

# Graduação em BD e SI

---

## Computação Cognitiva

### Case Zero

Prof. Roberto Santos

---

# **Uma Abordagem Prática de Negócio**



# Elementos Básicos Desta Apresentação



- **Definição de Projeto;**
- **Projeto de Data Analytics;**
- **Definição de Modelo.**

# O que é um Projeto?



Esforço **temporário** empreendido para criar um produto, serviço ou resultado, são atividades ou empreendimentos que têm início e **fim** programados.



Todos os projetos tem um início, um meio e um fim.

# Projeto X Processo



## PROJETO



**Tempo determinado**



**Resultado exclusivo**



**Trabalho inédito**

**vs.**



## PROCESSO



**Tempo indeterminado**



**Resultado padrão**



**Trabalho repetitivo**

# Projeto de Data Analytics

## Data Analytics



- 1 – Entendimento do Negócio;
- 2 – Entendimento dos Dados;
- 3 – Preparação dos dados;
- 4 – Aplicação de Técnicas (Modelagem);
- 5 – Avaliação de Desempenho;
- 6 – Implantação (Entrega).

## O que é um Modelo?

---

# Representação de Algo

# O que é um Modelo?



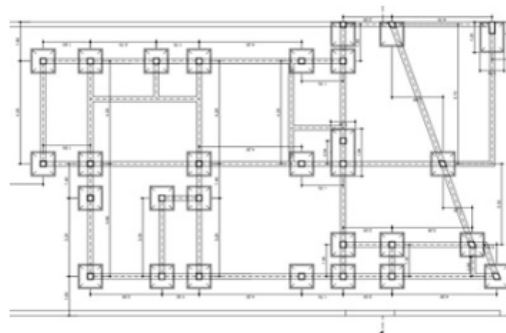


# O que é um Modelo?



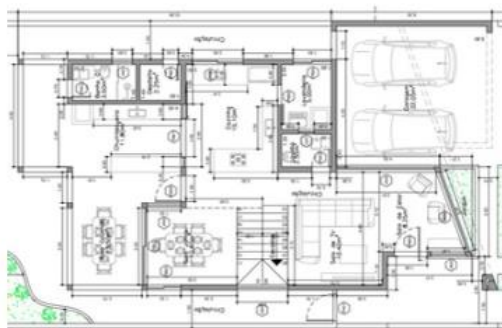
## Fachada em 3D

A maquete da fachada em 3D tem a função de apresentar o projeto. Elaborada em escala, ela tem exatamente as características da casa real.



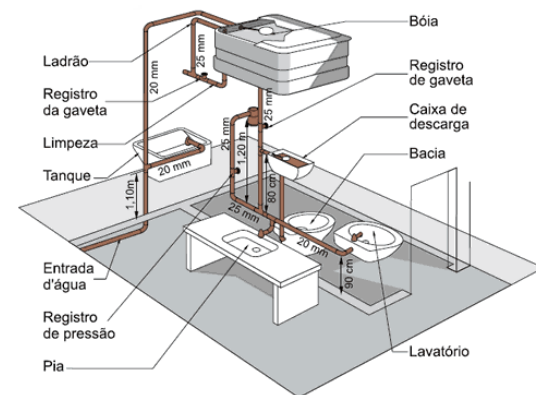
## Planta de Alicerce

Esta planta é uma prévia do projeto estrutural, facilitando a locação da construção no terreno e ajudando o engenheiro a planejar as estruturas da casa.

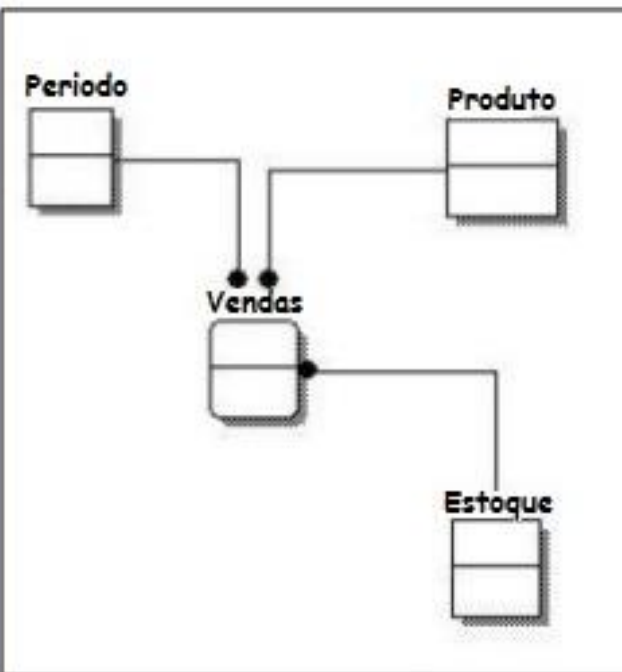


## Planta Baixa

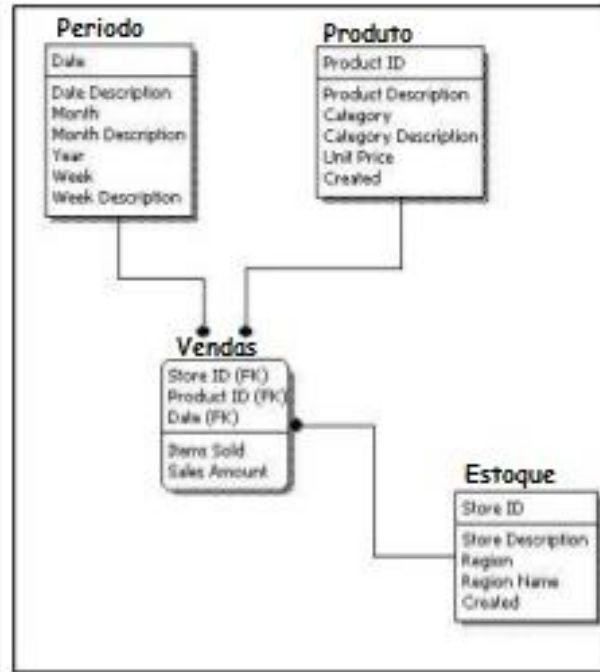
Um dos itens mais importantes do projeto, esse desenho apresenta todas as medidas espaciais no plano horizontal, como os vãos, portas, janelas, pilares entre outros.



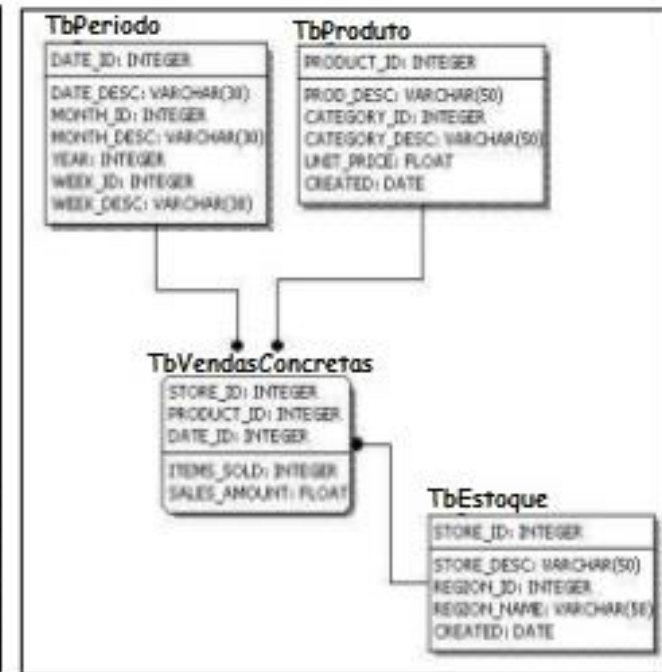
# O que é um Modelo?



**Modelo Conceitual**



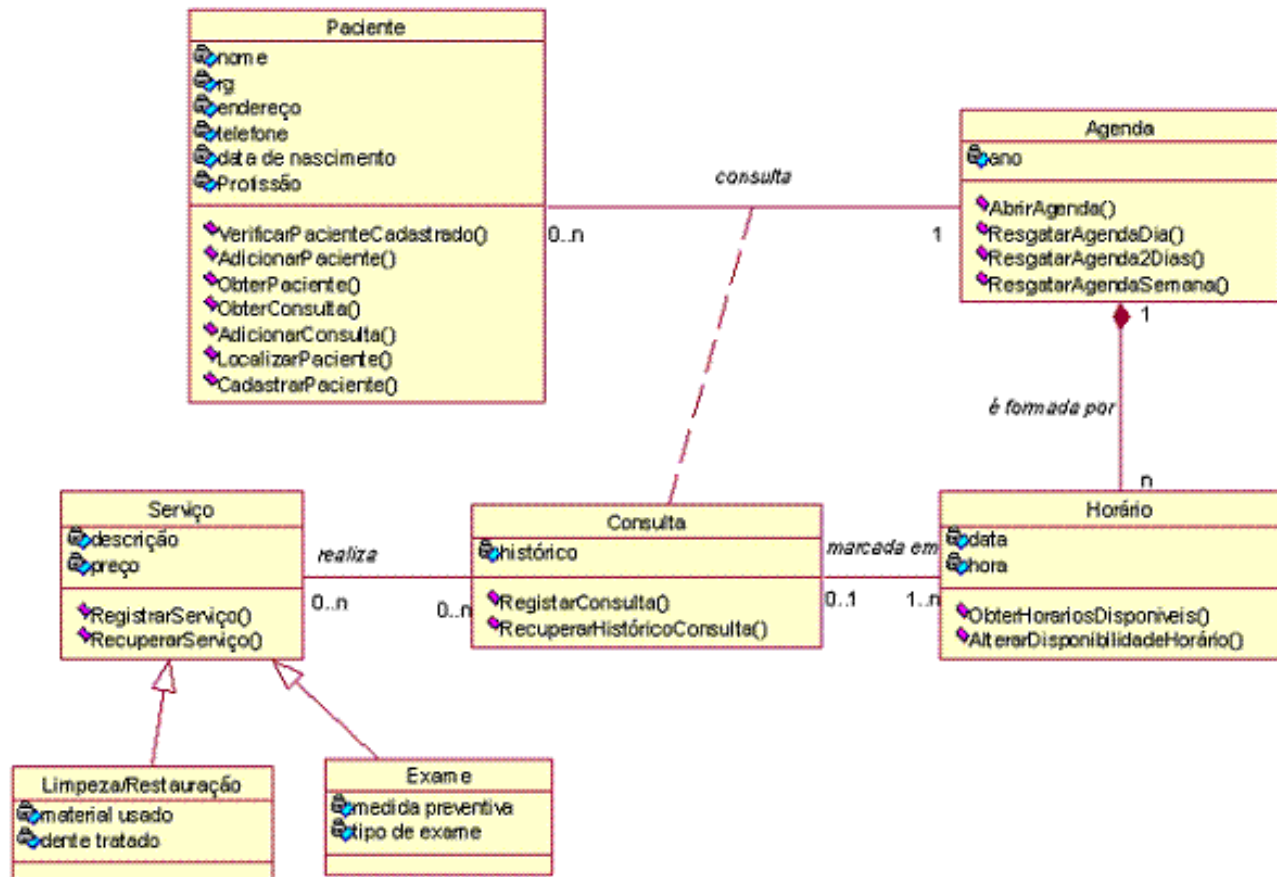
**Modelo Lógico**



**Modelo Físico**

Descreve os dados de um domínio de negócio e seus relacionamentos ...

# O que é um Modelo?



# O que é um Modelo?

## O que representa um modelo em Data Mining?

	A	B	C	D	E	F	G
1	Data	Motorista	Local da Entrega	KM Percorrido	Categoria Produto	Qtd. Paletes	Veículo
2	01/05/2015	Stewart	Loja SP	337	Eletrodomésticos	15	Truck
3	02/05/2015	John	Loja MG	262	Livros	15	Carreta 2 eixos
4	03/06/2015	Stewart	Loja RJ	257	Informática	13	Truck
5	04/06/2015	John	Loja SP	431	Livros	17	Carreta 3 eixos
6	05/07/2015	John	Loja MG	307	Livros	16	Carreta 3 eixos
7	06/06/2015	Carl	Loja MG	284	Eletrodomésticos	20	Truck
8	07/07/2015	Carl	Loja MG	250	Informática	15	Truck
9	08/08/2015	Carl	Loja MG	357	Informática	12	Truck
10	09/09/2015	Carl	Loja MG	346	Livros	20	Truck
11	10/10/2015	Carl	Loja MG	318	Eletrodomésticos	13	Truck
12	11/07/2015	Stewart	Loja SP	274	Eletrodomésticos	15	Truck
13	12/08/2015	John	Loja MG	427	Livros	15	Carreta 2 eixos
14	13/08/2015	Stewart	Loja RJ	327	Informática	13	Truck
15	14/09/2015	John	Loja SP	288	Livros	17	Carreta 3 eixos
16	15/10/2015	John	Loja MG	334	Informática	16	Carreta 3 eixos

*O conhecimento extraído de uma base de dados?*

# O que é um Modelo?

## Exemplos

### Equação da Regressão Múltipla

Os coeficientes do modelo de regressão múltipla são estimados a partir de dados amostrais

Modelo de regressão múltipla com k variáveis independentes

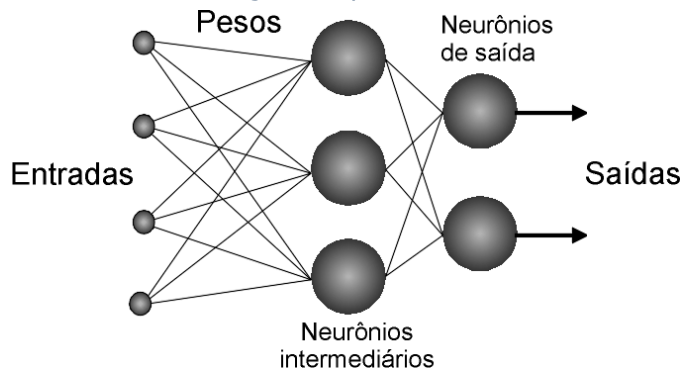
Valor estimado ou previsto de Y

Intercepto estimado

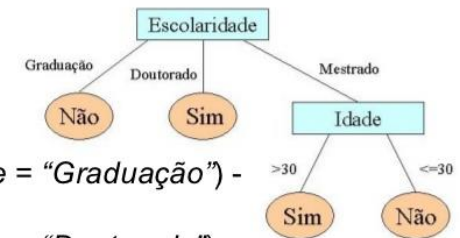
Coefficientes angulares estimados

$$\hat{Y}_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + b_k X_{ki}$$

### Modelo gerado por RNA



### Exemplo – árvore de decisão para regras de classificação



1. Se (*Escolaridade* = “*Graduação*”) - *Rico* = “*Não*”
2. Se (*Escolaridade* = “*Doutorado*”) - *Rico* = “*Sim*”
3. Se (*Escolaridade* = “*Mestrado*”) & (*Idade* = “*>30*”) - *Rico* = “*Sim*”
4. Se (*Escolaridade* = “*Mestrado*”) & (*Idade* = “*<=30*”) - *Rico* = “*Não*”

## Aplicações Reais

### Crédito e Cobrança

- Predição do risco de não pagamento de uma dívida;

### Fraude

- Predição do risco de fraude uma transação de cartão de crédito ou em uma determinada unidade consumidora de energia (Cat Score);

### Marketing

- Medição da propensão à compra/assinatura de um determinado produto ou serviço;

### Seguros

- Cálculo do risco de sinistro para apólices de seguros. Também pode ser usado para cálculo do valor do seguro;

## Aplicações Reais

### Varejo

- Market Basket Analysis - Técnica que faz uso de regras de associação para identificar combinação de produtos dentro das cestas de compras de clientes;

### Inteligência de Mercado

- Análise de Sentimentos - identificação de opinião expressada em um determinado texto em positiva ou negativa. Possibilita análise de produtos e serviços.

### Manutenção Preditiva

- Medição da probabilidade ou risco de falha de um determinado equipamento;

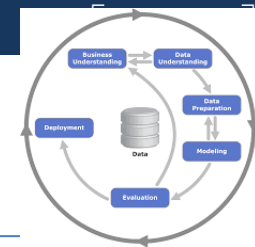
### Saúde

- Probabilidade de rejeição pós transplante de um determinado órgão como medula óssea ou fígado. Probabilidade de risco de doenças futuras.

---

# ***Uma Modelagem Genérica de Score***





1 – Entendimento do Negócio;

2 – Entendimento dos Dados;

3 – Preparação dos dados;

4 – Aplicação de Técnicas (Modelagem);

5 – Avaliação de Desempenho;

6 – Implantação (Entrega).



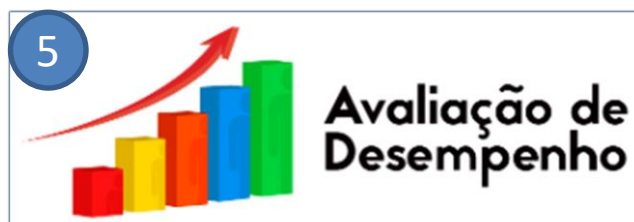
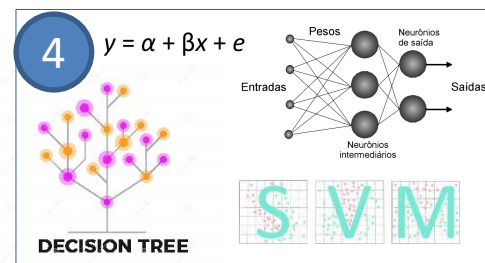
**2**

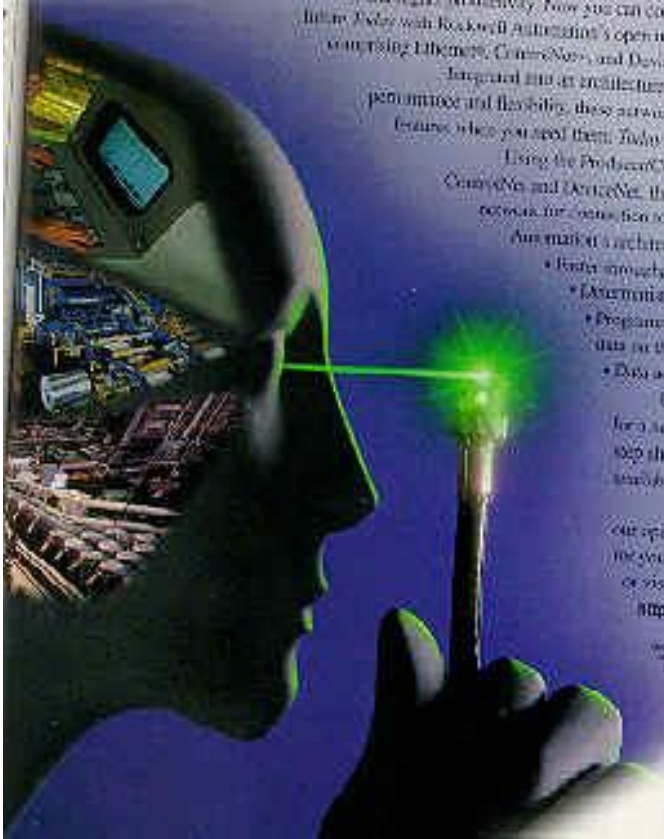
Histórico dos Empréstimos Aprovados Anteriores						
RENDA	PROFISSÃO	IDADE	COMPROMETIMENTO	RESULTADO	ROI Médio	
2000	APOSENTADO	35	MENOR QUE 30%	PAGOU	R\$	100,00
2000	AUTONOMO	19	MENOR QUE 20%	NÃO PAGOU	-R\$	200,00
3 M	1200 PROFESSOR	21	MENOR QUE 20%	NÃO PAGOU	-R\$	200,00
4 F	3000 AUTONOMO	35	MENOR QUE 30%	NÃO PAGOU	-R\$	200,00
5 M	1500 APOSENTADO	63	MENOR QUE 30%	PAGOU	R\$	100,00
6 F	2000 AUTONOMO	40	MENOR QUE 30%	PAGOU	R\$	100,00
7 F	2500 APOSENTADO	55	MENOR QUE 20%	PAGOU	R\$	100,00
8 F	3000 ESTUDANTE	22	MENOR QUE 30%	NÃO PAGOU	-R\$	200,00
9 M	4000 ESTUDANTE	25	MENOR QUE 20%	PAGOU	R\$	100,00
10 F	2500 APOSENTADO	33	MENOR QUE 30%	PAGOU	R\$	100,00



**3**

RENDA	APOSENTADO	AUTONOMO	IDADE<=22	REAL_01	REAL
0	0	1	0	1	PAGOU
0	0	1	0	0	NÃO PAGOU
1	0	0	0	1	PAGOU
0	1	0	0	1	PAGOU
0	0	0	1	0	NÃO PAGOU
0	0	0	0	1	PAGOU
0	1	0	0	1	PAGOU
0	0	1	1	0	NÃO PAGOU
0	1	0	0	1	PAGOU





# ***Case Zero***

---

***Qual o Problema?***

---

**Quem, onde, quando, o quê  
e o porquê**



## **IMPACTUS FINANCEIRA**

Dinheiro na mão é vendaval

**Você está sendo contratado para a área de empréstimos de curto prazo (até 5X) de uma financeira (com prejuízo).**

### **Processo Atual de Empréstimo:**

- 1 - Cliente solicita valor;
- 2 - Checa-se Serasa e SPC (Negativos);
- 3 - Verifica se o valor mensal é inferior à 30% da renda;
- 4 - Aprova/Nega

### **Objetivos:**

- Diminuir a inadimplência
- Dar Lucro

**Você precisa virar esse jogo!!!!**

---

***Como?***

# Credit Scoring

---

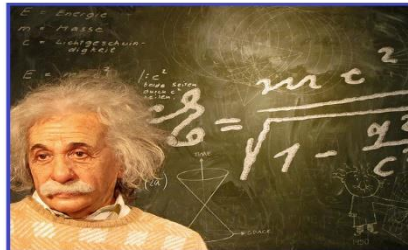
Aplicar algoritmos (técnicas)  
nos dados e gerar um **modelo**  
**(score)** para aprovar melhor.

# Credit Scoring

## Desenho da Solução em Produção

**Novas Propostas  
de Crédito**

**Modelo que calcula o Score**



**Ponto de Corte**

**Decisão**



**Aprovado**



**Negado**



# Credit Scoring

## Planejamento - Desenvolvimento do Modelo

*Histórico de propostas de crédito com tempo suficiente para saber se foi paga ou não.*



*Modelo*

$$y = \alpha + \beta x + e$$



# Credit Scoring

## Dados Históricos

Histórico dos Empréstimos Aprovados Anteriores							
Núm	SEXO	RENDA	PROFISSÃO	IDADE	COMPROMETIMENTO	RESULTADO	ROI Médio
1	M	2000	APOSENTADO	35	MENOR QUE 30%	PAGOU	R\$ 100,00
2	F	2000	AUTONOMO	19	MENOR QUE 20%	NÃO PAGOU	-R\$ 200,00
3	M	1200	PROFESSOR	21	MENOR QUE 20%	NÃO PAGOU	-R\$ 200,00
4	F	3000	AUTONOMO	35	MENOR QUE 30%	NÃO PAGOU	-R\$ 200,00
5	M	1500	APOSENTADO	63	MENOR QUE 30%	PAGOU	R\$ 100,00
6	F	2000	AUTONOMO	40	MENOR QUE 30%	PAGOU	R\$ 100,00
7	F	2500	APOSENTADO	55	MENOR QUE 20%	PAGOU	R\$ 100,00
8	F	3000	ESTUDANTE	22	MENOR QUE 30%	NÃO PAGOU	-R\$ 200,00
9	M	4000	ESTUDANTE	25	MENOR QUE 20%	PAGOU	R\$ 100,00
10	F	2500	APOSENTADO	33	MENOR QUE 30%	PAGOU	R\$ 100,00
11	M	1500	AUTONOMO	18	MENOR QUE 30%	NÃO PAGOU	-R\$ 200,00
12	M	1200	PROFESSOR	40	MENOR QUE 30%	NÃO PAGOU	-R\$ 200,00
13	M	3500	PROFESSOR	32	MENOR QUE 30%	PAGOU	R\$ 100,00
14	F	2500	PROFESSOR	20	MENOR QUE 10%	PAGOU	R\$ 100,00
15	M	1250	AUTONOMO	37	MENOR QUE 10%	PAGOU	R\$ 100,00
16	F	5000	ESTUDANTE	32	MENOR QUE 30%	PAGOU	R\$ 100,00
17	M	3500	AUTONOMO	31	MENOR QUE 30%	NÃO PAGOU	-R\$ 200,00
18	F	2000	APOSENTADO	55	MENOR QUE 30%	PAGOU	R\$ 100,00
19	M	1400	PROFESSOR	44	MENOR QUE 10%	PAGOU	R\$ 100,00
20	F	1100	APOSENTADO	65	MENOR QUE 30%	NÃO PAGOU	-R\$ 200,00

## Criação do Modelo

---

- Qual o tipo de resposta desejada? score
- Quais técnicas eu posso utilizar?
- Que ferramentas eu tenho disponível?
- Quantos registro posso utilizar?

## Ferramenta Principal Utilizada

---



## Avaliação de Desempenho

---

<b>TX APROVAÇÃO:</b>	<b>55%</b>
<b>TX ACERTO:</b>	<b>85%</b>
<b>ROI:</b>	<b>R\$ 800,00</b>

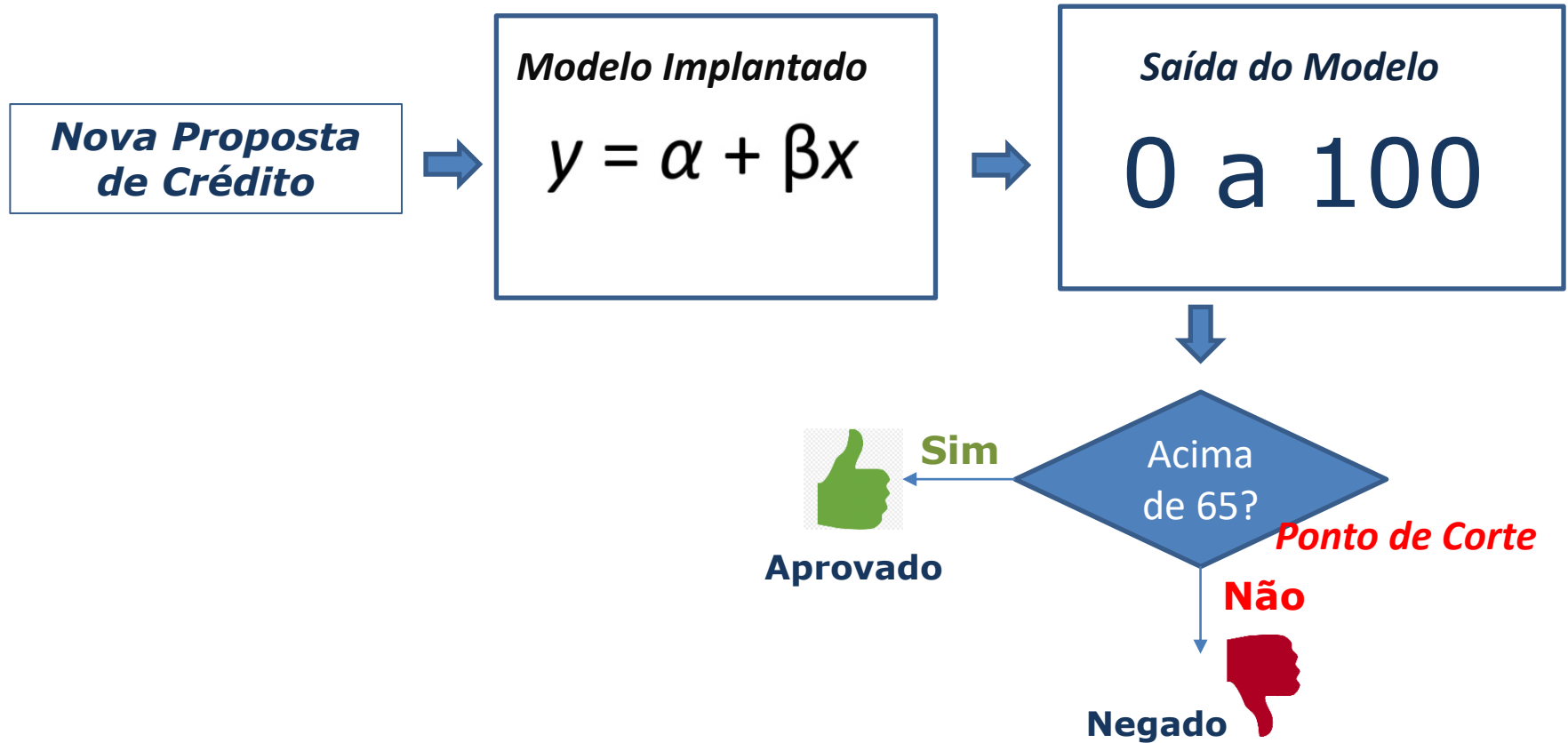
## Novas Decisões

Novos Casos						
Núm	SEXO	RENDA	PROFISSÃO	IDADE	COMPROMET. MÁX.	RESULTADO
101	M	3000	AUTONOMO	35	MENOR QUE 30%	?
102	F	4000	ESTUDANTE	19	MENOR QUE 20%	?
103	M	2500	AUTONOMO	21	MENOR QUE 20%	?
104	F	1500	APOSENTADO	35	MENOR QUE 30%	?
105	M	1200	PROFESSOR	63	MENOR QUE 30%	?
106	F	3500	APOSENTADO	40	MENOR QUE 30%	?



Não sei se o “cara” será bom ou mal, mas pode-se calcular o SCORE!!!

## Implantação do Modelo On-line



---

# Obrigado!

Prof. Roberto Santos  
[roberto.santos@faculdadeimpacta.com.br](mailto:roberto.santos@faculdadeimpacta.com.br)