

ABRAHAM LAREDO SICSÚ

# CREDIT SCORING

2ª EDIÇÃO

DESENVOLVIMENTO, IMPLANTAÇÃO E ACOMPANHAMENTO

COM PYTHON

Blucher

Abraham Laredo Sicsú

**CREDIT SCORING**  
Desenvolvimento, implantação  
e acompanhamento com Python

2ª Edição

*Credit scoring: desenvolvimento, implantação e acompanhamento*

1ª edição – 2010

© 2026 Abraham Laredo Sicsú

Editora Edgard Blücher Ltda.

*Publisher* Edgard Blücher

*Editor* Eduardo Blücher

*Coordenação editorial* Rafael Fulanetti

*Coordenação de produção* Ana Cristina Garcia

*Produção editorial* Juliana Midori Horie

*Diagramação* Guilherme Salvador

*Revisão de texto* Equipe editorial

*Capa* Juliana Midori Horie

*Imagem da capa* iStockphoto

# Blucher

Rua Pedroso Alvarenga, 1245, 4º andar

04531-934 – São Paulo – SP – Brasil

Tel.: 55 11 3078-5366

**contato@blucher.com.br**

www.blucher.com.br

Segundo o Novo Acordo Ortográfico, conforme 6. ed. do Vocabulário Ortográfico da Língua Portuguesa, Academia Brasileira de Letras, julho de 2021.

É proibida a reprodução total ou parcial por quaisquer meios sem autorização escrita da editora.

Todos os direitos reservados pela Editora Edgard Blücher Ltda.

Dados Internacionais de Catalogação na Publicação (CIP)

Heytor Diniz Teixeira, CRB-8/10570

Sicsú, Abraham Laredo

*Credit scoring : desenvolvimento, implantação e acompanhamento com Python / Abraham Laredo Sicsú. – 2. ed. – São Paulo : Blucher, 2026.*

192 p.

Bibliografia

ISBN 978-85-212-3133-2 (Impresso)

ISBN 978-85-212-3132-5 (Eletrônico – Epub)

ISBN 978-85-212-3131-8 (Eletrônico – PDF)

1. Crédito. 2. Risco de crédito. 3. Análise de dados. 4. Machine learning. 5. Python (Linguagem de programação de computador). I. Título.

CDU 336.77:004.43

Índice para catálogo sistemático:

1. Python (Linguagem de programação de computador) CDU 004.43.

# CONTEÚDO

<b>1. FUNDAMENTOS.....</b>	<b>19</b>
1.1 Risco de crédito e <i>credit scores</i> .....	19
1.2 Por que medir o risco de crédito? .....	21
1.3 Aplicações de <i>credit scoring</i> .....	22
1.4 Estimando o risco de crédito – ideia básica .....	22
1.5 Cálculo dos escores .....	23
1.6 Probabilidade e erros de decisão .....	24
1.7 Premissa básica em modelos de <i>credit scoring</i> .....	25
1.8 <i>Application scoring</i> e <i>behavioral scoring</i> .....	26
1.9 Modelos julgamentais e modelos quantitativos.....	26
1.9.1 Modelos julgamentais.....	26
1.9.2 Modelos quantitativos I – Modelos generalistas .....	27
1.9.3 Modelos quantitativos II – Modelos customizados.....	28
1.10 Roteiro para o desenvolvimento de um modelo de scoring.....	28
1.11 Uma nota de alerta .....	28
<b>2. PLANEJAMENTO E DEFINIÇÕES .....</b>	<b>31</b>
2.1 Objetivo do estudo.....	31

2.2	Análise do tipo de operação a ser considerada.....	32
2.3	Definição e segmentação do mercado-alvo .....	33
2.4	Definição de bom e mau pagador .....	34
2.5	Data de referência, período de performance, período histórico e safra ...	37
2.5.1	Data de referência .....	37
2.5.2	Período de performance.....	37
2.5.3	Período histórico.....	39
2.5.4	Período de concessão e safra.....	39
<b>3.</b>	<b>IDENTIFICAÇÃO DAS VARIÁVEIS PREVISORAS .....</b>	<b>41</b>
3.1	Variáveis discriminadoras potenciais.....	41
3.1.1	Tipos de variáveis.....	42
3.2	Agrupando informações em intervalos de tempo.....	43
3.3	Cuidados na identificação das variáveis potenciais.....	44
3.3.1	Definição operacional – uniformidade na interpretação .....	44
3.3.2	Confiabilidade das informações.....	46
3.3.3	Informações recentes .....	46
3.3.4	Disponibilidade ao longo do tempo.....	47
3.3.5	Variáveis aceitáveis pelos analistas ou pela empresa credora .....	47
3.3.6	Aspectos éticos e legais .....	48
3.4	Quantas variáveis utilizar?.....	49
Apêndice 3.1	Informações interessantes no desenvolvimento de modelos de <i>credit scoring</i> para pessoas físicas .....	50
Apêndice 3.2	Informações interessantes no desenvolvimento de modelos de <i>credit scoring</i> para pessoas jurídicas .....	52
<b>4.</b>	<b>AMOSTRAGEM E COLETA DE DADOS .....</b>	<b>55</b>
4.1	Amostragem.....	55
4.1.1	Unidade amostral .....	55
4.1.2	Mercado-alvo.....	56

4.2	Formas de amostragem.....	57
4.2.1	Amostragem aleatória simples (AAS).....	57
4.2.2	Amostragem aleatória estratificada.....	58
4.2.3	Dimensionamento das amostras .....	58
4.3	Amostras de desenvolvimento e teste .....	59
4.4	Aquisição dos dados.....	59
4.5	Cuidados especiais na coleta de dados .....	60
4.6	Dados a serem utilizados no Livro.....	62
<b>5.</b>	<b>ANÁLISE DOS DADOS .....</b>	<b>65</b>
5.1	Introdução.....	65
5.2	Análise univariada .....	66
5.2.1	Distribuições de frequências.....	66
5.2.2	Identificação e tratamento de inconsistências .....	70
5.2.3	Identificação e tratamento de <i>outliers</i> (dados discrepantes) .....	70
5.2.4	Identificação e tratamento de valores em branco ( <i>missing values</i> )... ..	72
5.2.5	“Síndrome de outros” .....	74
5.2.6	Problemas com siglas ou abreviações.....	74
5.3	Definição de novas variáveis .....	74
5.4	Transformações de variáveis .....	75
5.4.1	Por que transformar uma variável quantitativa? .....	75
5.4.2	Discretização uma variável quantitativa .....	77
5.4.3	Como discretizar variáveis quantitativas.....	79
5.5	Fusão não estatística de categorias de baixa frequência .....	81
5.6	Codificação de variáveis qualitativas .....	82
5.7	Concluindo a análise univariada.....	86
<b>6.</b>	<b>ANÁLISES BIVARIADAS.....</b>	<b>87</b>
6.1	Introdução.....	87
6.2	Fusão de categorias.....	92
6.2.1	Critérios utilizados para fusão de categorias .....	93

6.3	Pré-seleção de variáveis ( <i>feature selection</i> ) .....	96
6.3.1	Análise individual das variáveis.....	97
6.3.2	Análise conjunta de variáveis preditoras .....	99
6.4	Análise de multicolinearidade.....	100
6.5	Preparando o arquivo de trabalho (Kimsport6) .....	101
<b>7.</b>	<b>MODELO ESTATÍSTICO: REGRESSÃO LOGÍSTICA.....</b>	<b>103</b>
7.1	Introdução.....	103
7.2	Regressão logística .....	105
7.3	Cálculo da fórmula de escoragem com regressão logística .....	107
7.4	Análise dos sinais da fórmula de escoragem.....	108
7.5	Impacto de uma variável sobre <i>pmp</i> .....	109
7.6	Correção para amostragem estratificada .....	110
7.7	Conclusão .....	112
	Apêndice 7.1 Seleção de variáveis.....	112
<b>8.</b>	<b>MODELOS DE MACHINE LEARNING .....</b>	<b>115</b>
8.1	Introdução.....	115
8.2	Árvores de classificação .....	116
8.3	Random Forest .....	118
8.3.1	Aplicação de Random Forest ao nosso exemplo (Kimsport).....	119
8.3.2	Importância das variáveis no cálculo dos escores.....	120
8.4	XGBoost.....	121
8.4.1	Aplicação de Random XGBoost ao nosso exemplo (Kimsport) ....	122
8.4.2	Importância das variáveis no cálculo das probabilidades.....	124
<b>9.</b>	<b>ANÁLISE E VALIDAÇÃO DA FÓRMULA DE ESCORAGEM .....</b>	<b>127</b>
9.1	Introdução.....	127
9.2	Avaliação por analistas de crédito.....	128
9.3	Análise visual da distribuição de bons e maus pagadores por classes de frequência .....	129

9.4	Análise estatística do poder discriminador do modelo .....	130
9.4.1	KS: Índice de Kolmogorov-Smirnov (KS) .....	132
9.4.2	AUC (AUROC) – Area Under Receiver Operating Characteristic... ..	133
9.5	Análise estatística da calibração do modelo.....	135
9.6	Calibração do modelo .....	136
9.6.1	Ajuste manual .....	137
9.6.2	Método de Platt .....	137
9.6.3	Regressão isotônica .....	138
Apêndice 9.1	Cross validation.....	138
<b>10.</b>	<b>TOMADA DE DECISÃO: PONTOS DE CORTE E CLASSES DE RISCO ...</b>	<b>141</b>
10.1	Escore de crédito (SCR) .....	141
10.1.1	Definição de um ponto de corte .....	142
10.2	Indicadores de performance baseados em um ponto de corte .....	143
10.2.1	Matriz de classificação .....	144
10.2.2	Indicadores de performance.....	145
10.3	Dois pontos de corte e região cinza .....	148
10.4	Classes de Risco .....	149
<b>11.</b>	<b>APERFEIÇOANDO O MODELO .....</b>	<b>153</b>
11.1	Utilização das informações dos proponentes recusados ( <i>reject inference</i> ) .....	153
11.1.1	Alternativa 1: Considerar todos os recusados anteriormente como maus .....	156
11.1.2	Alternativa 2: Extrapolação I (extrapolação simples).....	156
11.1.3	Alternativa 3: Extrapolação II ( <i>parceling</i> ).....	157
11.1.4	Alternativa 4: Ponderação (ou <i>augmentation</i> ).....	159
11.2	Interação de variáveis .....	160
11.3	Balanceamento de amostras.....	161
11.3.1	Reamostragem .....	162
11.3.2	Algoritmos utilizados para contornar o desbalanceamento ....	163

11.3.3	Métricas para comparação dos resultados .....	164
11.3.4	Aplicação de balanceamento ao Kimsport modificado.....	165
<b>12.</b>	<b>IMPLANTAÇÃO DO MODELO .....</b>	<b>167</b>
12.1	Introdução.....	167
12.2	Filtros de crédito .....	167
12.3	Regras para interveniência ( <i>overrides</i> ).....	168
12.4	Documentação .....	169
12.4.1	Diário de bordo .....	170
12.5	SADC – Sistema de Apoio à Decisão de Crédito .....	170
12.6	Aspectos técnicos da implantação do <i>credit scoring</i> .....	171
12.6.1	Detalhamento para a área de informática .....	171
12.6.2	Segurança do sistema de cálculo dos escores.....	172
12.6.3	Homologação do sistema.....	173
12.6.4	Envolvimento de outras áreas.....	173
12.7	Educação e treinamento dos usuários .....	174
<b>13.</b>	<b>GESTÃO E MONITORAMENTO DO MODELO.....</b>	<b>177</b>
13.1	Introdução.....	177
13.2	Auditoria .....	177
13.2.1	Quando realizar as auditorias .....	178
13.2.2	Como realizar a auditoria.....	178
13.2.3	O que verificar?.....	178
13.2.4	Relatório da auditoria .....	179
13.3	Monitoramento da estabilidade populacional.....	180
13.3.1	Introdução.....	180
13.3.2	O que monitorar?.....	181
13.3.3	Periodicidade .....	181
13.3.4	Distribuições de referência .....	182
13.3.5	Amostragem e informações para monitoramento.....	183

---

13.3.6	Análise da estabilidade populacional ( <i>Population Stability Index</i> ) .....	184
13.3.7	Análise das distribuições das variáveis .....	185
13.3.8	Impacto da alteração da distribuição da variável na regressão logística .....	185
13.3.9	Análise da instabilidade populacional.....	187
13.4	Instabilidade dos escores – Ações de contingência .....	187
13.5	Monitoramento do desempenho de um modelo .....	188
13.5.1	Introdução.....	188
13.5.2	Amostragem para avaliação de desempenho .....	190
13.5.3	Indicadores de desempenho – sugestões .....	190
13.5.4	Exemplos de relatórios de monitoramento .....	190
13.6	Matriz de migração .....	191
13.6.1	Análise da matriz de migração.....	192

# CAPÍTULO 1

## FUNDAMENTOS

### 1.1 RISCO DE CRÉDITO E *CREDIT SCORES*

A concessão de crédito é uma decisão sob condições de incerteza. Em empréstimos, vendas a prazo, prestação de serviços etc., quer o crédito seja solicitado ou quer seja oferecido pelo credor, sempre existe a possibilidade de perda. Se o credor puder estimar a probabilidade de que essa perda ocorra, sua decisão será mais confiável.

Em termos gerais, quando houver perda em uma operação de crédito, diremos que o cliente foi “mau pagador”, ou simplesmente “mau”; caso contrário, diremos que o cliente foi “bom pagador”, ou simplesmente “bom”.<sup>1</sup> Eventualmente, poderemos definir clientes “intermediários”. Uma discussão sobre a forma de definir bons e maus clientes será apresentada no Capítulo 2.

O objetivo dos modelos de *credit scoring* é prever, na data da decisão do crédito, a probabilidade de que o crédito, se concedido, incorra em perda para o credor. À probabilidade de isso ocorrer, ou seja, à *probabilidade de perda em uma operação de crédito denominamos risco de crédito*.

A estimativa dessa probabilidade é função das características do solicitante de crédito e da operação. Há várias formas de estimar essa probabilidade. Este livro trata da estimação do risco de crédito por meio de técnicas estatísticas e *machine learning*.

O *credit score* é uma medida do risco de crédito. *Modelo de credit scoring* é a denominação genérica dada no mercado para a fórmula de cálculo dos escores de crédito.

---

<sup>1</sup> O fato de um cliente ficar inadimplente não o torna necessariamente um mau pagador. Como veremos, a definição de *mau pagador* ou *bom pagador* para o desenvolvimento de um modelo de escoragem é uma decisão que pode variar entre diferentes credores.

Os escores de crédito têm a finalidade única de quantificar o risco de crédito. A forma com que essa informação é utilizada para tomar a decisão de conceder ou não o crédito, para estruturar a operação definindo taxas, garantias, prazos etc., é atribuição dos gestores de crédito. Para operacionalizar um processo de concessão e gestão de crédito faz-se necessário, além do modelo de cálculo dos escores, uma política de crédito bem-definida, um sistema de informações gerenciais com dados do cliente, operação, formas de pagamento, políticas de cobrança etc.

A implantação e gestão de um processo automatizado de análise e decisão de crédito requerem a geração e atualização permanente de um sistema de informações gerenciais com dados sobre os solicitantes de crédito, as operações realizadas, os registros da forma de pagamento dos créditos concedidos e do processo de cobrança. Essa base de dados permitirá, por exemplo, para diferentes faixas de escores:

- analisar/monitorar o perfil dos solicitantes de crédito para orientar estratégias de marketing, como a segmentação do mercado ou a oferta de novos produtos de crédito;
- acompanhar e medir a ocorrência de inadimplência;
- rever a adequação dos limites de crédito dados aos clientes;
- avaliar a intensidade e forma de uso de determinado produto de crédito (cartão de crédito, por exemplo);
- atualizar os modelos para cálculo dos escores.

Os escores de crédito podem ser calculados considerando apenas as características do solicitante de crédito, o chamado *risco cliente*, ou, de forma mais ampla, considerando também características da operação (*risco cliente & operação*). Dada uma solicitação de crédito, quanto maior o número de informações disponíveis, mais confiável será a estimativa do risco. Por isso, é preferível calcular os escores considerando também as características da operação.

Os conceitos de *credit scoring* e *rating* são distintos. Enquanto o *credit scoring* é um processo puramente quantitativo, a determinação do *rating* de uma pessoa jurídica ou física depende em grande parte de avaliação subjetiva, podendo ou não contemplar métodos quantitativos como parte do processo. No entanto, é comum encontrarmos no mercado, sob a denominação de *ratings*, pontuações de empresas que na realidade são escores de crédito.

Apesar de *credit scoring* referir-se especificamente à avaliação do risco de crédito, é comum no mercado designar dessa forma outros modelos que podem ou não estar ligados ao crédito. Por exemplo, modelos para prever o cancelamento voluntário de um cliente do cartão de crédito (*churning*) ou para identificar o perfil de clientes que respondem à oferta de crédito por mala direta, são modelos que utilizam as mesmas técnicas estatísticas que o *credit scoring*, razão pela qual são inadequadamente denominados modelos de *credit scoring*.

Uma aplicação importante, ligada diretamente à gestão de crédito, são os modelos de *collection scoring* utilizados como instrumentos para orientar as estratégias de

cobranças. Os clientes inadimplentes são classificados em classes de acordo com seu escore de cobrança e, para cada classe, aplicam-se diferentes estratégias de cobrança. O uso desses modelos não só melhora o relacionamento com os clientes, pois permite aplicar a estratégia adequada a cada caso, como também reduz as despesas com cobranças desnecessárias. Ademais, auxilia na previsão de perdas e custos de cobrança.

## 1.2 POR QUE MEDIR O RISCO DE CRÉDITO?

O risco de uma solicitação de crédito pode ser avaliado de forma subjetiva ou medido de forma objetiva utilizando metodologia quantitativa. A avaliação subjetiva, apesar de incorporar a experiência do analista, não quantifica o risco de crédito. Dizer que uma empresa é de alto risco não é suficiente para estimar de maneira precisa as perdas ou ganhos esperados com a operação e, conseqüentemente, tomar a decisão mais adequada.<sup>2</sup>

Medir o risco de crédito de forma objetiva, utilizando técnicas quantitativas, apresenta uma série de vantagens:

- *consistência nas decisões*: se submetermos uma mesma solicitação de crédito a diferentes analistas, poderemos obter diferentes avaliações subjetivas, pois a experiência e o envolvimento com o cliente diferem entre eles. Ademais, um mesmo analista pode dar diferentes avaliações para uma mesma proposta se submetida em momentos diferentes. É humano! No entanto, isso não ocorrerá se aplicarmos um modelo quantitativo de *credit scoring*. Mantidas inalteradas as características da solicitação de crédito (do cliente e da operação), o escore será o mesmo, independentemente do analista, da agência ou da filial do credor;
- *decisões rápidas*: os recursos computacionais hoje disponíveis permitem que o escore seja computado quase que instantaneamente, logo após cadastrar os dados da solicitação. Centenas ou milhares de decisões podem ser tomadas em um dia, de forma segura e consistente! A pronta resposta a um cliente potencial é uma vantagem competitiva do credor;
- *decisões adequadas*: em função do risco quantificado, o credor poderá adotar diferentes regras de concessão do crédito;
- o conhecimento das probabilidades de perda permite calcular perdas e ganhos esperados com as operações. Isso permite precificar as operações de forma adequada;
- os clientes podem ser divididos em classes de risco conforme seu escore. Para cada classe, o credor pode (deve!) adotar diferentes regras de concessão de crédito, diferenciando, por exemplo, as taxas a aplicar. Ao reduzir essas taxas para clientes

---

<sup>2</sup> Por analogia, dizer que uma pessoa está com febre “alta” não é suficiente para recomendar o tratamento. Ela pode estar com 38 °C ou com 41 °C, o que, convenhamos, é significativamente diferente.

de baixo risco, terá como efeito a conquista de maior número de clientes, ou seja, de ampliação de mercado;

- *decisão à distância*: o credor não precisa alocar um analista de crédito em cada loja ou filial. O vendedor imputa os dados no ponto de venda e, logo após submeter essas informações, receberá a decisão de crédito em sua tela;
- *monitorar e administrar o risco de um portfólio de crédito*: sem a quantificação do risco individual esta tarefa é inviável. Para a avaliação do risco do portfólio são necessárias, além dos escores, outras medidas que não serão discutidas neste texto.

A utilização de medidas objetivas para o risco de crédito permite também:

- verificar o grau com que atendemos aos requisitos de órgãos reguladores;
- estabelecer uma linguagem comum entre os decisores de crédito; e
- definir níveis de alçada para concessão de crédito.

Apesar de não conhecermos trabalhos publicados no Brasil, comparando os resultados da aplicação de modelos de *credit scoring* com os resultados decorrentes da análise “tradicional” de crédito, informações que nos têm sido fornecidas por responsáveis pelas áreas de crédito de instituições financeiras sugerem que os resultados dos modelos de *credit scoring* são bem superiores!

### 1.3 APLICAÇÕES DE CREDIT SCORING

A aplicação dos modelos de *scoring* é fundamental para a decisão de crédito em massa. Nessas operações o número de solicitações de crédito a serem analisadas em um pequeno intervalo de tempo é muito alto. Alocar analistas de crédito experientes não compensaria em termos financeiros, pois os valores envolvidos em uma operação são relativamente pequenos e a relação custo operacional/benefício não é favorável. Isso, admitindo que dispuséssemos de um exército de bons analistas, o que sabemos não ser possível. Ademais, nesses casos, a negociação das taxas de juros e condições das operações com cada solicitante de crédito não é viável. A necessidade de um processo automatizado de análise e decisão de crédito é necessária e, portanto, os modelos de *scoring* aqui encontram seu nicho principal de aplicação.

Todavia, quando o valor do crédito a ser negociado é suficientemente alto, a alocação de um ou mais analistas experientes para analisar a operação com um maior nível de detalhamento é a opção recomendável. Nesses casos o escore de crédito ainda é importante por ser um indicador confiável do risco do cliente e da operação.

### 1.4 ESTIMANDO O RISCO DE CRÉDITO – IDEIA BÁSICA

A ideia que fundamenta as técnicas de *credit scoring* é simples. Para melhor entendê-la, vamos considerar um exemplo: suponhamos que no financiamento de automóveis o

credor analisa apenas três características dos solicitantes. O tipo de residência (própria ou alugada) na data da solicitação, se o solicitante possui protestos em aberto (sim ou não) na data da solicitação e se o automóvel objeto da operação é novo ou usado. A experiência do credor com clientes que receberam financiamentos no passado mostrou os seguintes resultados:

**Tabela 1.1** – Financiamento de automóveis

Auto	Residência	Protestos	Bons pagadores	Maus pagadores
Novo	Própria	Não	98,0%	2,0%
Novo	Própria	Sim	85,0%	15,0%
Novo	Alugada	Não	92,0%	8,0%
Novo	Alugada	Sim	79,0%	21,0%
Usado	Própria	Não	87,0%	13,0%
Usado	Própria	Sim	70,0%	30,0%
Usado	Alugada	Não	79,0%	21,0%
Usado	Alugada	Sim	61,0%	39,0%

Fonte: elaborada pelo autor.

Se um solicitante requisita financiamento para um veículo novo e na data da solicitação possui casa própria, mas tem protestos em aberto, então, com base em sua experiência, o credor estima que ao conceder o financiamento corre um risco de 15% de que o cliente se torne um mau pagador. Caberá ao analista de crédito decidir se concede ou não o crédito, a taxa de juros, prazos etc. em função dessa probabilidade.

O exemplo anterior mostra a lógica das técnicas quantitativas utilizadas para prever o risco de um futuro cliente. Com três variáveis, cada uma com apenas duas categorias, não necessitamos modelos mais refinados para estimar a probabilidade de ser um mau pagador. À medida que o número de variáveis aumenta, o número de combinações aumenta exponencialmente. Por exemplo, com 20 variáveis binárias (ou seja, cada variável apresenta apenas duas categorias), teremos 1.048.576 combinações possíveis. A tabela correspondente para analisar as probabilidades (equivalente à Tabela 1.1 acima) teria mais de um milhão de linhas, e a amostra necessária para dar estimativas confiáveis em cada uma dessas combinações teria de ser gigantesca. Problemas reais podem envolver algumas centenas de variáveis, a maior parte com mais de duas categorias. A análise só é viável com o auxílio de avançadas técnicas quantitativas.

## 1.5 CÁLCULO DOS ESCORES

Com certas técnicas quantitativas, os escores são calculados atribuindo-se pesos a variáveis que caracterizam o solicitante<sup>3</sup> e a operação. No passado costumava-se apresentar a fórmula de cálculo no formato de uma planilha denominada *scorecard*, provavelmente para facilitar que o analista calculasse “à mão” o escore do solicitante. O Quadro 1.1 apresenta um *scorecard* (com pesos fictícios) para calcular o escore de um solicitante.

Se um cliente reside em casa própria, tem 32 anos, é empresário, possui cartão de primeira linha e no ato da solicitação de crédito tem desabonos em aberto, seu escore será igual a  $30 + 15 + 10 + 40 - 40 = 55$ .

Várias são as técnicas quantitativas que podem ser utilizadas para calcular os escores. Hoje, provavelmente, as técnicas quantitativas mais utilizadas são as técnicas estatísticas de análise discriminante; em particular, a regressão logística.

**Tabela 1.2 – Scorecard**

Variável	Pesos			
	Própria	Alugada	Outras	
Tipo de residência	30	10	-15	
Idade (anos completos)	25 ou –	26 a 35	35-50	51 ou +
	10	15	30	50
Ocupação	Empresário	Funcionário	Autônomo	Em branco
	10	25	-15	-20
Possui cartão de crédito de 1ª linha	Sim	Não		
	40	0		
Desabonos	Sim	Não		
	-40	10		

Fonte: elaborada pelo autor.

## 1.6 PROBABILIDADE E ERROS DE DECISÃO

Em todo processo de tomada de decisão entre duas alternativas, podem-se cometer dois tipos de erro. No caso da decisão de crédito, um erro é recusar a solicitação de um solicitante que, apesar de ter perfil de alto risco, honraria seu compromisso caso recebesse o crédito. Outro erro é aprovar uma operação de crédito que irá implicar perdas no futuro.

Admitamos, por exemplo, que a estimativa da probabilidade de que um solicitante com determinado perfil se torne mau pagador seja igual a 0,80 (80% de chances). Isso significa que, a cada 100 tomadores de crédito com esse perfil, aproximadamente 80 se tornarão maus pagadores. É quase certo que negaríamos o crédito solicitado. Ao decidir dessa forma, estaríamos negando o crédito a 20% dos solicitantes que, mesmo apresentando esse perfil de alto risco, não se tornariam maus pagadores. Para esses 20%, a decisão seria incorreta e injusta. Esse tipo de erro tem um custo de difícil estimação: o custo de deixar um “bom” cliente insatisfeito. Além do custo relativo à não realização da operação, podemos ter o custo da eventual perda do cliente.

Por outro lado, admitamos que, para um determinado perfil, a probabilidade de um solicitante tornar-se mau pagador seja igual a 0,02, isto é, 2%. Nesse caso a tendência

seria conceder o crédito. Se concedermos crédito a clientes com esse perfil, a cada 100 clientes, teremos aproximadamente 2 maus pagadores. O custo desse erro é mais fácil de estimar que no caso anterior; contempla a perda total ou parcial do montante emprestado e os demais custos operacionais.

Em suma, ao tomar decisões com base em probabilidades, temos dois tipos de erro:

- erro I: negar o crédito a um solicitante que seria um bom pagador;
- erro II: conceder o crédito a um solicitante que será um mau pagador.

**Tabela 1.3** – Classificação das decisões de crédito

Realidade ↓	Decisão	
	Aprovar crédito	Negar crédito
Bom pagador	Correta	Erro I
Mau pagador	Erro II	Correta

Fonte: elaborada pelo autor.

## 1.7 PREMISSA BÁSICA EM MODELOS DE *CREDIT SCORING*

Para o desenvolvimento de modelos de *credit scoring*, baseamo-nos em amostras de clientes aos quais concedemos crédito no passado. Isso significa que estamos supondo que, no futuro, quando aplicarmos esses modelos, o comportamento dos novos solicitantes será igual ao dos clientes no passado.

Comportamento futuro ~ Comportamento passado
--

Ainda que esse “passado” seja recente (um ou dois anos antes da data de avaliação do risco desses novos solicitantes), as mudanças nas condições socioeconômicas do local podem ter mudado. Consequentemente, o comportamento dos clientes face aos compromissos assumidos também pode mudar.

No Brasil, antes de 1994, a premissa de que o futuro seria igual ao passado era frequentemente falsa. Era comum termos os famosos “pacotes econômicos” que, de um dia para outro, alteravam a operação do mercado de crédito e tinham sérias implicações sociais. Este autor, em várias situações, viu-se na necessidade de descartar amostras recém-coletadas, posto que muitos dos clientes considerados nessas amostras não tinham condições de tomar crédito em função das novas regras de mercado.

Como veremos no Capítulo 14, o monitoramento do modelo de *credit scoring*, enquanto estiver em operação, para detectar possíveis mudanças em relação às condições vigentes em seu desenvolvimento é muito importante para garantir a contínua eficácia do modelo.

## 1.8 APPLICATION SCORING E BEHAVIORAL SCORING

Os modelos de *credit scoring* podem ser aplicados a solicitantes de crédito com os quais o credor não teve experiência anterior ou a solicitantes que já tomaram crédito desse credor. Os modelos desenvolvidos para “novos” solicitantes de crédito são denominados modelos de *application scoring*. Os modelos desenvolvidos para clientes ou ex-clientes de crédito do credor são denominados modelos de *behavioral scoring*.

A diferença entre esses modelos reside nas variáveis utilizadas para estimar o risco de crédito. Nos modelos de *behavioral scoring*, utilizamos, além das informações consideradas no *application scoring*, informações relativas a créditos anteriores. Por exemplo, se um cliente honrou com pontualidade empréstimos anteriores, essa informação contribuirá de forma positiva no cálculo do score. Outrossim, se no passado o cliente atrasou várias parcelas, seu score será reduzido em razão dessa experiência negativa.

Os modelos de *behavioral scoring*, por incorporar essas preciosas informações sobre experiências de crédito com o cliente, tendem a fornecer modelos com maior poder de discriminação que os modelos de *application scoring*. Isso significa que a probabilidade de perda calculada com um modelo de *behavior scoring* tende a ser mais exata que a estimada com o modelo de *application scoring*. Devemos apenas ter cuidado ao aplicar modelos de *behavior scoring* considerando experiências de crédito em passado remoto, posto que as condições do proponente ou as regras do mercado de crédito podem ter mudado.

## 1.9 MODELOS JULGAMENTAIS E MODELOS QUANTITATIVOS

A fórmula para cálculo do score pode ser obtida de diferentes maneiras:

- Modelos qualitativos ou “julgamentais”
- Modelos quantitativos
  - Modelos de *bureau* (generalistas)
  - Modelos específicos (customizados)

### 1.9.1 MODELOS JULGAMENTAIS

Antes de discutir estes modelos, devemos ressaltar que não estamos nos referindo à análise de crédito “tradicional”, baseada na análise de documentos financeiros e outras informações do solicitante. Referimo-nos a fórmulas para o cálculo de scores cujas variáveis e respectivos pesos são definidos de forma subjetiva.

Para elaboração desses modelos, o analista baseia-se em sua experiência na área de crédito e em modelos quantitativos com os quais tenha trabalhado anteriormente. Existe um risco de modelagem inerente ao processo. A experiência anterior pode não valer para a concessão de crédito sob as novas condições. Características que para determinados

produtos de crédito podem ter efeito negativo no escore, para outros produtos podem ser neutras ou ter efeito positivo. Como já ressaltamos, a subjetividade do processo pode implicar que um mesmo analista, construindo a fórmula em dias diferentes, chegue a valores distintos para os pesos de cada variável e, até mesmo, divirja quanto às variáveis a incluir na fórmula.

Estes modelos são recomendados apenas como última opção, quando não dispomos de dados suficientes para desenvolver um modelo quantitativo. Por exemplo, para novos produtos de crédito ou quando o credor se instala em uma região na qual nunca atuou antes.

A única vantagem em trabalhar com estes modelos reside em permitir a automação das decisões e, portanto, aumentar a produtividade e a consistência das decisões. Mas isto não significa melhorar a qualidade das decisões; elas podem ser consistentemente ruins. Ademais, estes modelos não permitem quantificar a probabilidade de perda. Essa probabilidade é o elemento mais importante na avaliação do risco de crédito.

### 1.9.2 MODELOS QUANTITATIVOS I – MODELOS GENERALISTAS

Algumas empresas de informações comerciais disponibilizam modelos de scoring, corretamente desenvolvidos, porém generalistas. São modelos que estimam o risco de crédito de um indivíduo, PJ ou PF, sem considerar uma operação específica ou um mercado específico do credor. Os modelos generalistas são uma solução interessante no caso de novos produtos ou novos mercados, para os quais o credor não dispõe de base de dados. Permitem a automação da avaliação do risco de crédito e a estimativa da probabilidade de que o tomador de crédito se torne mau pagador no decorrer da operação. No entanto, essa estimativa, conforme explicaremos a seguir, não é, em geral, muito satisfatória.

Para elaboração desses modelos, a definição de mau pagador é, em geral, função da ocorrência de um ou mais desabonos (cheques sem fundo, protestos, pendências financeiras com instituições financeiras ou fornecedores etc.) em determinado período. Isso porque as empresas de informações comerciais não possuem informações completas sobre o desempenho de um cliente em uma operação de crédito. O credor, após solicitar ao *bureau* de crédito as informações relativas ao solicitante, não dá retorno quanto à sua decisão (conceder ou não o crédito) nem presta necessariamente informações relativas ao comportamento do tomador face ao crédito concedido (atrasos, rolagem de dívida, recuperação de crédito etc.). Essa definição de mau pagador pode diferir da definição adotada internamente pelo credor.

Em geral, um modelo genérico não foca o mercado específico do credor nem a operação de crédito a que se destina a avaliação do risco. Ainda que desenvolvidos para um mercado específico, varejo de roupas, por exemplo, clientes de diferentes lojas podem pertencer a classes sociais bastante distintas. O cliente de uma boutique que trabalha com produtos de grife certamente é diferente do cliente de uma cadeia de lojas de roupas populares. Certas características utilizadas no cálculo dos escores podem discriminar

tomadores de uma classe social e não ter o menor efeito em outra classe. Por exemplo, o tipo de residência não fará muita diferença para discriminar clientes de uma classe social alta, em que praticamente todos possuem casa própria. Na classe de renda mais baixa essa variável poderá fazer muita diferença no cálculo do escore. Em um modelo generalista, os solicitantes de uma ou da outra classe serão tratados da mesma forma.

### 1.9.3 MODELOS QUANTITATIVOS II – MODELOS CUSTOMIZADOS

Sem dúvida, estes são os modelos mais adequados para avaliar o risco de crédito. São desenvolvidos levando em consideração o mercado específico do credor, as características das operações a que se destina o modelo e, o que é muito importante, a sua definição de perda na operação. Eles permitem estimar com maior precisão o risco de crédito.

## 1.10 ROTEIRO PARA O DESENVOLVIMENTO DE UM MODELO DE SCORING

Nos capítulos seguintes, vamos apresentar as diferentes etapas no desenvolvimento, na implantação e na gestão de um modelo de escoragem. Os capítulos estão ordenados na sequência de trabalho para alcançar esses objetivos. Em síntese, as principais etapas são:

- planejamento e definições
- identificação das variáveis potenciais
- planejamento e seleção da amostra
- análise e tratamento dos dados
- cálculo da fórmula de escoragem
- análise e validação da fórmula
- ajuste final do modelo

### 1.11 UMA NOTA DE ALERTA

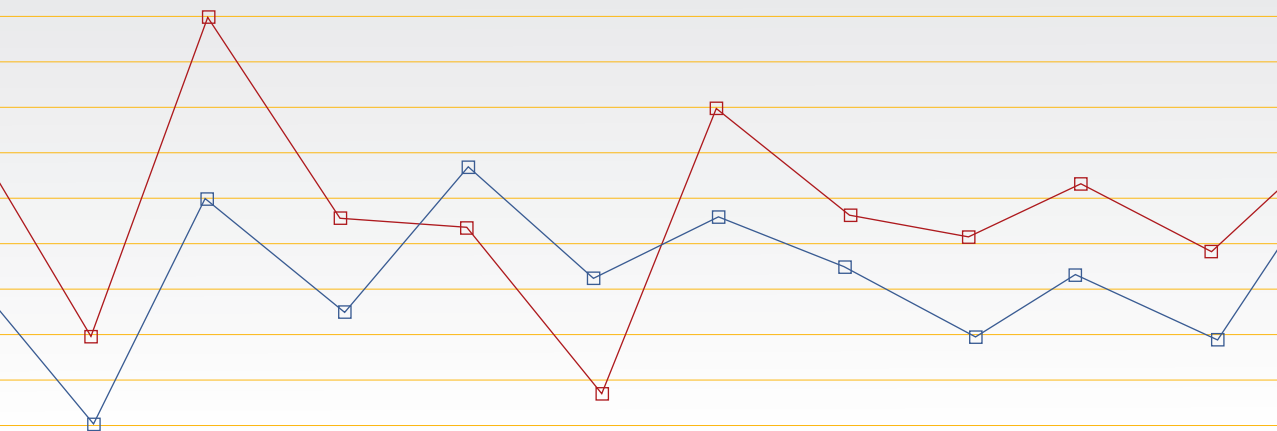
Antes de decidir desenvolver um modelo de *credit scoring* é fundamental avaliar se a medição do risco de crédito com modelos quantitativos é coerente com os objetivos e as estratégias da empresa. Essa questão é extremamente importante, pois já testemunhamos casos de empresas que desenvolveram excelentes modelos, mas nunca os implantaram.

A utilização de modelos de *credit scoring* representa uma transformação radical na cultura de crédito de uma empresa. Como toda mudança cultural, o número de barreiras, geradas consciente ou inconscientemente, são muitas e o comprometimento da alta direção é vital para o sucesso dessa transformação. Se a decisão de desenvolver o modelo não parte da alta direção da empresa, em função de metas e estratégias muito bem-definidas, dificilmente teremos sucesso. O comprometimento da direção do credor

contempla a participação efetiva nas definições dos parâmetros do modelo, o acompanhamento do desenvolvimento em reuniões periódicas, a criação de facilidades e a priorização das tarefas de implantação do sistema de concessão apoiado no *credit scoring* e o acompanhamento regular dos resultados obtidos. Não basta enviar um punhado de analistas a cursos de modelagem, implantar *softwares* avançados, contratar cientistas de dados, programadores etc.

Outro problema é saber se o que a empresa credora realmente necessita é um modelo de *credit scoring*. Nem tudo que requer métodos quantitativos é *credit scoring*! Às vezes, pode ser que o objetivo a ser alcançado não dependa do desenvolvimento desse modelo, mas, sim, de outro tipo de análise. Por exemplo, um caso interessante foi o de uma grande empresa varejista que desenvolveu um modelo de *credit scoring* para identificar regiões com baixa inadimplência. O *credit scoring* pode ter ajudado, mas uma simples análise exploratória de dados bastaria para suprir essa necessidade.

Finalmente, a empresa credora, ainda que contrate os serviços de uma consultoria para o desenvolvimento do projeto, deve envolver-se e participar das diferentes fases do projeto. Não é necessário ser cientista de dados para entender e contribuir com as diferentes etapas do projeto. O conhecimento do produto e do mercado em que o credor atua, da disponibilidade das variáveis que podem ser úteis no desenvolvimento do modelo, de restrições ou limitações impostas pela cultura do credor, entre outros aspectos, são extremamente relevantes para a obtenção de um sistema de concessão de crédito eficaz apoiado no modelo de *credit scoring*.



Este livro, fruto da longa experiência do autor, é um texto obrigatório para todos aqueles que se dedicam à avaliação do risco de crédito, uma vez que:

- descreve minuciosamente as diferentes etapas para desenvolver um modelo de *credit scoring*, desde o planejamento do modelo até a avaliação de sua eficácia;
- discute as principais dificuldades de sua implantação;
- destaca a importância do monitoramento do sistema, apresentando os principais testes estatísticos para esse fim e sugerindo relatórios para acompanhamento do desempenho do *credit scoring*;
- apresenta diferentes técnicas de *machine learning* para calcular o risco de crédito e inclui os respectivos códigos em Python.

A obra pode ser adotada em cursos de análise de crédito, especialmente aqueles que valorizem a mensuração do risco de crédito como elemento central para a tomada de decisões.

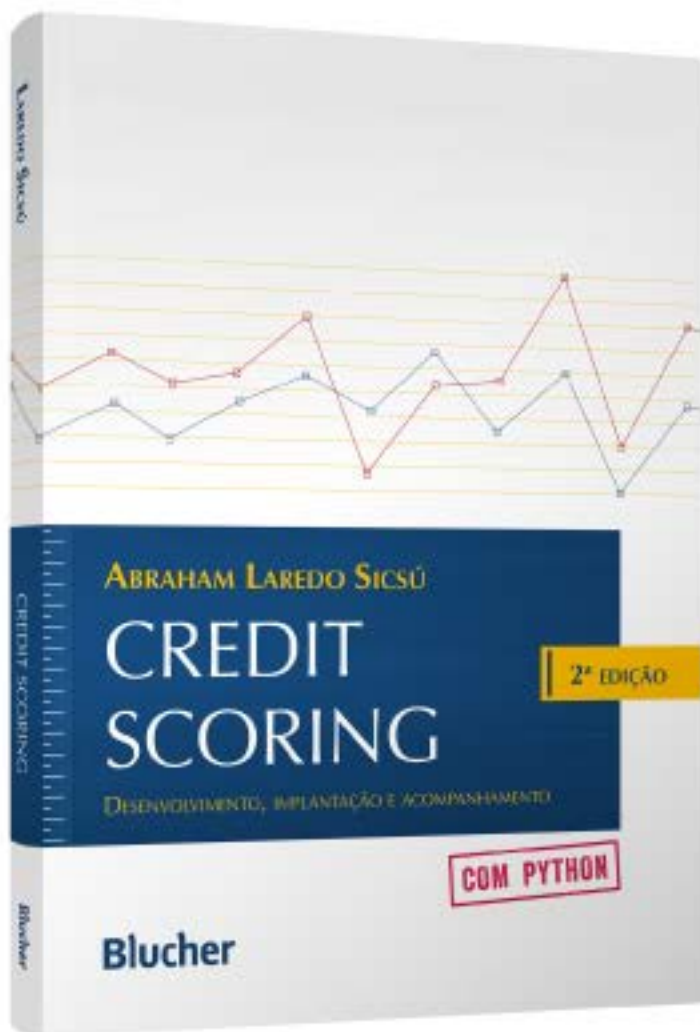


Material de apoio em:



[www.blucher.com.br](http://www.blucher.com.br)

**Blucher**



Clique aqui e:

[VEJA NA LOJA](#)

## Credit scoring

Desenvolvimento, implantação  
e acompanhamento com Python

---

Abraham Laredo Sicsú

ISBN: 9788521231332

Páginas: 192

Formato: 17 x 24 cm

Ano de Publicação: 2026

---