

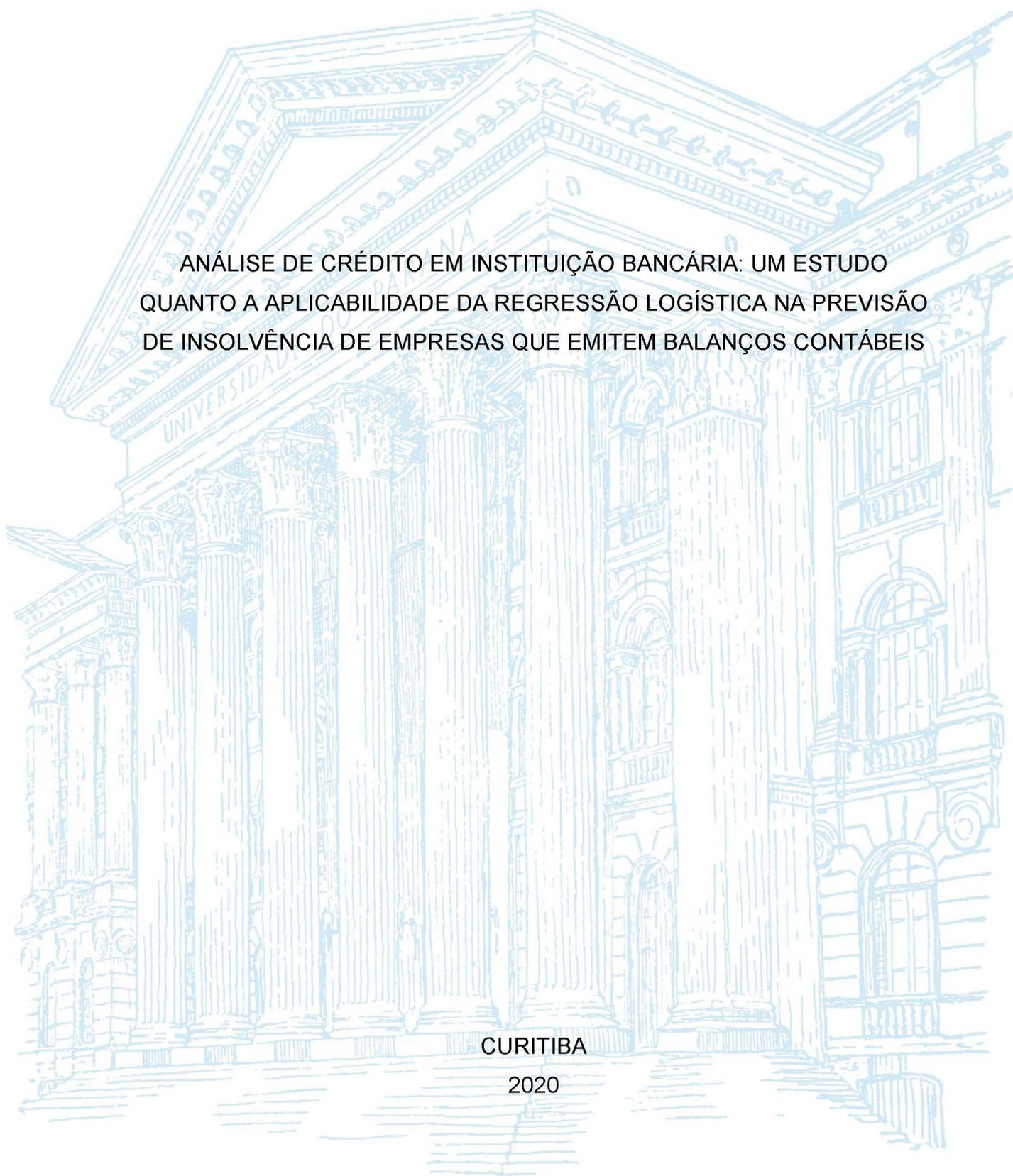
UNIVERSIDADE FEDERAL DO PARANÁ

FABRÍCIO BACH

ANÁLISE DE CRÉDITO EM INSTITUIÇÃO BANCÁRIA: UM ESTUDO
QUANTO A APLICABILIDADE DA REGRESSÃO LOGÍSTICA NA PREVISÃO
DE INSOLVÊNCIA DE EMPRESAS QUE EMITEM BALANÇOS CONTÁBEIS

CURITIBA

2020



FABRÍCIO BACH

ANÁLISE DE CRÉDITO EM INSTITUIÇÃO BANCÁRIA: UM ESTUDO
QUANTO A APLICABILIDADE DA REGRESSÃO LOGÍSTICA NA PREVISÃO
DE INSOLVÊNCIA DE EMPRESAS QUE EMITEM BALANÇOS CONTÁBEIS

Monografia apresentada ao curso de Pós-Graduação em Contabilidade e Finanças, Setor de Ciências Sociais Aplicadas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Especialista em Contabilidade e Finanças.

Orientador: Prof. Dr. Claudio Marcelo Edwards Barros.

CURITIBA

2020

TERMO DE APROVAÇÃO

FABRÍCIO BACH

ANÁLISE DE CRÉDITO EM INSTITUIÇÃO BANCÁRIA: UM ESTUDO QUANTO A APLICABILIDADE DA REGRESSÃO LOGÍSTICA NA PREVISÃO DE INSOLVÊNCIA DE EMPRESAS QUE EMITEM BALANÇOS CONTÁBEIS

Monografia apresentada ao curso de Pós-Graduação em Contabilidade e Finanças, Setor de Ciências Sociais Aplicadas, Universidade Federal do Paraná, como requisito parcial à obtenção do título de Especialista em Contabilidade e Finanças.

Prof. Dr. Claudio Marcelo Edwards Barros

Orientador – Departamento de Contabilidade e Finanças, UFPR

Curitiba, 29 de julho de 2020.

AGRADECIMENTOS

Agradeço primeiramente a instituição Universidade Federal do Paraná, precisamente o curso de Administração por proporcionar toda estrutura necessária ao desenvolvimento do meu apreço por finanças durante minha formação como bacharel.

Ao curso de Pós-graduação em Contabilidade e Finanças pela gama e qualidade dos professores, estrutura e organização.

Ao professor orientador Dr. Claudio Marcelo Edwards Barros por engrandecer a importância da estatística no âmbito das finanças, pela motivação das aulas e ânimo pelo conteúdo ministrado.

Aos amigos que se formaram durante o curso, sendo eles: Danielli, Fernanda, Kelaine, Leonardo e Michele, tornando os sábados mais leves, produtivos e propiciando diversas aplicações do conteúdo em suas respectivas áreas de atuação no mercado.

RESUMO

Este estudo teve como principal objetivo verificar a aplicabilidade da regressão logística no auxílio a análise de crédito de empresas que emitem balanços contábeis, com base em dados de uma grande instituição bancária. Os modelos do estudo não tem o objetivo de prever a insolvência das empresas, entretanto visam compreender sua aplicabilidade a uma situação real e possível auxílio aos analistas de crédito. A base de dados utilizada corresponde a dados contábeis de 2016 a 2018, bem como dados comportamentais das empresas ao final de 2018, a partir de uma amostra de 126 empresas do segmento de transporte de cargas. Foram selecionadas as variáveis independentes e divididas entre variáveis categóricas e contínuas, com o objetivo de auxiliar na identificação da probabilidade de as empresas apresentarem vencimentos por mais de 90 dias com a instituição bancária nos 12 meses subsequentes. A técnica de regressão logística foi aplicada paralelamente, junto a dois softwares estatísticos para base comparativa. Os resultados da aplicação da regressão logística contribuíram para uma avaliação hierárquica sobre quais variáveis possuem maior influência quanto a probabilidade de insolvência das empresas analisadas, qual sua correlação para a probabilidade de evidenciar vencimentos e quais variáveis podem estatisticamente contribuir para tal previsão.

Palavras-chave: Análise de Crédito. Regressão Logística. Previsão de Insolvência.

ABSTRACT

This study had as main objective to verify the applicability of the logistic regression in the aid to credit analysis of companies that issue balance sheets, based on data from a large banking institution. The models in the study are not intended to predict the insolvency of companies, however, they aim to understand its applicability to a real situation and possible assistance to credit analysts. The database used corresponds to accounting data from 2016 to 2018, as well as behavioral data of companies at the end of 2018, from a sample of 126 companies in the cargo transportation segment. Independent variables were selected and divided between categorical and continuous variables in order to help identify the likelihood that companies will have debts over 90 days with the bank in the subsequent 12 months. The logistic regression technique was applied in parallel with two statistical softwares for comparative basis. The results of the application of logistic regression contributed to a hierarchical assessment of which variables have the greatest influence on the probability of insolvency of the analyzed companies, what their correlation is with the probability of evidencing debts and which variables can statistically contribute to such forecast.

Keywords: Credit Analysis. Logistic Regression. Insolvency Forecast.

LISTA DE FIGURAS

FIGURA 1 – COMPOSIÇÃO DA AMOSTRA.....	18
FIGURA 2 – MODELO FINAL APÓS MÉTODO <i>BACKWARD STEPWISE</i> PELO IBM SPSS®.....	24
FIGURA 3 – CASOS CORRETAMENTE PREVISTOS <i>SOFTWARE</i> IBM SPSS® ...	24

LISTA DE TABELAS

TABELA 1 – VARIÁVEIS INDEPENDENTES UTILIZADAS NA PESQUISA.....	19
TABELA 2 – PARÂMETROS DAS VARIÁVEIS CATEGÓRICAS	20
TABELA 3 – RESULTADOS DO MODELO COMPLETO PELO SOFTWARE IBM SPSS®.....	22
TABELA 4 – SEQUÊNCIA DO MÉTODO <i>BACKWARD STEPWISE</i> PELO IBM SPSS®.....	23
TABELA 5 – RESULTADOS DO MODELO COMPLETO PELO SOFTWARE GRETl®	25
TABELA 6 – SEQUÊNCIA DO MÉTODO <i>STEPWISE</i> PELO GRETL®.....	26
TABELA 7 – MODELO FINAL APÓS MÉTODO <i>STEPWISE</i> PELO GRETL®.....	27

SUMÁRIO

1	INTRODUÇÃO	11
1.1	PROBLEMAS DE PESQUISA	11
1.2	OBJETIVOS	11
1.2.1	Objetivo geral	12
1.2.2	Objetivos específicos.....	12
1.3	JUSTIFICATIVA	12
2	REVISÃO DE LITERATURA	14
2.1	PREVISÃO DE INSOLVÊNCIA	14
2.2	RISCO DE CRÉDITO	15
2.3	REGRESSÃO LOGÍSTICA.....	15
3	METODOLOGIA	17
3.1	UNIVERSO DA PESQUISA	17
3.2	AMOSTRA E COLETA DE DADOS	17
3.3	AS VARIÁVEIS.....	18
3.3.1	Variável dependente.....	18
3.3.2	Variáveis independentes	18
4	APRESENTAÇÃO DOS RESULTADOS	22
4.1	APLICAÇÃO 1: PARÂMETROS APLICADOS AO SOFTWARE IBM SPSS®.....	22
4.1.1	Modelo completo antes da aplicação do método <i>Backward Stepwise</i>	22
4.1.2	Modelo após a aplicação do método <i>Backward Stepwise</i>	23
4.2	APLICAÇÃO 2: PARÂMETROS APLICADOS AO SOFTWARE GRETL®.....	24
4.2.1	Modelo completo antes da aplicação do método <i>Stepwise</i>	25
4.2.2	Modelo após a aplicação do método <i>Stepwise</i>	26
4.3	INTERPRETAÇÃO DOS RESULTADOS	27
5	CONSIDERAÇÕES FINAIS	29
5.1	RECOMENDAÇÕES PARA TRABALHOS FUTUROS	29
	REFERÊNCIAS	31
	APÊNDICE 1 – MÉTODO STEPWISE COMPLETO ATRAVÉS DO	34
	SOFTWARE IBM SPSS® STATISTICS SUBSCRIPTION.....	34
	APÊNDICE 2 – MODELO COMPLETO ATRAVÉS DO SOFTWARE	

GRETL®.....	38
APÊNDICE 3 – MODELO FINAL APÓS MÉTODO <i>STEPWISE</i> PELO	
GRETL®.....	39

1 INTRODUÇÃO

1.1 PROBLEMAS DE PESQUISA

Cada vez mais a modelagem de dados vem sendo aplicada à análise de crédito ou *credit scoring* nas instituições financeiras, as quais tem recorrido cada vez menos à análises e julgamentos subjetivos conforme observado por Cintra et al. (2007 apud Pires et al., 2014), por meio de modelos preditivos estatísticos, reduzindo assim a necessidade de um maior corpo de funcionários para realização das análises manuais pela mesa de crédito, aumentando a velocidade de resposta ao cliente, acelerando os ganhos financeiros da instituição e ao mesmo tempo oferecendo aos clientes um melhor atendimento, característica relevante em um período de ascensão das *fintechs* no mercado, que prezam tanto pela experiência do cliente e vem alterando o conceito de instituição financeira como era comumente conhecido, além de otimizar seu principal objetivo, a redução do risco de crédito.

Todavia, tanto na teoria quanto na prática, observa-se ampla abordagem do tema em relação a variáveis comportamentais relacionadas principalmente às pessoas físicas e pequenas empresas, através de variáveis como histórico de pagamentos, restrições financeiras, faturamento médio e endividamento, para a geração e aplicação dos modelos de *credit scoring*. Tais modelos se utilizam de regressão linear e regressão logística.

Já em relação às empresas de grande porte, as quais demandam maior nível de profissionalização em suas demonstrações contábeis, na geração de balanços patrimoniais e demonstrações de resultado, conseqüentemente empresas que logram de maior apetite ao risco e tomam um maior volume de recursos junto aos bancos, o que se observa são aplicações de modelos provenientes da análise discriminante conforme menciona Pinheiro et al. (2007).

1.2 OBJETIVOS

A questão a ser observada trata da possibilidade de desenvolvimento de um modelo de previsão de insolvência que agregue tanto dados contábeis quanto dados comportamentais, como auxílio à análise de crédito de forma automatizada para empresas ou grupos econômicos de grande porte. O estudo será adequado ao

processo de análise de crédito do Departamento de Crédito Alto Valor, localizado em Curitiba-Paraná, pertencente à grande instituição bancária de nível nacional.

Qual a aplicabilidade da regressão logística na criação de um modelo de previsão de insolvência que combine dados contábeis e comportamentais para automatizar a análise de crédito – acerca de empresas de grande porte - no departamento de crédito de uma instituição bancária?

1.2.1 Objetivo geral

Analisar a possibilidade de automatização da análise de crédito de uma instituição bancária, em relação a empresas de grande porte, através de regressão logística, relacionando dados contábeis e comportamentais. Os modelos do estudo não tem o objetivo de prever a insolvência das empresas, entretanto visam compreender sua aplicabilidade a uma situação real e possível auxílio aos analistas de crédito.

1.2.2 Objetivos específicos

- Classificar as rubricas do balanço patrimonial e da demonstração de resultados consideradas mais relevantes para análise de crédito de empresas de grande porte;
- Classificar as variáveis comportamentais consideradas mais importantes para análise de crédito de empresas de grande porte;
- Realizar a interação entre dados contábeis e comportamentais em um mesmo modelo executado em *software* estatístico;
- Analisar os relatórios gerados pelo modelo e verificar a possibilidade de aplicação na instituição bancária.

1.3 JUSTIFICATIVA

Saber se um cliente provavelmente honrará seus compromissos é uma informação imprescindível na hora de tomar uma decisão com vistas à concessão de crédito. Com isso, pode-se demonstrar que as instituições financeiras poderiam ter um acréscimo nos lucros se, na concessão de crédito, os critérios fossem mais

precisos. De posse da classificação fornecida por um modelo de previsão de risco de crédito, a empresa pode ter um diagnóstico preliminar do provável comportamento do novo cliente, aprovando ou não a concessão do crédito (VASCONCELOS, 2004).

Através do estudo aplicado a um contexto real vivenciado por instituição bancária nos dias de hoje, o projeto colabora no aperfeiçoamento à tomada de decisão de crédito sobre empresas de grande porte, contribuindo para que a análise ocorra de forma automatizada para clientes com uma maior exposição em riscos na instituição.

Essa possibilidade de elevação de alçada para concessões de crédito automatizadas via sistema, através do modelo a ser analisado neste projeto, geraria ganhos para a instituição financeira quanto a maior velocidade de resposta às propostas de crédito, portanto elevando a qualidade do atendimento aos clientes, além de que o tempo poupado pela análise automatizada pode ser redistribuído aos analistas da mesa de crédito, direcionando os esforços para propostas mais complexas e de valores superiores, que consequentemente geram maior retorno financeiro a instituição e, assim, valoriza-se a capacidade técnica dos analistas.

Em um mercado de ascensão das *fintechs*, a instituição tem a possibilidade de manter sua vantagem competitiva neste novo cenário, e os clientes são beneficiados com maior qualidade e rapidez no atendimento ao crédito, fomentando a economia.

O projeto permite conciliar estudos de autores tanto da análise discriminante quanto da regressão logística, analisando sua aplicabilidade a um contexto recente e tangível do departamento de crédito de uma grande instituição bancária, localizado em Curitiba-Paraná, explorando a contemporaneidade do tema.

2 REVISÃO DE LITERATURA

2.1 PREVISÃO DE INSOLVÊNCIA

O tema previsão de insolvência passou a apresentar estudos relevantes somente a partir de 1930, no período pós crise, evidenciado pela alta quantidade de falências no período. Entretanto, a aplicação da estatística só teve início a partir dos anos 60, influenciando diretamente na maior segurança e confiança acerca dos modelos. Segundo Thomas (2000), o crescimento do uso de modelos na década de 60 transformou os negócios no mercado americano.

Segundo Martins (2003), o primeiro estudo sobre previsão de insolvência foi o trabalho sobre a *Comparison of the Ratios of the Successful Industrial Enterprises with those of Failed Companies*, elaborado no ano de 1932 por Fitzpatrick, por meio de suas respectivas demonstrações financeiras.

Segundo Pinheiro et al. (2007), entre os modelos de previsão de insolvência destacam-se na literatura os modelos, cronologicamente, desenvolvidos por Elizabetsky (1976), Kanitz (1978), Altman, Baidya e Dias (1979), Silva (1982), Sanvicente e Minardi (1998) e Scarpel (2000). Dos modelos acima, seis deles baseiam-se na análise discriminante e um modelo, o de Scarpel (2000), baseia-se em regressão logística. Segundo Bouroche e Saporta (1980), “O método estatístico de análise discriminante considera um conjunto de indivíduos no qual se observa um caráter qualitativo que toma q ($q \geq 2$) modalidades.

Porém, cabe destacar em relação à análise discriminante algumas restrições, conforme cita Brito e Assaf Neto (2008) “A análise discriminante linear está baseada em pressupostos bastante restritivos, como a normalidade das variáveis independentes e a igualdade das matrizes de variância-covariância dos grupos de interesse. Essas suposições podem não ser válidas em muitas situações práticas de análise de risco de crédito, principalmente quando há variáveis independentes de natureza não métrica.”.

Para este estudo, será utilizado o conceito de insolvência de Lev (1978), o qual cita que o estado de insolvência de uma empresa ocorre quando da incapacidade de pagar suas obrigações financeiras na data de seu vencimento, bem como quando seus ativos forem inferiores ao valor dos seus passivos.

2.2 RISCO DE CRÉDITO

Para que se tomem boas decisões na concessão de crédito, é necessário que se avalie de forma adequada o risco de inadimplência dos respectivos proponentes. Silva (2003) descreve o risco de crédito como a probabilidade de que o recebimento não ocorra.

Tal avaliação foi cada vez mais aprimorada a partir dos modelos aplicados aos segmentos de atacado, ou seja, principalmente a concessão de crédito a pessoas físicas. Dessa forma, grande parte dos estudos referentes ao tema se baseiam em *behavior score* e variáveis como gênero, idade, estado civil, nível de escolaridade, tempo em que é correntista do banco, emprego e demais variáveis conforme pode-se observar nos trabalhos de Yap et al. (2011), Bijak e Thomas (2012), Mavri et al. (2008), Abdou (2009).

Porém, quando os proponentes do risco se tratam de empresas, Blatt (1999) menciona que não é possível definir a maneira correta de estabelecer uma linha de crédito, pois cada analista possui sua maneira de analisar as empresas, tornando difícil uma definição exata ao assunto.

Diante disso, um diferencial em relação a análise de empresas é a análise de balanços. Matarazzo (2010) informa que as técnicas de análise de balanços possibilitam grande número de informações sobre a empresa.

O cenário acima evidencia a importância do julgamento por parte dos analistas baseando-se em sua experiência em relação a concessão de crédito, e também o quanto a análise de balanços contribui, ofertando uma gama de informações de significativa contribuição para auxiliar a decisão do crédito.

A análise financeira dos balanços também se baseia no uso de índices-padrão. Segundo Gitman (2004), os principais indicadores são os de liquidez, rentabilidade, endividamento, atividade e valor de mercado.

2.3 REGRESSÃO LOGÍSTICA

Regressão logística é uma técnica de análise multivariada e trata da “[...] obtenção da probabilidade de que uma observação pertença a um conjunto determinado, em função do comportamento das variáveis independentes.” (GIMENES e OPAZO, 2003). Nela, a variável dependente é binária, ou seja, através

do uso de variáveis *dummy*, assumindo o valor 0 para determinada característica ou 1 para outra característica. Já as variáveis independentes podem ser tanto contínuas como categóricas. Na regressão logística, os parâmetros do modelo são estimados utilizando-se o método de estimação da máxima verossimilhança (HOSMER; LEMESHOW, 1989). Na estimação dos coeficientes da função logística, esta técnica procura maximizar a verossimilhança de que um evento ocorra (HAIR JR et al., 2009).

Segundo diversos autores como Rosa (2000), Ohtoshi (2003), Crook et al. (2007), Bijak e Thomas (2012), Herszkowicz, (2000), a técnica de regressão logística está entre as mais utilizadas no mercado para modelos de risco de crédito.

Essa técnica trabalha basicamente com os conceitos de que a variável dependente, variável dicotômica que é afetada por um conjunto de fatores como por exemplo: financeiros, sócio demográficos, entre outros, é por si só binária, assumindo um valor de 0 para uma determinada característica e valor de 1 para outra característica, por exemplo mau pagador sendo 0 e 1 para bom pagador. “ A regressão logística se enquadra na classe de métodos estatísticos multivariados de dependência, pois relaciona um conjunto de variáveis independentes com uma variável dependente categórica.” (HAIR et al, 2009). Brito e Assaf Neto, (2008) menciona que o fato de os resultados da regressão logística poderem ser interpretados em termos de probabilidade, favorece sua utilização.

O modelo da regressão logística se dá por:

$$Z = \alpha + \beta_1.X_1 + \beta_2.X_2 + \dots + \beta_{28}.X_{28}$$

Onde Z refere-se ao logaritmo natural da chance de ocorrência de um evento, logo para Z temos $\ln(p/1-p)$, ou seja, a chance de ocorrência do evento 1.

E a probabilidade de ocorrência do evento de interesse pode ser obtida por:

$$p = \frac{e^Z}{1 + e^Z}$$

Essa função descreve a probabilidade da variável dependente ser igual a 1, dado o comportamento das variáveis independentes X.

3 METODOLOGIA

A seguir apresentam-se informações quanto ao universo da pesquisa, a coleta de dados qualitativos através de processo empírico observação-participante, a coleta documental de dados quantitativos primários, a definição das variáveis independentes e dependente a serem utilizadas, e os procedimentos de regressão logística executados através de dois *softwares* estatísticos distintos sendo IBM SPSS® *Statistics Subscription* e Gretl®, utilizados paralelamente como base comparativa. A escolha dos *softwares* se deu pela vasta aplicação e notória utilização do sistema IBM SPSS® *Statistics Subscription* no mercado, bem como pelo *layout* amigável e plataforma aberta do Gretl®.

3.1 UNIVERSO DA PESQUISA

Para a realização do estudo foram coletados dados históricos datados de 2016 a 2018 referentes à empresas do segmento de transporte de cargas, fornecidos por instituição bancária, que concordou em fornecer estes dados desde que mantida a confidencialidade em relação à própria instituição, bem como as empresas pesquisadas. Portanto, no desenvolvimento do estudo, não serão divulgados nomes das empresas ou da instituição bancária.

3.2 AMOSTRA E COLETA DE DADOS

A amostra de empresas utilizadas na atual pesquisa refere-se a um único segmento, transporte de cargas, com o objetivo de blindar esse estudo das complexas variações que existem para cada segmento de atividade no âmbito da análise de crédito. Essa afirmação foi definida com base em processo empírico observação-participante junto ao departamento de crédito da instituição bancária, o qual evidencia grande diferença de características e análise entre os diferentes ramos de atividade.

Foram coletados dados referentes a 126 empresas do segmento de transporte de cargas. Estes dados foram subdivididos em dois grupos sendo: dados advindos dos balanços e demonstrativos contábeis de 2016 a 2018 das empresas, e

dados advindos do histórico comportamental destas empresas junto a instituição bancária até o fim de 2018.

3.3 AS VARIÁVEIS

3.3.1 Variável dependente

A variável dependente para a realização da regressão logística, conforme inicialmente definido neste estudo segundo Lev (1978), o qual cita que “[...] o estado de insolvência de uma empresa ocorre quando da incapacidade de pagar suas obrigações financeiras na data de seu vencimento.”, portanto foi coletado o histórico de vencimentos dessas 126 empresas conforme FIGURA 1, no período de 12 meses, com início em janeiro de 2019 e término em dezembro de 2019, sendo 1 quando a empresa apresentou vencimentos por mais de 90 dias com o banco nos últimos 12 meses e 0 quando não apresentou vencidos por mais de 90 dias.

FIGURA 1 – COMPOSIÇÃO DA AMOSTRA

	Observado	Previsto		Porcentagem correta	
		Historico Venc nos ult 12 meses			
		0	1		
Etapa 0	Historico Venc nos ult 12 meses	0	98	0	100,0
		1	28	0	,0
Porcentagem global					77,8

FONTE: Dados do IBM SPSS® Statistics

3.3.2 Variáveis independentes

Em relação aos dados advindos do histórico comportamental das empresas junto a instituição bancária, através do procedimento de observação-participante, foram definidas as variáveis tomadas como mais relevantes pelos analistas de crédito da instituição e que não possuem relação com os balanços contábeis

emitidos pelas empresas. São variáveis utilizadas para análise de crédito de empresas que possuem uma exposição abaixo de cinco milhões de reais em crédito liberado junto a instituição bancária em estudo, em que não há a obrigatoriedade de demonstrativos contábeis segundo as diretrizes da instituição.

As variáveis pré-selecionadas foram: Tempo de Atividade, Tempo de Relacionamento com a Instituição Bancária, Endividamento da Empresa Junto ao Banco Central e se possui Restritivo Grau 4¹ junto a instituição bancária.

Já em relação aos dados advindos dos balanços contábeis, a dificuldade de obter os dados foi um dos fatores limitadores na seleção dos indicadores para o estudo. Assim, o critério para definição das variáveis usadas no estudo foi sua disponibilidade para as empresas da amostra.

As variáveis pré-selecionadas foram: Índice de Endividamento, Patrimônio Líquido, Lucro Líquido, Crescimento da Receita, Geração Média de Caixa nos Últimos 3 anos, Margem Líquida Média nos Últimos 3 anos, Crescimento da Margem EBITDA, Margem EBITDA em 2018, Liquidez Seca, Liquidez Corrente, Liquidez Geral.

No tratamento das variáveis, houve inclusão da variável Alavancagem, compreendendo o Endividamento da Empresa Junto ao Banco Central, dividido pela Receita Média Mensal da empresa, índice este estruturado a partir da importância atribuída a este indicador pelos analistas de crédito conforme observação-participante.

A TABELA 1 a seguir apresenta as variáveis coletadas e separadas entre contínuas e categóricas.

TABELA 1 – VARIÁVEIS INDEPENDENTES UTILIZADAS NA PESQUISA

(continua)

Dados Contábeis (Balanço e DRE)		Dados Históricos	
Contínua	Categórica	Contínua	Categórica
• Índice de Endividamento	• Patrimônio Líquido	• Tempo de Atividade	• Restritivo Grau 4

¹ Restritivo Grau 4: a instituição bancária aplica um restritivo grau 4 a empresas que, nos últimos 5 anos, já apresentaram vencidos acima de 90 dias com a organização.

<ul style="list-style-type: none"> • Liquidez Seca • Liquidez Corrente • Liquidez Geral 	<ul style="list-style-type: none"> • Lucro Líquido • Média Margem Líquida Últimos 3 Anos • Geração Média de Caixa Últimos 3 Anos • Crescimento da Receita • Crescimento Margem EBITDA • Margem EBITDA em 2018 	<ul style="list-style-type: none"> • Tempo de Relacionamento • Alavancagem 	
--	---	--	--

FONTE: Elaborado pelo autor.

Em relação as variáveis categóricas, foram definidos os seguintes parâmetros para geração das variáveis *dummy* conforme TABELA 2:

TABELA 2 – PARÂMETROS DAS VARIÁVEIS CATEGÓRICAS

(continua)

Variável Categórica	Parâmetro
Patrimônio Líquido	1 para empresas com Patrimônio Líquido positivo, 0 para negativo
Lucro Líquido	1 para empresas que apresentaram lucro líquido positivo em média nos últimos três anos e 0 para empresas que evidenciaram prejuízo líquido médio nesse mesmo período
Média Margem Líquida Últimos 3 Anos	1 para margem média positiva e 0 para negativa
Geração Média de Caixa Últimos 3 Anos	1 para geração média positiva e 0 para negativa
Crescimento de Receita	1 para empresas que aumentaram a Receita de 2016 para 2018, 0 para empresas que reduziram a receita nesse período
Crescimento Margem EBITDA	1 para empresas que apresentaram crescimento na margem EBITDA de 2016 para 2018 e 0 para que apresentaram redução nesse mesmo período

Margem EBITDA em 2018	1 para empresas que terminaram 2018 com margem EBITDA positiva e 0 para negativa
Restritivo Grau 4	1 para empresas que já possuíam restritivo grau 4 junto ao banco até 12/2018, 0 para empresas que não possuíam restritivo grau 4 nesse período

FONTE: Elaborado pelo autor.

4 APRESENTAÇÃO DOS RESULTADOS

Para a aplicação do procedimento de regressão logística foi utilizado o método matemático *backward stepwise* função “*Backward: LR*” no IBM SPSS® *Statistics Subscription*, o qual inicia-se com um modelo que contém todas as variáveis consideradas, posteriormente são removidas as variáveis menos significantes estatisticamente, uma após a outra, gerando-se um novo modelo a cada variável removida. O mesmo método foi aplicado junto ao *software* Gretl® para comparação dos resultados.

Portanto, inicialmente serão incluídas todas as variáveis para elaboração do modelo. Depois, serão gradativamente excluídas, por nível de significância, restando ao final apenas as variáveis mais estatisticamente significantes.

4.1 APLICAÇÃO 1: PARÂMETROS APLICADOS AO SOFTWARE IBM SPSS®

4.1.1 Modelo completo antes da aplicação do método *Backward Stepwise*

A regressão logística binária, através do método *Backward Stepwise*, foi aplicada ao *software* IBM SPSS® *Statistics Subscription* utilizando-se os dados coletados das 126 transportadoras e as 15 variáveis, portanto na primeira etapa considera-se um modelo com todas as variáveis conforme a TABELA 3 abaixo:

TABELA 3 – RESULTADOS DO MODELO COMPLETO PELO SOFTWARE IBM SPSS®

(continua)

Variável	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. p/ EXP(B)	
							Inferior	Superior
Índice de endividamento	1,121	,516	4,716	1	,030	3,068	1,115	8,441
Geração média de caixa últ. 3 anos	-,365	,642	,324	1	,569	,694	,197	2,441
Margem EBITDA 2018	1,150	,997	1,331	1	,249	3,160	,448	22,303
Liquidez seca	,115	,827	,019	1	,890	1,121	,222	5,670
Liquidez corrente	-,242	,806	,090	1	,764	,785	,162	3,811
Liquidez geral	,065	,418	,024	1	,877	1,067	,470	2,420
Tempo de atividade	-,027	,033	,645	1	,422	,974	,912	1,039
Tempo de relacionamento	,029	,039	,566	1	,452	1,030	,954	1,112
Alavancagem	,025	,042	,344	1	,557	1,025	,944	1,113
Margem líquida últ. 3 anos	-,009	,755	,000	1	,991	,991	,226	4,356
Patrimônio líquido	,565	,959	,347	1	,556	1,760	,269	11,531
Crescimento da receita	,369	,815	,205	1	,651	1,446	,293	7,142

Crescimento margem EBITDA	-,729	,566	1,662	1	,197	,482	,159	1,462
Restritivo grau 4	2,798	,746	14,063	1	,000	16,420	3,803	70,888
Constante	-3,684	1,527	5,825	1	,016	,025		

FONTE: Elaborado pelo autor a partir das saídas do IBM SPSS® *Statistics Subscription*

O indicador “B” da tabela aponta a constante para cada variável independente, indicando a correlação positiva ou negativa com a variável dependente. Já o indicador “Sig.” ou valor-p, evidencia o grau de significância da variável. Para interpretação do valor-p, quanto menor o indicador, maior a significância da variável independente em relação a variável dependente.

Conforme observa-se na TABELA 3, grande parte das variáveis não é estatisticamente significante. Entretanto, a visualização do modelo completo nos permite, em primeiro momento, elencar qual a correlação de cada variável independente frente a nossa variável dependente. No modelo completo acima, observa-se que a variável com maior significância é “Restritivo Grau 4”, e a variável com menor significância é “Margem Líquida nos Últimos 3 anos”.

4.1.2 Modelo após a aplicação do método *Backward Stepwise*

O método *Backward Stepwise* realiza a gradativa exclusão das variáveis menos significantes no modelo, e a sequência se deu conforme TABELA 4 abaixo:

TABELA 4 – SEQUÊNCIA DO MÉTODO *BACKWARD STEPWISE* PELO IBM SPSS®

Variável removida no passo 2	Margem líquida Ult. 3 anos
Variável removida no passo 3	Liquidez geral
Variável removida no passo 4	Liquidez seca
Variável removida no passo 5	Crescimento da receita
Variável removida no passo 6	Alavancagem
Variável removida no passo 7	Geração média de caixa Ult. 3 anos
Variável removida no passo 8	Patrimônio líquido
Variável removida no passo 9	Liquidez corrente
Variável removida no passo 10	Tempo de relacionamento
Variável removida no passo 11	Tempo de atividade
Variável removida no passo 12	Crescimento margem EBITDA
Variável removida no passo 13	Margem EBITDA 2018

FONTE: Elaborado pelo autor a partir das saídas do IBM SPSS® *Statistics Subscription*

Restando, ao final do modelo, as duas variáveis com maior significância conforme evidenciado na FIGURA 2 abaixo:

FIGURA 2 – MODELO FINAL APÓS MÉTODO *BACKWARD STEPWISE* PELO IBM SPSS®

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para EXP(B)	
								Inferior	Superior
Etapa 13	Índice de Endividamento	,600	,317	3,571	1	,059	1,822	,978	3,394
	Rest grau 4	2,634	,654	16,218	1	,000	13,927	3,865	50,185
	Constante	-2,302	,406	32,097	1	,000	,100		

FONTE: Dados do IBM SPSS® *Statistics Subscription*

A FIGURA 3 abaixo demonstra a porcentagem de casos corretamente previstos pelo modelo final:

FIGURA 3 – CASOS CORRETAMENTE PREVISTOS SOFTWARE IBM SPSS®

	Observado		Previsto		Porcentagem correta	
			Historico Venc nos ult 12 ...	0		1
Etapa 1	Historico Venc nos ult 12 ...	0		94	4	95,9
Etapa 13	Historico Venc nos ult 12 ...	1	15		13	46,4
	Porcentagem global					84,1

FONTE: Dados do IBM SPSS® *Statistics Subscription*

4.2 APLICAÇÃO 2: PARÂMETROS APLICADOS AO SOFTWARE GRETl®

Paralelamente ao *software* IBM SPSS® *Statistics Subscription*, aplicamos também a regressão logística, através do método *stepwise*, ao *software* estatístico Gretl® com o objetivo de realizarmos uma comparação entre ambos os *softwares*.

O *software* Gretl® permite uma maior interação com os dados uma vez que o método *stepwise* é aplicado manualmente através da análise da significância de cada variável independente para o modelo. A exclusão das variáveis independentes menos significantes é realizada pelo próprio usuário do *software*, o qual retira gradualmente do modelo as variáveis com maior valor-p, gerando um novo modelo subsequente.

4.2.1 Modelo completo antes da aplicação do método *Stepwise*

A TABELA 5 abaixo demonstra o modelo completo aplicado ao *software* Gretl®:

TABELA 5 – RESULTADOS DO MODELO COMPLETO PELO SOFTWARE GRETl®

Variável	Coefficiente	Erro Padrão	Z	p-valor
Constante	-3,68448	1,36681	-2,696	0,0070 ***
Índice de endividamento	1,12119	0,559401	2,004	0,0450 **
Geração média de caixa	-0,365480	0,641605	-0,5696	0,5689
Margem EBITDA 2018	1,15047	1,13818	1,011	0,3121
Liquidez seca	0,114592	0,516224	0,2220	0,8243
Liquidez corrente	-0,242196	0,471543	-0,5136	0,6075
Liquidez geral	0,0646630	0,210996	0,3065	0,7592
Tempo de atividade	-0,0267625	0,0296075	-0,9039	0,3660
Tempo de relacionamento	0,0294179	0,0363509	0,8093	0,4184
Alavancagem	0,0246391	0,0262508	0,9386	0,3479
Margem líquida últ. 3 anos	-0,00856462	0,962873	-0,008895	0,9929
Patrimônio líquido	0,565111	1,24175	0,4551	0,6490
Crescimento da receita	0,368771	0,721792	0,5109	0,6094
Crescimento margem EBITDA	-0,729233	0,536940	-1,358	0,1744
Restritivo grau 4	2,79848	0,708417	3,950	7,80e-05 ***

Média var. dependente	0,222222	D.P. var. dependente	0,417399
R-quadrado de McFadden	0,258066	R-quadrado ajustado	0,033323
Log de verossimilhança	-49,51888	Critério de Akaike	129,0378
Critério de Schwarz	171,5820	Critério Hannan-Quinn	146,3221

Número de casos 'corretamente previstos'	107 (84,9%)
F (beta'x) na média das variáveis independentes	0,163
Teste de razão de verossimilhança: Qui-quadrado (2)	34,4482

FONTE: Elaborado pelo autor a partir das saídas do *software* Gretl®

Para tornar mais apropriada a avaliação da significância dos coeficientes, foram considerados os erros robustos para que os valores-p dos coeficientes “B” dos regressores levassem em conta os efeitos da heterocedasticidade e da autocorrelação dos resíduos. Por essa razão, a variável “Lucro Líquido” foi omitida, assim como foi omitida pelo *software* IBM SPSS®.

Do mesmo modo que o modelo completo evidenciado através do *software* IBM SPSS® *Statistics Subscription*, o modelo gerado pelo *software* Gretl® mantém indicadores muito próximos. Sendo a variável com maior significância “Restritivo Grau 4”, e a variável com menor significância “Margem Líquida nos Últimos 3 Anos”.

4.2.2 Modelo após a aplicação do método *Stepwise*

A sequência de exclusão das variáveis menos significativas através do *software* Gretl® se deu da seguinte forma conforme TABELA 6 abaixo:

TABELA 6 – SEQUÊNCIA DO MÉTODO *STEPWISE* PELO GRETl®

Variável removida no passo 2	Margem líquida Ult. 3 anos
Variável removida no passo 3	Liquidez seca
Variável removida no passo 4	Liquidez geral
Variável removida no passo 5	Crescimento da receita
Variável removida no passo 6	Patrimônio líquido
Variável removida no passo 7	Geração média de caixa Ult. 3 anos
Variável removida no passo 8	Alavancagem
Variável removida no passo 9	Tempo de relacionamento
Variável removida no passo 10	Tempo de atividade
Variável removida no passo 11	Liquidez corrente
Variável removida no passo 12	Crescimento margem EBITDA
Variável removida no passo 13	Margem EBITDA 2018

FONTE: Elaborado pelo autor a partir das saídas do *software* Gretl®

No método *stepwise* aplicado ao *software* Gretl®, observamos apenas algumas alterações na ordem de exclusão de algumas variáveis em relação ao SPSS®, entretanto o resultado final mantém-se o mesmo, sendo as variáveis “Restritivo Grau 4” e “Índice de Endividamento” as com maior significância em relação a variável dependente, ou seja, em relação a probabilidade da empresa apresentar ou não vencidos nos 12 meses subsequentes. Cabe salientar que as constantes para cada variável apresentaram exatamente os mesmos valores entre ambos os *softwares* estatísticos, conforme observa-se na TABELA 5 e TABELA 3.

TABELA 7 – MODELO FINAL APÓS MÉTODO STEPWISE PELO GRETl®

Variável	Coefficiente	Erro Padrão	Z	p-valor
Constante	-2,30181	0,387620	-5,938	2,88e-09 ***
Índice de endividamento	0,599806	0,315884	1,899	0,0576 *
Restritivo grau 4	2,63386	0,664149	3,966	7,32e-05 ***

Média var. dependente	0,222222	D.P. var. dependente	0,417399
R-quadrado de McFadden	0,213464	R-quadrado ajustado	0,168515
Log de verossimilhança	-52,49579	Crítério de Akaike	110,9916
Crítério de Schwarz	119,5004	Crítério Hannan-Quinn	114,4485

Número de casos 'corretamente previstos'	106 (84,1%)
F (beta'x) na média das variáveis independentes	0,166
Teste de razão de verossimilhança: Qui-quadrado (2)	28,4944

FONTE: Dados do Gretl®

Observa-se, através do modelo final acima, exatamente as mesmas variáveis e constantes do modelo final obtido pelo SPSS®. Prevendo corretamente 84,1% dos casos através de uma variável comportamental “Restritivo Grau 4” e uma variável financeira advinda dos dados de balanço “Índice de Endividamento”.

4.3 INTERPRETAÇÃO DOS RESULTADOS

Primeiramente, o método *stepwise* realiza por si só a avaliação dos valores-p, que indica se as variáveis independentes são ou não estatisticamente significantes (quanto menor o valor-p, mais significativa a variável), realizando a gradativa exclusão das variáveis menos significantes do modelo. Portanto, considera-se que as variáveis finais “Restritivo Grau 4” e “Índice de Endividamento” conseguem discriminar satisfatoriamente a probabilidade de ocorrência ou não de vencimentos acima de 90 dias nos próximos 12 meses das transportadoras, clientes da instituição bancária, conforme podemos observar nas FIGURA 2 e TABELA 7.

Sobre a variável “Restritivo Grau 4”, este restritivo é aplicado às empresas que, nos últimos 5 anos, apresentaram riscos vencidos acima de 90 dias junto à instituição, trata-se de uma marcação por parte da instituição bancária de que aquela empresa já apresentou problemas no passado. Em relação a variável “Índice

de Endividamento”, refere-se ao capital de terceiros sobre o ativo total das empresas, ou seja, os passivos circulante e não-circulante divididos pelo ativo total.

Outra análise relevante refere-se ao número de casos corretamente previstos, onde ambos os softwares demonstram que o modelo final prevê corretamente 84,1% dos casos conforme FIGURA 3 e TABELA 7.

Aplicando o modelo gerado após a metodologia supracitada, temos:

$$Z = - 2,302 + 2,634 \times \text{Rest Grau 4} + 0,5998 \times \text{Índice de Endividamento}$$

Sabemos, conforme TABELA 2, que para a variável Restritivo Grau 4, 0 (zero) assume que a empresa não possui este restritivo com o banco e 1 (um) para empresas que possuem este restritivo. Dessa forma, ao analisarmos sua constante definida pelo modelo, quanto maior o nível dessa variável, maior o nível da variável Z, pois a correlação é positiva, ou seja, quanto maior o nível da variável Restritivo Grau 4, o Z assume 1, onde a empresa tende a apresentar vencidos acima de 90 dias nos 12 meses subsequentes. A mesma correlação ocorre com a variável Índice de Endividamento, onde o peso 0,5998 da constante evidencia que, quanto maior o nível dessa variável, o Z assume 1.

Aplicando o modelo a um caso real na empresa doravante denominada Alpha Transportes (nome fictício), temos que a empresa possui um Índice de Endividamento de 1,4010 e possui Restritivo Grau 4 junto a instituição bancária. Ao calcularmos o Z, temos:

$$Z = -2,302 + 2,634 \times 1 + 0,5998 \times 1,4010, \text{ portanto } Z = 1,1723$$

Logo, a probabilidade da variável dependente ser 1, ou seja, da empresa apresentar vencimentos acima de 90 dias nos próximos 12 meses, é de:

$$1/(1+\exp(-1,1723)) = 76,36\% \text{ de chances.}$$

Diante disso, a empresa Alpha Transportes possui 76,36% de chances de apresentar vencimentos acima de 90 dias nos próximos 12 meses junto a instituição bancária, de acordo com o modelo final.

5 CONSIDERAÇÕES FINAIS

Por meio deste estudo, verificou-se a aplicabilidade de modelo de regressão logística sobre dados contábeis e comportamentais de uma amostra de 126 empresas do segmento de transporte de cargas, clientes de grande instituição bancária de âmbito nacional. Foram obtidas informações contábeis de 2016 a 2018, sendo o ano de 2019 utilizado como referência quanto a presença ou não de vencimentos acima de 90 dias com a instituição.

O estudo permitiu elucidar e discutir as aplicações da regressão logística e sua utilidade na área de mesa de crédito. Apesar do grande número de variáveis utilizadas no modelo, apenas duas ao final se mostraram relevantes na classificação das empresas em relação a probabilidade de apresentarem ou não vencimentos junto a instituição bancária nos próximos 12 meses.

Observou-se que, apesar da limitada amostra de empresas, o modelo e os métodos utilizados permitem acessível compreensão quanto a relação de cada variável frente a previsão do risco de inadimplência das transportadoras, permitindo também uma hierarquização das variáveis através do seu respectivo valor-p, o que possibilita o cruzamento das informações obtidas pelo modelo com os conceitos subjetivos de análise utilizados pelos analistas, fomentando o debate acerca de quais variáveis devem possuir maior relevância no momento de uma análise.

Não obstante, o modelo apresentou uma taxa de 84,1% dos casos corretamente previstos em ambos os *softwares* estatísticos, indicando que as duas variáveis finais, aplicadas ao modelo *Logit*, permitem um indicador de probabilidade fidedigno que pode de fato auxiliar a análise de crédito ao ser aplicado em conjunto com o julgamento dos analistas.

5.1 RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Como limitações deste trabalho, evidencia-se o tamanho da amostra utilizada mediante as fortes restrições de segurança de dados aplicadas às grandes instituições bancárias do país. Estudos futuros podem explorar a ampliação de amostras de dados possibilitando maior gama de aplicações e estudos quanto a cada variável de maneira mais contundente, permitindo novas variações de modelos, bem como empresas de outros segmentos de atuação. Cabe salientar que

uma base de dados mais robusta permite a aplicação de testes acerca dos modelos gerados, questão que não foi observada neste estudo.

REFERÊNCIAS

- ABDOU, H. A. Genetic programming for credit scoring: the case of Egyptian public sector banks. **Expert Systems with Applications**, v. 36, n. 9, p. 11402-11417, nov. 2009.
- ARAÚJO, E. A., ULISES, C., & CARMONA, D. M. Desenvolvimento de Modelos Credit Scoring com Abordagem de Regressão de Logística para a Gestão da Inadimplência de uma Instituição de Microcrédito. **Contabilidade Vista & Revista**, 18(3), 107–131. 2007.
- BIJAK, K.; THOMAS, L. C. Does segmentation always improve model performance in credit scoring? **Expert Systems with Applications**, v. 39, n. 3, p. 2433-2442, fev. 2012. <<http://dx.doi.org/10.1016/j.eswa.2011.08.093>>
- BLATT, A. Avaliação de Riscos e Decisões de Crédito: um enfoque prático. São Paulo: **Nobel**, 1999.
- BOUROCHE, J.M. SAPORTA, G. Análise de dados. Rio de Janeiro: **Zahar**, 1980. 117 p.
- BRITO, G. A. S.; ASSAF NETO, A. Modelo de classificação de risco de crédito de empresas. **Rev. contab. finanç.**, São Paulo, v. 19, n. 46, p. 18-29, Apr. 2008.
- CROOK, J. N., EDELMAN, D. B.; THOMAS, L. C. Recent developments in consumer credit risk assessment. **European Journal of Operational Research**, v.183, n.3, p.1447–1465, 2007
- EDITH, L., & PINHEIRO, T. Validação de Modelos Brasileiros de Previsão de Insolvência. **Contabilidade Vista & Revista**, 18(4), 83–103. 2007.
- GIMENES, R. M. T.; URIBE-OPAZO, M. A. Previsão de Insolvência de Cooperativas Agropecuárias, por meio de modelos multivariados. Disponível em: <http://www.fae.edu/publicacoes/pdf/revista_da_fae/fae_v4_n3/previsao_de_insolvencia.pdf>
- GITMAN, L. J. Princípios de administração financeira. 10. ed. **Pearson**, 1946 – Ed. 2004.
- GONÇALVES, E. B., GOUVÊA, M. A., & MANTOVANI, D. M. N. Análise de risco de crédito com o uso de regressão logística. **Revista Contemporânea de Contabilidade**, 10(20), 96–123. 2013. <<https://doi.org/10.5007/2175-8069.2013v10n20p139>>
- GOUVÊA, M. A., GONÇALVES, E. B., & MANTOVANI, D. M. N. Aplicação De Regressão Logística E Algoritmos Genéticos Na Análise De Risco De Crédito. **Revista Universo Contábil**, 84–102. 2012 <<https://doi.org/10.4270/ruc.2012214>>

HAIR JR., J. F. et al. Multivariate data analysis. 5th ed. New Jersey: **Prentice Hall**, 1998.

HAIR JR., J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. Análise multivariada de dados. São Paulo: **Bookman**, 2009.

HERSZKOWICZ, F. Credit scoring: A Aplicação de Métodos Estatísticos na Avaliação de Risco de Crédito. 2000. 90p. Monografia – **Faculdade de Economia, Administração e Contabilidade da USP**, São Paulo.

HOSMER, D. W.; LEMESHOW, D. W. Applied logistic regression. New York: **John Wiley & Sons**, 1989

LEV. B. Análisis de estados financieros, um nuevo enfoque. **Esic**, Madrid. 1978.

LIMA, F. G., FONSECA, C. V. C., SILVEIRA, R. L. F., & ASSAF NETO, A. Os Determinantes dos Ratings de Crédito dos Bancos Brasileiros. **Revista de Administração Contemporânea**, 22(2), 178–200. 2018. <<https://doi.org/10.1590/1982-7849rac2018160373>>

MARTINS, M. S. A previsão de insolvência pelo modelo de Cox: uma contribuição para a análise de companhias abertas brasileiras. Porto Alegre: UFRGS, 2003. Dissertação (Mestrado em Administração), **Programa de Pós-Graduação em Administração, Universidade Federal do Rio Grande do Sul**. 2003. Disponível em: Acesso em: 10 abr. 2006.

MATARAZZO, Dante C. Análise financeira de balanços: abordagem gerencial. 7. ed. São Paulo: **Atlas**, 2010.

MAVRI, M.; ANGELIS, V.; IOANNOU, G.; GAKI, E.; KOUFODONTIS, I. A two-stage dynamic credit scoring model, based on customers' profile and time horizon. **Journal of Financial Services Marketing**, v. 13, n. 1, p. 17-27, maio, 2008. <<http://dx.doi.org/10.1057/fsm.2008.2>>

OLIVEIRA, G. DE, SOARES, G., & CAMARGOS, M. A. De. Determinantes do Rating de Crédito de Companhias Brasileiras. **Contabilidade Vista & Revista**, 23(3), 109–143. 2012.

ONUSIC, L. M., CASA NOVA, S. P. DE C., & ALMEIDA, F. C. de. Modelos de previsão de insolvência utilizando a análise por envoltória de dados: aplicação a empresas brasileiras. **Revista de Administração Contemporânea**, 11(spe2), 77–97. 2007. <<https://doi.org/10.1590/s1415-65552007000600005>>

PIRES, V. M. Comparação Entre As Classificações De Risco De Crédito Estimadas Pelos Modelos Estruturais E Não Estruturais: Um Estudo Com Empresas Brasileiras Comparison Between the Classifications of Credit Risk Estimated By Structural and Non Structural Models: a S. 55(51), 35–53. 2014.

ROSA, P. T. M. Modelos de Credit Scoring: Regressão Logística, CHAID e REAL. 2000. Dissertação (Mestrado em Estatística) – Programa de Pós-graduação em Estatística, Instituto de Matemática e Estatística da Universidade de São Paulo, São Paulo, 2000.

SILVA, J. P. Gestão e análise de risco de crédito. 4 ed. São Paulo: **Atlas**, 2003.

THOMAS, L. C. A Survey of Credit and Behavioral Scoring: forecasting financial risk of lending to consumers, **International Journal of Forecasting**, Edinburg- U.K, v.16, p. 149-172, 2000.

VASCONCELOS, R. S. Modelos de Escoragem de Crédito aplicados a Empréstimo Pessoal com Cheque. Rio de Janeiro, 2004. Dissertação (Mestrado em Finanças e Economia Empresarial) – **Escola de Pós-Graduação em Economia, Fundação Getúlio Vargas**.

YAP, B. W.; ONG, S. H.; HUSAIN, N. H. M. Using data mining to improve assessment of credit worthiness via credit scoring models. **Expert Systems with Applications**, v. 38, n. 10, p. 13274-13283, set 2011.
<<http://dx.doi.org/10.1016/j.eswa.2011.04.147>>

**APÊNDICE 1 – MÉTODO STEPWISE COMPLETO ATRAVÉS DO
SOFTWARE IBM SPSS® STATISTICS SUBSCRIPTION**

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para EXP(B)	
								Inferior	Superior
Etapa 1ª	Índice de Endividamento	1,121	,516	4,716	1	,030	3,068	1,115	8,441
	Geracao media de Caixa ult 3 anos	-,365	,642	,324	1	,569	,694	,197	2,441
	Margem EBITDA 18	1,150	,997	1,331	1	,249	3,160	,448	22,303
	Liquidez Seca	,115	,827	,019	1	,890	1,121	,222	5,670
	Liquidez Corrente	-,242	,806	,090	1	,764	,785	,162	3,811
	Liquidez Geral	,065	,418	,024	1	,877	1,067	,470	2,420
	Tempo de Atividade	-,027	,033	,645	1	,422	,974	,912	1,039
	Tempo de relacionamento	,029	,039	,566	1	,452	1,030	,954	1,112
	Alavancagem	,025	,042	,344	1	,557	1,025	,944	1,113
	Margem Liquida ult 3 anos	-,009	,755	,000	1	,991	,991	,226	4,356
	Patrimonio Liquido	,565	,959	,347	1	,556	1,760	,269	11,531
	Crescimento da Receita	,369	,815	,205	1	,651	1,446	,293	7,142
	Crescimento Margem Ebitda	-,729	,566	1,662	1	,197	,482	,159	1,462
	Rest grau 4	2,798	,746	14,063	1	,000	16,420	3,803	70,888
	Constante	-3,684	1,527	5,825	1	,016	,025		
	Etapa 2ª	Índice de Endividamento	1,122	,514	4,770	1	,029	3,070	1,122
Geracao media de Caixa ult 3 anos		-,367	,625	,345	1	,557	,693	,204	2,358
Margem EBITDA 18		1,149	,990	1,346	1	,246	3,156	,453	21,986
Liquidez Seca		,114	,825	,019	1	,890	1,121	,222	5,652
Liquidez Corrente		-,242	,805	,090	1	,764	,785	,162	3,806
Liquidez Geral		,065	,418	,024	1	,877	1,067	,470	2,421
Tempo de Atividade		-,027	,033	,655	1	,418	,974	,913	1,039
Tempo de relacionamento		,029	,039	,567	1	,452	1,030	,954	1,112
Alavancagem		,025	,042	,344	1	,558	1,025	,944	1,113
Patrimonio Liquido		,562	,907	,384	1	,536	1,753	,297	10,367

	Crescimento da Receita	,368	,813	,205	1	,651	1,445	,294	7,115
	Crescimento Margem Ebitda	-,729	,566	1,664	1	,197	,482	,159	1,461
	Rest grau 4	2,799	,744	14,162	1	,000	16,431	3,824	70,599
	Constante	-3,687	1,515	5,918	1	,015	,025		
Etapa 3ª	Índice de Endividamento	1,110	,507	4,789	1	,029	3,035	1,123	8,205
	Geracao media de Caixa ult 3 anos	-,364	,624	,340	1	,560	,695	,204	2,363
	Margem EBITDA 18	1,142	,991	1,328	1	,249	3,133	,449	21,845
	Liquidez Seca	,129	,785	,027	1	,869	1,138	,245	5,295
	Liquidez Corrente	-,221	,761	,084	1	,772	,802	,180	3,563
	Tempo de Atividade	-,027	,033	,675	1	,411	,973	,913	1,038
	Tempo de relacionamento	,030	,039	,587	1	,443	1,030	,955	1,112
	Alavancagem	,024	,042	,330	1	,566	1,024	,944	1,111
	Patrimonio Liquido	,563	,905	,388	1	,533	1,757	,298	10,348
	Crescimento da Receita	,351	,805	,190	1	,663	1,420	,293	6,883
	Crescimento Margem Ebitda	-,726	,566	1,646	1	,200	,484	,160	1,467
	Rest grau 4	2,782	,735	14,336	1	,000	16,152	3,826	68,181
	Constante	-3,641	1,479	6,060	1	,014	,026		
Etapa 4ª	Índice de Endividamento	1,106	,507	4,763	1	,029	3,023	1,119	8,162
	Geracao media de Caixa ult 3 anos	-,362	,624	,337	1	,561	,696	,205	2,365
	Margem EBITDA 18	1,150	,986	1,361	1	,243	3,160	,457	21,837
	Liquidez Corrente	-,110	,204	,291	1	,590	,896	,600	1,337
	Tempo de Atividade	-,027	,033	,676	1	,411	,973	,912	1,038
	Tempo de relacionamento	,030	,039	,588	1	,443	1,030	,955	1,112
	Alavancagem	,024	,042	,320	1	,572	1,024	,944	1,111
	Patrimonio Liquido	,562	,901	,389	1	,533	1,754	,300	10,264
	Crescimento da Receita	,343	,806	,181	1	,670	1,409	,290	6,842
	Crescimento Margem Ebitda	-,721	,565	1,630	1	,202	,486	,161	1,471
	Rest grau 4	2,789	,733	14,469	1	,000	16,260	3,864	68,415
	Constante	-3,632	1,479	6,033	1	,014	,026		
Etapa 5ª	Índice de Endividamento	1,148	,500	5,273	1	,022	3,153	1,183	8,404
	Geracao media de Caixa ult 3 anos	-,378	,624	,367	1	,545	,685	,202	2,328

	Margem EBITDA 18	1,247	,961	1,684	1	,194	3,480	,529	22,891
	Liquidez Corrente	-,108	,200	,292	1	,589	,897	,606	1,329
	Tempo de Atividade	-,028	,032	,764	1	,382	,972	,912	1,036
	Tempo de relacionamento	,032	,038	,703	1	,402	1,033	,958	1,114
	Alavancagem	,019	,039	,234	1	,629	1,019	,944	1,101
	Patrimonio Liquido	,608	,891	,466	1	,495	1,837	,321	10,524
	Crescimento Margem Ebitda	-,729	,564	1,671	1	,196	,482	,160	1,457
	Rest grau 4	2,750	,723	14,479	1	,000	15,649	3,795	64,527
	Constante	-3,468	1,410	6,046	1	,014	,031		
Etapa 6ª	Indice de Endividamento	1,096	,484	5,131	1	,023	2,993	1,159	7,729
	Geracao media de Caixa ult 3 anos	-,316	,609	,269	1	,604	,729	,221	2,406
	Margem EBITDA 18	1,146	,945	1,472	1	,225	3,146	,494	20,042
	Liquidez Corrente	-,127	,212	,357	1	,550	,881	,582	1,335
	Tempo de Atividade	-,029	,032	,803	1	,370	,971	,911	1,035
	Tempo de relacionamento	,033	,039	,727	1	,394	1,033	,958	1,114
	Patrimonio Liquido	,655	,890	,541	1	,462	1,924	,336	11,010
	Crescimento Margem Ebitda	-,700	,560	1,565	1	,211	,497	,166	1,487
	Rest grau 4	2,818	,712	15,679	1	,000	16,747	4,151	67,571
	Constante	-3,330	1,388	5,758	1	,016	,036		
Etapa 7ª	Indice de Endividamento	1,103	,481	5,253	1	,022	3,013	1,173	7,737
	Margem EBITDA 18	1,046	,928	1,270	1	,260	2,847	,461	17,567
	Liquidez Corrente	-,117	,197	,350	1	,554	,890	,605	1,310
	Tempo de Atividade	-,027	,032	,730	1	,393	,973	,914	1,036
	Tempo de relacionamento	,032	,038	,699	1	,403	1,032	,958	1,113
	Patrimonio Liquido	,657	,887	,549	1	,459	1,929	,339	10,971
	Crescimento Margem Ebitda	-,679	,557	1,485	1	,223	,507	,170	1,512
	Rest grau 4	2,844	,708	16,129	1	,000	17,186	4,289	68,863
	Constante	-3,541	1,328	7,113	1	,008	,029		
Etapa 8ª	Indice de Endividamento	,953	,424	5,054	1	,025	2,595	1,130	5,957
	Margem EBITDA 18	1,356	,835	2,638	1	,104	3,882	,755	19,945
	Liquidez Corrente	-,108	,184	,342	1	,558	,898	,626	1,288
	Tempo de Atividade	-,029	,032	,799	1	,371	,971	,912	1,035
	Tempo de relacionamento	,029	,039	,579	1	,447	1,030	,955	1,111

	Crescimento Margem Ebitda	-,702	,555	1,600	1	,206	,496	,167	1,471
	Rest grau 4	2,751	,688	16,009	1	,000	15,665	4,070	60,293
	Constante	-3,076	1,149	7,175	1	,007	,046		
Etapa 9ª	Indice de Endividamento	1,028	,415	6,127	1	,013	2,796	1,239	6,309
	Margem EBITDA 18	1,345	,847	2,524	1	,112	3,839	,730	20,186
	Tempo de Atividade	-,028	,033	,716	1	,397	,973	,913	1,037
	Tempo de relacionamento	,032	,039	,660	1	,417	1,032	,956	1,114
	Crescimento Margem Ebitda	-,691	,556	1,547	1	,214	,501	,169	1,489
	Rest grau 4	2,785	,690	16,296	1	,000	16,198	4,190	62,615
	Constante	-3,340	1,099	9,230	1	,002	,035		
Etapa 10ª	Indice de Endividamento	1,001	,413	5,866	1	,015	2,720	1,210	6,113
	Margem EBITDA 18	1,289	,839	2,360	1	,124	3,629	,701	18,789
	Tempo de Atividade	-,008	,022	,142	1	,707	,992	,950	1,036
	Crescimento Margem Ebitda	-,652	,549	1,411	1	,235	,521	,178	1,527
	Rest grau 4	2,710	,673	16,191	1	,000	15,025	4,014	56,236
	Constante	-3,189	1,082	8,689	1	,003	,041		
Etapa 11ª	Indice de Endividamento	1,007	,412	5,968	1	,015	2,738	1,220	6,144
	Margem EBITDA 18	1,302	,840	2,404	1	,121	3,676	,709	19,053
	Crescimento Margem Ebitda	-,672	,546	1,514	1	,219	,511	,175	1,489
	Rest grau 4	2,709	,673	16,198	1	,000	15,017	4,014	56,180
	Constante	-3,380	,966	12,233	1	,000	,034		
Etapa 12ª	Indice de Endividamento	,866	,389	4,963	1	,026	2,378	1,110	5,095
	Margem EBITDA 18	,979	,787	1,550	1	,213	2,662	,570	12,441
	Rest grau 4	2,631	,662	15,789	1	,000	13,885	3,793	50,830
	Constante	-3,327	,946	12,362	1	,000	,036		
Etapa 13ª	Indice de Endividamento	,600	,317	3,571	1	,059	1,822	,978	3,394
	Rest grau 4	2,634	,654	16,218	1	,000	13,927	3,865	50,185
	Constante	-2,302	,406	32,097	1	,000	,100		

a. Variável(is) inserida(s) no passo 1: Índice de Endividamento, Geracao media de Caixa ult 3 anos, Margem EBITDA 18, Liquidez Seca, Liquidez Corrente, Liquidez Geral, Tempo de Atividade, Tempo de relacionamento, Alavancagem, Margem Liquida ult 3 anos, Patrimonio Liquido, Crescimento da Receita, Crescimento Margem Ebitda, Rest grau 4.

APÊNDICE 2 – MODELO COMPLETO ATRAVÉS DO SOFTWARE GRETl®

Modelo 1: Logit, usando as observações 1-126

Variável dependente: HistoricoVencnosult12meses

Erros padrão QML

Omitido devido a colinearidade exata: LucroLiquido

	coeficiente	erro padrão	z	p-valor	
const	-3,68448	1,36681	-2,696	0,0070	***
IndicedeEndivida~	1,12119	0,559401	2,004	0,0450	**
GeracaomediadeCa~	-0,365480	0,641605	-0,5696	0,5689	
MargemEBITDA18	1,15047	1,13818	1,011	0,3121	
LiquidezSeca	0,114592	0,516224	0,2220	0,8243	
LiquidezCorrente	-0,242196	0,471543	-0,5136	0,6075	
LiquidezGeral	0,0646630	0,210996	0,3065	0,7592	
TempodeAtividade	-0,0267625	0,0296075	-0,9039	0,3660	
Tempoderelaciona~	0,0294179	0,0363509	0,8093	0,4184	
Alavancagem	0,0246391	0,0262508	0,9386	0,3479	
MargemLiquidault~	-0,00856462	0,962873	-0,008895	0,9929	
PatrimonioLiquido	0,565111	1,24175	0,4551	0,6490	
CrescimentodaRec~	0,368771	0,721792	0,5109	0,6094	
CrescimentoMarge~	-0,729233	0,536940	-1,358	0,1744	
Restgrau4	2,79848	0,708417	3,950	7,80e-05	***
Média var. dependente	0,222222	D.P. var. dependente		0,417399	
R-quadrado de McFadden	0,258066	R-quadrado ajustado		0,033323	
Log da verossimilhança	-49,51888	Critério de Akaike		129,0378	
Critério de Schwarz	171,5820	Critério Hannan-Quinn		146,3221	

Número de casos 'corretamente previstos' = 107 (84,9%)

f(beta*x) na média das variáveis independentes = 0,163

Teste de razão de verossimilhança: Qui-quadrado(14) = 34,4482 [0,0018]

		Previsto	
		0	1
Efetivo	0	94	4
	1	15	13

**APÊNDICE 3 – MODELO FINAL APÓS MÉTODO *STEPWISE* PELO
GRET[®]**

Modelo 13: Logit, usando as observações 1-126
Variável dependente: HistoricoVencnosult12meses
Erros padrão QML

	coeficiente	erro padrão	z	p-valor
const	-2,30181	0,387620	-5,938	2,88e-09 ***
IndicedeEndivida~	0,599806	0,315884	1,899	0,0576 *
Restgrau4	2,63386	0,664149	3,966	7,32e-05 ***
Média var. dependente	0,222222	D.P. var. dependente		0,417399
R-quadrado de McFadden	0,213464	R-quadrado ajustado		0,168515
Log da verossimilhança	-52,49579	Critério de Akaike		110,9916
Critério de Schwarz	119,5004	Critério Hannan-Quinn		114,4485

Número de casos 'corretamente previstos' = 106 (84,1%)
f(beta'x) na média das variáveis independentes = 0,166
Teste de razão de verossimilhança: Qui-quadrado(2) = 28,4944 [0,0000]

		Previsto	
		0	1
Efetivo	0	93	5
	1	15	13