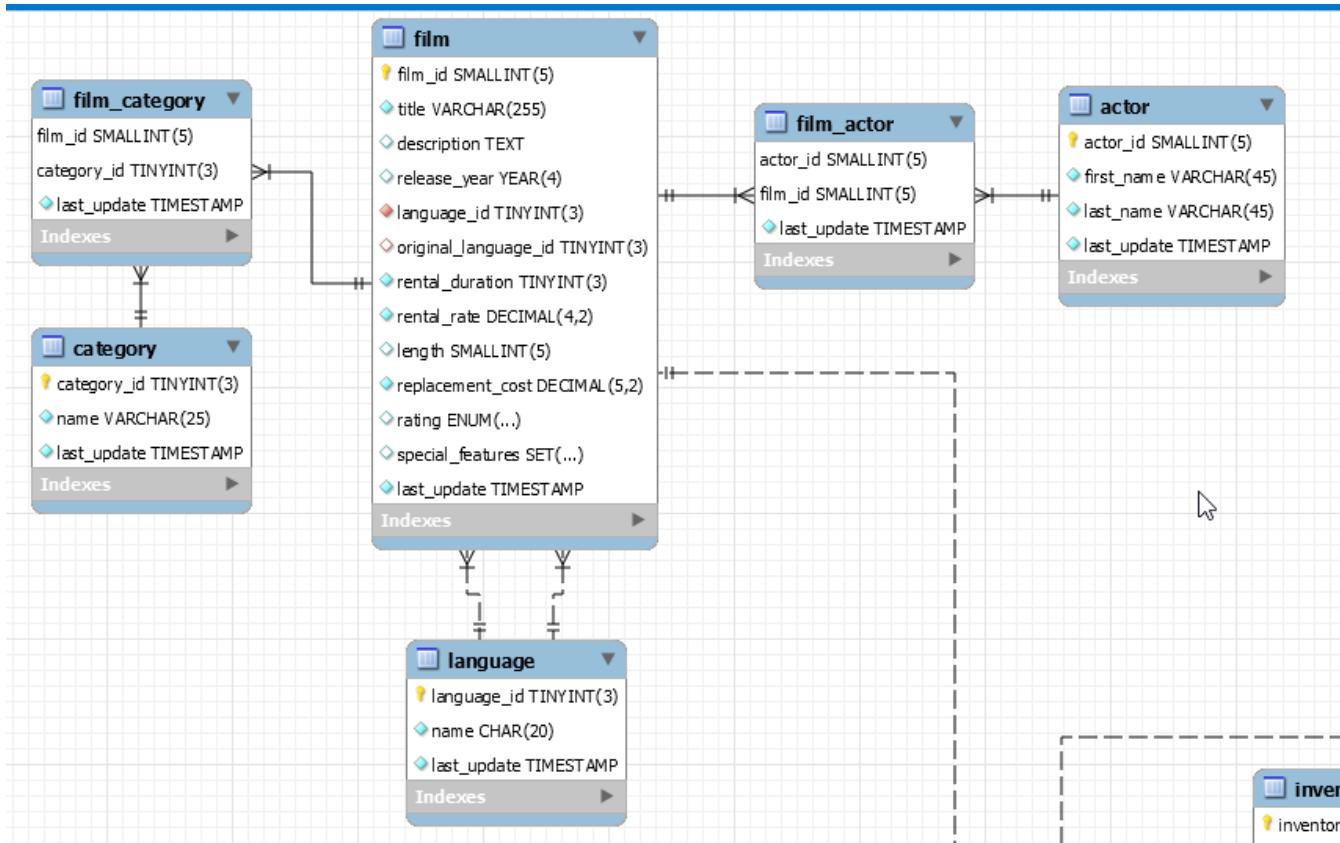
Relações Possíveis



Símbolos Básicos Utilizados na Notação Crow's Foot

03/09/2019

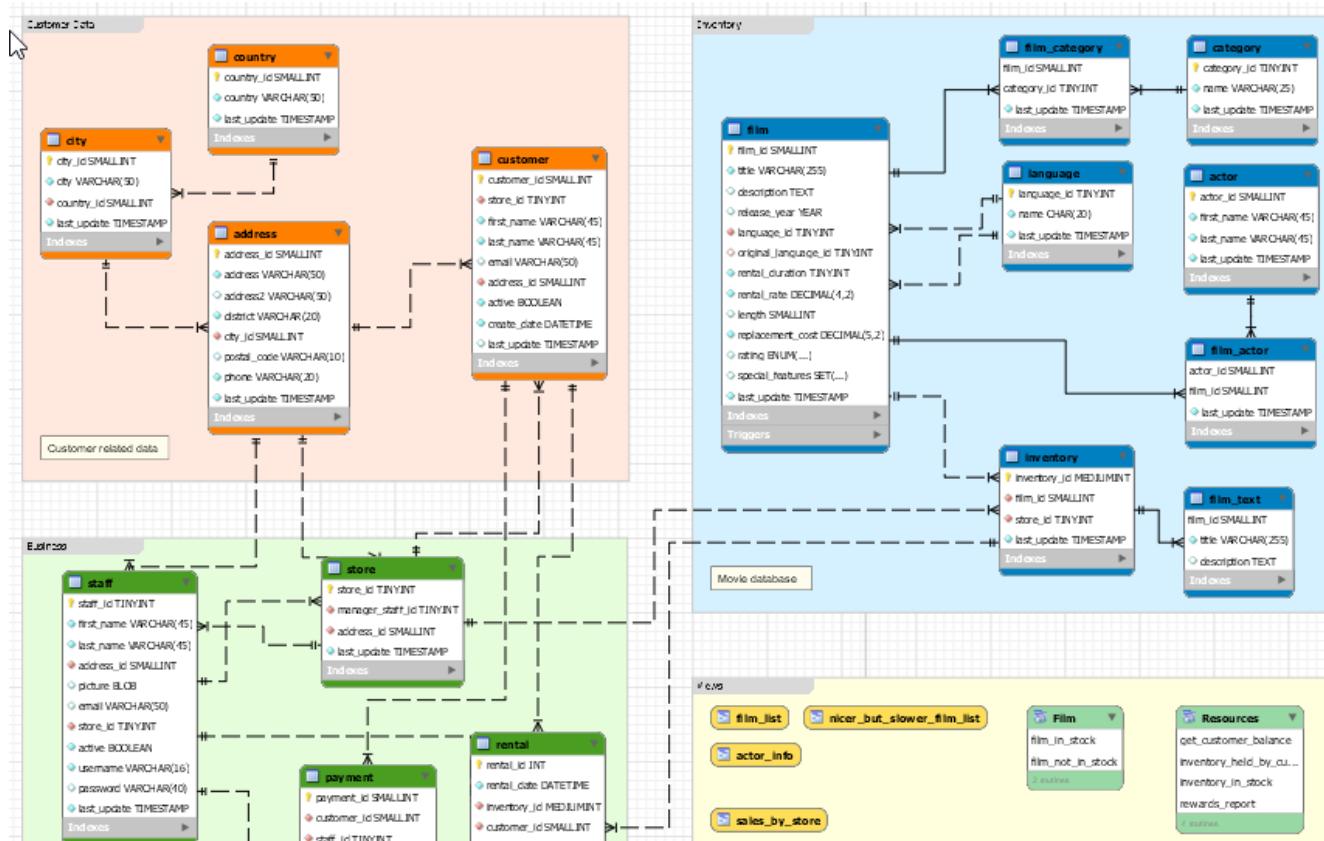
# Exemplo 3 - Solução Possível



Solução possível para o Exemplo dos Filmes

03/09/2019

# Exemplo 3 - Adicionando Mais Tabelas



Solução com mais Tabelas

03/09/2019

# Acessando um Banco de Dados (Via linha de comando):

- É possível acessar bases de dados via prompt de comandos:

```
MySQL 8.0 Command Line Client - Unicode
+----+-----+-----+-----+-----+
| ID | Name | CountryCode | District | Population |
+----+-----+-----+-----+-----+
208	Salvador	BRA	Bahia	2302832
209	Belo Horizonte	BRA	Minas Gerais	2139125
210	Fortaleza	BRA	CearÁa	2097757
211	BrasÁlia	BRA	Distrito Federal	1969868
212	Curitiba	BRA	ParanÁa	1584232
213	Recife	BRA	Pernambuco	1378087
214	Porto Alegre	BRA	Rio Grande do Sul	1314032
215	Manaus	BRA	Amazonas	1255049
+----+-----+-----+-----+-----+
10 rows in set (0.00 sec)

mysql> select * from city where Name = "Porto Alegre";
+----+-----+-----+-----+
| ID | Name | CountryCode | District | Population |
+----+-----+-----+-----+
| 214 | Porto Alegre | BRA | Rio Grande do Sul | 1314032 |
+----+-----+-----+-----+
1 row in set (0.00 sec)

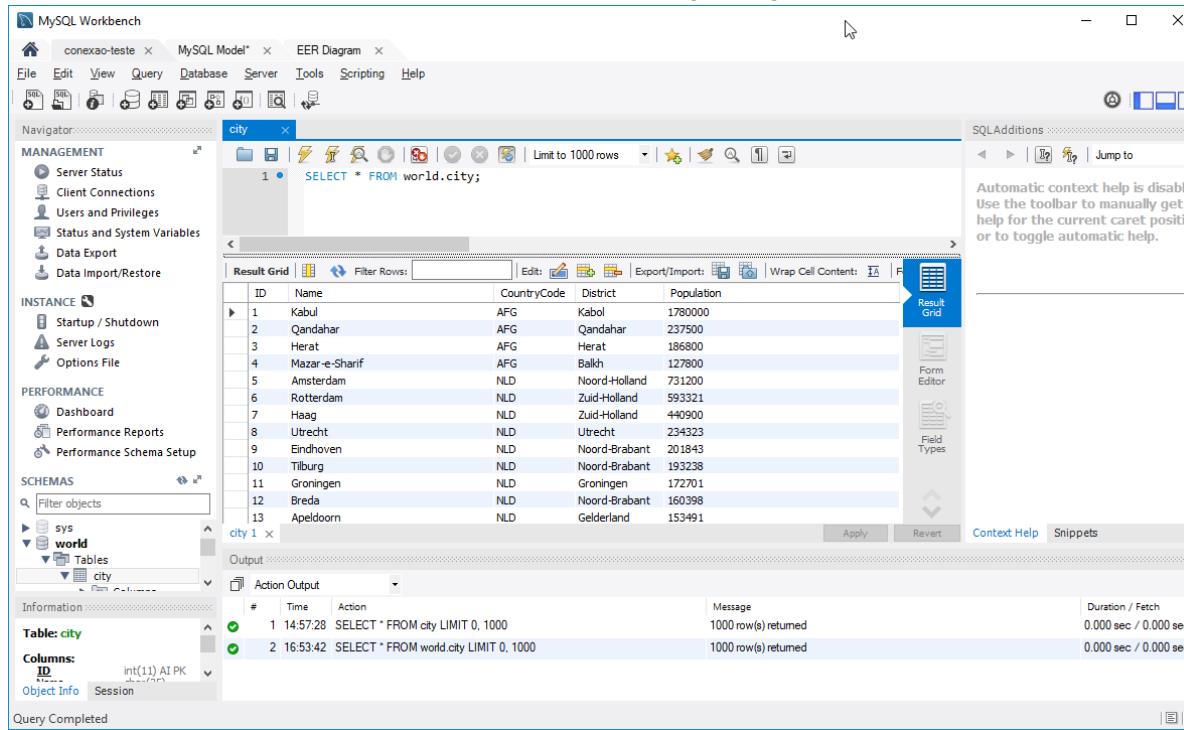
mysql> select sum(Population) from city where CountryCode = "BRA";
+-----+
| sum(Population) |
+-----+
| 85876862 |
+-----+
1 row in set (0.00 sec)

mysql> select * from country where city = "BRA";
ERROR 1054 (42S22): Unknown column 'city' in 'where clause'
mysql>
```

03/09/2019

# Acessando um Banco de Dados (Via Software):

– Porém é mais comum que o acesso seja feito via software: Acessando o Mysql Via Software



03/09/2019

# Esquemas e Instâncias de Bases de Dados

Uma distinção conceitual importante é a descrição do banco de dados e o próprio banco de dado:

- O **Esquema** do banco de dados trata-se a descrição do mesmo. O esquema de um banco de dados define todas as suas entidades, atributos e relações. Embora o esquema de um banco de dados possa mudar, o mesmo muda muito menos do que suas instâncias.
- Uma **instância** do banco de dados é o banco de dados em si, contendo todas as suas informações.

03/09/2019

# Execução do Trabalho

- Iniciar o Projeto da Base de Dados:
- Para a entrega do trabalho, será obrigatório apresentar o projeto da base de dados utilizando o software gratuito:

<https://dbdiagram.io>

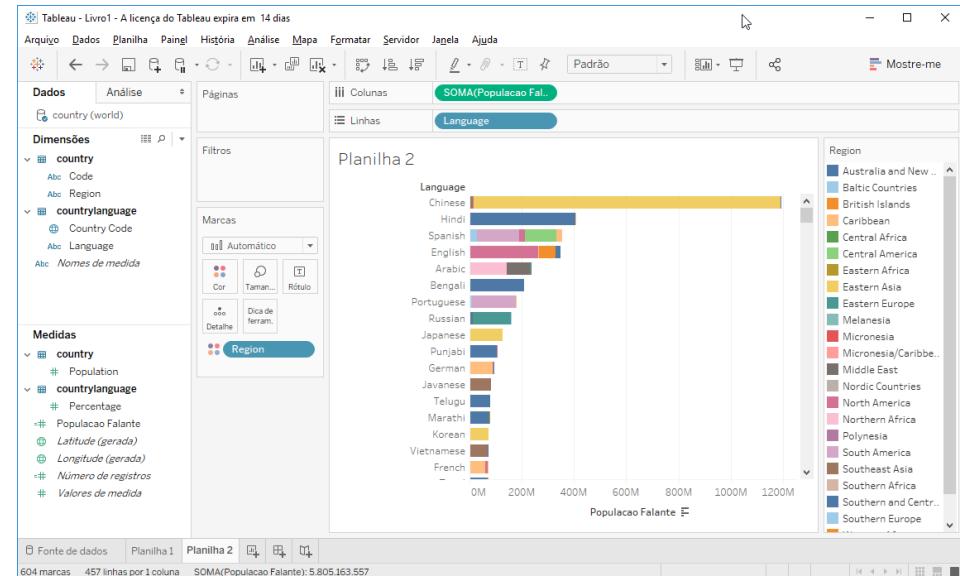
03/09/2019

# Dicas para a Realização do Exercício

- Começar listando apenas três entidades mais importantes, criar as tabelas de modo independente;
- Se necessário, criar as entidades que as interligam;
- Não deletar exemplo gerado pelo software – utilizar relações como referência!
- O Diagrama utilizado para o trabalho final deve possuir no mínimo 7 entidades (tabelas).

03/09/2019

```
-- Copiando estrutura para tabela avaliabrazil.avaliacao
CREATE TABLE IF NOT EXISTS `avaliacao` (
 `idavaliacao` int(11) NOT NULL AUTO_INCREMENT,
 `datacadastro` datetime NOT NULL DEFAULT CURRENT_TIMESTAMP,
 `idusuario` int(11) NOT NULL,
 `idinstituicao` int(11) NOT NULL,
 `status` tinyint(1) NOT NULL DEFAULT '0',
 `idinstrumento` int(11) NOT NULL,
 PRIMARY KEY (`idavaliacao`),
 UNIQUE KEY `idavaliacao_UNIQUE` (`idavaliacao`),
 KEY `idusuario_idx` (`idusuario`),
 KEY `idinstituicao_idx` (`idinstituicao`),
 KEY `idinstrumento_idx` (`idinstrumento`),
 CONSTRAINT `idinstituicao` FOREIGN KEY (`idinstituicao`) REFERENCES `instituicao` (`idinstituicao`) ON UPDATE CASCADE,
 CONSTRAINT `idinstrumento` FOREIGN KEY (`idinstrumento`) REFERENCES `instrumento` (`idinstrumento`) ON UPDATE CASCADE,
 CONSTRAINT `idusuario` FOREIGN KEY (`idusuario`) REFERENCES `usuario` (`idusuario`) ON UPDATE CASCADE
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```



# SQL E BI

# SQL

O SQL (**Structured Query Language**) é um padrão de linguagem de programação dedicado à banco de dados. A partir de 1970, a IBM desenvolveu a linguagem Sequel, a qual evoluiu e foi padronizada em 1986 pelo ANSI (American National Standards Institute) e pelo ISO. [@Cardoso2013].

A maior parte dos Sistemas Gerenciadores de Banco de Dados (SGDBs) adota o mesmo padrão, com algumas variações.

03/09/2019

# Mais sobre o SQL

- O SQL é uma linguagem **declarativa** (o que é muito diferente de outras linguagens de programação procedurais);
- Para o SQL, dizemos apenas o resultado que queremos, não importando (para nós), como este resultado será obtido.

## Exemplo: Linguagem Procedural:

- “Abra a tabela de alunos e comece pela primeira linha. Inicialize uma variável que tem a maior nota da turma. Passe para o próximo aluno. Se a nota do aluno for maior do que a nota anterior, atualize a variável com a maior nota da turma. Caso contrário, simplesmente passe para o próximo aluno. Ao finalizar a lista, apresente a maior nota.”

## Exemplo: Linguagem Declarativa:

- “Me mostre a maior nota na tabela de notas.”

03/09/2019

# Composição da Linguagem SQL

- **DDL (Data Definition Language)**: Permite determinar qual será a estrutura do banco de dados (exemplos: CREATE, ALTER, DROP ).
- **DML (Data Manipulation Language)**: Permite realizar a manipulação dos dados (inserção, edição, exclusão de dados). (exemplos: INSERT, UPDATE, DELETE ).
- **DQL (Data Query Language)**: Permite consultar os dados (exemplos: SELECT ).
- **DCL (Data Control Language)**: Permite controlar a autorização e acesso dos usuários aos dados(exemplos: GRANT, REVOKE ).
- **DTL (Data Transaction Language)**: Permite controlar a execução de transações com o banco de dados (exemplos: BEGIN, COMMIT, ROLLBACK ).

03/09/2019

# DDL - Data Definition Language

- Os comandos da DLL tratam diretamente da criação e edição da base de dados. Exemplos:

CREATE DATABASE - Cria uma Base de Dados.

USE DATABASE - Inicia o acesso uma Base de Dados.

DROP DATABASE - Exclui a Base de Dados.

03/09/2019

# DDL - Data Definition Language

Exemplos de Criação de uma Tabela:

```
CREATE TABLE
 (tipo de dado [NOT NULL] , ... ,
PRIMARY KEY (,< nome do atributo2> , ...)
FOREIGN KEY () REFERENCES);
```

03/09/2019

# DML - Data Manipulation Language

Pode-se utilizar os comandos da DML para selecionar dados. Para os usuários do banco de dados, é comum utilizarmos o comando **SELECT** para selecionar dados:

```
SELECT NOME, SOBRENOME FROM TABELAA
WHERE COLUNAA = "
```

03/09/2019

# O que é um “DUMP” de uma base de dados:

- Um **dump** de uma base de dados contém a definição de todas as suas tabelas, e também pode conter informações da base de dados;
- Contém código de DML (criação de tabelas), bem como de DML (inserção de dados), porém não possui instruções de seleção de dados;
- Cada um dos SGDBs possui mecanismos avançados de backup do banco de dados, porém realizar uma exportação ou importação de um **dump** é uma maneira comum de realizar um backup ou transferência de uma base de dados SQL para outro servidor.

03/09/2019

# O que é um “DUMP” de uma base de dados:

## Exemplo de um DUMP

```
-- Copiando estrutura para tabela avaliabrasil.avaliacao
CREATE TABLE IF NOT EXISTS `avaliacao` (
 `idavaliacao` int(11) NOT NULL AUTO_INCREMENT,
 `datacadastro` datetime NOT NULL DEFAULT CURRENT_TIMESTAMP,
 `idusuario` int(11) NOT NULL,
 `idinstituicao` int(11) NOT NULL,
 `status` tinyint(1) NOT NULL DEFAULT '0',
 `idinstrumento` int(11) NOT NULL,
 PRIMARY KEY (`idavaliacao`),
 UNIQUE KEY `idavaliacao_UNIQUE` (`idavaliacao`),
 KEY `idusuario_idx` (`idusuario`),
 KEY `idinstituicao_idx` (`idinstituicao`),
 KEY `idinstrumento_idx` (`idinstrumento`),
 CONSTRAINT `idinstituicao` FOREIGN KEY (`idinstituicao`) REFERENCES `instituicao`(`idinstituicao`) ON UPDATE CASCADE,
 CONSTRAINT `idinstrumento` FOREIGN KEY (`idinstrumento`) REFERENCES `instrumento`(`idinstrumento`) ON UPDATE CASCADE,
 CONSTRAINT `idusuario` FOREIGN KEY (`idusuario`) REFERENCES `usuario`(`idusuario`) ON UPDATE CASCADE
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```

03/09/2019

# Demonstração - Criando e Utilizando um Banco de Dados

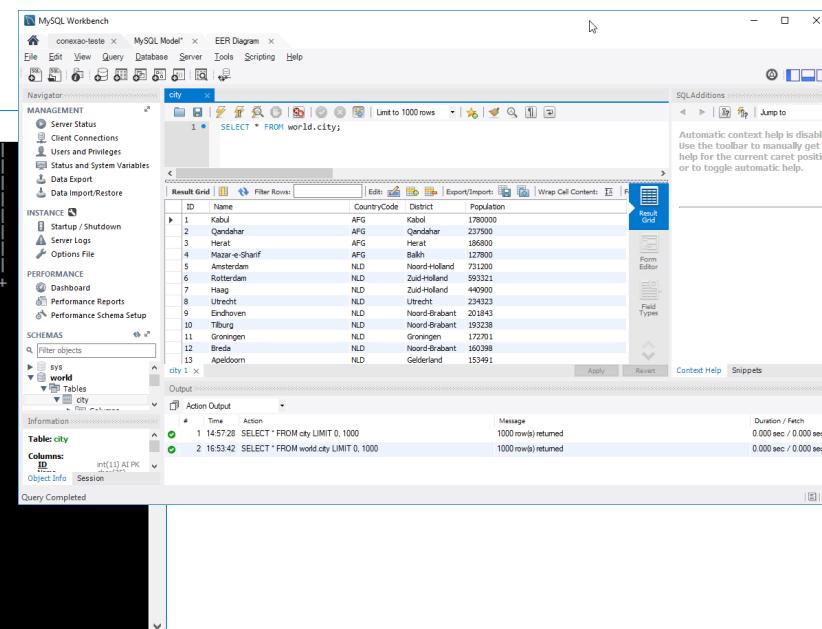
– Demonstração utilizando o MySQL.

```
MySQL 8.0 Command Line Client - Unicode
+----+-----+-----+-----+-----+
208	Salvador	BRA	Bahia	2302832
209	Belo Horizonte	BRA	Minas Gerais	2139125
210	Fortaleza	BRA	CearÁa	2097757
211	BrasAlia	BRA	Distrito Federal	1969868
212	Curitiba	BRA	ParanÁa	1584232
213	Recife	BRA	Pernambuco	1378087
214	Porto Alegre	BRA	Rio Grande do Sul	1314032
215	Manaus	BRA	Amazonas	1255049
+----+-----+-----+-----+-----+
10 rows in set (0.00 sec)

mysql> select * from city where Name = "Porto Alegre";
+----+-----+-----+-----+
| ID | Name | CountryCode | District | Population |
+----+-----+-----+-----+
| 214 | Porto Alegre | BRA | Rio Grande do Sul | 1314032 |
+----+-----+-----+-----+
1 row in set (0.00 sec)

mysql> select sum(Population) from city where CountryCode = "BRA";
+-----+
| sum(Population) |
+-----+
| 85876862 |
+-----+
1 row in set (0.00 sec)

mysql> select * from country where city < "BRA";
ERROR 1054 (42S22): Unknown column 'city' in 'where clause'
mysql>
```



03/09/2019

# Tableau - Versão para Estudantes

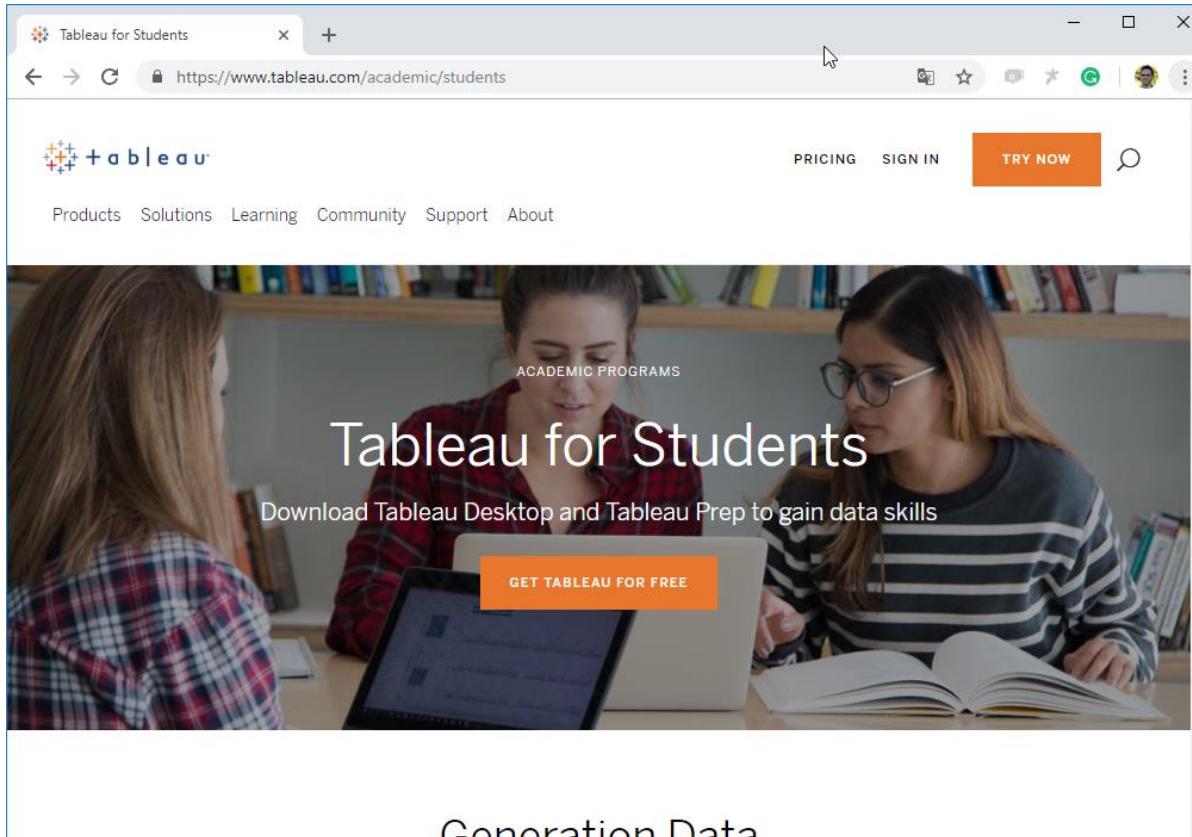


Tableau - Versão Estudantil

03/09/2019

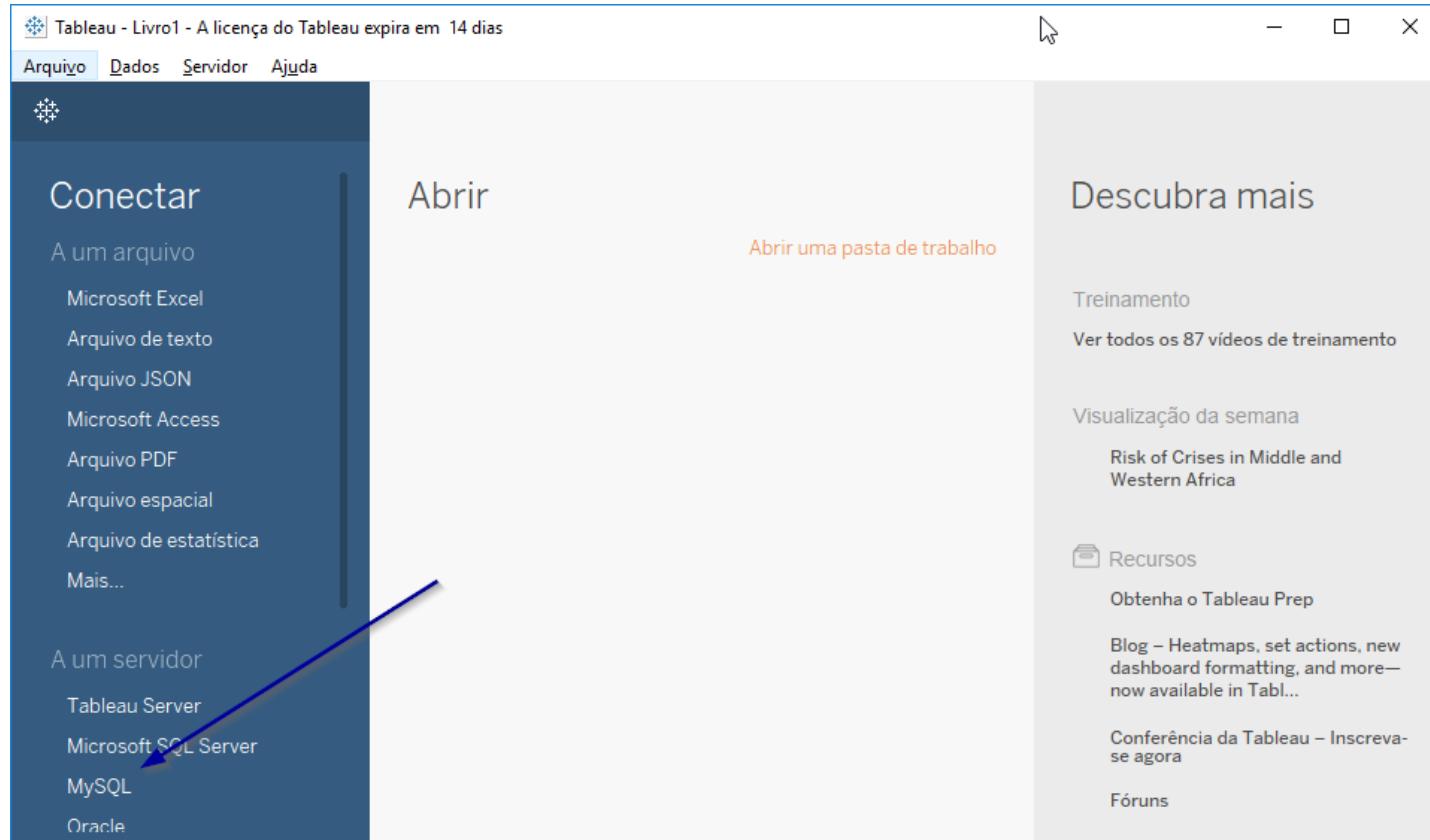
# Conectando ao Mysql

The screenshot shows the Tableau Data Source interface. At the top, it says "Tableau - Livro1 - A licença do Tableau expira em 14 dias". The menu bar includes Arquivo, Dados, Servidor, Janela, and Ajuda. On the left, there's a toolbar with icons for new connection, previous connection, next connection, and refresh. Below that is a "Conexões" section with "localhost" selected, and an "Adicionar" button. Under "Banco de dados", a dropdown menu is open, showing "Selecionar banco de dados" at the top, followed by "Digite o nome do banco de dados" (with a yellow background), and a list of MySQL databases: information\_schema, mysql, performance\_schema, sakila, sys, and world. The "world" database is currently selected. To the right, there's a tree view showing a connection to "localhost". At the bottom, there are buttons for "Fonte de dados" (Data Source), "Planilha 1" (Sheet 1), and three plus signs for adding new sources or sheets. The status bar at the bottom right shows "linhas" (rows).

Tableau

03/09/2019

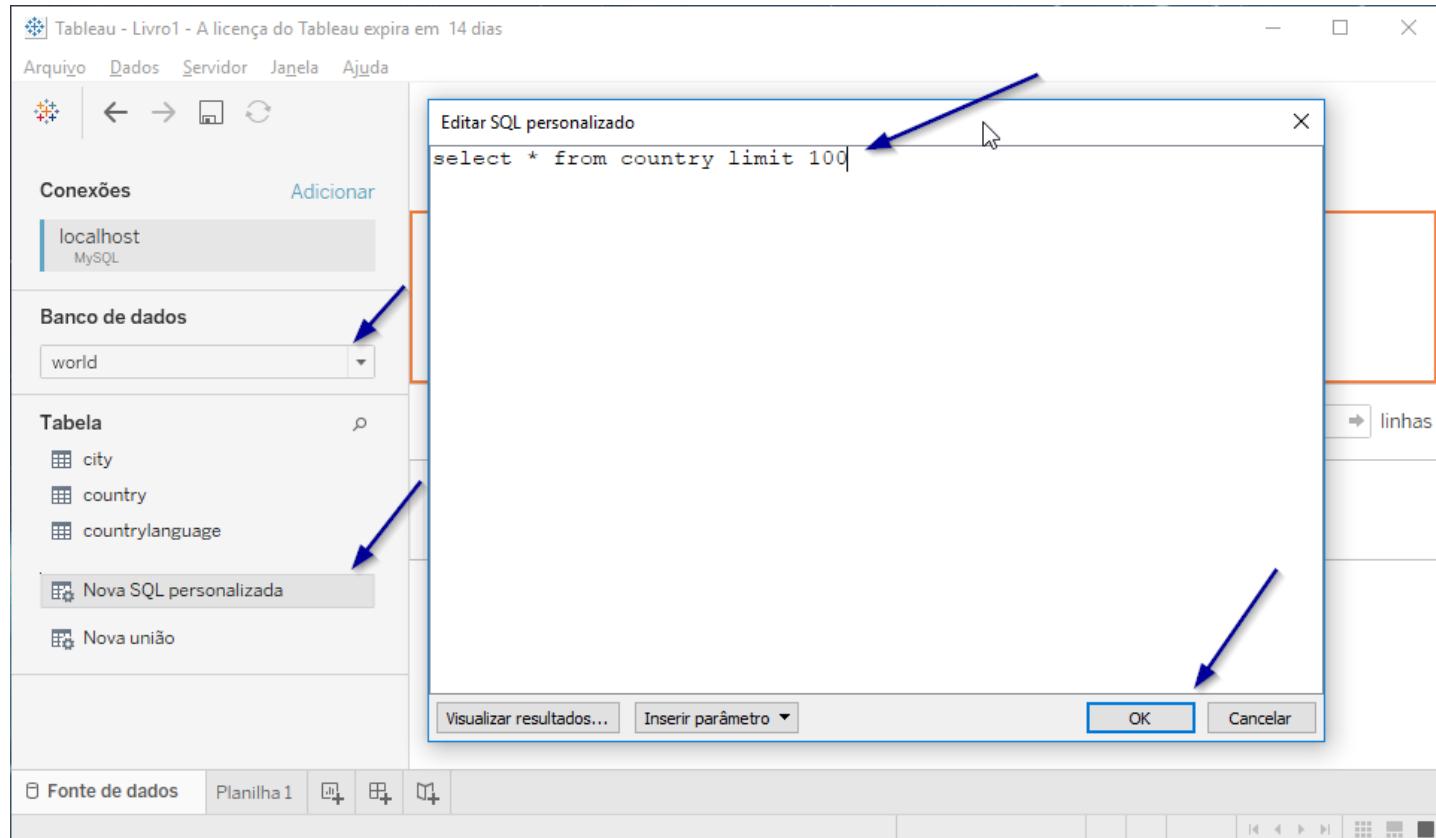
# Alguns Passos depois:



Tableau

03/09/2019

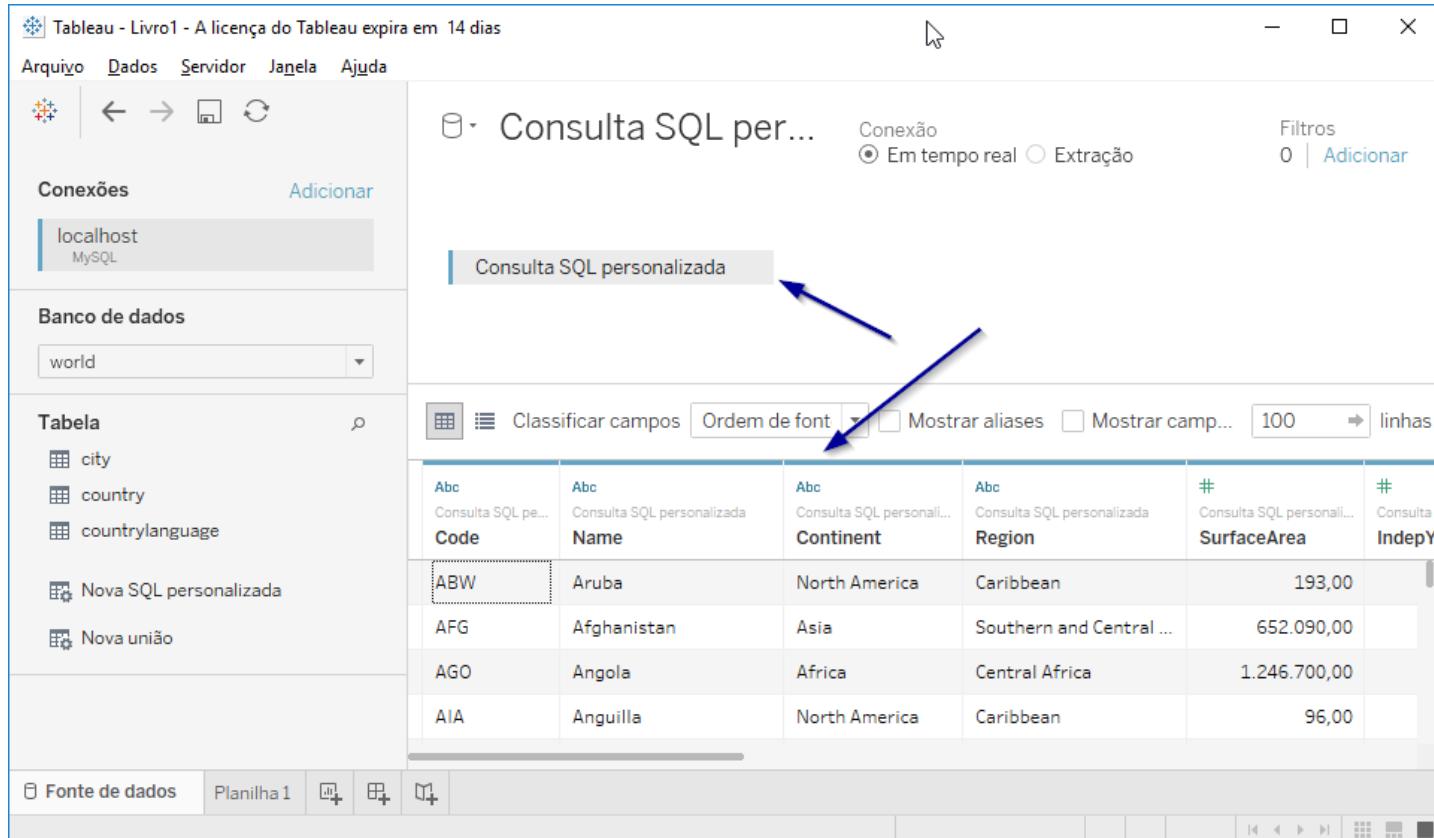
# Utilizando SQL no Tableau



Tableau

03/09/2019

# Resultados da Consulta



The screenshot shows the Tableau interface with a custom SQL query results table. The query is titled "Consulta SQL per...". The results table has columns: Code, Name, Continent, Region, SurfaceArea, and IndepYe. The data includes rows for ABW, Aruba; AFG, Afghanistan; AGO, Angola; and AIA, Anguilla.

| Code | Name        | Continent     | Region                   | SurfaceArea  | IndepYe |
|------|-------------|---------------|--------------------------|--------------|---------|
| ABW  | Aruba       | North America | Caribbean                | 193,00       |         |
| AFG  | Afghanistan | Asia          | Southern and Central ... | 652.090,00   |         |
| AGO  | Angola      | Africa        | Central Africa           | 1.246.700,00 |         |
| AIA  | Anguilla    | North America | Caribbean                | 96,00        |         |

Tableau

03/09/2019

# Uma Consulta Simples...

Quero saber qual é a população somada por região, para isso preciso apenas consultar uma única tabela (a tabela de países). Em SQL, uma consulta também é chamada de **query**.

03/09/2019

The screenshot shows the Tableau desktop application interface. At the top, the title bar reads "Tableau - Livro1 - A licença do Tableau expira em 14 dias". The menu bar includes "Arquivo", "Dados", "Servidor", "Janela", and "Ajuda". On the left, the "Conexões" pane shows a connection to "localhost MySQL" with the "world" database selected. The "Tabela" pane lists tables: "city", "country", "countrylanguage", "Nova SQL personalizada", and "Nova união". The main workspace displays a connection named "country (world)" with the status "Conexão Em tempo real". Below it is a preview of the "country" table with columns: "Code", "Region", "#", and "Code2". The preview shows five rows of data: ABW, Caribbean, 103.000, AW; AFG, Southern and Central ..., 22.720.000, AF; AGO, Central Africa, 12.878.000, AO; AIA, Caribbean, 8.000, AI; ALB, Southern Europe, 3.401.200, AL. The "Orden de fonte" button is highlighted with a blue arrow. The bottom navigation bar includes buttons for "Fonte de dados", "Planilha 1", and other data source icons.

Tableau

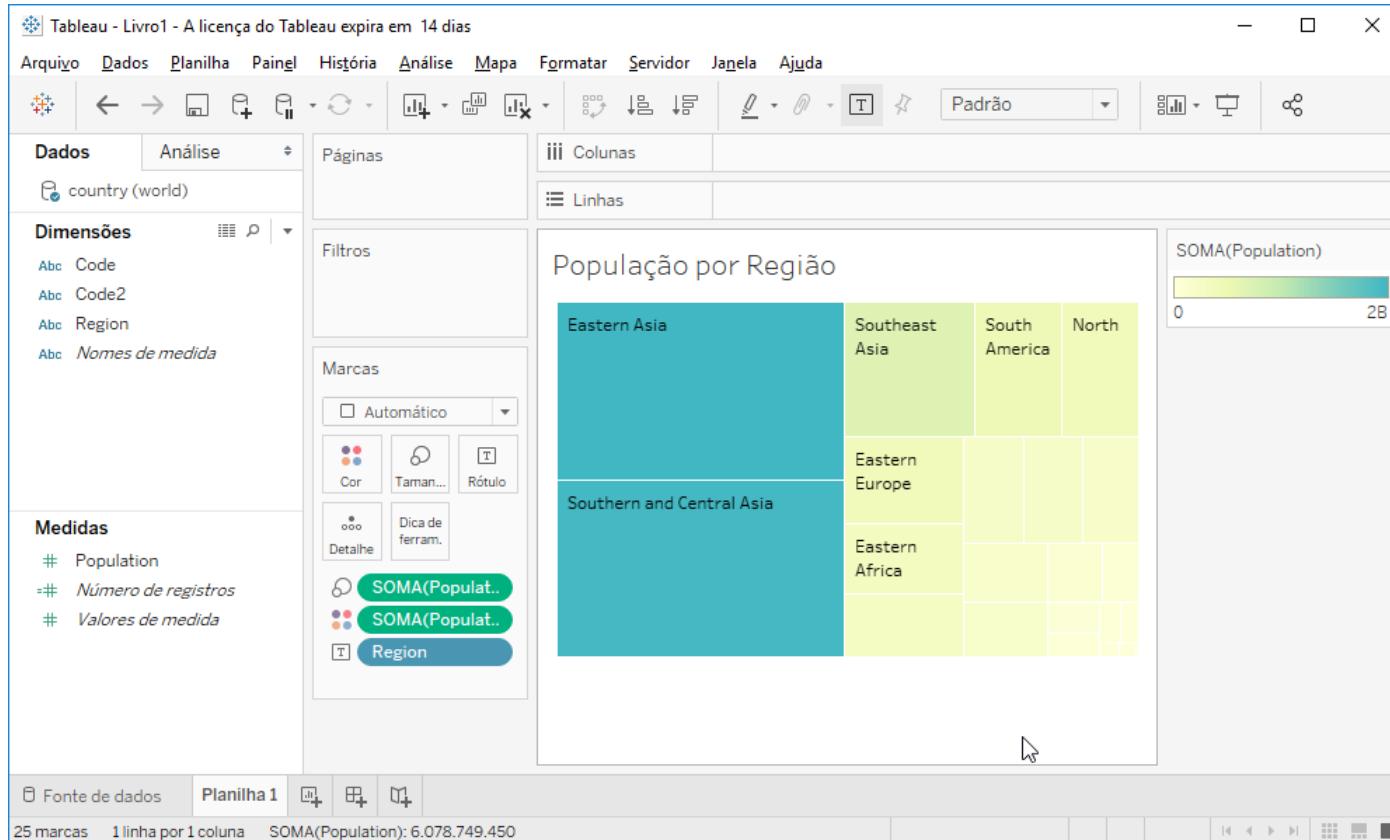
03/09/2019

E agora podemos criar uma ou mais visualizações para esta consulta:

Ao selecionar as variáveis que queremos utilizar, o Tableau sugere possíveis visualizações:  
Tableau

03/09/2019

# E pronto!



Tableau

03/09/2019

# Exemplo 2 - Utilizando Dados Relacionais

- Uma das vantagens da utilização de um banco de dados relacional é executar consultas utilizando dados distribuídos em diversas tabelas, relacionados por **chaves estrangeiras**.
- **Chaves Estrangeiras** são colunas em uma tabela que referenciam outras tabelas. Exemplo: A coluna CountryCode é uma **chave estrangeira** na tabela countryLanguage.
- Para esta tarefa, podemos utilizar consultas SQL com o comando `JOIN`. Estes comandos permitem rapidamente resgatar informações por meio da relação entre tabela. Neste caso, queremos saber qual é o número de pessoas aproximado que fala uma determinada língua em uma região.

03/09/2019

# Exemplo 2 - Utilizando Dados Relacionais

- O percentual está armazenado na tabela **countryLanguage**, e a região está armazenada na tabela **country**.
- No Excel, este comando poderia corresponder ao uso da função “PROCV”, porém é muito mais útil, pois permite utilizar consultas com “n” tabelas **ao mesmo tempo**.

03/09/2019

# Passo 1 - Realizar uma Consulta que resgata os dados consolidados

country (world)

Conexão  
Em tempo real



Tableau

03/09/2019

# Passo 2 - Adicionando Campo Calculado

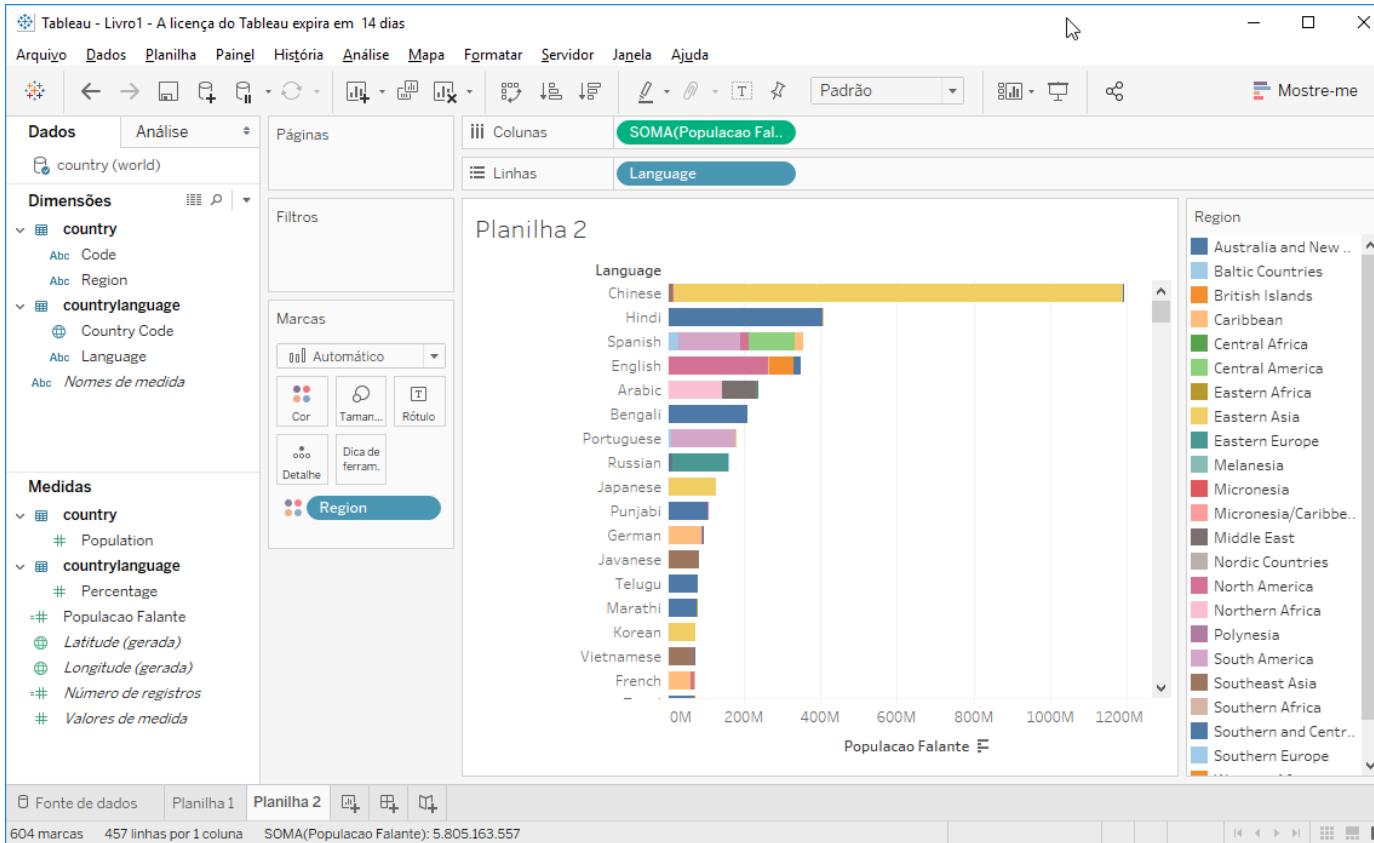
The screenshot shows the Tableau Data Source interface. On the left, the 'Conexões' panel lists 'localhost MySQL' as the connection. The 'Banco de dados' panel shows 'world' selected. The 'Tabela' panel lists 'city', 'country', 'countrylanguage', 'Nova SQL personalizada', and 'Nova união'. The main area displays the schema for 'countrylanguage':

| Abc<br>country<br>Code | Abc<br>country<br>Region | #<br>country<br>Population | @<br>countrylanguage<br>Country Code | Abc<br>countrylanguage<br>Language | #<br>countrylanguage<br>Percentage |
|------------------------|--------------------------|----------------------------|--------------------------------------|------------------------------------|------------------------------------|
| ABW                    | Caribbean                | 103.000                    | ABW                                  | Dutch                              | 5,300                              |
| ABW                    | Caribbean                | 103.000                    | ABW                                  | English                            | 9,500                              |
| ABW                    | Caribbean                | 103.000                    | ABW                                  | Papiamento                         | 76,700                             |
| ABW                    | Caribbean                | 103.000                    | ABW                                  | Spanish                            | 7,400                              |
| AFG                    | Southern and Central ... | 22.720.000                 | AFG                                  | Balochi                            | 0,900                              |

Tableau

03/09/2019

# Passo 3 - Visualizando Informações



Tableau

03/09/2019

# Desafio 1: Vendas para Universidades

- Etapa 1: Utilize o arquivo “Exemplo Dados Tableau” para construir um dashboard de visualização de dados que apresente:
  - Quais clientes compram mais em função de peso ou valor monetário?
  - Quais vendedores vendem mais?
  - Quais segmentos compram mais?
  - Crie um dashboard onde seja possível filtrar as informações por mês.

03/09/2019

# Desafio 2: Façam o mesmo para o seu trabalho da Disciplina

- Com base no seu projeto de banco de dados, crie uma única planilha no excel na qual cada aba será uma tabela do seu banco de dados;
- Insira dados de demonstração para teste;
- Crie um dashboard do tableau para visualizar as informações de seu sistema;
- Inclua estas visualizações no projeto da disciplina.
- Obs: Baixe o Tableau, ative a versão estudantil e finalize em casa:

<http://tableau.com>

03/09/2019



# SISTEMAS DE INFORMAÇÃO INDUSTRIAL CRM, ERP, SCM

# Sistemas Empresariais

Principais Sistemas Empresariais  
[@Obrien2012]:

- **CRM:** Obtenção e Retenção do relacionamento com o cliente.
- **ERP - Enterprise Resource Planning:** Planejamento Integrado de Recursos.
- **SCM - Supply Chain Management:** Gestão da Cadeia de Suprimentos.

03/09/2019

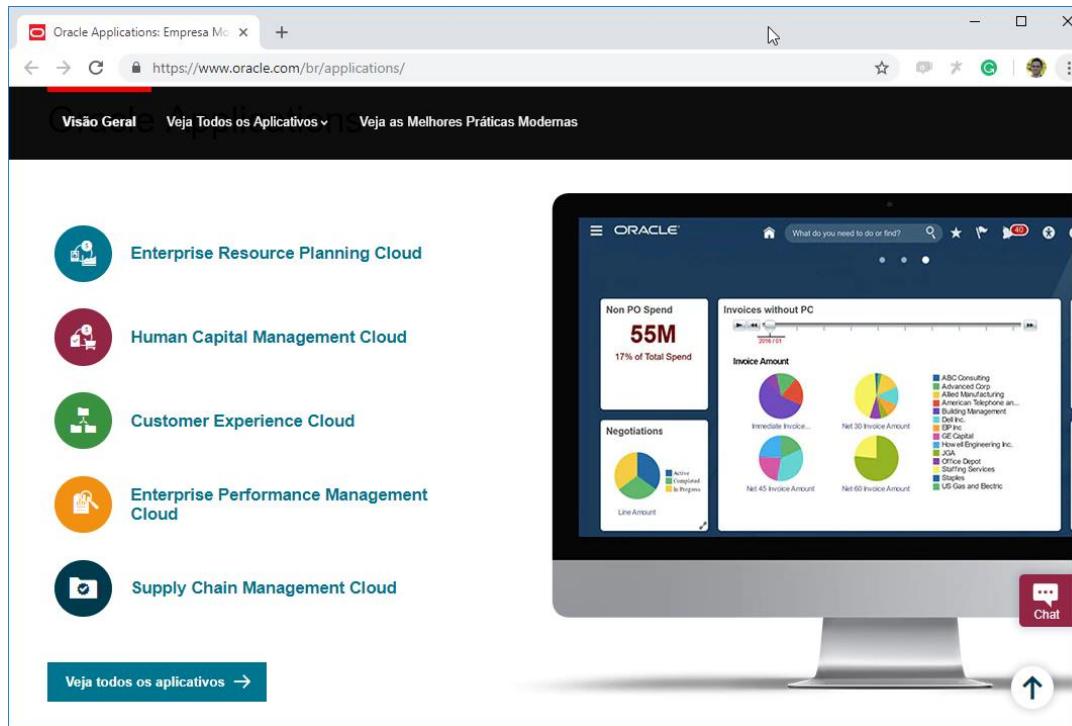
# Evolução dos Sistemas CRM, ERP, SCM

- Os Sistemas ERP tem evoluído há mais de 30 anos, e a competição neste ramo é intensa (seguramente, há mais de 100 ERPs disponíveis ao redor do mundo);
- Softwares Open Source: Além das alternativas mais conhecidas (SAP, Oracle, etc.), também existem alternativas Open Source ( [Dolibarr](#), [ERPNext](#), ), [Odoo](#), normalmente focadas em pequenas e médias empresas.

03/09/2019

# Exemplos (Oracle)

Como a Oracle organiza seus Sistemas:



03/09/2019

# Exemplos (SAP)

Como a SAP organiza seus Sistemas:

The screenshot shows the SAP Software Solutions | Business website at <https://www.sap.com/index.html>. The main navigation bar includes links for Products, Industries, Support, Training, Community, Developer, Partner, and About. A search icon and user profile icons are also present. The page content is organized into several sections:

- ERP and Digital Core**: SAP S/4HANA, Cloud ERP, ERP for Small and Midsize Enterprises, Finance.
- CRM and Customer Experience**: Customer Data, Marketing, Commerce, Sales, Service.
- Network and Spend Management**: Supplier Management, Strategic Sourcing, Procurement, Services Procurement and External Workforce, Selling and Fulfillment, Travel and Expense.
- Digital Supply Chain**: Supply Chain, Manufacturing, R&D / Engineering, Asset Management.
- HR and People Engagement**: Core HR and Payroll, Time and Attendance Management, Recruiting and Onboarding, Learning and Development, Performance and Compensation, Workforce Planning and Analytics.
- Digital Platform**: SAP Cloud Platform, SAP HANA and Databases, Data Management, Enterprise Information Management.
- Analytics**: SAP Analytics Cloud, Business Intelligence, Enterprise Planning, Predictive Analytics.
- Intelligent Technologies**: SAP Leonardo, Internet of Things, Machine Learning, Blockchain.
- Cloud**: Intelligent Enterprise, Small and Midsize Enterprises, more... (with a blue button labeled "Get Us").

A footer link "All products" is visible at the bottom left, and the URL "https://www.sap.com/products.html" is at the bottom center. The address bar shows the full URL.

03/09/2019

# Relações entre os Sistemas empresariais:

Uma visão Geral das principais aplicações de Sistemas Empresariais [@Obrien2012, pg. 261]:



03/09/2019

# CRM - Gestão do Relacionamento com o Cliente

- **O que é:** O CRM é uma categoria de sistemas que integra e automatiza processos de **atendimento ao cliente, marketing, etc..;**
- **Principais Usuários:** Os principais usuários do CRM são os profissionais de marketing e vendas.

03/09/2019

# CRM - Gestão do Relacionamento com o Cliente

- **Fontes de Informação:** As informações dos CRMs podem ser obtidas de diversas **fontes de contato** com os clientes (site da empresa, apps, lojas, terminais de autoatendimento, etc.).
- **Principais Processos:** O CRM usualmente irá lidar com os processos que possuem relação direta com o cliente (orçamentos, vendas, acompanhamento do leads, acompanhamento dos pedidos, campanhas de marketing).

03/09/2019

# CRM - Alguns Exemplos

## Odoo:

The screenshot shows the Odoo CRM Pipeline interface. The pipeline is divided into five stages: New, Qualified, Proposition, Negotiation, and Won. Each stage has a progress bar at the top indicating the number of leads or opportunities.

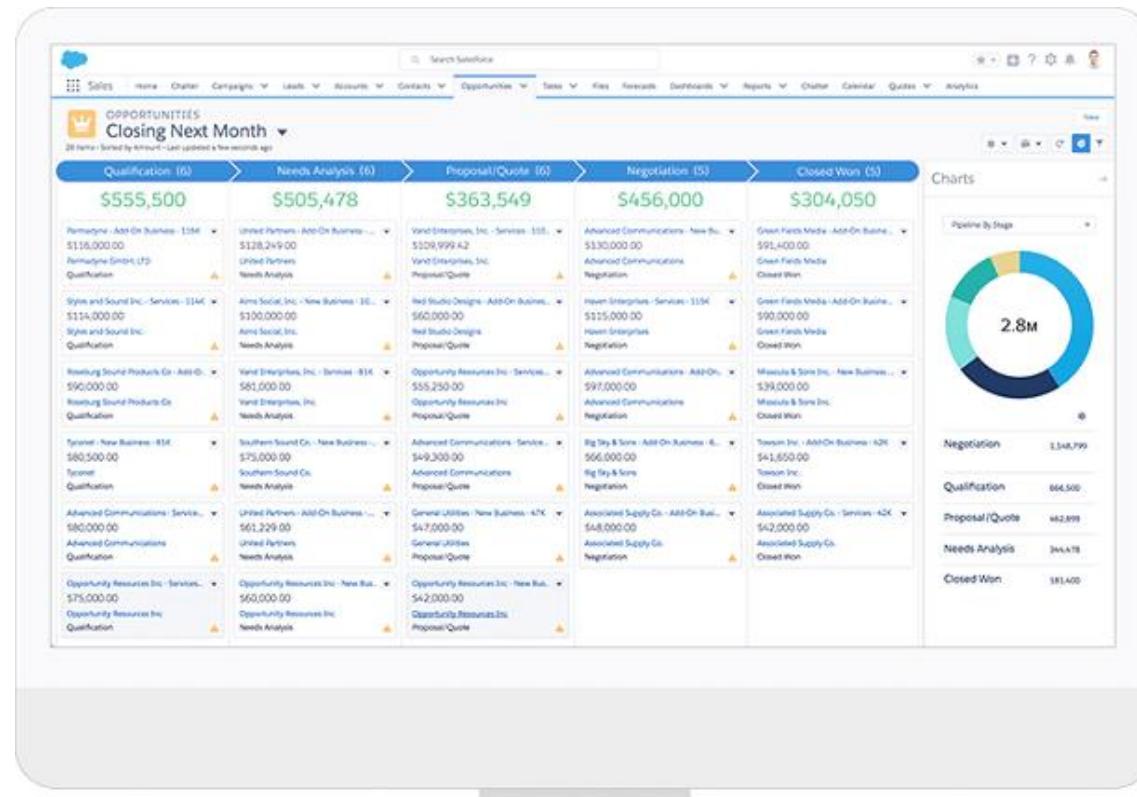
- New:** \$69,600 total value. Includes leads like "Interest in your Graphic Design Project" and "Script to import external data".
- Qualified:** \$61,000 total value. Includes leads like "Trainee's training plan in your Organization" and "Pricing Information of Services".
- Proposition:** \$24,500 total value. Includes leads like "Would appreciate more information about your products" and "Need to customize the solution".
- Negotiation:** \$77,000 total value. Includes leads like "Interest in your customizable PCs" and "Want to subscribe to your online solution".
- Won:** \$25,600 total value. Includes leads like "Interest in your products" and "Need new design for my website".

Each lead card displays the lead name, product type, value, and a rating from 1 to 5 stars. The interface includes a search bar, filters, and a configuration menu at the top.

03/09/2019

# CRM - Alguns Exemplos

## Salesforce:



03/09/2019

# ERP - Enterprise Resource Planning

- **O que é:** O ERP é um conjunto integrado de módulos, funcionando como uma “espinha dorsal” da gestão de informações industriais. Os ERPs orquestram diversas funções da empresa, incluindo o processamento de pedidos, controle e gerenciamento do estoque, planejamento da produção, distribuição, finanças e contabilidade;
- **ERP e SGDBs:** As informações dos ERPs são armazenadas em um banco de dados central, no qual todas as informações são mantidas atualizadas.

03/09/2019

# ERPs - Características (desejáveis)

- Idealmente, incorporam as melhores práticas de negócios;
- São compostos por diferentes módulos;
- Utilizam Bancos de Dados Corporativos (usualmente, bases SQL);
- Usualmente, são pacotes comerciais, genéricos;
- São Integrados.

03/09/2019

# MRP, MRP II e ERP

- **MRP (~1970)**: Permitiu a execução do planejamento de materiais, permitindo a determinação de lotes de pedidos, estrutura dos produtos (Bill of Materials - BOM), tempos de atendimento, controle de compra, determinação de lotes de segurança.
- **Objetivo do MRP**: Com base em uma definição dos produtos finais a produzir, o que, quanto e quando produzir ou comprar os semi-acabados, componentes e matéria-prima.
- **MRP II (~1985)**: Questão que o MRP II busca responder: Há **capacidade** para realizar o plano de produção sugerido pelo MRP? Os recursos humanos e equipamentos são suficientes para atender a este plano?
- **ERP**: Os ERPS, idealmente, englobam as funções do MRP I e II, e incorporam outras funções relacionadas aos diversos outros setores funcionais (finanças, contabilidade, etc.).

03/09/2019

# Benefícios (geralmente) Associados ao ERP:

[@Obrien2012] apresenta uma série de benefícios associados à implementação de ERPs:

- **Qualidade e eficácia:** A estrutura de informações propiciada pelos ERPs geralmente tende a suportar de modo positivo os processos das empresas;
- **Redução de Custos:** Em comparação ao uso de sistemas de informação não integrados, os ERPs tendem à reduzir os custos com suporte e TI.
- **Informações para Suporte à Decisão:** O ERP pode disponibilizar informações rapidamente, incluindo informações oriundas de diversos setores.
- **Agilidade empresarial:** Em um sistema sem integração, é comum que as informações sejam armazenadas e tratadas em “silos”.

03/09/2019

# Principais Malefícios Associados aos ERPs:

Alguns aspectos são lógicamente fraquezas da adoção de sistemas integrados de gestão genéricos:

- “**Lock-in**”: Quanto maior for o número de funções de uma empresa vinculados a um único sistema, maior será a dependência da empresa em relação a este sistema.
- **O Dilema das “melhores práticas”**: Se o ERP possui “as melhores práticas de classe mundial”, e seu concorrente pode comprar mesmo ERP, logo sua vantagem competitiva tende a se dissipar gradativamente.
- **Riscos e Complexidade na Implantação**: A implantação de sistemas ERP tende a custar um valor significativo, e o custo do fracasso pode significar a paralização das operações da empresa.

03/09/2019

# ERP - Tradeoffs a Avaliar durante a Seleção, Implantação e Manutenção

- **Implantação Incremental vs Implantação Abrangente:** A implantação incremental pode diminuir os riscos relacionados à mudança no sistema, enquanto a implantação abrangente tem o potencial de reduzir o período total de implantação.
- **Desenvolvimento Interno vs Externo:** Fornecedores externos tendem a desenvolver sistemas de modo mais eficiente, porém podem implicar no “dilema das melhores práticas”.
- **Abordagem Waterfall vs Ágil:** Como indicado nas aulas anteriores, abordagens focalizadas em documentação abrangente e focalização em requisitos bem-definidos podem burocratizar o processo de desenvolvimento, enquanto abordagens ágeis podem ser prejudiciais em ambientes regulados / de alta criticidade.

03/09/2019

# ERP - Tradeoffs a Avaliar durante a Seleção, Implantação e Manutenção

- **Customização:** Customização não é uma atividade trivial em software, pois gera necessidade de manutenção.
- **Fonte de Direção Sugerida para avaliar estes tradeoffs:** Estratégia Organizacional e Estratégia de Operações.

03/09/2019

# Principais ERPs (a nível Global):

Principais ERPs  
avaliados pela [Gartner](#),  
em 2017:

Magic Quadrant

Figure 1. Magic Quadrant for Cloud Core Financial Management Suites for Midsize, Large and Global Enterprises



Source: Gartner (June 2017)

03/09/2019

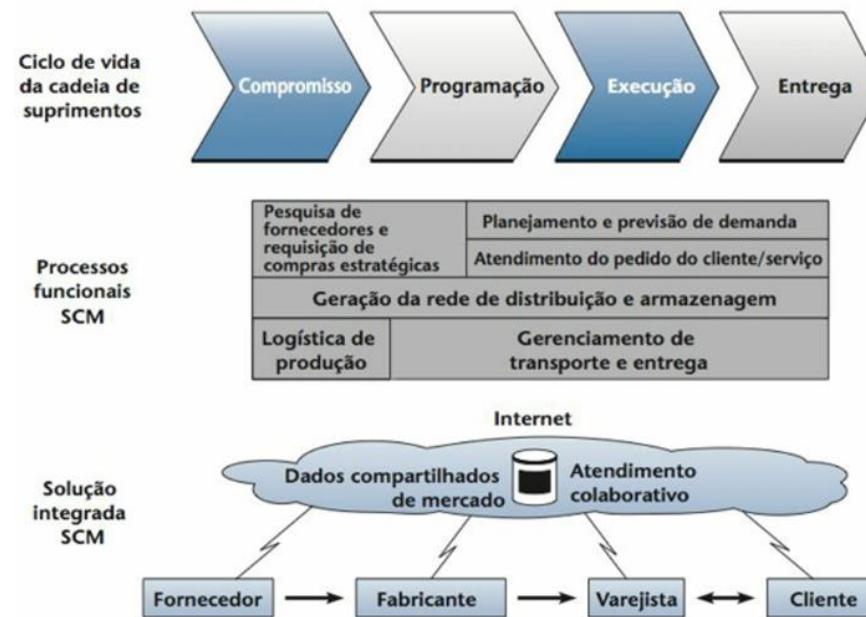
# SCM - Supply Chain Management (Systems)

- **O que é:** São sistemas que controlam os principais processos relacionando a empresa e sua cadeia de suprimentos, à montante e à jusante (podendo, portanto, englobar os fornecedores, distribuidores e revendedores da empresa).
- Idealmente, um sistema para SCM deve permitir a **coordenação** entre as empresas envolvidas na cadeia de suprimentos.

03/09/2019

# SCM - Supply Chain Management (Systems)

Processos relacionados aos sistemas SCM  
[@Obrien2012, pg. 275]:



03/09/2019

# Fechamento e Próximos Passos

- Realização do trabalho;
- Aula com Alexandre Balestrin (Elipse Software) no dia 13/11;
- Apresentação dos Trabalhos e Revisão para a Prova (final de SCM) (20/11).

03/09/2019

# Referências

- Alessio Ishizaka, & Philippe Nemery. (2013). *Multi-Criteria Decision Analysis - Methods and Software*. <https://doi.org/10.1002/9781118644898>
- Alves, J. R. X., & Alves, J. M. (2015). Definição de localidade para instalação industrial com o apoio do método de análise hierárquica (AHP). *Production*, 25(1), 13–26. <https://doi.org/10.1590/S0103-65132014005000023>
- Amadeu, C. V. (2015). *Banco de Dados*. Pearson.
- Auschitzky, E., Hammer, M., & Rajagopaul, A. (2014). How big data can improve manufacturing. Retrieved August 29, 2016, from <http://www.mckinsey.com/business-functions/operations/our-insights/how-big-data-can-improve-manufacturing>
- Bandeira, D. (2015). *Banco de dados : desenvolvimento de aplicações*. Editora UNISINOS.
- Briozo, R. A., & Musetti, M. A. (2015). Método multicritério de tomada de decisão: aplicação ao caso da localização espacial de uma Unidade de Pronto Atendimento – UPA 24 h. *Gestão & Produção*, 22(4), 805–819. <https://doi.org/10.1590/0104-530X975-13>
- Cardoso, V. (2009). *Linguagem SQL*. Saraiva.
- Cardoso, V. (2013). *Linguagem SQL, fundamentos e práticas*. Saraiva.
- Cohn, M. (2011). *Desenvolvimento de Software com SCRUM - Aplicando Métodos Ágeis com Sucesso*. Bookman.
- Corrêa, H. L., GIANESI, I. N., & CAON, M. (2007). *Planejamento, programação e controle da produção : MRP II/ERP - conceitos, uso e implantação:base para SAP, Oracle Applications e outros softwares integrados de gestão*. Atlas.
- Davenport, T. H. (2006). Competing on analytics. *Harvard Business Review*, 22, 5–20.
- Ellwange, C., Santos, C. P., & Levandowski, J. (2013). Aplicação do Método AHP para Avaliação da Usabilidade de Sistemas Application Of The AHP Method For Assessing The Usability Of System, 4–19.

03/09/2019

# Referências

- Glur, C. (n.d.). AHP - R Shiny App. Retrieved August 16, 2018, from <https://ipub.com/dev-corner/apps/ahp/>
- GORDON, Steven R., GORDON, J. R. (2006). Capítulo 1 - Gestão da Informação numa Economia Global. In *Sistemas de Informação - Uma Abordagem Gerencial*.
- Hillier, F. S., & Hillier, M. S. (2013). Capítulo 15 - Análise de Decisão. In *Introdução à Pesquisa Operacional* (p. 300). Bookman.
- Keefer, D. L., Kirkwood, C. W., & Corner, J. L. (2004). Perspective on Decision Analysis Applications, 1990–2001. *Decision Analysis*, 1(1), 4–22. <https://doi.org/10.1287/deca.1030.0004>
- Keeney, R. L. (1982). Feature Article—Decision Analysis: An Overview. *Operations Research*, 30(5), 803–838. <https://doi.org/10.1287/opre.30.5.803>
- Marins, C. S., Souza, D. de O., & Barros, M. da S. (2009). O Uso Do Método De Análise Hierárquica ( AHP ) Na Tomada De Decisões Gerenciais - Um Estudo. *Xli Sbpo*, 11.
- MIT Sloan Management Review. (2016). *Lessons from Becoming a Data-Driven Organization*.
- O'BRIEN, James A., MARAKAS, G. M. (2012). Desenvolvendo Soluções de Negócio/TI. In *Administração de Sistemas de Informação*. Bookman.
- O'BRIEN, James A., MARAKAS, G. M. (2012). Módulo III - Aplicações de Negócios. In *Administração de Sistemas de Informação*. Bookman.
- O'BRIEN, James A., MARAKAS, G. M. (2012). Módulo II - Tecnologias da Informação. In *Administração de Sistemas de Informação*. Bookman.
- Process, H. (2003). O Método AHP – Analytic Hierarchy Process.
- Sommerville, I. (2011). *Engenharia de Software* (9th ed.). Pearson.
- Souza, V. (2011). *Engenharia de requisitos de software*.
- Wazlawick, R. S. (2013). *Engenharia de Software: Conceitos e Práticas*. Elsevier.
- Winig, L. (2016). GE'S Big Bet on Data and Analytics. *MIT Sloan Management Review*. Retrieved from <http://sloanreview.mit.edu/case-study/ge-big-bet-on-data-and-analytics/>

03/09/2019



# ANÁLISE DE DECISÃO MULTI-CRITÉRIO - AHP

# Como tomar decisões quando...

- Não se sabe qual será o cenário a se concretizar?
- Há informações sobre a probabilidade dos cenários?
- O valor esperado do dinheiro não é igual para dois decisores?

03/09/2019

# Como tomar decisões quando...

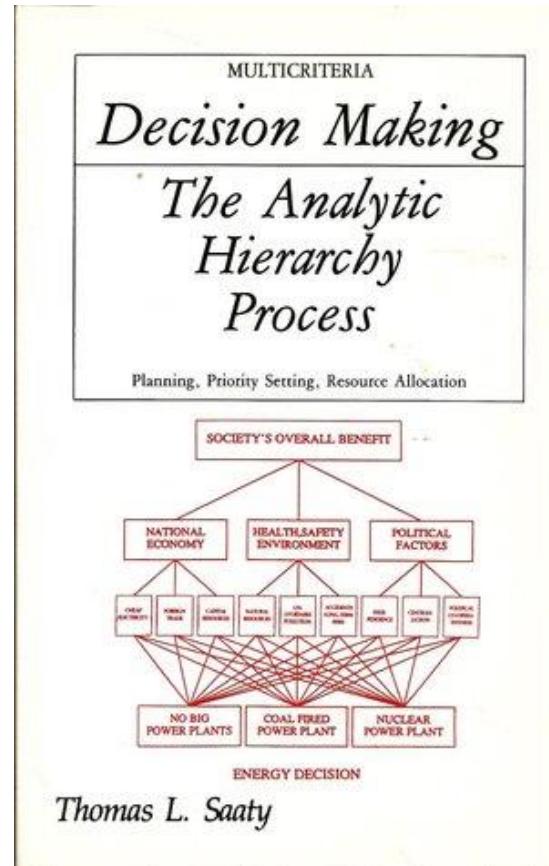
- Existem diversos critérios a considerar?
- Existe mais de um decisor envolvido na decisão, o qual tem preferências distintas?
- Existem critérios intangíveis a serem considerados na decisão?
- As preferências dos decisores são subjetivas?



03/09/2019

# AHP – Analytic Hierarchy Process

- Método Criado Thomas Saaty (1926 – 2017);
- Matemático, professor na Universidade de Pittsburgh.



Thomas Saaty

03/09/2019

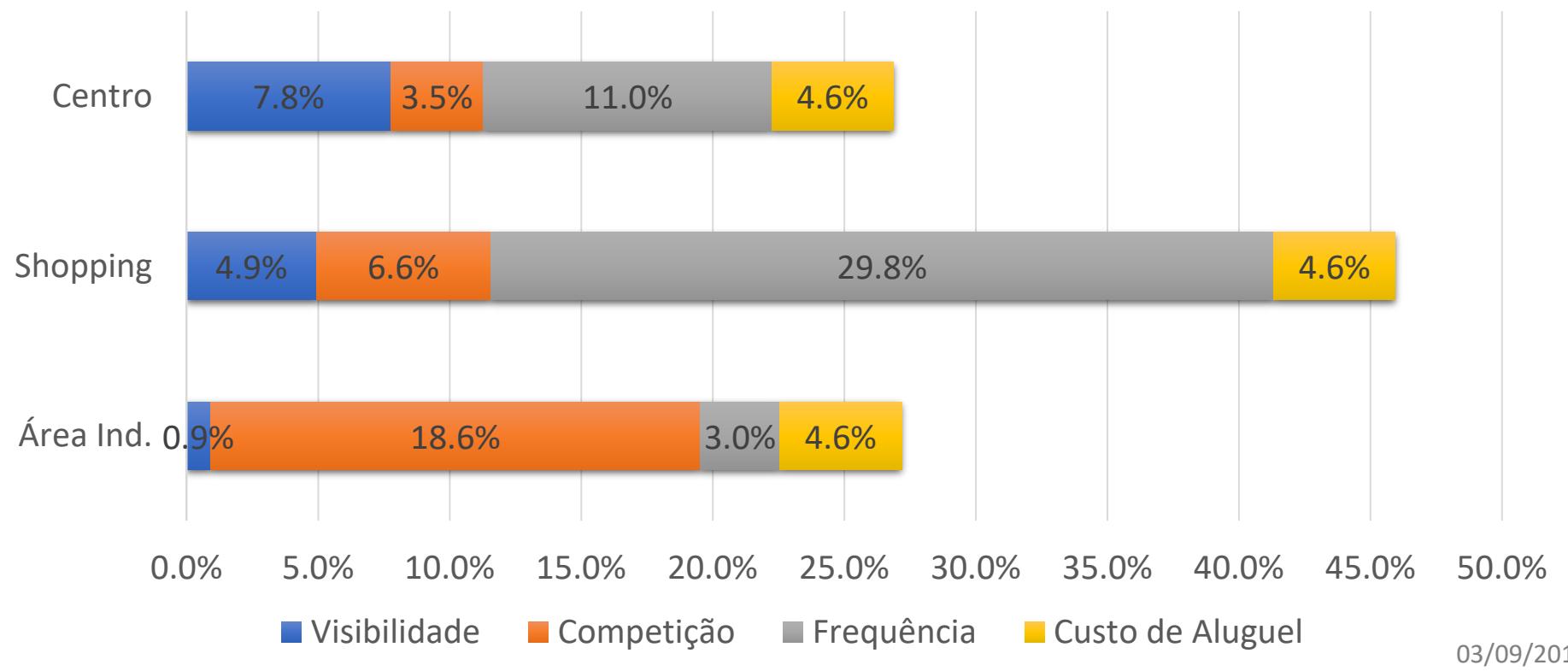
# Exemplo 1 – Definição do Local de uma Loja

- Uma empresa deseja abrir uma nova loja de esportes em uma das três seguintes localizações:
  - Um Shopping: O shopping possui alta concentração de pessoas. O shopping possui valores de aluguel mediano.
  - O centro de uma cidade: O centro possui um alto volume de movimentação, porém os aluguéis não são baratos. Há alta competição entre as lojas no centro..
  - Uma nova área industrial: Aluguéis são baratos, mas o movimento também pode ser menor.

03/09/2019

# Objetivo – Comparar as Opções considerando múltiplos critérios

Comparação de Opções - Comparação Geral



03/09/2019

# 4 Passos para a Análise AHP

1. Estruturar a Decisão de modo hierárquico, considerando o objetivo global, critérios e alternativas.
2. Calcular as prioridades utilizando comparações par-a-par.
3. Avaliar a consistência dos resultados.
4. Avaliar o quanto sensível a decisão é em relação às mudanças de parâmetros.

Os passos 3 e 4 são opcionais, embora altamente recomendados.

03/09/2019

# Exemplo 1 – Definição do Local de uma Loja

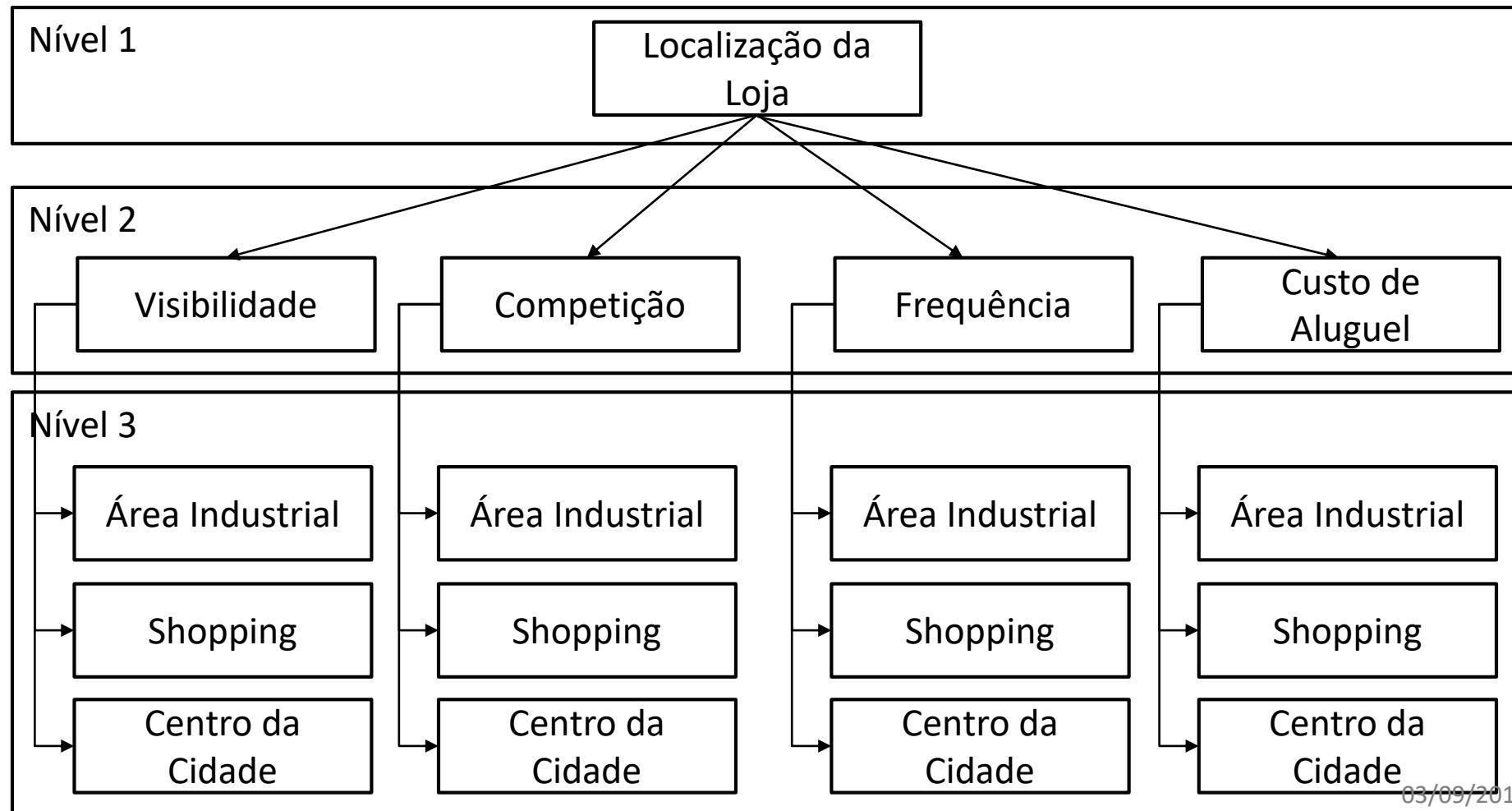
## Passo 1 – Estruturação da Decisão

- Critérios definidos para a escolha do Local:
  - Visibilidade: Probabilidade de uma pessoa notar a loja.
  - Competição: Nível de Competição na Área;
  - Frequência: Número médio de clientes em lojas similares na área;
  - Custo do Aluguel: Custo médio do aluguel por metro quadrado.

03/09/2019

# Exemplo 1 – Definição do Local de uma Loja

## Passo 1 – Estruturação da Decisão

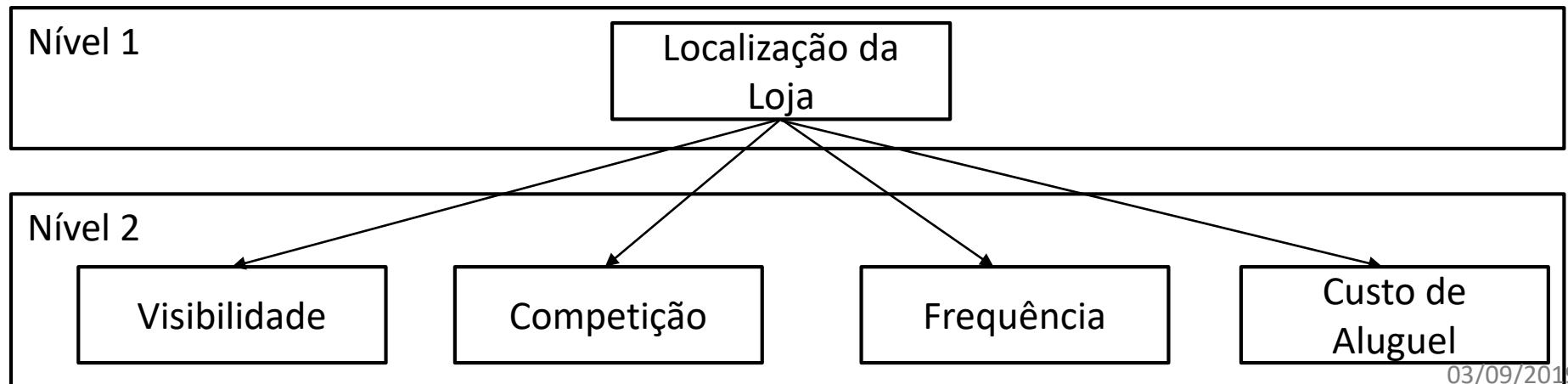


03/09/2019

# Exemplo 1 – Definição do Local de uma Loja

## Passo 1 – Estruturação da Decisão

- É necessário no mínimo três níveis de hierarquia para representar um problema no AHP.
- Cada nível inferior é priorizado em função de seu nível superior.  
Exemplo:
  - “O quê é mais importante para definir uma boa localização da loja? Visibilidade, Competição, Frequência ou Custo do Aluguel?”

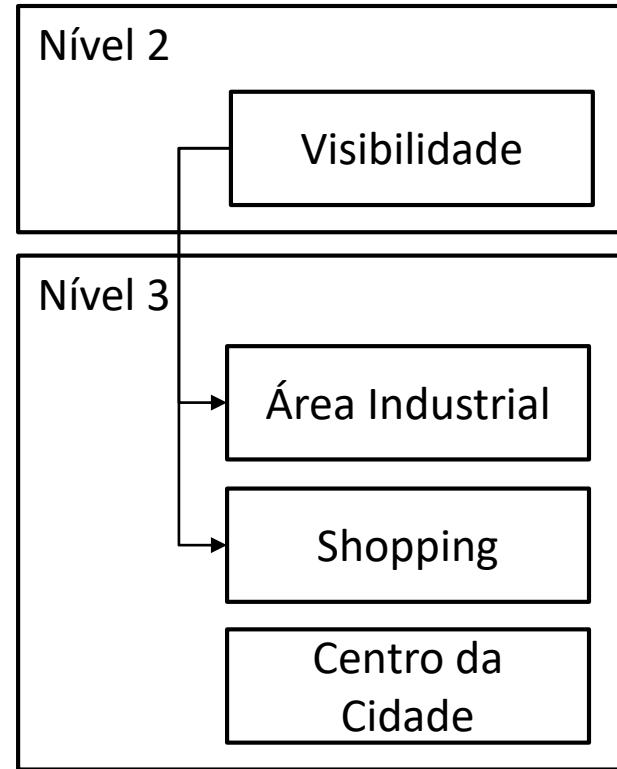


03/09/2019

# Exemplo 1 – Definição do Local de uma Loja

## Passo 1 – Estruturação da Decisão

- É necessário no mínimo três níveis de hierarquia para representar um problema no AHP.
- Cada nível inferior é priorizado em função de seu nível superior.  
Exemplo:
  - “Qual das alternativas é preferível no critério Visibilidade, e até que ponto?”



03/09/2019

# Escala Saaty de Comparações

| Grau de Importância | Definição                                  |
|---------------------|--------------------------------------------|
| 1                   | <b>Igualmente Preferível</b>               |
| 2                   | Igualmente a Moderadamente Preferível      |
| 3                   | <b>Moderadamente Preferível</b>            |
| 4                   | Moderadamente a Fortemente Preferível      |
| 5                   | <b>Fortemente Preferível</b>               |
| 6                   | Fortemente a Muito Fortemente Preferível   |
| 7                   | <b>Muito Fortemente Preferível</b>         |
| 8                   | Muito Fortemente a Extremamente Preferível |
| 9                   | <b>Extremamente Preferível</b>             |

03/09/2019

# Comparações Pareadas

- A técnica AHP utiliza **comparações pareadas**, em todos os seus níveis. Para tanto, é necessário perguntar:
  - O quanto o a Competição é preferível em relação à Visibilidade?
  - Ao responder **uma pergunta**, **preenche-se duas células**.
  - R: 4 Moderadamente a Fortemente Preferível.
  - Registra-se a resposta desta pergunta em uma matriz de comparações:

Em relação à ...



|                  | Visibilidade | Competição | Frequência | Custo do Aluguel |
|------------------|--------------|------------|------------|------------------|
| Visibilidade     | 1            | 1/4        | 1/5        | 2                |
| Competição       | 4            | 1          | 1/2        | 1                |
| Frequência       | 5            | 2          | 1          | 4                |
| Custo do Aluguel | 1/2          | 1          | 1/4        | 1                |

03/09/2019

# Comparações Pareadas – Organizando a Coleta de Dados

- Uma maneira de Coletar os dados é dispor as comparações em um instrumento de coleta de dados com o seguinte formato:
- “Em relação ao critério “Custo do Aluguel”, como você compararia as seguintes opções?”

|         |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |         |
|---------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---------|
| Local 1 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Local 2 |
| Local 1 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Local 3 |
| Local 1 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Local 4 |
| Local 1 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Local 5 |
| Local 1 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Local 6 |

Exemplo contido em: BRIOZO, Rodrigo Amancio; MUSSETTI, Marcel Andreotti. Método multicritério de tomada de decisão: aplicação ao caso da localização espacial de uma Unidade de Pronto Atendimento – UPA 24 h. **Gestão & Produção** v. 22, n. 4, p. 805–819 , 2015. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0104-530X2015000400805&lng=pt&tlang=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-530X2015000400805&lng=pt&tlang=pt)>.

03/09/2019

# Fórmulas – Método Aproximado

Prioridade Média Local:

$$PML_i = \frac{\sum_{k=1}^n \frac{p_{ij}}{\sum_{i=1}^n p_{ij}}}{n}$$

Índice de Consistência (CI):

$$CI = \frac{\lambda_{max} - n}{n - 1}$$

$$\lambda_{max} = \frac{\sum_{i=1}^n \lambda_i}{n}$$

$$\lambda_i = \frac{\sum_{j=1}^n PML_j * p_{ij}}{PML_i}$$

Razão de Consistência (RC):

$$RC = \frac{CI}{RI}; RI = f(n)$$

| n  | 3    | 4   | 5    | 6    | 7    | 8    | 9    | 10   |
|----|------|-----|------|------|------|------|------|------|
| RI | 0,58 | 0,9 | 1,12 | 1,24 | 1,32 | 1,41 | 1,45 | 1,49 |

03/09/2019

# Passo 2 - Calculando o AHP no Excel – Exemplo 1

Preencher Matrizes de Comparação e Calcular Prioridade Média Local (PML):

|                  | Visibilidade | Competição | Frequência | Custo do Aluguel |
|------------------|--------------|------------|------------|------------------|
| Visibilidade     | 1.00         | 0.25       | 0.20       | 2.00             |
| Competição       | 4.00         | 1.00       | 0.50       | 1.00             |
| Frequência       | 4.00         | 2.00       | 1.00       | 4.00             |
| Custo do Aluguel | 0.50         | 1.00       | 0.25       | 1.00             |
| Soma             | 9.50         | 4.25       | 1.05       | 8.00             |

|                  | Visibilidade | Competição | Frequência | Custo do Aluguel | Prioridade Média Local |
|------------------|--------------|------------|------------|------------------|------------------------|
| Visibilidade     | 0.11         | 0.06       | 0.10       | 0.25             | 12.9%                  |
| Competição       | 0.42         | 0.24       | 0.26       | 0.13             | 25.9%                  |
| Frequência       | 0.42         | 0.47       | 0.51       | 0.50             | 47.6%                  |
| Custo do Aluguel | 0.05         | 0.24       | 0.13       | 0.13             | 13.5%                  |
| Soma             | 1.00         | 1.00       | 1.00       | 1.00             | 1.00                   |

$0,11 = 1 / 9,50$   
Em comparação aos outros Critérios, a Frequência sempre recebeu uma alta prioridade. Por isso, sua Prioridade é maior do que a dos outros critérios

03/09/2019

# Passo 2 - Calculando o AHP no Excel – Exemplo 1

## Fórmulas Utilizadas:

|    | A                | B              | C              | D              | E                | F                      |
|----|------------------|----------------|----------------|----------------|------------------|------------------------|
| 1  |                  |                |                |                |                  |                        |
| 2  |                  | Visibilidade   | Competição     | Frequência     | Custo do Aluguel |                        |
| 3  | Visibilidade     | 1              | =1/4           | =1/5           | 2                |                        |
| 4  | Competição       | 4              | 1              | =1/2           | 1                |                        |
| 5  | Frequência       | 4              | 2              | 1              | 4                |                        |
| 6  | Custo do Aluguel | =1/2           | 1              | =1/4           | 1                |                        |
| 7  | Soma             | =SOMA(B3:B6)   | =SOMA(C3:C6)   | =SOMA(D3:D6)   | =SOMA(E3:E6)     |                        |
| 8  |                  |                |                |                |                  |                        |
| 9  |                  | Visibilidade   | Competição     | Frequência     | Custo do Aluguel | Prioridade Média Local |
| 10 | Visibilidade     | =B3/B\$7       | =C3/C\$7       | =D3/D\$7       | =E3/E\$7         | =MÉDIA(B10:E10)        |
| 11 | Competição       | =B4/B\$7       | =C4/C\$7       | =D4/D\$7       | =E4/E\$7         | =MÉDIA(B11:E11)        |
| 12 | Frequência       | =B5/B\$7       | =C5/C\$7       | =D5/D\$7       | =E5/E\$7         | =MÉDIA(B12:E12)        |
| 13 | Custo do Aluguel | =B6/B\$7       | =C6/C\$7       | =D6/D\$7       | =E6/E\$7         | =MÉDIA(B13:E13)        |
| 14 | Soma             | =SOMA(B10:B13) | =SOMA(C10:C13) | =SOMA(D10:D13) | =SOMA(E10:E13)   | =SOMA(F10:F13)         |
| 15 |                  |                |                |                |                  |                        |

03/09/2019

# Passo 2 - Calculando o AHP no Excel – Exemplo 1

Passo 2 – Uma segunda forma (mais compacta) de realizar todos os cálculos é exibida abaixo:

| Matriz de Comparações - Critérios |              |            |            |                  | n =                    | 4      |
|-----------------------------------|--------------|------------|------------|------------------|------------------------|--------|
|                                   | Visibilidade | Competição | Frequência | Custo do Aluguel | Prioridade Média Local |        |
| <b>Visibilidade</b>               | 1.00         | 0.25       | 0.20       | 2.00             |                        | 12.9%  |
| <b>Competição</b>                 | 4.00         | 1.00       | 0.50       | 1.00             |                        | 25.9%  |
| <b>Frequência</b>                 | 4.00         | 2.00       | 1.00       | 4.00             |                        | 47.6%  |
| <b>Custo do Aluguel</b>           | 0.50         | 1.00       | 0.25       | 1.00             |                        | 13.5%  |
| <b>Soma</b>                       | 9.50         | 4.25       | 1.95       | 8.00             |                        | 100.0% |

03/09/2019

# Passo 2 - Calculando o AHP no Excel – Exemplo 1

Passo 2 – Uma segunda forma (mais compacta) de realizar todos os cálculos é exibida abaixo:

| A | B                                 | C            | D            | E            | F                |
|---|-----------------------------------|--------------|--------------|--------------|------------------|
| 1 |                                   |              |              |              |                  |
| 2 | Matriz de Comparações - Critérios |              |              |              | n = 4            |
| 3 |                                   | Visibilidade | Competição   | Frequência   | Custo do Aluguel |
| 4 | Visibilidade                      | 1            | =1/4         | =1/5         | 2                |
| 5 | Competição                        | 4            | 1            | =1/2         | 1                |
| 6 | Frequência                        | 4            | 2            | 1            | 4                |
| 7 | Custo do Aluguel                  | =1/2         | 1            | =1/4         | 1                |
| 8 | Soma                              | =SOMA(B4:B7) | =SOMA(C4:C7) | =SOMA(D4:D7) | =SOMA(E4:E7)     |
| 9 |                                   |              |              |              | =SOMA(F4:F7)     |

03/09/2019

# Passo 2 - Calculando o AHP no Excel – Exemplo 1

- Este passo é repetido para cada matriz de priorização, utilizando os mesmos procedimentos de calculo. Para este caso, serão necessárias 5 matrizes de priorização:
- Matriz para a priorização dos Critérios;
- Matriz para a Priorização das Alternativas vs ..
  - Visibilidade;
  - Competição;
  - Frequência;
  - Custo do Aluguel.

03/09/2019

# Passo 2 - Calculando o AHP no Excel – Exemplo 1

- Este passo é repetido para cada matriz de priorização. Para este caso, serão necessárias 5 matrizes de priorização:

| Matriz de Comparações - Visibilidade |           |          |        | n =                    | 3      |
|--------------------------------------|-----------|----------|--------|------------------------|--------|
|                                      | Área Ind. | Shopping | Centro | Prioridade Média Local |        |
| Área Ind.                            | 1.00      | 0.14     | 0.14   |                        | 6.7%   |
| Shopping                             | 7.00      | 1.00     | 1.00   |                        | 46.7%  |
| Centro                               | 7.00      | 1.00     | 1.00   |                        | 46.7%  |
| Soma                                 | 15.00     | 2.14     | 2.14   |                        | 100.0% |

| Matriz de Comparações - Competição |           |          |        | n =                    | 3      |
|------------------------------------|-----------|----------|--------|------------------------|--------|
|                                    | Área Ind. | Shopping | Centro | Prioridade Média Local |        |
| Área Ind.                          | 1.00      | 0.20     | 0.20   |                        | 9.1%   |
| Shopping                           | 5.00      | 1.00     | 1.00   |                        | 45.5%  |
| Centro                             | 5.00      | 1.00     | 1.00   |                        | 45.5%  |
| Soma                               | 11.00     | 2.20     | 2.20   |                        | 100.0% |

| Matriz de Comparações - Frequência |           |          |        | n =                    | 3      |
|------------------------------------|-----------|----------|--------|------------------------|--------|
|                                    | Área Ind. | Shopping | Centro | Prioridade Média Local |        |
| Área Ind.                          | 1.00      | 0.20     | 0.20   |                        | 9.1%   |
| Shopping                           | 5.00      | 1.00     | 1.00   |                        | 45.5%  |
| Centro                             | 5.00      | 1.00     | 1.00   |                        | 45.5%  |
| Soma                               | 11.00     | 2.20     | 2.20   |                        | 100.0% |

| Matriz de Comparações - Custo de Aluguel |           |          |        | n =                    | 3      |
|------------------------------------------|-----------|----------|--------|------------------------|--------|
|                                          | Área Ind. | Shopping | Centro | Prioridade Média Local |        |
| Área Ind.                                | 1.00      | 0.20     | 0.33   |                        | 11.5%  |
| Shopping                                 | 5.00      | 1.00     | 1.00   |                        | 48.0%  |
| Centro                                   | 3.00      | 1.00     | 1.00   |                        | 40.5%  |
| Soma                                     | 9.00      | 2.20     | 2.33   |                        | 100.0% |

2019

# Passo 2 - Calculando o AHP no Excel – Exemplo 1

- Após calcular todas as matrizes de priorização, é possível calcular a prioridade global dos Fatores pelo método distributivo:

Agregação dos Resultados

| Opções       | Visibilidade  | Competição    | Frequência    | Custo de Aluguel | Prioridade Final |
|--------------|---------------|---------------|---------------|------------------|------------------|
| <b>Pesos</b> | <b>13.62%</b> | <b>28.75%</b> | <b>43.75%</b> | <b>13.88%</b>    |                  |
| Área Ind.    | 6.7%          | 9.1%          | 9.1%          | 11.5%            | 9.09%            |
| Shopping     | 46.7%         | 45.5%         | 45.5%         | 48.0%            | 45.97%           |
| Centro       | 46.7%         | 45.5%         | 45.5%         | 40.5%            | 44.94%           |

$$9,09 = 6,67 * 13,62 + 9,1 * 28,75 + 9,1 * 43,74$$

Decisão: Instalar fábrica no Shopping, ou ainda no Centro, porém não na Área Industrial.

03/09/2019

# Passo 2 - Calculando o AHP no Excel – Visualizando os Resultados

Agregação dos Resultados

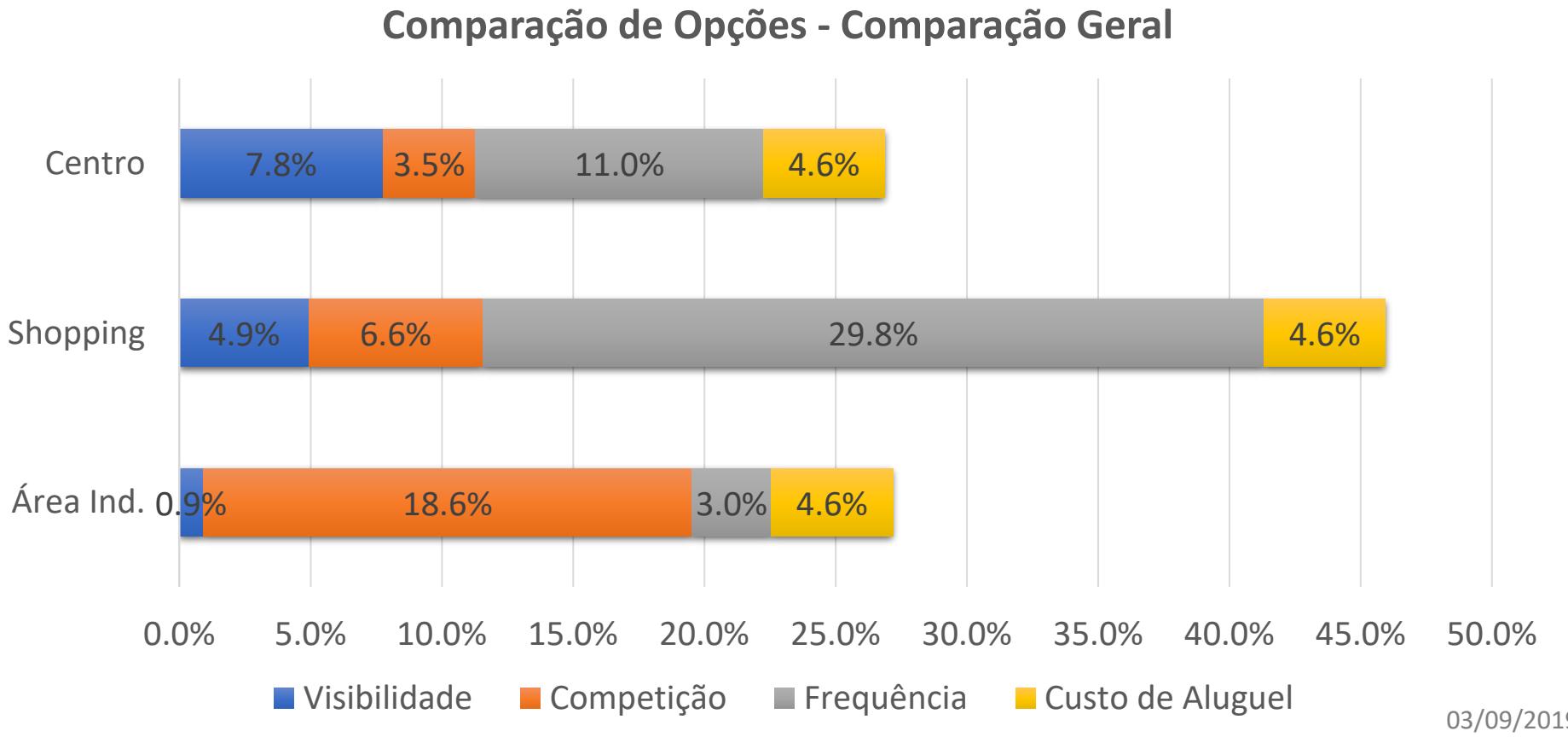
| Opções       | Visibilidade  | Competição    | Frequência    | Custo de Aluguel | Prioridade Final |
|--------------|---------------|---------------|---------------|------------------|------------------|
| <b>Pesos</b> | <b>13.62%</b> | <b>28.75%</b> | <b>43.75%</b> | <b>13.88%</b>    |                  |
| Área Ind.    | 6.6%          | 64.8%         | 6.9%          | 33.3%            | 27.18%           |
| Shopping     | 36.3%         | 23.0%         | 68.0%         | 33.3%            | 45.95%           |
| Centro       | 57.1%         | 12.2%         | 25.1%         | 33.3%            | 26.87%           |

Agregação dos Resultados - Distribuição da Utilidade

| Opções    | Visibilidade | Competição | Frequência | Custo de Aluguel | Prioridade Final |
|-----------|--------------|------------|------------|------------------|------------------|
| Área Ind. | 0.9%         | 18.6%      | 3.0%       | 4.6%             | 27.2%            |
| Shopping  | 4.9%         | 6.6%       | 29.8%      | 4.6%             | 45.9%            |
| Centro    | 7.8%         | 3.5%       | 11.0%      | 4.6%             | 26.9%            |

03/09/2019

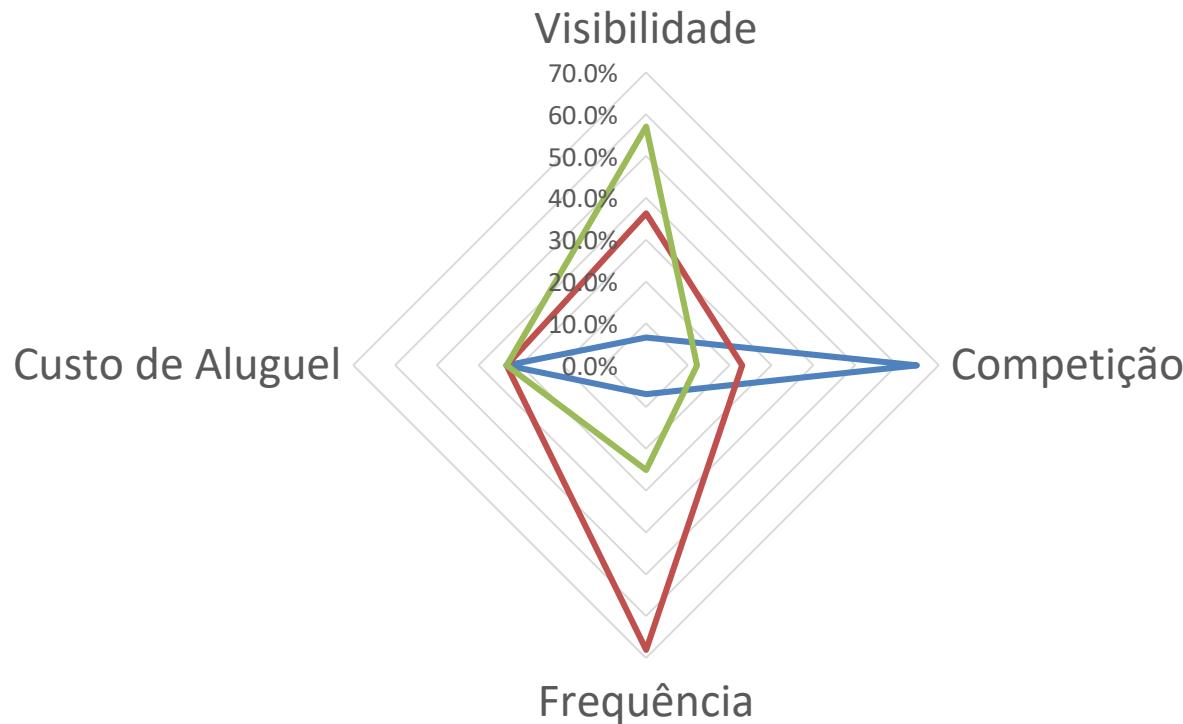
# Passo 2 - Calculando o AHP no Excel – Visualizando os Resultados



# Passo 2 - Calculando o AHP no Excel – Visualisando os Resultados

Comparação da Performance das Opções

— Área Ind. — Shopping — Centro



03/09/2019

# Verificação de Consistência

- a) O Shopping é **duas vezes** mais visível do que o centro da cidade.
- b) O centro da cidade é **três vezes** mais visível do que a área industrial.
- c) O shopping é **quatro vezes** mais visível do que a área industrial.

**O que há de errado com a terceira afirmação?**

03/09/2019

# Verificação de Consistência

- a) O Shopping é **duas vezes** mais visível do que o centro da cidade. ( $S = 2 \times C$ )
- b) O centro da cidade é **três vezes** mais visível do que a área industrial. ( $C = 3 \times A$ )
- c) O shopping é **quatro vezes** mais visível do que a área industrial.

**O que há de errado com a terceira afirmação?**

- $S = 2 \times C \rightarrow S = 2 \times 3 \times A$
- $S = 6 \times A$

03/09/2019

# Verificação de Consistência

- Considerando as inconsistências que podem haver na resposta às perguntas, o AHP permite que haja no máximo **10 % de inconsistência**, comparada à inconsistência existente no preenchimento de 500 matrizes preenchidas aleatoriamente.

03/09/2019

# Verificação de Consistência

- Calculo da Inconsistência, no Excel (realizar para cada matriz):

$$CI = \frac{\lambda_{max} - n}{n - 1}$$

$$\lambda_{max} = \frac{\sum_{i=1}^n \lambda_i}{n}$$

$$\lambda_i = \frac{\sum_{j=1}^n PML_j * p_{ij}}{PML_i}$$

| Análise de Consistência<br>- Prioridades Gerais |                             |             |
|-------------------------------------------------|-----------------------------|-------------|
| Lâmbda Critério                                 | Média - LambdaMax           | 4.226       |
| 4.209                                           | Índice de Concistencia (CI) | 0.075236054 |
| 4.138                                           | RI - Inconsist. Aleatória.  | 0.9         |
| 4.206                                           | RC - Razão de Inconsist.    | 8.36%       |
| 4.350                                           | Aceitável (RC > 0,1) ?      | VERDADEIRO  |

Uma razão de inconsistência menor do que 10 % é tolerável. Um valor maior do que este sugere a necessidade de reavaliar as avaliações.

03/09/2019

# Verificação de Consistência

- Fórmulas no Excel (observar arquivo postado no moodle):

| C  | D                | E                | F                | G                                           | H                                            | I                           | J               |
|----|------------------|------------------|------------------|---------------------------------------------|----------------------------------------------|-----------------------------|-----------------|
| 1  |                  |                  |                  |                                             |                                              |                             |                 |
| 2  |                  |                  | n = 4            |                                             |                                              |                             |                 |
| 3  | Competição       | Frequência       | Custo do Aluguel | Prioridade Média Local                      | Análise de Consistência - Prioridades Gerais |                             |                 |
| 4  | =1/4             | =1/5             | 2                | =SOMARPRODUTO(B4:E4;1/\$B\$8:\$E\$8)/\$F\$2 | Lâmbda Critério                              | Média - LambdaMax           | =MÉDIA(H4:H7)   |
| 5  | 1                | =1/2             | 1                | =SOMARPRODUTO(B5:E5;1/\$B\$8:\$E\$8)/\$F\$2 | =SOMARPRODUTO(B4:E4;\$B\$9:\$E\$9)/F4        | Índice de Concistencia (CI) | =(J3-F2)/(F2-1) |
| 6  | 2                | 1                | 4                | =SOMARPRODUTO(B6:E6;1/\$B\$8:\$E\$8)/\$F\$2 | =SOMARPRODUTO(B5:E5;\$B\$9:\$E\$9)/F5        | RI - Inconsist. Aleatória.  | 0.9             |
| 7  | 1                | =1/4             | 1                | =SOMARPRODUTO(B7:E7;1/\$B\$8:\$E\$8)/\$F\$2 | =SOMARPRODUTO(B6:E6;\$B\$9:\$E\$9)/F6        | RC - Razão de Inconsist.    | =J4/J5          |
| 8  | =SOMA(C4:C7)     | =SOMA(D4:D7)     | =SOMA(E4:E7)     | =SOMA(F4:F7)                                | =SOMARPRODUTO(B7:E7;\$B\$9:\$E\$9)/F7        | Aceitável (RC > 0,1) ?      | =J6<0.1         |
| 9  | =TRANSPOR(F4:F7) | =TRANSPOR(F4:F7) | =TRANSPOR(F4:F7) |                                             |                                              |                             |                 |
| 10 |                  |                  |                  |                                             |                                              |                             |                 |

03/09/2019

# Mais Possibilidades com o AHP

- Análise de Sensibilidade;
- ANP – AHP em Rede;
- Múltiplos decisiores;
- Hierarquias mais complexas (ex.: 5 níveis de critérios).

03/09/2019

# Exemplo

- Item 3 da Lista 2:



03/09/2019

# Softwares para a Análise AHP

- Expert Choice: <https://expertchoice.com/>
- Transparent Choice:  
<https://www.transparentchoice.com>
- Pacote R ahp: <https://cran.r-project.org/package=ahp>

03/09/2019

# Referências - AHP

- ALESSIO ISHIZAKA; PHILIPPE NEMERY. **Multi-Criteria Decision Analysis - Methods and Software**. 2013.
- MARINS, C. S.; SOUZA, D. DE O.; BARROS, M. DA S. O Uso Do Método De Análise Hierárquica ( AHP ) Na Tomada De Decisões Gerenciais - Um Estudo. **Xli Sbpo**, p. 11, 2009.
- ALVES, J. R. X.; ALVES, J. M. Definição de localidade para instalação industrial com o apoio do método de análise hierárquica (AHP). **Production**, v. 25, n. 1, p. 13–26, 2015. Disponível em:  
[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0103-65132015000100013&lng=pt&tlang=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-65132015000100013&lng=pt&tlang=pt)..

03/09/2019