



UFC

**UNIVERSIDADE FEDERAL DO CEARÁ
INSTITUTO DE CULTURA E ARTE
DEPARTAMENTO DE COMUNICAÇÃO SOCIAL
GRADUAÇÃO EM PUBLICIDADE E PROPAGANDA**

GABRIELLE DESIREE GONÇALVES SILVA

***CUSTOMER SUCCESS: MARKETING PREDITIVO COMO ESTRATÉGIA DE
SUSTENTABILIDADE PARA EMPRESA DE RECEITA RECORRENTE***

FORTALEZA

2022

GABRIELLE DESIREE GONÇALVES SILVA

***CUSTOMER SUCCESS: MARKETING PREDITIVO COMO ESTRATÉGIA DE
SUSTENTABILIDADE PARA EMPRESA DE RECEITA RECORRENTE***

Monografia apresentada ao Curso de Graduação em Publicidade e Propaganda da Universidade Federal do Ceará, como requisito à obtenção do título de bacharel em Publicidade e Propaganda.

Orientador: Prof. Me. Janice Leal de Carvalho Vasconcelos

FORTALEZA

2022

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- S58c Silva, Gabrielle Desiree Gonçalves.
Customer success : marketing preditivo como estratégia de sustentabilidade para empresa de receita recorrente / Gabrielle Desiree Gonçalves Silva. – 2022.
100 f. : il. color.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Instituto de Cultura e Arte, Curso de Comunicação Social (Publicidade e Propaganda), Fortaleza, 2022.
Orientação: Profa. Ma. Janice Leal de Carvalho Vasconcelos.
1. Sucesso do Cliente. 2. Ciência de Dados. 3. Marketing Preditivo. I. Título.

CDD 070.5

GABRIELLE DESIREE GONÇALVES SILVA

***CUSTOMER SUCCESS: MARKETING PREDITIVO COMO ESTRATÉGIA DE
SUSTENTABILIDADE PARA EMPRESA DE RECEITA RECORRENTE***

Monografia apresentada ao Curso de Graduação em Publicidade e Propaganda da Universidade Federal do Ceará, como requisito à obtenção do título de bacharel em Publicidade e Propaganda.

Aprovada em: 29/06/2022.

BANCA EXAMINADORA

Prof. Me. Janice Leal de Carvalho Vasconcelos (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Me. Soraya Madeira da Silva
Universidade Federal do Ceará (UFC)

Prof. Esp. Thaisy Barbosa
Universidade de Fortaleza (UNIFOR)

À Neumona,
consequimos.

AGRADECIMENTOS

À Universidade Federal do Ceará, pelo comprometimento em formar cidadãos e profissionais com diferencial competitivo e pelo excepcional quadro de mestres e doutores que nos iluminam durante toda a jornada da graduação.

À Universidade Estadual do Ceará, pelo excelente projeto de extensão gratuito que forma tecnicamente profissionais na área de tecnologia.

À Prof. Me. Janice Leal, pela notável orientação e que sempre me fez gostar um pouco mais da área de administração e *marketing*.

Às professoras participantes da banca examinadora Soraya Madeira e Thaisy Barbosa, pelo tempo e colaboração em minha jornada acadêmica e profissional.

A Nemésio, Isabel, Iúri e Isabelle, que da sua maneira, contribuem em meus principais traços de desenvolvimento como pessoa.

A Giulia, João e Benício, a vida ficou mais alegre com a chegada de vocês.

A Socorro e Lourdes, que tanto me ensinaram sobre amor e doação.

A Neuman, Tereza e Nertan, os quais admiro como grandes mentores de vida, comprometimento e inteligência.

A Rebeca e Davi, que cresceram e dividiram a infância comigo, suas presenças são fundamentais até hoje.

A Ataíde Júnior pela parceria, experiências e conselhos trocados há mais de uma década.

À Daniela, Jéssica, Lourena, Letícia, Rayanna e Bruna, a força e o sucesso de vocês me dão ânimo e sentido para minhas próprias conquistas.

A Wilker, pela confiança, admiração, desenvolvimento mútuo e pelas muitas oportunidades de ouvir e fazer música.

À Sophia e Mímia, que me cuidam, ouvem e ensinam por meio da sua imensa e divertida sabedoria felina.

“O mais importante é prever para onde os clientes estão indo e chegar lá primeiro.”
(KOTLER, 2000, p. 181).

RESUMO

O presente estudo monográfico fundamenta modelos da ciência de dados que possibilitam estratégias preditivas de *marketing*, considerando fundamentos das áreas de experiência do consumidor, gestão de relacionamento e sucesso do cliente como estratégia de sustentabilidade para empresas de receita recorrente. A partir de indicadores que visam garantir a sustentabilidade do negócio e relacionamento, como compras previstas, probabilidade de evasão e valor do tempo de vida do cliente, o estudo desenvolve uma metodologia de modelos matemáticos e estatísticos denominados “*Buy ‘til you die models*” que consideram a base de dados de transações históricas para prever o futuro do relacionamento entre empresa e cliente. O objetivo geral desta monografia é desenvolver análises preditivas para negócios *B2B* com tecnologias relacionadas à ciência de dados. Adicionalmente, seus objetivos específicos são, por meio da linguagem *Python*, possibilitar análises preditivas de compras, probabilidade de evasão (*churn*) e valor do cliente (CLV) para toda a extensão base de dados (*big data operation*). Para isso, foram aplicadas duas metodologias: RFM e BG/NBD vinculado ao modelo Gamma-Gamma. As metodologias possibilitaram a modelagem estatística dos indicadores de sucesso para toda a base de clientes pelos próximos quatro anos. Assim, foi desenvolvido um plano de ações estratégicas e relacionamento com base nas previsões obtidas.

Palavras-chave: SUCESSO DO CLIENTE; CIÊNCIA DE DADOS; MARKETING PREDITIVO.

ABSTRACT

The monographic study presents fundamentals of data science that enable predictive marketing strategies, considering as fundamentals the areas of consumer experience, management of relationship and customer success as a sustainability strategy for recurring companies. Based on indicators that seek to guarantee the sustainability of the business and the relationship, such as the forecast of purchases, the probability of evasion and the value of the customer's useful life, the study develops a methodology of mathematical and statistical models called "Buy 'til you die models" that model a database of historical transactions to predict the future of the company-customer relationship. The overall objective of this monograph is to develop predictive analytics for B2B businesses with technologies related to data science. Additionally, its specific objectives are, through the Python language, to enable estimates of purchases, probability of evasion (churn) and customer value (CLV) for the entire database extension (big data operation). For this, two methodologies were applied: RFM and BG/NBD linked to the Gamma-Gamma model. The methodologies made it possible to model statistical indicators of success for the entire customer base for the next four years. Thus, a strategic and relationship action plan was developed based on the forecasts obtained.

Keywords: CUSTOMER SUCCESS; DATA SCIENCE; PREDICTIVE MARKETING.

LISTA DE FIGURAS

Figura 1	Volume de publicações e pesquisas acadêmicas em temas sobre práticas de 20 administração de clientes.....	
Figura 2	Pirâmide de Maslow da hierarquia das necessidades.....	21
Figura 3	<i>Processos vitais da empresa evoluindo para o modelo de Customer Experience e Customer Success.....</i>	22
Figura 4	<i>Processos vitais da empresa evoluindo para o modelo de Customer Experience e Customer Success.....</i>	23
Figura 5	<i>O caminho dos 5As do cliente.....</i>	24
Figura 6	<i>Diversos canais e fontes de informação para tomada de decisão antes, durante e após a compra, fazendo com que as empresas se tornem Omnichannel.....</i>	26
Figura 7	Interface do sistema de CRM <i>Salesforce</i>	28
Figura 08	<i>Pedidos da base instalada de clientes versus pedidos de novos clientes.....</i>	31
Figura 09	<i>New Marketing Funnel.....</i>	32
Figura 10	<i>Hierarquia do valor do cliente.....</i>	33
Figura 11	<i>O Customer Lifecycle do Cliente faz parte do modelo de relacionamento. Na parte de cima do diagrama estão as ações comerciais potenciais da empresa.....</i>	34
Figura 12	Calculadora de <i>lifetime value</i>	40
Figura 13	Cálculo simplificado de LTV.....	40
Figura 14	Cálculo de <i>Churn</i>	42
Figura 15	<i>Negative churn</i>	43
Figura 16	Aplicações do <i>marketing</i> preditivo.....	46
Figura 17	A pirâmide da gestão do conhecimento.....	46
Figura 18	<i>Jupyter Notebook</i> . Importação de bibliotecas <i>Python</i> para análises	50

	preditivas.....	
Figura 19	<i>Jupyter Notebook</i> . Importação da base de dados e campos de dados.....	51
Figura 20	<i>Jupyter Notebook</i> . Exclusão de transações não identificáveis pelo ID.....	52
Figura 21	<i>Jupyter Notebook</i> . Exclusão de transações de devolução e restrição da análise apenas à transações de quantidades acima de um item.....	52
Figura 22	<i>Jupyter Notebook</i> . Padronização do formato de datas e outras revisões.....	53
Figura 23	<i>Jupyter Notebook</i> . Dados totais após tratamento inicial e exclusão de colunas irrelevantes.....	53
Figura 24	<i>Jupyter Notebook</i> . Visão geral da base de dados após tratamento e inserção da coluna “faturamento”.....	54
Figura 25	<i>Jupyter Notebook</i> . Teste de indicadores com cliente específico.....	55
Figura 26	<i>Jupyter Notebook</i> . Treinamento de máquina e teste ampliado para toda a base de dados.....	56
Figura 27	<i>Jupyter Notebook</i> . Visão geral das principais médias contemplando toda a base de dados.....	57
Figura 28	<i>Jupyter Notebook</i> . Frequência média contemplando toda a base de clientes.	58
Figura 29	<i>Jupyter Notebook</i> . Função <code>sumary_data_from_transaction_data()</code> que estrutura a base de dados a ser utilizada no modelo BG/NBD.....	60
Figura 30	<i>Jupyter Notebook</i> . Ajuste da base de dados à distribuição para derivação de seus parâmetros.....	60
Figura 31	<i>Jupyter Notebook</i> . Treinamento do modelo X transações atuais que demonstra que o aplicativo está equilibrado.....	61
Figura 32	<i>Jupyter Notebook</i> . Visão geral de estatísticas descritivas.....	62
Figura 33	<i>Jupyter Notebook</i> . Função <code>bgf.predict</code> para previsões de clientes específicos.....	65
Figura 34	<i>Jupyter Notebook</i> . Função <code>predict_purch</code> para previsões de toda a base de clientes.....	65
Figura 35	<i>Jupyter Notebook</i> . Top 10 clientes pela previsão de compras nos próximos	66

	90 dias.....	
Figura 36	<i>Jupyter Notebook</i> . Função <i>conditional_probability_alive</i>	67
Figura 37	<i>Jupyter Notebook</i> . Clientes com probabilidade menor de 90% de recompra	67
Figura 38	<i>Jupyter Notebook</i> . Probabilidade de <i>churn</i> de cliente específico (ID Cliente 13093).....	68
Figura 39	<i>Jupyter Notebook</i> . Descrição de previsões de toda a base de clientes.....	69
Figura 40	<i>Jupyter Notebook</i> . <i>Dataframe</i> de visualização de indicadores por cliente....	70
Figura 41	Visualização de dados preditivos em <i>dashboard</i> hospedado no <i>Data Studio</i>	71
Figura 42	Mapeamento da jornada do cliente: captação e <i>onboarding</i>	73
Figura 43	Mapeamento da jornada do cliente:energização, proteção e fidelização.....	74
Figura 44	Mapeamento da jornada do cliente: retenção e recuperação.....	75

LISTA DE GRÁFICOS

Gráfico 1	<i>Jupyter Notebook</i> . Visualização da frequência x densidade da base de 58 clientes.....	58
Gráfico 2	<i>Jupyter Notebook</i> . Visualização da recência x densidade da base de 58 clientes.....	58
Gráfico 3	<i>Jupyter Notebook</i> . Visualização da idade do cliente “T” x densidade da base de 59 de clientes.....	59
Gráfico 4	Comparativo da frequência de compras do modelo preditivo X compras reais no período.....	61
Gráfico 5	<i>Frequência de transações</i> . Treinamento do modelo X transações atuais que demonstra que o aplicativo está equilibrado.....	62
Gráfico 6	<i>Compras previstas X Compras reais no período</i> . Treinamento do modelo X compras atuais que demonstra que o aplicativo está equilibrado.....	63
Gráfico 7	<i>Compras previstas nos próximos 90 dias</i>	64
Gráfico 8	<i>Compras previstas nos próximos 180 dias</i>	64
Gráfico 9	<i>Jupyter Notebook</i> . Probabilidade de <i>churn</i>	66

LISTA DE TABELAS

Tabela 1	Diferenças entre <i>customer success</i> e suporte ao cliente.....	30
Tabela 2	<i>Diferença entre a empresa com foco em si mesma e empresa com Foco no Valor do Cliente</i>	38

LISTA DE ABREVIATURAS E SIGLAS

CS	<i>Customer Success</i>
CX	<i>Customer Experience</i>
SaaS	<i>Software as a service</i>
B2B	<i>Business to business</i>
B2C	<i>Business to customer</i>
LTV	<i>Lifetime value</i>
CLV	<i>Customer lifetime value</i>
CRM	<i>Customer Relationship Management</i>
MRR	<i>Monthly Recurring Revenue</i>
NPS	<i>Net Promoter Score</i>
CSAT	<i>Customer satisfaction</i>
RFM	<i>Recency, frequency and monetary value model</i>
BG/NBD	<i>Beta-geometric/negative binomial distribution</i>

SUMÁRIO

1	INTRODUÇÃO	14
2	O SUCESSO DO CLIENTE COMO UMA ESTRATÉGIA DE MARKETING	17
2.1	Customer experience management	22
2.1.1	<i>Mapa de jornada do cliente</i>	26
2.2	Customer relationship management	28
2.3	Customer success	31
2.3.1	<i>Ciclo de vida do cliente</i>	34
2.3.1.1	NPS.....	36
2.3.2	<i>Valor do cliente</i>	38
2.3.2.1	Customer lifetime value (CLV).....	39
2.3.2.2	Churn.....	42
3	Data Science aplicado ao sucesso do cliente	43
3.1	<i>Marketing preditivo</i>	45
3.2	<i>Biblioteca lifetime e modelos matemáticos</i>	48
3.2.1.1	Modelo RFM.....	48
3.2.1.2	Beta-geometric/Negative Binomial (BG/NBD) e	49
	Gamma-Gamma.....	
4	MÉTODO DE PESQUISA	51
4.1	Análises preditivas em Python	51
4.1.1	<i>Base de dados e tratamento inicial</i>	51
4.1.2	<i>Aplicação do modelo RFM</i>	55
4.1.2.1	<i>Modelo RFM - Função calibration_and_holdout_data para aprendizado de</i> <i>máquina</i>	55
4.1.3	<i>Aplicação do modelo BG/NBD</i>	59
4.1.3.1	<i>Modelo BG/NBD - Função BetaGeoFitter para previsão de indicadores</i>	59
4.1.3.2	<i>Modelo BG/NBD aplicado à base de dados em integridade</i>	62
5	APRESENTAÇÃO E ANÁLISE DE DADOS DA PESQUISA	64
5.1	Compras previstas	64
5.2	Probabilidade de churn	66
5.3	Previsão monetária do lifetime value (CLV)	68

5.4	Visualização de dados	70
5.5	Ações estratégicas	71
5.5.1	<i>Mapeamento da jornada do cliente</i>	72
5.5.2	<i>Canais de comunicação</i>	74
5.5.3	<i>Plano tático</i>	75
6	CONCLUSÃO E TRABALHOS FUTUROS	77
	REFERÊNCIAS	79
	APÊNDICE A – CÓDIGO DE PROGRAMAÇÃO	81

1 INTRODUÇÃO

Os clientes, de fato, são a essência de todo negócio. Fidelizá-los e retê-los por muito tempo são imperativos para negócios de receita recorrente (aqueles em que o faturamento é baseado na venda de produtos em uma frequência recorrente, por exemplo, mensalmente) e, por isso, são necessários investimentos reais para torná-los bem sucedidos com seu produto ou serviço prestado.

Nesse contexto, surge no Vale do Silício uma filosofia de negócios denominada *Customer Success*, uma unidade organizacional focada na experiência do cliente com intuito de maximizar a retenção e o valor do seu tempo de vida (CRUZ, 2021), reduzindo possíveis evasões. A redução da evasão na carteira de clientes, por si, já justifica a implantação de uma operação voltada ao sucesso do cliente.

Ao longo dos anos, essa filosofia vem demonstrando sua eficiência por meio de indicadores que expressam seu potencial financeiro e também não financeiro, agregando maior valor de mercado às empresas que a adotam, que em maioria são empresas *start ups* e/ou empresas de tecnologia (Steinman et al., 2020).

Em 2020, com o avanço da pandemia do vírus Covid-19, organizações em todo o mundo passaram por um processo de aceleração digital, no qual processos internos, experiência de compra e o relacionamento com o cliente foi adaptado para acontecer em canais digitais que, por sua vez, adequaram-se para satisfazer clientes com jornadas mais simplificadas, rápidas e com comunicação coerente.

Portanto, investimentos em tecnologia foram realizados para reinventar os setores de *marketing* e vendas, com intuito de reestruturá-los. E há um fator comum entre todos os canais em ambiente digital: todas as interações entre empresa-cliente deixam dados históricos que podem e devem convergir para resultar em informação e inteligência em prol da organização sobre seus clientes.

Nesse sentido, a tecnologia tem um papel protagonista ao funcionar como um orientador preditivo para ações de *marketing* voltadas ao cliente (Kotler et al., 2021), melhorando assim as entregas de relacionamento e experiência, que consideram o estágio do ciclo de vida em que o cliente se encontra, a jornada mapeada que é proposta para cada grupo de clientes, assim como o planejamento de ações que tem o objetivo de reter e fidelizar a base de clientes (Madruga, 2021).

Esta pesquisa objetiva identificar os benefícios da implementação de estratégias com foco no sucesso do cliente, fundamentadas a partir de conceitos da Experiência do Consumidor (do inglês, *Customer Experience* ou CX) e do Gerenciamento do Relacionamento com o Cliente (do inglês, *Customer Relationship Management* ou CRM) aplicados ao contexto de empresas de receita recorrente e ao *marketing* preditivo, operacionalizados por soluções relacionadas à ciência de dados.

Os objetivos específicos definidos para essa monografia, que visam garantir a execução do objetivo geral do estudo, é desenvolver uma aplicação de *marketing* preditivo para negócios *B2B* por meio da linguagem de programação *Python*, que possibilita previsões pelos próximos quatro anos dos indicadores de frequência de compras, probabilidade de evasão dos clientes e valor do tempo de vida do cliente a partir do histórico de transações dos últimos dois anos.

No Capítulo 1 é introduzido o tema, proporcionando uma visão geral de suas características, o cenário acadêmico sobre práticas de administração de clientes e a contribuição das técnicas da disciplina de *marketing*.

O Capítulo 2 aborda conceitos e técnicas da área de *Customer Experience*, como o mapeamento da jornada do cliente, *Customer Relationship Management* e *Customer Success* como uma estratégia adotada para o planejamento de *marketing*, compreendendo seus indicadores quantitativos e monetários. Ainda, levanta quantitativamente os estudos acerca do tema já produzidos e a relevância do tema.

No Capítulo 3 são fundamentados os conceitos relacionados ao *marketing* preditivo e aplicação de metodologias da ciência de dados para análise de dados históricos e modelagem matemática e estatística com intuito de prever indicadores de sucesso e monetários para períodos futuros.

O Capítulo 4 apresenta um detalhamento acerca da metodologia dos modelos estatísticos que são aplicados à base de dados considerada para esta pesquisa com intuito de possibilitar a previsão dos indicadores de *marketing*.

No Capítulo 5, apresenta-se a análise dos dados obtidos por meio da modelagem de dados via linguagem de programação, separando os resultados em compras previstas para o

período, probabilidade de *churn* (ou evasão do cliente) e previsão do valor monetário do seu tempo de vida, bem como a apresentação de ações estratégicas de *marketing* proposto orientado por esses resultados.

Por fim, no Capítulo 6, apresenta-se a conclusão, seguida pelas referências bibliográficas que embasam o estudo e pelo apêndice no qual consta o código de programação desenvolvido pela autora.

2 O SUCESSO DO CLIENTE COMO UMA ESTRATÉGIA DE MARKETING

O tema desta pesquisa é *Customer Success*, em português, Sucesso do Cliente e os benefícios deste em empresas *B2B (Business to Business)* que são as que contam com o modelo de negócio de receita recorrente, que consiste em compras periódicas mas sem vínculo contratual. O setor organizacional abreviado por “CS” surge no início dos anos 2000 em empresas SaaS (*Software as a Service* - venda de *softwares* como um serviço estratégico para as empresas), por exemplo, a pioneira em sua adoção e integração, a *Sales Force*. (Steinman, Murphy e Mehta, 2017).

O futuro, de fato, já chegou. Nos últimos anos tornaram-se um importante diferencial e fator de sustentabilidade financeira os investimentos em “transformação digital” - uma cultura organizacional que permeia processos operacionais a tomadas de decisão. No geral, são softwares e sistemas que otimizam a mineração, tratamento e visualização de dados para as tomadas de decisão, assim como a agilidade de tarefas e assertividade de estratégias.

Esse mercado e metodologia cresceu exponencialmente após a aceleração digital vivida em 2020 (Kotler et al., 2021) com a pandemia de COVID-19 e regimes *officeless*, adotados por organizações em todo o mundo que respondem à uma demanda *omnichannel* (experiência satisfatória do cliente entre os canais de contato com a organização) de um contexto em que o consumidor torna seus hábitos cada vez mais digitais, seja através de plataformas de e-commerce, redes sociais, aplicativos em geral ou de relacionamento, assim como o torna mais consciente sobre suas decisões de compra e jornadas de experiência com as marcas (Kotler et al., 2021).

Por conta disso, a necessidade de personalizar as experiências, baseadas no conhecimento e dados de cada cliente é essencial para atender as expectativas do público que, nesse contexto, giram em torno de 3 premissas: personalização, facilidade e velocidade (Zendesk, 2022). Por meio dessas, as jornadas de experiência e ofertas mais adequadas a cada perfil de cliente e a fluidez do processo de compra ou troca, evitam quaisquer ruídos e garantem a satisfação geral do mesmo.

Nesse sentido, em meados de 2020, o *Customer Success* começa a ganhar relevância no mercado brasileiro, quando as empresas, cada vez mais com hábitos digitais, tomam o Sucesso do Cliente como prioridade para a transformação e diferencial do negócio.

O CS alinhado à transformação digital, segundo Steinman et al., “*agrega expertise, inovação e tecnologias Data Science - como análise de big-data e sofisticada business intelligence (inteligência de negócio) para acelerar o time-to-value (prazo de entrega do valor prometido/esperado) e, em última instância, o sucesso*” (p.12, 2016).

Outro ponto relevante, é a evolução da participação do CRM (*customer relationship management* ou gestão do relacionamento com clientes) como “*uma automação da força de vendas para algo semelhante a uma plataforma de clientes, abrangendo vendas, serviços, marketing, análise de dados, aplicativos e internet das coisas (internet of things - IoT)*¹” (Steinman et al., 2016). Por isso, os dados agregados aos *softwares* de CRM transformam-se de simples análises de histórico de compras reativas para possibilitar previsões e ações proativas.

O *Customer Success* trabalha orientado por métricas que visam garantir a saúde do cliente, focando em retenção e fidelização dos mesmos, o que prolonga sua vida útil junto à organização e reduz, por consequência, as taxas de evasão e os custos de aquisição de clientes. Atualmente, algumas empresas no Brasil já aderem a esse novo conceito, porém em grande maioria dentro do segmento de tecnologia, startups e empresas de receita recorrente. O *Customer Success* é um diferencial que busca satisfação, relacionamento, soluções e, a médio prazo, estabilizar o faturamento mensal e sucesso da empresa; que, apesar do seu caráter *B2B*, deve ser otimizado e implementado em empresas cujo preocupam-se em seu desenvolvimento sustentável ao longo dos anos.

Em “*Uma perspectiva da indústria / acadêmica sobre a gestão do sucesso do cliente*”, Bryan Hochstein, Deva Rangarajan, Nick Mehta e David Kocher (2020) abordam o antigo problema de serviço, em que os clientes *B2B* deixam de perceber valor e ficam insatisfeitos, o que leva à rotatividade e redução da lucratividade da empresa fornecedora. Especificamente, concentram em como o atendimento proativo dos clientes pode reduzir as taxas de rotatividade, garantindo que os clientes realizem o faturamento mensal previsto como meta comercial, que deve ser recorrente e suficiente para a sustentabilidade do negócio. A questão da rotatividade de clientes está se tornando cada vez mais crítica à medida que as empresas fazem a transição para ofertas de serviços digitais e por assinatura, o que reduz muito os custos de troca. Nesses modelos, que dependem de receitas recorrentes, os relatórios

¹ *Internet das coisas (Internet of things - IoT): a conexão de objetos dentro do espaço físico da empresa, como sensores, softwares ou totens, conectados por meio da internet, trocando ou obtendo dados dos clientes que interagem com eles.*

da indústria demonstram que os esforços tradicionais de atendimento e suporte ao cliente não são suficientes para retê-los.

O principal fator de rotatividade de esforços crescentes de serviço, segundo os autores, é a incapacidade das empresas de se envolverem de forma proativa com os clientes, gerando ativamente valor que resultem na perspectiva de sucesso para o cliente. Para superar esse problema, o gerenciamento do sucesso do cliente, que define-se operacionalmente como “*o envolvimento relacional proativo (versus reativo) dos clientes para garantir que o potencial de valor das ofertas de produtos seja realizado pelo cliente*” (Steinman et al., 2020, p.34), está se tornando uma prática comum da indústria. O gerenciamento do sucesso do cliente é necessário porque dentro de relacionamentos dinâmicos, os clientes que não estão devidamente engajados (ou seja, ajudados na manutenção do valor conforme as ofertas se expandem e mudam) podem se sentir negligenciados, o que faz com que migrem para estados de relacionamento negativos.

Dado seu foco em reter clientes, *Customer Success* voltado ao gerenciamento de relacionamento *business to business (B2B)* está sendo rapidamente adotado por empresas que oferecem soluções de serviço completo, ou seja, aquelas que para além do produto, oferecem maneiras com que o cliente se satisfaça e atinja sucesso, como é o caso desta pesquisa, ao analisar estratégias voltadas para negócios de receita recorrente.

Para CRUZ (2021), em “*Customer success: os benefícios de uma cultura com foco no sucesso do cliente em uma indústria de alimentos do interior do Rio Grande do Sul*”, os conceitos relacionados à filosofia de *Customer Success*, consistem, basicamente, em uma cultura com foco no sucesso do cliente. *Customer Success* possui métricas que visam garantir a saúde do cliente (ou seja, seu *lifetime* como cliente ativo do negócio), focando na retenção e fidelização dos mesmos, o que prolonga sua vida útil junto à organização e reduz, por consequência, as taxas de evasão e os custos de aquisição de clientes.

Apesar de pouco desenvolvido academicamente no Brasil, como observado na Figura 01, o tema e métodos possuem forte relevância e pouco conteúdo escrito por autores nacionais, além de pouco aplicado de maneira prática no mercado.



Figura 01: Volume de publicações e pesquisas acadêmicas em temas sobre práticas de administração de clientes. Fonte: Pesquisa Popular Press (Factiva) | Academic Press (Web of Science), apud. Zeithaml, Valarie A. et al., 2020.

Para isso, o *marketing* tem contribuição ativa para alcançar os objetivos de sucesso do cliente. Define-se *marketing* como a atividade humana dirigida para satisfazer necessidades e desejos por meio da troca (Kotler, 2003), por isso conhecer o seu cliente é essencial para alcançar o objetivo de satisfazê-lo. Afinal, “quanto mais o *marketing* entende o seu público-alvo e as transformações em seus hábitos, necessidades e desejos, maior a probabilidade de saber o que pode ser oferecido [...] para que ocorra a troca” (Gabriel, 2010, p. 29).

Nesse sentido, os desejos humanos obedecem a uma hierarquia de necessidades proposta pelo psicólogo Abraham Maslow em 1943 na obra “*A teoria da motivação humana*”, como observado na figura 02:

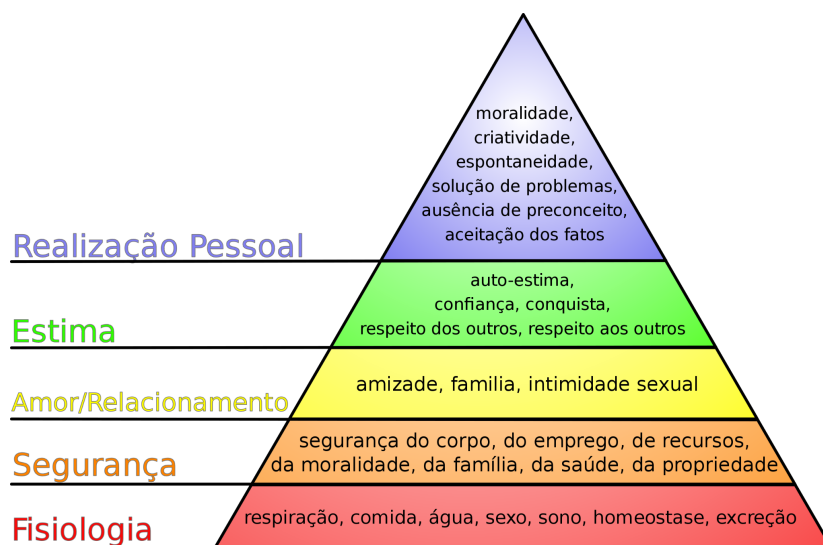


Figura 02: Pirâmide de Maslow da hierarquia das necessidades. Fonte: Wikipédia, 2022.

Ainda segundo Gabriel, “o marketing não cria necessidades [...] mas influencia e/ou cria desejos de produtos para satisfazer as necessidades” (2010, p.30). Por isso, a definição de uma estratégia de *marketing* é essencial para desenvolver inteligência em captar, reter e fidelizar clientes, sejam esses negócios *Business to Business (B2B)* ou *Business to Customer* (abreviado para B2C, referente a negócios varejistas).

Planejamento de *marketing* é, portanto, um processo de análise do ambiente - macro, micro e interno e a definição de alocação de recursos para atingimento de certos objetivos, determinando a melhor combinação entre praça, preço, produto e promoção, ajustando-os à variáveis controláveis, minimizando ou anulando ameaças e maximizando oportunidades (Gabriel, 2010).

O plano de *marketing* deve ainda contemplar os aspectos estratégicos e táticos, sendo o primeiro relacionado à análise e oportunidades e o segundo às ações operacionais determinadas, como definição de canais, contratações a serem feitas etc. Além disso, segmentar a base de clientes é essencial, pois cada público requer uma estratégia específica, afinal, cada público tem seu objetivo ao interagir com uma marca e seus produtos ou serviços, é em função deles que o plano de *marketing* deve ser implementado. (IBIDEM, 2010)

Conforme Martha Gabriel reitera em sua obra “*Marketing na Era Digital*”: “*marketing digital não existe, o que existe é marketing e o seu planejamento estratégico é o que determinará que plataformas ou tecnologias serão usadas - digitais ou não*” (2010, p.

105), por isso, a lógica que será desenvolvida neste trabalho de monografia deve ser constantemente interpretada como *marketing*, apesar da multidisciplinaridade de conhecimentos aqui tratados.

Na próxima seção, serão apresentados os fundamentos da experiência do cliente, que subsidiarão ações estratégicas e táticas que atenderão ao objetivo deste estudo.

2.1 Customer Experience Management

Segundo estudos dos psicólogos Leaf Van Boven e Thomas Gilovich publicados no *Journal of Personality and Social Psychology*, “*experiências trazem mais felicidade do que bens materiais*” (Madruga, 2021, p. 8). Neste estudo, os especialistas compreendem que existe uma diferença entre comprar um bem ou serviço e viver uma experiência, tendo em vista que essa última engaja um relacionamento, em grande parte afetivo, com o consumidor.

Dessa forma, a estruturação para integrar e elevar as experiência do cliente em canais presenciais e remotos, atualmente, são desafios vividos pelas áreas de *marketing*, negócios e operações em grande parte das empresas brasileiras. (Madruga, 2021). Em 2020, com o surgimento da pandemia de Covid-19, fez com que empresas em todo o mundo voltassem investimentos às áreas de *Customer Experience* e *Customer Success*, que contam com “*o princípio omnichannel e com estratégias proativas de relacionamento*” (Madruga, 2021, pág. 3), conforme a figura 03, que explicita as principais *expertises* envolvidas no processo:



Figura 03: *Processos vitais da empresa evoluindo para o modelo de Customer Experience e Customer Success* (Madruga, 2021, p. 3).

Esta metodologia de desenvolvimento do negócio voltado à área de relacionamento com o cliente é denominada como uma cultura *Customer Oriented Organization*, na qual a área de negócios e relacionamento se aliam para projetar o *Customer Experience*, prevendo e respeitando toda a jornada do cliente, pontos de contato (produto, compra, assistência técnica e outros), além da memória que os clientes terão desse relacionamento. (Madruga, 2021, p.6).

Esta projeção considera todos os canais de relacionamento, sejam eles presenciais ou digitais, que devem ser interpretados de forma mais ampla, também como “*medidas para otimizar a experiência dos funcionários e o estabelecimento de foco na jornada do cliente*” (IBIDEM, p. 4). Conforme a figura 04 ilustra, *Customer Experience* é a integração de diferentes *expertises* que tem como intuito proporcionar a melhor experiência, independente do canal de comunicação, levando em consideração os objetivos iniciais de captar, reter e fidelizar clientes.

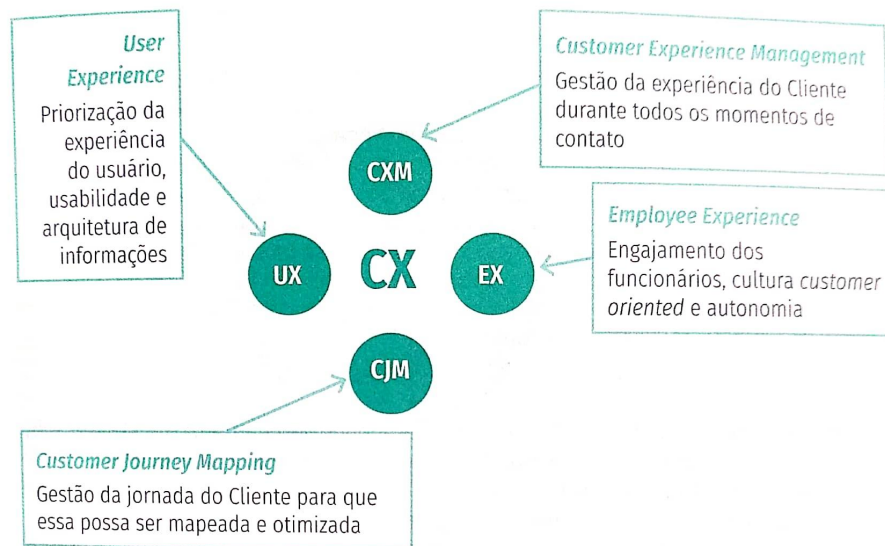


Figura 04: *CX possui a propriedade de integrar UX, CXM, CJM e EX.* (Madruga, 2021).

O objetivo de uma cultura organizacional centrada na CX, quando bem gerenciada, é o aumento da conversão em vendas, qualidade de atendimento, retenção de consumidores e maior valor do cliente (Madruga, 2021). Portanto, na prática, planejar o *Customer Experience*

tem como finalidade reduzir o índice de evasão de clientes (*churn*), aumentar a sua satisfação e engajamento, aumentar o valor comprado (estratégias *up-sell*) ou aumentar o mix de produtos comprados (estratégias *cross-sell*).

A CX pode ser considerada como um das principais contribuições para o faturamento do negócio e fidelização do cliente:

Um terço dos clientes conectados está disposto a pagar mais por uma CX excepcional, segundo uma pesquisa da Salesforce. Um estudo da PwC também mostrou que quase três em cada quatro clientes dizem que uma ótima CX vai fidelizá-los. E os consumidores pagariam um preço premium de até 16% mais por uma CX melhor (Kotler et al., 2021, p.123).

Para os autores, a experiência do cliente vai além de um atendimento ou compra, permeia pela comunicação da marca, ponto de venda, atendimento, produto, conversas com outros clientes e serviços de assistência técnica (Kotler et al., 2021). Por isso, sugerem a metodologia dos 5 As, que considera todas a jornada do cliente (ver figura 05).



Figura 05: *O caminho dos 5As do cliente.* Adaptado de Kotler et al., 2021, p.125.

Nesse sentido, a jornada ideal de todo negócio consiste em guiar o cliente entre as fases de assimilação à apologia, o que significa um cliente leal e fidelizado que mantém compras recorrentes.

O especialista em *marketing* global Gary DeAsi realizou em 2019 uma pesquisa com cerca de 700 profissionais de CX e *marketing* em todo o mundo para compreender como empresas líderes de segmento melhoram as jornadas de seus cliente, experiência e, conseqüentemente, o impacto em seu negócio. (2019, não paginado, apud. CRUZ, 2021).

Segundo o relatório nomeado “*The State of Customer Journey Management and CX Measurement*”, “87% dos entrevistados afirmaram que proporcionar uma experiência excepcional ao cliente é muito ou extremamente importante para os negócios. Apenas 2% disseram que não é tão importante” (IBIDEM, p.27), enquanto 80% afirmam que uma estratégia baseada na jornada do cliente é relevante para sua satisfação, retenção e valor de vida útil. (DeAsi, 2019, não paginado, apud. CRUZ, 2021, p.28)

Portanto, conclui-se que existe um consenso entre profissionais em todo o mundo que a estruturação da jornada de experiência do cliente contribui diretamente para o sentimento de satisfação, assim como influenciam positivamente os indicadores de retenção e prolongamento da ,vida útil.

2.1.1 Mapa de jornada do cliente

O mapa de jornada do cliente é a ferramenta que operacionaliza e guia a CX. A ferramenta considera todos os pontos de contato (do inglês, *touch points*) do cliente com o negócio e os planeja para entregar experiências satisfatórias e emocionalmente positivas (Madruga, 2021).

Segundo Madruga, 2021, o mapa de jornada do cliente é:

- Gestão integrada de todas as formas de atendimento presencial, remoto e automático;
- Atuação em tempo real para controle e aumento do *First Call Resolution (FCR)* e a taxa de resolutividade;
- Atuação para que a Jornada do Cliente seja contínua, eliminando “espaços em branco” nas chamadas e mensagens, reduzindo transferências e interrupções;
- Projeto em conjunto com áreas de negócios e gerenciamento de réguas de relacionamento ativas;
- Gerenciamento constante da oferta de experiência de qualidade e de agilidade para o Cliente por meio do controle e da integração do autoatendimento (p.36).

Portanto, a CX deve considerar a omnicanalidade (comunicação que atua de forma linear em todos os pontos de contato, sejam *online* ou *offline*). Para isso, considera-se todos os

canais de comunicação possíveis do negócio antes, durante e depois da compra (ver figura 06):

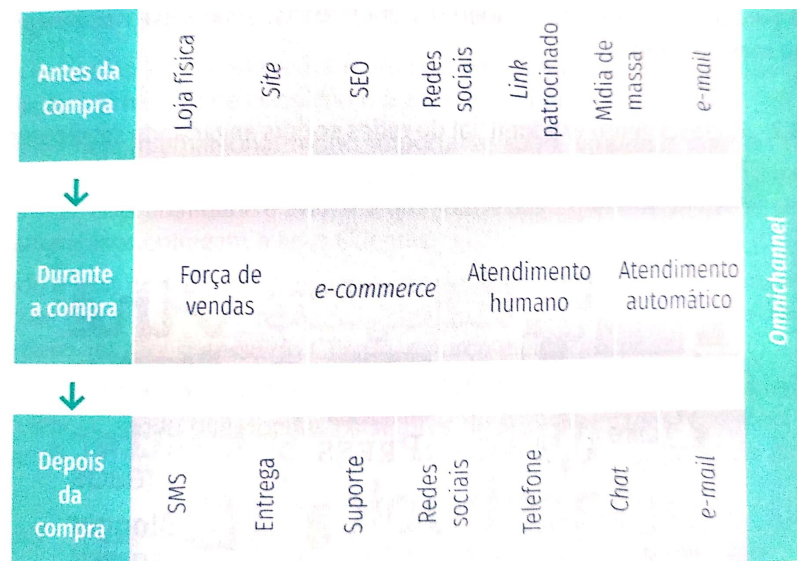


Figura 06: Diversos canais e fontes de informação para tomada de decisão antes, durante e após a compra, fazendo com que as empresas se tornem Omnichannel. Adaptado de Madruga, 2021, p.61.

Os canais de contato devem ser otimizados por meio da tecnologia. Alguns deles, como vistos na figura 06, contam com a automação de *chatbots*, e-mails, gravações que guiam o atendimento via telefone etc. Em “*Marketing 5.0*”, Kotler et al (2021, p.126) reiteram que “os papéis das pessoas e das máquinas têm importância igual em uma *CX híbrida*”. A participação de ambos os atores (tecnologia e interação humana) se complementam, afinal, a primeira inova, acelera, automatiza e possibilita decisões mais assertivas, enquanto a participação humana possibilita uma percepção sensível e ações que fogem do programado para a inteligência artificial.

2.2 Customer relationship management

O primeiro sistema de CRM é atribuído à *Siebel Systems*, empresa estadunidense fundada em 1993 por Thomas Siebel, consultor de vendas da *Oracle* (Madruga, 2021). O CRM surge como uma tecnologia que proporciona uma interface para o relacionamento entre *marketing*, atendimento e público consumidor e um complemento dos sistemas *Enterprise Resource Planning (ERP)*. Enquanto este último funciona como um *back-office* de custos internos, insumos e operações, o CRM surge como um complemento *front-office*, retendo e

cruzando dados sobre clientes, compras e possibilitando análises de perfis de consumo (Madruga, 2021).

No estudo *Customer relationship management: implementation process perspective*, os autores Mishra e Mishra reiteram que:

[...] à medida que mais organizações percebem a importância de tornar-se centradas no cliente na era competitiva atual, mais elas adotam o CRM como estratégia central de negócio [...] pode ajudar as organizações a gerenciar suas interações com os clientes de forma mais eficaz, a fim de manter a competitividade (2009, apud., Demo, Gisela et al., 2015, p. 130).

Por isso, o CRM tem como prerrogativa utilizar os dados concentrados para gerar informações relevantes acerca da base de clientes, oferecendo valor por meio de ofertas segmentadas e possibilitando relacionamentos mais duradouros (Demo, Gisela et al., 2015).

O CRM de vendas integra o histórico de clientes e as oportunidades de vendas automaticamente, fazendo com que a equipe de vendas foque apenas na atividade-fim (Kotler et al., 2021). Portanto, o CRM automatiza essas *expertises* que guiam a equipe de vendas para que foquem nos atendimentos aos clientes ideais.

O sistema de CRM possibilita, além da administração da base de dados e compras das empresas, visualizações de dados e acompanhamento em tempo real do faturamento, além de comparativos com históricos de venda, conforme pode ser observado na figura 05, que retrata a visualização interna de um *dashboard* desenvolvido dentro do *Salesforce*, um dos principais e mais avançados sistemas de CRM disponíveis no mercado.

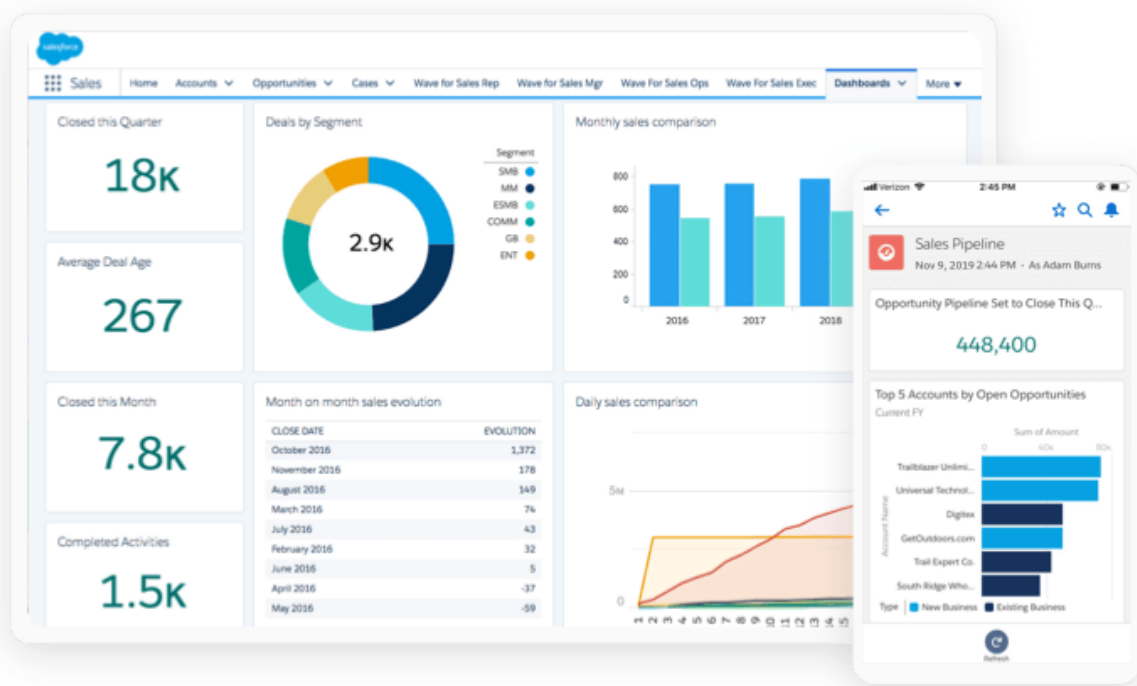


Figura 07: Interface do sistema de CRM *Salesforce*. (Fonte: [Salesforce](https://www.salesforce.com), 2022).

Apesar do amplo potencial, o CRM também é muitas vezes subutilizado, ganhando também uma designação não tecnológica no contexto brasileiro. Restringindo-se, por vezes, para traçar envios de comunicação para o cliente, seja digital (quando feito a partir de e-mails, sms, *whatsapp* ou redes sociais), seja humanizado (por meio do contato de um atendente), esta tecnologia é desenvolvida em perspectiva *omnichannel*, possibilitando todo o mapeamento da jornada do cliente (Madruga, 2021).

Ainda, realizar previsões ainda não é comum em muitas empresas, pois a maior parte das equipes de vendas baseiam-se na intuição, o que não é uma estratégia assertiva, se considerar-se o todo (Kotler et al., 2021). Nesse sentido, apesar das funcionalidades do sistema de CRM, que agrupa dados e disponibiliza informações, ainda sim existe um trabalho humano em interpretá-los e providenciar ações, como uma ligação do time de atendimento, um e-mail *marketing* programado com ofertas, entre outros.

A aplicação do CRM em conjunto com as prerrogativas do *marketing* preditivo permite que a equipe de *marketing* e vendas atuem com previsões mais precisas, priorizando as melhores oportunidades em vendas ou recuperação de clientes (Kotler et al., 2021).

Nesse sentido, o CRM deve ser implementado em etapas com foco na experiência do consumidor. Nesta monografia, trata-se de uma proposta com foco na experiência da empresa consumidora (como é o caso de negócios *B2B*). A principal entrega que o negócio deve entregar para satisfazer o cliente, nesta perspectiva portanto, é prospectar seu sucesso e alavancar seu faturamento (Steinman et al., 2020).

Sucesso, para uma empresa que contrata uma funcionalidade ou produto de outra, é que aquele produto otimize suas vendas, afinal. O *RD Station*, por exemplo, possibilita aos seus clientes o envio de *e-mail marketing* e criação de *landing pages* com intuito de que seus usuários (clientes) fechem mais vendas.

Tendo como exemplo as empresas atendidas pelo *RD Station*, assim como as revendedoras *Natura*, possuem pontos em comum: ambas desejam alcançar sucesso em suas vendas e ambos proporcionam receita recorrente (compras periódicas) à empresa fornecedora. Portanto, o faturamento e orçamento mensal de empresas como as citadas depende diretamente da recompra periódica de suas soluções sem a formalização de um contrato.

Por isso, estratégias de CRM integradas à perspectiva de *customer success* garantem recompra, aumento da vida útil do cliente (*CLV*) e evitam a inativação (evasão ou *churn*). Nesse sentido, na próxima seção, introduz-se a cultura CS, que proporciona uma metodologia com foco em fidelização e retenção de clientes, aqui aplicadas ao contexto de empresas *B2B* e na seção 2.5.1 introduz-se essa perspectiva relacionando-a ao *marketing* preditivo e seus benefícios.

2.3 Customer success

Segundo Steinman et al., “*customer success é uma cultura organizacional que foca na experiência do cliente, com o objetivo de maximizar a retenção e CLV. Somente se essas tarefas forem executadas com eficácia, a empresa por assinatura será capaz de sobreviver [...]*” (2020, p.39). Portanto, investimentos em CS são essenciais para reduzir e gerenciar a evasão de clientes, promover o aumento do valor de compra com os atuais clientes e melhorar a experiência e satisfação dos mesmos.

Investimentos na saúde do cliente (aqui tratada sob a perspectiva da administração, na qual uma “boa saúde” representa o período ativo em que o cliente faz transações com a empresa), por certo, existem na maior parte das empresas, contudo, para as estratégias serem executadas com maestria, exigem competências profissionais e metodológicas específicas

(IBIDEM, p.46). *Customer success* também pode ser compreendido, de forma mais ampla, como uma metodologia para o processo pós-venda.

Por definição, *CS* não deve ser CRM ou suporte ao cliente. É um setor organizacional indutor de receita, com indicadores voltados ao sucesso do cliente, focado na análise de dados e preditivo, prevendo quem procurar ao invés de ser procurado, conforme exposto na tabela 01:

	<i>Customer success</i>	Suporte ao cliente
Finanças	Indutor de receita	Centro de custos
Ação	Proativo	Reativo
Métricas	Voltado para o sucesso	Voltado para a eficiência
Modelo	Focado na análise de dados	Intensivo em pessoas
Objetivos	Preditivo	Responsivo

Tabela 01: Diferenças entre *customer success* e suporte ao cliente. (Adaptado de Steinman, Murphy e Mehta, 2020, p.50).

Conforme demonstrado por Steinman, Murphy e Mehta na figura 08, a maior parte receita de uma empresa tende, ao longo dos anos, depender da base já instalada de clientes fidelizados, enquanto o faturamento advindo de novos clientes representa um percentual menor, enquanto captações de novos negócios exigem, rotineiramente, maiores esforços em tempo de prospecção e dinheiro.

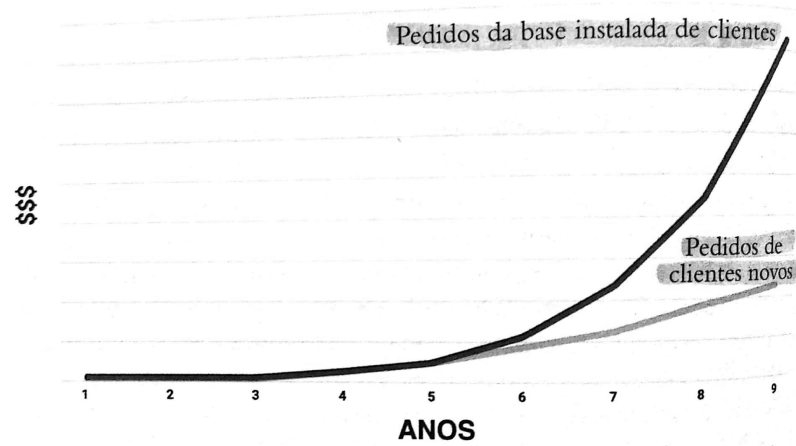


Figura 08: *Pedidos da base instalada de clientes versus pedidos de novos clientes.* (Steinman, Dan et al. 2020, p.56).

Portanto, a metodologia do sucesso do cliente aliado ao *customer experience* propõe o acompanhamento da saúde do cliente, orientado a dados, assim como mapeamento da sua jornada e todos os pontos de contato.

Segundo Gary DeAsi, no artigo “*Why the customer journey is your new marketing funnel*”, após passar pelas etapas do tradicional funil de vendas, o consumidor passa para um novo funil, agora invertido, que considera as etapas de pós-venda, como descrito na figura 09 (DeAsi, 2016, não paginado, apud. CRUZ, 2021, p.30).

Compreendendo que nem todos os clientes seguem a mesma jornada e têm interações com os mesmos pontos de contato, o novo funil com 10 estágios possibilita a visão geral de etapas que devem ser mapeadas na estratégia de sucesso.

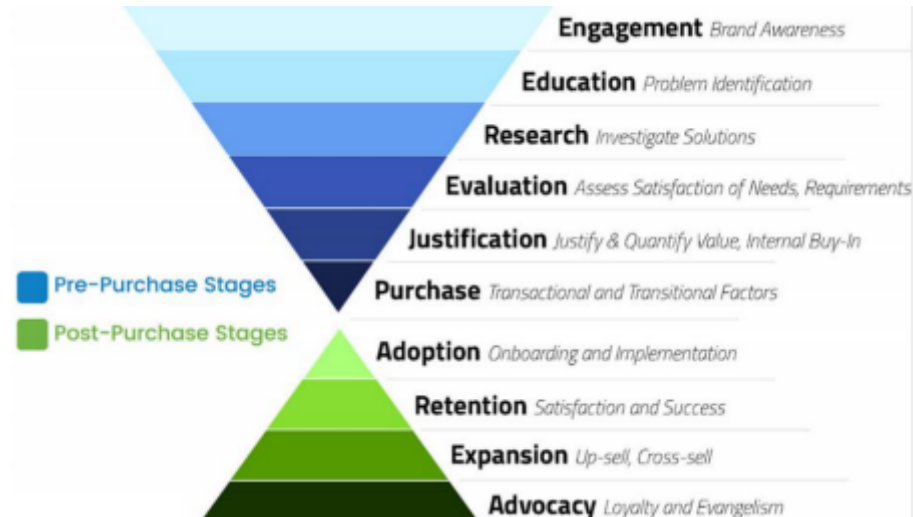


Figura 09: *New Marketing Funnel*. Adaptado de DeAsi, 2016, não paginado, apud. CRUZ, 2021, p.30.

Nesse sentido, cada estágio do funil é uma etapa da jornada do cliente:

- Engagement (Engajamento): consiste em lançar uma ampla rede e agregar valor ao público. O objetivo é que os clientes encontrem a marca/produto/serviço;
- Education (Educativa): o objetivo desse estágio é que os clientes percebam que têm um problema para o qual precisam de solução;
- Research (Pesquisa): objetiva explicitar os principais benefícios, capacidades e diferenciadores do produto/serviço, reafirmando o valor e a necessidade do mesmo para o indivíduo;
- Evaluation (Avaliação): a partir deste estágio, o foco está no produto/serviço. Aqui fala-se de proposições de valor, recursos, capacidades e especificações técnicas. Essa fase pode ser determinada pela interrogativa: “Este produto resolve meus problemas e atende às minhas necessidades e exigências específicas?”;
- Justification (Justificativa): nessa fase, o questionamento passa a ser: “Por que eu realmente preciso disso agora?”. E então inicia-se a análise sobre o valor que o cliente e potenciais envolvidos percebem sobre produto ou serviço. Pode haver sete a vinte pessoas envolvidas em qualquer decisão de compra;
- Purchase (Compra): esse estágio é grande e muitas vezes pode ser subestimado. Conforme o processo de compra se aproxima, algumas novas perguntas, preocupações, objeções e bloqueadores podem surgir. A partir desse estágio, é finalizada a fase “pré-compra” e inicia-se a fase “pós-compra”;
- Adoption (Adoção): primeiro estágio da “pós-compra”, que é caracterizado pela integração e implementação do produto/serviço;
- Retention (Retenção): nesse estágio é necessário garantir que os clientes estejam satisfeitos, atingindo suas metas com o produto/serviço. Segundo DeAsi (2016), “Os clientes devem ver que você realmente se importa e deseja ajudá-los a ter sucesso. Quanto mais pudermos agregar valor e mostrar amor ao cliente aqui, mais nos prepararemos para que o cliente devolva o amor nos estágios de Expansion e Advocacy.”. Conforme Mehta, Murphy e Steinman (2018, p. 72), “Este é o cerne do customer success.”;
- Expansion (Expansão): caso a operação tenha obtido sucesso em todos os estágios anteriores, nesse estágio é possível já ter obtido a confiança do cliente. O trabalho do time de Sucesso do Cliente é atuar para que o cliente perceba claramente os benefícios e volte a procurar a empresa em caso de novas necessidades ou aumento de demanda;
- Advocacy (leia-se Lealdade): um advogado da marca pode ser definido como um cliente que fala favoravelmente sobre uma marca ou produto e transmite

mensagens positivas (“boca a boca”) sobre a marca para outras pessoas (DeAsi, 2016, apud. CRUZ, 2021, p.30-31).

Para isso, são definidas três frentes para otimizar o esforço com a base de clientes, considerando seu valor, são elas: *high touch*, *low touch* e *tech touch*, estratégias de atendimento ao cliente que obedecem uma hierarquia pelo seu valor para o negócio (ver figura 10).



Figura 10: *Hierarquia do valor do cliente*. Adaptado de Steinman et al., 2020, p.68.

A primeira, idealizada para clientes de alto *ticket* médio (valor monetário por compra efetuada), defende uma comunicação de “alto toque”, humanizada e com maior frequência. A segunda, voltada para clientes de *ticket* médio intermediário, defende um acompanhamento humanizado espaçado em períodos trimestrais e segmentado por grupos de clientes, mas também digital, para trabalhar a constância da satisfação desse nicho de consumidores. Por último, comumente a maior parte da base de clientes, *tech touch* define uma comunicação rápida, por meio de canais digitais e muitas vezes automatizado, para acompanhamento dos clientes de baixo *ticket* médio, por intermédio de e-mails, webinars, comunidades, grupos de usuários e/ou eventos, por exemplo. (Steinman et al., 2020).

Portanto, o *customer success* é aplicável a qualquer escala de empresa ou base de clientes, afinal, existem segmentações que, focadas em análise de dados e orientadas ao sucesso, otimizam esforços com comunicação e geram satisfação.

2.3.1 *Ciclo de vida do cliente*

A jornada prevista na estratégia de CX ganha mais especificidade quando integrada às estratégias de *customer success*. Dessa perspectiva, a estratégia de CS considera cinco etapas gerais de pós-venda em que os clientes seguem para alcançar seu sucesso e satisfação, são elas: *onboarding*, *energização*, *proteção*, *reconquista* e *recuperação* (ver figura 11).

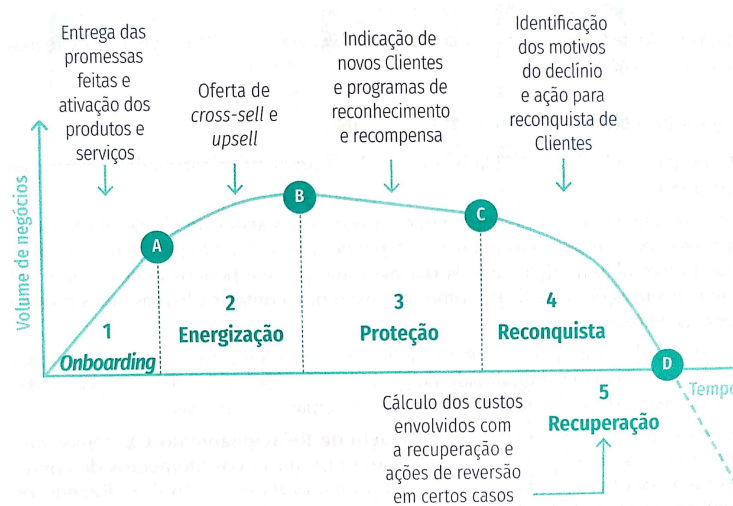


Figura 11: O *Customer Lifecycle* do Cliente faz parte do modelo de relacionamento. Na parte de cima do diagrama estão as ações comerciais potenciais da empresa. Adaptado de Madruga, 2021, p.53.

Estas etapas são denominadas de “ciclo de vida do cliente” ou *customer lifecycle*, em inglês, e consideram o planejamento de *marketing*, vendas e relacionamento para cada etapa da vida do cliente junto ao negócio.

- *Onboarding*: a etapa de *onboarding* ou boas-vindas, em livre tradução, prevê a jornada inicial na qual a empresa entrega o valor prometido na fase de captação do novo cliente. Há um acompanhamento especial desse cliente, afinal ele está em fase de experimentação do produto/serviço. (Madruga, 2021)
- *Energização*: após período inicial definido pela estratégia do negócio, o cliente entrará em uma fase de energização, na qual está pronto para expandir seu relacionamento com a empresa. Aqui, são consideradas ofertas de *cross-sell* e *upsell*. (IBIDEM, 2021)
- *Proteção*: nesta etapa, o cliente atingiu um nível ideal de aquisição de produtos/serviços e segue para uma jornada de fidelização. (IBIDEM, 2021)

- *Reconquista*: a fase de reconquista acontece quando o cliente esfria em relação à sua frequência e faturamento de compras. (IBIDEM, 2021) Esta fase tem como intuito tratar de maneira preventiva uma possível evasão e deve ser monitorada de forma preditiva, conforme se verá adiante na seção 3.1 que apresenta estratégias de *marketing* preditivo (Kotler et al., 2021).
- *Recuperação*: os esforços na etapa de consideração consistem em análises, por parte dos gestores, dos clientes evasores e avaliações, caso a caso, de possíveis negociações para o retorno dos mesmos (Madruga, 2021).

Segundo o artigo “*Customer Journey – Quando, como e porque mapear a jornada de seus clientes*” (2018) de Eduardo Tavares, autor do blog *Customer Success Brasil*, empresas brasileiras iniciaram investimentos no mapeamento da CX, contudo com falhas de implementação e operação:

Na maioria dos casos, após esse mapeamento inicial algumas coisas vêm à tona, sendo as mais comuns:

- Descobri que depois do *onboarding* nada é feito com o cliente;
- Estamos mandando muito e-mail para o cliente;
- Estamos tratando todos os clientes da mesma maneira;
- Atualmente todas as ações que temos são reativas; (TAVARES, 2018, não paginado, apud., CRUZ, 2021, p.32).

Portanto, reforça-se que as estratégias de comunicação devem ser traçadas considerando os períodos do ciclo de vida do cliente e a hierarquia de valores dos mesmos. Dessa forma, a CX se torna funcional, operacionalizável e mapeável e o valor de sucesso para o cliente vai sendo percebido no desenvolvimento de cada uma das etapas seguidas.

A seguir serão apresentados indicadores de sucesso voltados para a estratégia de CX, que compreendem o engajamento e satisfação da base de clientes.

2.3.1.1 NPS

Net Promoter Score (NPS) é um indicador que mede o engajamento e excelência da experiência da base de clientes.

Segundo Madruga, “o *Net Promoter Score* é um mecanismo prático utilizado para entender o que os Clientes querem e voltar toda e empresa para corresponder efetivamente à demanda” (2021, p.146). O NPS consiste em um feedback pós-venda no qual o cliente

responde à uma única pergunta: “Em uma escala de 0 a 10, o quanto você recomendaria nossa empresa a um amigo ou familiar?”.

A partir das respostas colhidas, divide-se os respondentes entre cliente promotores (aqueles que pontuam a experiência entre 9 e 10), clientes neutros (aqueles que pontuam entre 7 e 8) e detratores (os que pontuam entre 0 a 6). Em que:

- Promotores: são aqueles clientes que avaliam com notas entre 9 e 10 e representam a parcela de clientes leais, que recomendam o produto/serviço à amigos e familiares através do marketing boca-boca;
- Neutros: são aqueles clientes que avaliam com notas entre 7 e 8 e representam clientes passivamente satisfeitos, os quais podem facilmente migrar para outra categoria (positiva ou negativamente);
- Detratores: são aqueles clientes que avaliam com notas entre 0 e 6 e representam um fracasso grave na operação. (CRUZ, 2021, p.53).

Madruga propõe uma tabulação das respostas recebidas dadas pelo cálculo: “*Quantidade de promotores - Quantidade de detratores / Total de respondentes*” (2021, p.150). O resultado da análise deve ser comparado com o *benchmark* do segmento e pode ser classificado como:

- Abaixo de 0 ponto: **Crítico** (requer intervenções urgentes e importantes);
- De 0 a 49 pontos: **Aperfeiçoamento** (requer intervenções importantes);
- 50 a 74 pontos: **Qualidade** (requer esforços para elevação do índice);
- 75 a 100 pontos: **Excelência** (requer esforços para manter o índice). (Madruga, 2021, p.151).

A pergunta principal também pode vir acompanhada de uma pergunta secundária, que aborda o motivo pelo qual o indivíduo deu a nota inicial (Madruga, 2021).

Segundo o artigo *Building loyalty at B2B Companies*:

Os clientes B2B que são "promotores" têm um valor médio da vida útil normalmente três a oito vezes o valor dos "detratores", dependendo do segmento e do setor. Os promotores ficam mais tempo com a empresa, compram mais produtos, geralmente custam menos para servir e têm maior probabilidade de encaminhar o fornecedor a colegas e amigos (Bain & Company, 2019a, não paginado, apud., CRUZ, 2021, p.56).

Já no artigo *Reliable, Trusted Currency/Metric* afirma que:

[...] os clientes promotores aumentam suas compras mais rapidamente que os detratores, são menos sensíveis ao preço e exigem menos em vendas, marketing e

custos de publicidade do que outros clientes (BAIN & COMPANY, 2019b, não paginado, apud., CRUZ, 2021, p.56).

Portanto, a medição do NPS funciona como uma rápida perspectiva de excelência ou probabilidade de evasão do cliente, mas deve ser analisado em conjunto com demais dados para uma leitura positiva e correta, sendo um passo essencial dado em rumo à melhoria da CX.

Na seção a seguir, serão apresentados os principais indicadores da estratégia de CS voltados ao valor do cliente, que serão o ponto chave para a implementação do *marketing* preditivo.

2.3.2 Valor do cliente

Valor do cliente, ou *customer value* em inglês, para Madruga está “*intimamente ligado aos benefícios gerados pelo processo de relacionamento [...] desde que conduzidos por profissionais que dominem o CX*” (2021, p.186). Este valor é um processo contínuo, que não deve ser compreendido apenas como a receita gerada pelo cliente, afinal empresas com foco na CX dão oportunidade para que os clientes participem das decisões, gerando valor por meio da interação com a linha de frente do negócio (IBIDEM, 2021).

No artigo “*The commitment-trust theory of relationship marketing*”, sobre a relação de troca entre empresa e consumidor, considera que compartilhar valores é fundamental para geração de confiança entre empresa-cliente. Afinal, os clientes avaliam, julgam, orientam discussões e guiam ações em prol do sucesso da organização (Morgan; Hunt; 1994, apud., Madruga, 2021).

Nesse sentido, pode-se identificar diferenças culturais das empresas, principalmente quando considera-se que as estratégias das empresas B2B, foco do estudo, tendem a focar em metas de produtividade, eficiência, qualidade do produto, gerenciamento de estoque, entre outros, enquanto o relacionamento do cliente se torna reativo (CRUZ, 2021).

Por isso, Madruga (2021) diferencia características, em perspectiva cultural, entre empresas que tem foco voltado ao produto de empresas *Customer Centric*:

Empresas com foco em si mesmas	Empresas com foco no valor do Cliente
Crença na eficiência da produção, na redução de custos e no produto	Identificação e monitoramento dos valores que são importantes no relacionamento
Busca resolver questões de estoque excedente por meio de promoções agressivas	Utiliza ferramentas com os Clientes para prever eficientemente a demanda
Os produtos são projetados e comercializados sem a participação do Cliente	O cliente é envolvido em diversas etapas do produto: da concepção à customização
O relacionamento com o Cliente é visto como mero contato do produto com o Cliente	O relacionamento é priorizado como um grande diferencial competitivo com a criação de valores compartilhados
A recompra a curto prazo é a maior garantia que a estratégia foi bem-sucedida	O engajamento e a longevidade indicam o sucesso das estratégias

Tabela 02: *Diferença entre a empresa com foco em si mesma e empresa com Foco no Valor do Cliente.* Adaptado de Madruga, 2021, p.187.

Uma cultura *Customer Centric* apresenta benefícios mútuos para a cliente e empresa. Na perspectiva do primeiro, há o engajamento, experiências positivas, satisfação com produtos e atendimento, economia de tempo e dinheiro, além do sentimento de participação ativo e menos reclamações. Na perspectiva do segundo, há o resultado em vendas, retenção dos clientes, indicações de novos negócios, equipe interna envolvida e eliminação de custos e tomadas incorretas de decisões (Madruga, 2021).

Nesse sentido, apresenta-se a seguir como medir o valor do cliente por meio de indicadores estratégicos sugeridos pelo CS.

2.3.2.1 Customer lifetime value (CLV)

Para Madruga, o valor de vida do cliente ou valor vitalício é “*uma representação, ao longo do tempo, da lucratividade que um Cliente pode fornecer à empresa mediante trocas constantes*” (2021, p.189). Segundo o autor, o indicador nomeado *lifetime value (LTV)* vem sofrendo uma tendência de mudança de nomenclatura para *customer lifetime value (CLV)*, que inclui o termo “cliente” na expressão, contudo pode ser encontrado na literatura de ambas as formas (Madruga, 2021).

O CLV considera a média de compras anuais, a margem de lucro média, os custos com *marketing e vendas*, taxa média de desconto, valor médio da compra, a taxa de aquisição de novos clientes e a taxa média de retenção dos mesmos (Madruga, 2021). Portanto, é um indicador de valor monetário a longo prazo, que possibilita otimização das estratégias de comunicação e vendas com base no mapa de jornada do cliente, como visto no tópico 2.2.1 desta monografia.

Frederick Reichheld, na obra “*Princípios da lealdade*”, defende ainda que os lucros por cliente podem aumentar anualmente e que podem ser compreendidos em cinco frentes lucrativas: “*lucros por melhor preço, lucro por indicações, lucro por redução de custos (aprendizado), lucro por aumento de compras (cross e upsell) e lucro original do produto/serviço*” (Reichheld, 2002, apud., Madruga, 2021, p.191).

Por isso, aumentar a longevidade do cliente por meio da estratégia de relacionamento é essencial para a sustentabilidade monetária de negócios B2B.

Segundo Gallo (2014) no artigo *How valuable are your customers?*, existem várias maneiras de se calcular o CLV, contudo a matemática que processa os modelos estatísticos é complexa (Gallo, 2014, apud., CRUZ, 2021). A autora criou uma calculadora de CLV, conforme observado na figura 12:

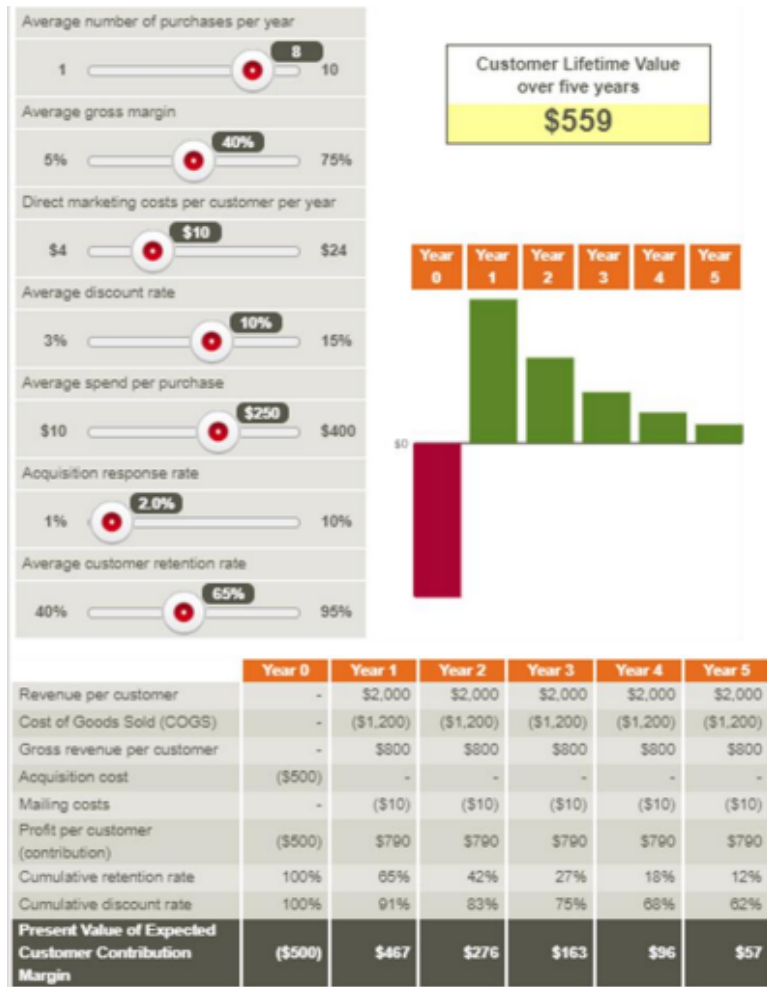


Figura 12: Calculadora de *lifetime value*. (Fonte: Gallo, 2014, não paginado, apud., CRUZ, 2021, p.40)

Ainda sobre o cálculo do CLV, Otavio Andrade, no artigo *LTV: o que é e como calcular o Lifetime Value* (2018) sugere um cálculo simplificado do indicado, conforme indicado na figura 13:

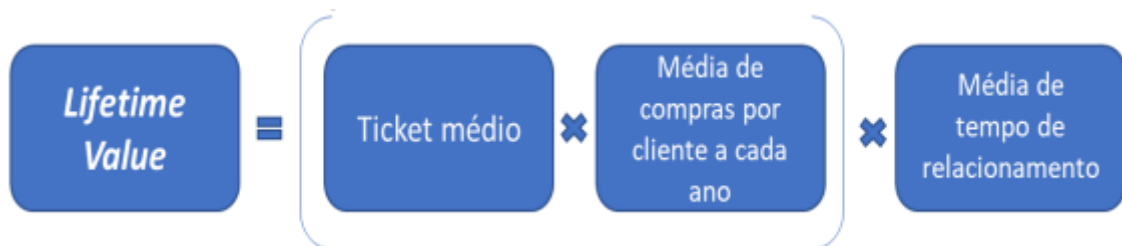


Figura 13: Cálculo simplificado de LTV. Fonte: Adaptado de Andrade (2018, não paginado), apud., CRUZ, 2021, p.41.

As possibilidades de cálculo do CLV existentes são limitadas e voltadas para o tratamento *small data* (pequena quantidade de dados), tendo em vista que a primeira, apresentada por Amy Gallo, prevê o indicador para um cliente específico, enquanto o método proposto por Otavio Andrade possibilita, de forma simplista e não exata, a medição do indicador.

Esta monografia concentra-se na modelagem matemática do CLV, possibilitada por meio de soluções da ciência de dados para *big data* (manipulação de grandes bases de dados), conforme o objetivo específico e que será apresentado no Capítulo 3 desta pesquisa.

Existem, ainda, uma série de técnicas para implementação do indicador de CLV, que consideram valores emocionais e racionais do cliente, o *Service Level Agreements* (que estabelece processos de trabalho e mapeamento dos esforços para a equipe interna) e diretrizes para a diretoria e presidência.

Essas soluções de procedimento não serão exploradas no presente estudo, tendo em vista que o foco do mesmo são os modelos matemáticos envolvidos que possibilitam a estratégia de uma perspectiva orientada a dados.

2.3.2.2 Churn

Segundo Steinman et al., *churn* são “todas as ações que reduzem a receita recorrente anual da empresa [...] evasão é simplesmente o valor monetário [...] que foi perdido” (2020, p.37). Portanto, *churn* é toda ação na qual clientes abandonam a empresa, ou conseguem grandes parcelas de descontos para reativação, ou compram menos produtos no período. Dessa forma, para controlar e gerenciar o *churn* é necessário o desenvolvimento de técnicas para retenção dos clientes mais rentáveis do negócio (CRUZ, 2021).

Churn rate é a denominação para a taxa percentual que aponta o abandono de clientes em um determinado período e pode ser calculada ao “dividir o número de clientes que deixaram a empresa em um determinado período pelo número de clientes no início desse período, multiplicando por 100 para que a taxa se apresente em porcentagem” (BALIEIRO, 2019, apud., CRUZ, 2021, p.49).

$$\text{CHURN} = \frac{\text{NÚMERO DE CLIENTES QUE CANCELARAM O SERVIÇO NO MÊS}}{\text{NÚMERO DE CLIENTES NO INÍCIO DO MÊS}}$$

Figura 14: Cálculo de *Churn*. Fonte: Wagner, 2015, não paginado, apud., CRUZ, 2021, p.49.

Outra métrica relevante ao considerar a evasão é o *churn* de receita ou *MRR Churn Rate* (abreviação de *Monthly Recurring Revenue Churn Rate*, que se refere ao modelo de receita recorrente mensal), que é definida “como o montante monetário que se deixou de ganhar naquele mês em função dos clientes desistentes (clientes que “deram” *churn*). Para isso, basta multiplicar o número de clientes perdidos pelo seu ticket médio” (Balieiro, 2019, apud., CRUZ, 2021, p.49).

O *MRR Churn Rate* também pode ser medido em percentual, dado pelo cálculo: “*MRR CHURN % = (MRR Churn/MRR último mês) x 100*” (Klemz, 2018, apud., CRUZ, 2021, p.50). Portanto, enquanto o *Churn rate* indica a quantidade de clientes que se evadiram em um período, o *MRR Churn rate* indica o valor monetário perdido no período por esse montante de clientes. (CRUZ, 2021).

Para Steinman et al., uma taxa anual entre o intervalo de 5% a 7% representa um cenário saudável para a empresa (2020). Contudo, Klemz reitera que “É quase impossível estimar um valor que seja verdadeiro para todo tipo de negócio. Em segmentos específicos temos alguns *benchmarks* [estudos comparativos com outros negócios do mesmo segmento” (2018, apud., CRUZ, 2021, p.51).

O objetivo de monitoramento do *churn* consiste em identificar o cliente que tem a probabilidade de abandonar o relacionamento com o negócio, compreendendo qual deles justificam financeiramente o esforço para ações de retenção e a compreensão do motivo pelo qual os clientes lucrativos resolvem sair (Ikeda, 2006, apud., CRUZ, 2021). Por isso, é necessário prever quais são os clientes com maior propensão a abandonar o negócio e definir ações estratégicas proativas para evitar a evasão (Neslin et al, 2006, apud., CRUZ, 2021).

Nesse ponto, a análise de dados tem um papel importante, pois é uma maneira de identificar padrões de comportamento aos clientes prestes a cancelar o contrato (Ward e Ryals, 2001, apud., CRUZ, 2021). Nesse sentido, o CRM tem uma participação ativa e

ampliada, pois deixa de analisar eventos passados e passa a trabalhar com a predição de eventos futuros (Neslin et al., 2006, apud., CRUZ, 2021), por meio de modelos estatísticos.

Com o amadurecimento do negócio, assim com a estratégia voltada para a CX e gerenciamento do *churn rate*, “a empresa atinge o patamar ideal de *churn* e é possível focar mais esforços no *up-selling*: vender mais para os clientes atuais” (Wagner, 2015, apud., CRUZ, 2021, p.52). Dessa forma, o negócio é levado para atingir o *negative churn*, que segundo o autor, é quando a estratégia de *up-selling* supera a taxa de evasão (ver figura 15).



Figura 15: *Negative churn*. Fonte: Wagner, 2015, não paginado, apud., CRUZ, 2021, p.52).

Este é o equilíbrio ideal para negócios sustentáveis. Nesta monografia, serão apresentadas soluções possíveis por meio da ciência de dados para decisões proativas e antecipadas para o gerenciamento do *churn*. Dessa forma, serão utilizados modelos matemáticos que possibilitam a previsão em larga escala da base de contatos de um negócio de receita recorrente.

Na próxima seção, contextualiza-se *expertises* da área da ciência de dados que possibilitarão análises preditivas que são aliadas intrínsecas da estratégia de sucesso do cliente.

3 Data science aplicado ao sucesso do cliente

Assim como o *marketing*, a ciência de dados contribui, por meio de análises de bancos de dados, com o conhecimento acerca do consumidor, seus comportamentos e preferências, assim como aspectos pessoais. Essa *expertise* dá-se o nome de engenharia social, que pode definir-se como “esforços para influenciar atitudes e comportamentos dos indivíduos através da obtenção e análise de dados pessoais deles” (Gabriel, 2010, p.95). Dentre todas as linguagens de programação disponíveis, *Python* é uma das quais é

considerada de alto nível, interpretada de *script*, imperativa e orientada a objetos, funcional, de tipagem dinâmica e forte, que foi desenvolvida por Guido van Rossum, em 1991. (BORGES, 2014).

Python possui uma série de bibliotecas com funções específicas para cada finalidade, sejam modelagens estatísticas para engenharia, sejam análises voltadas à área de *marketing*.

As bibliotecas de ciência de dados que atendem aos requisitos de *marketing* são chamadas de *Buy 'Til You Die Models: BTYD*², usada para “*modelar as características de compra e hábitos dos clientes, que é então usado para prever o valor de vida útil do cliente*” - CLV ou LTV (Ganti, 2019, p.2, traduzido pela autora).

Os modelos de rotatividade de clientes que encontram-se com mais frequência na ciência de dados são aplicáveis à configurações contratuais, nas quais os clientes assinam contratos de prazo limitado com provedores de internet, redes de telefonia móvel, entre outros serviços baseados em assinatura. As empresas em contextos contratuais podem desenvolver uma visão mais clara dos clientes que têm propensão a *churn*, por exemplo.

A análise preditiva permite que as empresas antecipem o movimento do mercado e clientes antes que ele ocorra (Kotler et al., 2021). Até então, profissionais e empresas tentam prever de forma intuitiva resultados para suas estratégias com base nos históricos de compra, enquanto na análise preditiva, a maior parte do trabalho é realizada pela inteligência artificial (IA) e máquinas de aprendizado (IBIDEM, 2021).

Para empresas contratuais, os dados brutos de seus contratos mostram um risco elevado sempre que um prazo de contrato chegar ao fim. As equipes de vendas e *marketing* costumam entrar em contato com um cliente antes que ele comece a pensar em renovar ou cancelar sua assinatura.

² *Buy Til You Die Models (BTYD)*: do inglês, “Modelos Compre Até Morrer”. São modelos estatísticos lineares e distributivos aplicados ao *marketing* que analisam a probabilidade de evasão do cliente.

Em contrapartida, em um ambiente de negócios não contratual (receita recorrente) esse tipo de *insight* não é claro. As equipes de vendas e *marketing* não podem prever o risco elevado de um cliente em seus calendários. A maioria dos modelos de classificação que normalmente aplicam-se a problemas de rotatividade de clientes em configurações de negócios contratuais não são aplicáveis a configurações não contratuais. As datas de *churn* não são transparentes.

Dessa forma, se analisa o comportamento de compra dos clientes fazendo suposições distributivas. Um modelo de probabilidade também permite prever o volume de compras. Para Kotler et al. (2021), “*os dados passados são alimentados em um motor de aprendizado de máquina para revelar padrões específicos, o que é chamado de ‘modelo preditivo’ [...] o profissional de marketing pode rever desfechos futuros, como quem tem mais probabilidade de comprar [...]*” (p.260).

Por isso, as empresas conseguem traçar estratégias mais proativas e otimização de investimentos futuros, descobrindo quais clientes, apesar de novos, poderão se tornar valiosos, fazendo com que a decisão de investir recursos no crescimento de clientes específicos é mais assertiva (IBIDEM, 2021), possibilitando estratégias cada vez mais segmentadas e únicas para cada cliente.

3.1 Marketing preditivo

Segundo Kotler et al. (2021), “*a análise preditiva usa e examina dados históricos*” (p.160), o que possibilita a descoberta, por meio do *big data*, padrões que orientam ações e propostas que influenciam o resultado de forma preventiva ou proativa, o que dialoga com Steiman et al. (2020), porque reitera que “*a capacidade de predição [...] vai mais longe - ao prever a quem procurar antes de ser procurado*” (p.54).

A análise preditiva possui soluções para gerenciamento do cliente, produtos e lançamentos, assim como gestão da marca (ações de *branding*), conforme é possível observar na figura 16:



Figura 16: Aplicações do *marketing* preditivo. Adaptado de Kotler et al., 2021, p.261.

Denominado *customer equity*, a gestão de clientes tem a possibilidade de medir o valor vitalício do cliente (*lifetime value*) considerando períodos futuros. Nesse sentido, o intuito das tecnologias da ciência de dados é preparar *insights*, segmentações de clientes de perfis semelhantes e comportamentos baseados nos históricos de compras para que o setor de *marketing* tenha a possibilidade de traçar estratégias que conduzam a jornada do cliente à ofertas *up-selling* ou *cross-selling*. (Kotler et al., 2021)

A figura 17 ilustra “*A pirâmide da gestão de conhecimento*”, que defende que qualquer estratégia de experiência do cliente aliada a dados não descarta a intervenção humana.



Figura 17: *A pirâmide da gestão do conhecimento*. Adaptado de Kotler et al., 2021, p.128.

A partir de perfis de CLV, estratégias *one to one* (segmentadas para clientes específicos), ações comerciais, ações e canais de relacionamento *high touch*, *low touch* ou *tech touch* podem ser melhor planejados para que a experiência ideal satisfaça o cliente conforme seu valor para o negócio e faça com que esses clientes evoluam ao longo do tempo.

Corroborando com as estratégias de sucesso do cliente, considera-se que “*clientes com contribuição mais lucrativa terão acesso a uma equipe exclusiva de assistência ao cliente, ao passo que os demais terão acesso a uma interface digital automatizada*” (IBIDEM, p.163).

O profissional de *marketing* precisa ter uma compreensão das ideias por trás dos modelos preditivos, afinal o projeto deve ser direcionado pelo setor de *marketing* que deve possuir habilidades em tecnologia e utilizar setores parceiros, como o setor de Tecnologia da Informação (TI), como um suporte operacional e não líder do processo (IBIDEM, 2021).

3.1.1 Biblioteca *lifetime* e modelos matemáticos

Neste estudo, portanto, apresenta-se a *lifetime*, biblioteca *Python* exclusiva para análises de negócios de compra recorrente e que possibilita, por exemplo, a predição da frequência de transações ou quantidade de clientes evasores durante um período. A *lifetime* utiliza-se de dois modelos matemáticos avançados: o modelo RFM, *Beta-Geometric/Negative Binomial (BG/NBD)* combinado com o modelo *Gamma-Gamma* para estimar o valor monetário da vida útil do cliente.

3.1.1.1 Modelo RFM

Para Kotler e Armstrong “*atrair clientes é uma tarefa importante, mas reter clientes é mais importante, pois perder um cliente significa perder o todo o fluxo de compra que o cliente faria ao longo da vida*” (2006, apud., WEI et al., 2010, p.4199). Por isso, atrair e reter os clientes mais rentáveis é o ponto de partida para o modelo RFM.

O modelo RFM (recência, frequência e valor monetário, em livre tradução) é uma técnica de segmentação baseada em um modelo aplicado para analisar o comportamento de um cliente e, em seguida, fazer previsões estatísticas com base no histórico do banco de dados (WEI et al., 2010).

Neste modelo, cada cliente é classificado em uma categoria, geralmente em uma escala de 1 a 5 (quanto maior o número, melhor o resultado). “O segmento de 20% superior é codificado como 5, enquanto o próximo segmento de 20% é codificado como 4 e assim por diante. Finalmente, a recência [frequência e valor monetário] de cada cliente no banco de dados é indicada por um número de 5 a 1” (Hughes, 1996; Kahan, 1998; Tsai and Chiu, 2004, apud., WEI et al., 2010, p.4200, traduzido pela autora).

Portanto, quanto maior a classificação do cliente, mais provável é que ele volte a fazer negócios com uma empresa. Nesse sentido, o modelo aumenta a precisão da classificação e extração de regras de classificação melhorando o CRM para as empresas (WEI et al., 2010).

Essencialmente, o modelo RFM corrobora o ditado de *marketing* de que “80% dos negócios vêm de 20% dos clientes” (Investopedia, 2022).

A definição para cada um desses indicadores é:

- Recência: o número em dias do período de tempo entre a penúltima e a última (mais recente) compra de um cliente;
- Frequência : o número de compras em determinado período;
- Valor Monetário: receita gerada pelo cliente no período específico;
- Idade do cliente (T) : o número de períodos de tempo desde a primeira compra do cliente.

As previsões feitas pelo modelo Beta Geométrico/Distribuição Negativa Binomial, apresentado na seção a seguir, baseiam-se nas métricas do RFM.

3.1.1.2 Beta-geometric/Negative Binomial (BG/NBD) e Gamma-Gamma

O modelo *beta-geometric/NBD* compreende que em negócios de receitas recorrentes o número de transações não é determinado por uma tendência, afinal, faturamento e taxa de evasão (*churn*) são variáveis aleatórias. Por isso, o modelo possibilita processos que determinam o que influencia o consumo dos clientes, individualmente e em larga escala, sendo uma evolução do Pareto/NBD, de mais fácil implementação em empresas. (Fader, Hardie e Lee (2005, 275).

O primeiro modelo, Pareto/NBD, foi proposto em 1987 por Schmittlein, Morrison e Colombo em “*Counting Your Customers: Who-Are They and What Will They Do Next?*”. O modelo matemático observa o comportamento do cliente em períodos “*dropout*”, que levam em conta o abandono ou o espaçamento entre transações de compra (IBIDEM, 275).

Este tempo de abandono é modelado por meio do modelo de tempo de Pareto e comportamento de compra repetida para clientes ativos, denominado modelo *Poisson-Gamma*.

Segundo Fader, Hardie e Lee (2005, 276, traduzido pela autora):

- O número de pedidos que um cliente fará em um período de tempo segue uma distribuição de *Poisson* com taxa de transação λ . Essa distribuição de contagem *de Poisson* é equivalente à suposição de que o tempo entre transações segue uma distribuição exponencial com a mesma taxa de transação λ .
- A demanda varia independentemente entre os clientes: heterogeneidade em λ . A variação em λ é uma variável aleatória que segue uma distribuição *Gamma* com parâmetro de forma r e escala α .
- Após qualquer compra, um cliente pode ficar inativo com probabilidade p e se afastar do negócio. O risco de *churn* segue uma distribuição geométrica.
- O risco de *churn* varia independentemente entre os clientes: heterogeneidade na p . A variação da probabilidade de *churn* ou *dropout* p é uma variável aleatória que segue uma distribuição *Beta*.

Enquanto isso, o modelo matemático Gamma-Gamma estima valores prováveis de compra por transação a nível do cliente (Fader et al., 2013). Além dessa, bibliotecas principais de *Python* também serão importadas e utilizadas, como a *numpy*, *pandas*, *matplotlib* e *seaborn*.

A maior parte da matemática envolvida é processada pela linguagem de programação e bibliotecas aqui citadas, por isso, neste estudo será apresentada a metodologia de programação em torno da base de dados de clientes do negócio não contratual (recorrente), assim como sua aplicação para chegar aos indicadores previstos na estratégia de *Customer Success* já contextualizada.

4 MÉTODO DE PESQUISA

4.1 Análises preditivas em Python

Para possibilitar análises preditivas do faturamento mensal, *lifetime value* do cliente e indicador de evasão (*churn*) como proposto, utilizou-se o ambiente de programação gratuito *Jupyter Notebook*. Nesse sentido, importou-se todas as bibliotecas de programação que serão utilizadas para a análise, conforme demonstrado na figura 18:

```
In [3]: #Importação das bibliotecas que serão utilizadas para a análise da base de dados que está em Excel

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from datetime import timedelta
import seaborn as sns

from sklearn.metrics import mean_absolute_percentage_error

from lifetimes import BetaGeoFitter, GammaGammaFitter
from lifetimes.utils import \
    calibration_and_holdout_data, \
    summary_data_from_transaction_data, \
    calculate_alive_path
from lifetimes.plotting import \
    plot_frequency_recency_matrix, \
    plot_probability_alive_matrix, \
    plot_period_transactions, \
    plot_history_alive, \
    plot_cumulative_transactions, \
    plot_calibration_purchases_vs_holdout_purchases, \
    plot_transaction_rate_heterogeneity, \
    plot_dropout_rate_heterogeneity

import warnings
warnings.filterwarnings("ignore")

%matplotlib inline
sns.set(rc={'image.cmap': 'coolwarm'})

pd.set_option("display.precision",2)
np.set_printoptions(precision=2, suppress=True)
pd.options.display.float_format = '{:,.0f}'.format
```

Figura 18: *Jupyter Notebook*. Importação de bibliotecas *Python* para análises preditivas. Imagem da autora.

4.1.1. Base de dados e tratamento inicial

A base de clientes das empresas é um documento muitas vezes sigiloso, principalmente por resguardo da Lei Geral de Proteção de Dados (LGPD)³. Por isso, optou-se por aplicar uma base de dados de venda em *Excel*, disponibilizada para estudo pelo [Repositório de Aprendizado de Máquina da UC Irvine](#) (2015). Esta base atende aos dados necessários para modelagem realizada na metodologia deste estudo.

³ A Lei Geral de Proteção de Dados Pessoais (LGPD), Lei nº 13.709/2018, é a legislação brasileira que regula as atividades de tratamento de dados pessoais por organizações.

Essa base de dados contém cerca de 542.000 registros dos anos de 2010 a 2011 de um negócio *B2B* britânico, conforme explicita a figura 19. Registra-se, ainda, a presença das colunas “País”, “ID Cliente”, “Preço unitário (por venda)”, “InvoiceDate” (do inglês, data de compra), “Quantidade”, “Descrição”, “StockCode” (do inglês, “nº de estoque”) e “InvoiceNo (do inglês, “número da compra”).

```
#Importação (abertura) da base de dados
df0 = pd.read_excel(r"Users\gabri\OneDrive\Documentos\Online Retail.xlsx")
```

InvoiceNo	StockCode	Descrição	Quantidade	InvoiceDate	Preço unitário	ID Cliente	País
0	538385 85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	2010-12-01 08:28:00	3	17,850	United Kingdom
1	538385 71053	WHITE METAL LANTERN	6	2010-12-01 08:28:00	3	17,850	United Kingdom
2	538385 84406B	CREAM CUPID HEARTS COAT HANGER	8	2010-12-01 08:28:00	3	17,850	United Kingdom
3	538385 84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	2010-12-01 08:28:00	3	17,850	United Kingdom
4	538385 84029E	RED WOOLLY HOTTIE WHITE HEART.	6	2010-12-01 08:28:00	3	17,850	United Kingdom
...
541904	581587 22813	PACK OF 20 SPACEBOY NAPKINS	12	2011-12-09 12:50:00	1	12,680	France
541905	581587 22899	CHILDREN'S APRON DOLLY GIRL	6	2011-12-09 12:50:00	2	12,680	France
541906	581587 23254	CHILDRENS CUTLERY DOLLY GIRL	4	2011-12-09 12:50:00	4	12,680	France
541907	581587 23255	CHILDRENS CUTLERY CIRCUS PARADE	4	2011-12-09 12:50:00	4	12,680	France
541908	581587 22138	BAKING SET 9 PIECE RETROSPOT	3	2011-12-09 12:50:00	5	12,680	France

541909 rows x 8 columns

Figura 19: *Jupyter Notebook*. Importação da base de dados e campos de dados. Imagem da autora.

Com intuito de tratar a base de dados antes de estruturar a análise preditiva, alguns procedimentos de limpeza foram adotados, como a exclusão de clientes sem identificação do ID (figura 20), exclusão de transações de devolução e restrição para tratamento apenas de transações acima de uma quantidade (figura 21), padronização do formato da data e outras revisões de variáveis (figura 22).

```
In [14]: #Conferência se o campo "ID do cliente" está vazio, afinal, precisamos trabalhar apenas com os clientes identificáveis
# para compreender de forma real os indicadores sugeridos
df0[df0.isnull().any(axis=1)]
```

Out[14]:

	InvoiceNo	StockCode	Descrição	Quantidade	InvoiceDate	Preço unitário	ID Cliente	Pais
	622	536414	22139	NaN	56	2010-12-01 11:52:00	0	NaN United Kingdom
	1443	536544	21773	DECORATIVE ROSE BATHROOM BOTTLE	1	2010-12-01 14:32:00	3	NaN United Kingdom
	1444	536544	21774	DECORATIVE CATS BATHROOM BOTTLE	2	2010-12-01 14:32:00	3	NaN United Kingdom
	1445	536544	21788	POLKADOT RAIN HAT	4	2010-12-01 14:32:00	1	NaN United Kingdom
	1446	536544	21787	RAIN PONCHO RETROSPOT	2	2010-12-01 14:32:00	2	NaN United Kingdom
...
	541536	581498	85099B	JUMBO BAG RED RETROSPOT	5	2011-12-09 10:28:00	4	NaN United Kingdom
	541537	581498	85099C	JUMBO BAG BAROQUE BLACK WHITE	4	2011-12-09 10:28:00	4	NaN United Kingdom
	541538	581498	85150	LADIES & GENTLEMEN METAL SIGN	1	2011-12-09 10:28:00	5	NaN United Kingdom
	541539	581498	85174	S/4 CACTI CANDLES	1	2011-12-09 10:28:00	11	NaN United Kingdom
	541540	581498	DOT	DOTCOM POSTAGE	1	2011-12-09 10:28:00	1,714	NaN United Kingdom

135080 rows x 8 columns

```
In [16]: # Deleta-se as linhas que não se pode identificar o cliente
df1 = df0.copy()
df1 = df1[pd.notnull(df1["ID Cliente"])]
```

Figura 20: *Jupyter Notebook*. Exclusão de transações não identificáveis pelo ID. Imagem da autora.

```
In [20]: #Para prevenir que compras devolvidas de clientes contabilizem, exclui-se os valores numéricos que estejam negativos
df1[df1["Quantidade"] <= 0]
```

Out[20]:

	InvoiceNo	StockCode	Descrição	Quantidade	InvoiceDate	Preço unitário	ID Cliente	Pais
	141	C536379	D	Discount	-1	2010-12-01 09:41:00	28	14,527 United Kingdom
	154	C536383	35004C	SET OF 3 COLOURED FLYING DUCKS	-1	2010-12-01 09:49:00	5	15,311 United Kingdom
	235	C536391	22556	PLASTERS IN TIN CIRCUS PARADE	-12	2010-12-01 10:24:00	2	17,548 United Kingdom
	236	C536391	21984	PACK OF 12 PINK PAISLEY TISSUES	-24	2010-12-01 10:24:00	0	17,548 United Kingdom
	237	C536391	21983	PACK OF 12 BLUE PAISLEY TISSUES	-24	2010-12-01 10:24:00	0	17,548 United Kingdom
...
	540449	C581490	23144	ZINC T-LIGHT HOLDER STARS SMALL	-11	2011-12-09 09:57:00	1	14,397 United Kingdom
	541541	C581499	M	Manual	-1	2011-12-09 10:28:00	225	15,498 United Kingdom
	541715	C581568	21258	VICTORIAN SEWING BOX LARGE	-5	2011-12-09 11:57:00	11	15,311 United Kingdom
	541716	C581569	84978	HANGING HEART JAR T-LIGHT HOLDER	-1	2011-12-09 11:58:00	1	17,315 United Kingdom
	541717	C581569	20979	36 PENCILS TUBE RED RETROSPOT	-5	2011-12-09 11:58:00	1	17,315 United Kingdom

8905 rows x 8 columns

```
In [22]: # Restringir análise apenas à transações com quantidade positiva
df1 = df1[df1["Quantidade"] > 0]
```

Figura 21: *Jupyter Notebook*. Exclusão de transações de devolução e restrição da análise apenas à transações de quantidades acima de um item. Imagem da autora.


```
In [30]: # Restringir análise apenas à transações com quantidade positiva
df1 = df1[df1["Quantidade"] > 0]

# Na coluna "Data Transação" temos o dia e hora que o cliente comprou. Como apenas a data será utilizada na análise, remove-se
# os Algarismos relativos à hora, pois não serão necessários.
df1["Data Transação"] = pd.to_datetime(df1["Data Transação"]).dt.date #normalize()
#df1.set_index("InvoiceDate", inplace=False)

# A biblioteca Pandas interpreta o "ID Cliente" como uma variável numérica. Será necessário redefinir
# como um tipo de objeto/string e não como número.
df1["ID Cliente"] = df1["ID Cliente"].astype(np.int64).astype(object)

# review the categorical variables
df1.describe(include='object').T
```

```
Out[30]:
```

	count	unique	top	freq
Transação	397924	18536	578339	542
Cód Estoque	397924	3685	85123A	2035
Descrição	397924	3877	WHITE HANGING HEART T-LIGHT HOLDER	2028
Data Transação	397924	305	2011-11-08	3423
ID Cliente	397924	4339	17841	7847
País	397924	37	United Kingdom	354345

Figura 22: *Jupyter Notebook*. Padronização do formato de datas e outras revisões. Imagem da autora.

Após o tratamento inicial e com uso da função “*describe*” (ver figura 23), tem-se uma visão geral que totalizou a base em 4.339 clientes únicos, com cerca de 397.924 transações totais de venda e o país com mais transações é o Reino Unido.

Foram excluídas as colunas “Transação”, “Cód Estoque” e “Descrição” por não fornecerem informações significativas. A coluna “País” apresenta dados relevantes para segmentar geograficamente os clientes, contudo a base de dados já soma quase 400.000 dados, por isso, foi removida com intuito de facilitar a manipulação e trabalhou-se com a perspectiva global de cada um dos clientes identificados pelo número ID, conforme a figura 23.

```
In [49]: # O conjunto de dados compreende 4.339 IDs de clientes exclusivos e 305 datas de compra dentro do período analisado
# O país (Localidade) com mais transações é o Reino Unido.

# Os números das faturas, códigos de estoque e descrições não fornecem informações significativas para a análise,
# portanto, exclui-se.

#A coluna "país" apresenta dados relevantes para a segmentação de cliente, contudo, por trabalhar-se com uma quantidade
# de mais de 400.000 dados será removida, para facilitar a manipulação. Os clientes serão analisados de uma perspectiva global.

# Deleta-se colunas que não são úteis
try:
    d1 = df1.drop(["Transação", "Cód Estoque", "Descrição", "País"], axis=1, inplace=True)
except:
    pass
df1.describe(include='object').T
```

```
Out[49]:
```

	count	unique	top	freq
Data Transação	397924	305	2011-11-08	3423
ID Cliente	397924	4339	17841	7847

Figura 23: *Jupyter Notebook*. Dados totais após tratamento inicial e exclusão de colunas irrelevantes. Imagem da autora.

Para determinar os valores monetários das transações – e, mais à frente, também os *LTV* de todos os clientes – insere-se uma nova coluna “Faturamento” multiplicando as quantidades pelos preços unitários correspondentes, conforme retrata a figura 24:

```
In [32]: # Para determinar os valores monetários das transações - e posteriormente também os LTV de todos os clientes -
# insere-se uma nova coluna "Faturamento" multiplicando as quantidades pelos preços unitários correspondentes.

df1["Faturamento"] = df1["Quantidade"] * df1["Preço unitário"]
df1
```

```
Out[32]:
```

	Transação	Quantidade	Preço unitário	ID Cliente	Faturamento
0	538385	6	3	17850	15
1	538385	6	3	17850	20
2	538385	8	3	17850	22
3	538385	6	3	17850	20
4	538385	6	3	17850	20
...
541904	581587	12	1	12880	10
541905	581587	6	2	12880	13
541906	581587	4	4	12880	17
541907	581587	4	4	12880	17
541908	581587	3	5	12880	15

397924 rows x 5 columns

Figura 24: *Jupyter Notebook*. Visão geral da base de dados após tratamento e inserção da coluna “faturamento”. Imagem da autora.

4.1.2. Aplicação do modelo RFM

Para demonstração da aplicação do modelo RFM, escolheu-se aleatoriamente um cliente para desenvolvimento de um *dataframe* de transações de um cliente específico, neste caso o 14527.

Utilizou-se, ainda, o indicador “Idade *T*”, que representa o período máximo desde a primeira compra do cliente com o negócio.

```

In [53]: # Recência, Frequência e Idade "T"
dfx = df1[df1["ID Cliente"] == 14527]
xmax_date = dfx["Data Transação"].max()
xmin_date = dfx["Data Transação"].min()

# Recência:
print("Data da primeira compra:", xmin_date)
print("Data da última compra:", xmax_date)
xrec = (xmax_date - xmin_date).days
print("Recência:", xrec) # recência é o tempo entre a primeira e última compra

# Idade "T":
xmaxall_date = df1["Data Transação"].max()
print("Data máxima do histórico:", xmaxall_date)
xage = (xmaxall_date - xmin_date).days # Idade "T"
print("T:", xage)

# Frequência:
xfreq = len(dfx[dfx["Quantidade"] > 0].groupby("Data Transação"))-1 # Frequência: períodos entre compras
print("Frequência:", xfreq)

Data da primeira compra: 2010-12-05
Data da última compra: 2011-12-07
Recência: 367
Data máxima do histórico: 2011-12-09
T: 369
Frequência: 53

```

Figura 25: *Jupyter Notebook*. Teste de indicadores com cliente específico. Imagem da autora.

Conforme expressa a figura 25, já pode-se observar dados específicos como a data de primeira compra, 01/12/2010; data da última compra, 07/12/2011; sua recência, 371 dias; tempo do cliente na empresa, 373 dias e a frequência de suas compras, 53 dias.

Na próxima seção, será conduzida a modelagem desses indicadores em perspectiva de testagem e treinamento do *machine learning* (*aprendizagem de máquina*) para contemplar análise de toda a base de clientes.

4.1.2.1 Modelo RFM - Função *calibration and holdout data* para aprendizado de máquina

A função “*calibration and holdout data*”, da biblioteca *lifetime*, tem o intuito de dividir a base de dados para testar e treinar o aplicativo que está sendo desenvolvido e, nesse sentido, precisa de uma grande quantidade de dados para melhorar o aprendizado de máquina, que seria limitado caso fosse considerado um curto período de tempo.

Dessa forma, testes foram aplicados e definiu-se que a análise seguirá considerando que o período de inativação da empresa *B2B* em questão seja 240 dias, ao invés de 90 dias como rotineiramente ocorre, contudo, limitaria o *machine learning*.

Esta função divide a base de dados entre o conjunto de dados de treinamento e de teste. Em paralelo, calcula-se os valores de *recência*, *frequência*, *faturamento* e *idade* “*T*” para cada cliente simultaneamente. Esta função cria um novo *dataframe*, que foi atribuído à variável “*df_ch*”. Os dados de origem originais, “*df1*”, foram organizados pela identificação

do cliente e número da fatura; enquanto “*df_ch*” usa os IDs exclusivos do cliente como seu índice e resume os registros de fatura por cliente.

```
In [55]: # Etapa de Treinamento/Teste (calibration/holdout)
t_holdout = 240 # dias reservados para o período de espera

max_date = df1["Data Transação"].max() # data final de observação
print("Data fim da observação:", max_date)

max_cal_date = max_date - timedelta(days=t_holdout) # data final escolhida para o período de calibração
print("Data fim do período de calibração:", max_cal_date)

df_ch = calibration_and_holdout_data(
    transactions = df1,
    customer_id_col = "ID Cliente",
    datetime_col = "Data Transação",
    monetary_value_col = "Faturamento",
    calibration_period_end = max_cal_date,
    observation_period_end = max_date,
    freq = "D")

print("Comportamento do cliente no período de calibração e espera")
pd.options.display.float_format = '{:,.0f}'.format
df_ch

Data fim da observação: 2011-12-09
Data fim do período de calibração: 2011-04-13
Comportamento do cliente no período de calibração e espera
```

```
Out[55]:
```

ID Cliente	frequency_cal	recency_cal	T_cal	monetary_value_cal	frequency_holdout	monetary_value_holdout	duration_holdout
12346	0	0	85	0	0	0	240
12347	2	121	127	556	4	25	240
12348	2	110	118	297	1	103	240
12350	0	0	70	0	0	0	240
12352	3	34	56	422	3	20	240
...
18270	0	0	26	0	1	29	240
18272	0	0	6	0	5	18	240
18273	0	0	17	0	2	76	240
18280	0	0	37	0	0	0	240
18283	2	53	97	105	11	3	240

2299 rows x 7 columns

Figura 26: *Jupyter Notebook*. Treinamento de máquina e teste ampliado para toda a base de dados. Imagem da autora.

Utiliza-se a função “*describe*” para análise geral das principais médias da base de dados (figura 27). Nesse sentido, já é perceptível alguns insights sobre o comportamento de compra dos clientes desse negócio, são eles:

- 75% dos clientes mantiveram, pelo menos, 01 (uma) compra durante o período de 240 dias;
- O máximo de compras feitas por um cliente durante o período foi 35 transações;

```
In [57]: df_ch.describe()
```

```
Out[57]:
```

	frequency_cal	recency_cal	T_cal	monetary_value_cal	frequency_holdout	monetary_value_holdout	duration_holdout
count	2,299	2,299	2,299	2,299	2,299	2,299	2,299
mean	1	31	79	176	3	28	240
std	2	44	43	438	5	113	0
min	0	0	0	0	0	0	240
25%	0	0	42	0	1	3	240
50%	0	0	82	0	2	16	240
75%	1	65	124	252	4	23	240
max	35	133	133	8,030	96	4,307	240

Figura 27: *Jupyter Notebook*. Visão geral das principais médias contemplando toda a base de dados. Imagem da autora.

A figura 28 contempla frequências médias e o percentual de clientes da base que adere a cada classificação. Reitera-se que clientes que representam frequência zero são aqueles que fizeram a primeira compra no período e ainda não retornaram. Aqui, percebe-se que o maior percentual de clientes está entre zero e cinco compras. Uma minoria que corresponde à soma de 0,032 mantém uma frequência maior que seis compras durante o período, chegando ao máximo de 17 transações.

```
In [58]: # treinamento: frequência
pd.options.display.float_format = '{:,.3f}'.format
x = df_ch["frequency_cal"].value_counts(normalize=True)
x = x.nlargest(15)
print("Frequência:")
x.sort_index(ascending=True)
```

```
Out[58]:
```

Frequência:	0.000	0.591
1.000	0.184	
2.000	0.092	
3.000	0.051	
4.000	0.027	
5.000	0.017	
6.000	0.009	
7.000	0.008	
8.000	0.004	
9.000	0.004	
10.000	0.003	
11.000	0.001	
13.000	0.001	
14.000	0.001	
17.000	0.001	

Name: frequency_cal, dtype: float64

Figura 28: *Jupyter Notebook*. Frequência média contemplando toda a base de clientes. Imagem da autora.

Nos gráficos 01, 02 e 03, utilizou-se os *displots* da biblioteca *seaborn* para possibilitar a visualização dos indicadores de frequência, recência e idade do cliente em dias *versus* densidade da base de clientes.

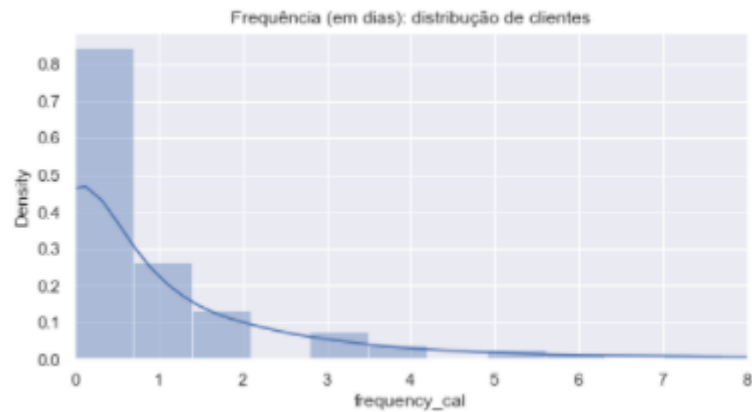


Gráfico 01: *Jupyter Notebook*. Visualização da frequência x densidade da base de clientes. Imagem da autora.

Observa-se, por meio do gráfico 01, que a maior quantidade de clientes tende à recomprar uma ou duas vezes.



Gráfico 02: *Jupyter Notebook*. Visualização da recência x densidade da base de clientes. Imagem da autora.

Em relação à recência (gráfico 02), que mede o número de dias entre a primeira e última compra, percebe-se que a maior parte da base de clientes tende a zero, indicando a alta adesão de novos clientes no período analisado.

Já em relação à idade do cliente “T” (gráfico 03), que vai do primeiro dia de compra até o período final setado, analisa-se que a maior parte da base analisada possui 120 ou mais dias de histórico com a empresa. Contudo, o equilíbrio entre as colunas de dias que representam de zero a 90 dias aponta que a taxa de ativação de novos clientes se mantém estável, indicando tendência positiva de sustentabilidade.



Gráfico 03: *Jupyter Notebook*. Visualização da idade do cliente “T” x densidade da base de clientes. Imagem da autora.

4.1.3 Aplicação do modelo BG/NBD

4.1.3.1 Modelo BG/NBD - Função BetaGeoFitter para previsão de indicadores

Para estruturar a base de dados para aplicação do modelo proposto, utilizou-se a função `Summary_data_from_transaction_data()`, da *lifetime*, para obter o perfil RFM (frequência, recência, idade “T” e faturamento) de toda a base de clientes, conforme mostra a figura 29:

```
In [62]: # determinar recência, frequência, idade do cliente T, faturamento de cada cliente
```

```
df_rft = summary_data_from_transaction_data(
    transactions = df1,
    customer_id_col = "ID Cliente",
    datetime_col = "Data Transação",
    monetary_value_col = "Faturamento",
    observation_period_end = max_date,
    freq = "D")

pd.options.display.float_format = '{:,.0f}'.format
df_rft
```

```
Out[62]:
```

ID Cliente	frequency	recency	T	monetary_value
12346	0	0	325	0
12347	6	385	387	600
12348	3	283	358	301
12349	0	0	18	0
12350	0	0	310	0
...
18280	0	0	277	0
18281	0	0	180	0
18282	1	119	126	78
18283	13	334	337	153
18287	2	159	201	538

```
4339 rows x 4 columns
```

Figura 29: *Jupyter Notebook*. Função `summary_data_from_transaction_data()` que estrutura a base de dados a ser utilizada no modelo BG/NBD. Imagem da autora.

```
In [116]: # Modelo BG/NBD
bgf = BetaGeoFitter(penalizer_coef=1e-06)
bgf.fit(
    frequency = df_rft["frequency"],
    recency = df_rft["recency"],
    T = df_rft["T"],
    weights = None,
    verbose = True,
    tol = 1e-06)
pd.options.display.float_format = '{:,.3f}'.format
bgf.summary

Optimization terminated successfully.
Current function value: -2.709794
Iterations: 38
Function evaluations: 39
Gradient evaluations: 39
```

```
Out[116]:
```

	coef	se(coef)	lower 95% bound	upper 95% bound
r	0.827	0.027	0.774	0.879
alpha	68.910	2.618	63.779	74.041
a	0.002	0.008	-0.010	0.014
b	2.967	6.267	-9.317	15.250

Figura 30: *Jupyter Notebook*. Ajuste da base de dados à distribuição para derivação de seus parâmetros. Imagem da autora.

O próximo passo foi testar o modelo programado para entender se ele é coerente e preditivo, a partir de histórico de transações anteriores, para com as transações atuais, que compreendem os 240 dias definidos anteriormente. Esse treinamento é dado pela função *plot_period_transactions* que gera um gráfico comparativo entre o modelo e transações reais. Conforme observado na figura 31, a partir do gráfico, o modelo (em laranja) é equilibrado com as transações que se concretizaram no período (em azul).

```
In [119]: # treinamento: o modelo reflete dados reais de forma suficientemente próxima?

# frequência de transações repetidas: previsão X real
fig = plt.figure(figsize=(20, 20))
plot_period_transactions(bgf);

<Figure size 1440x1440 with 0 Axes>
```

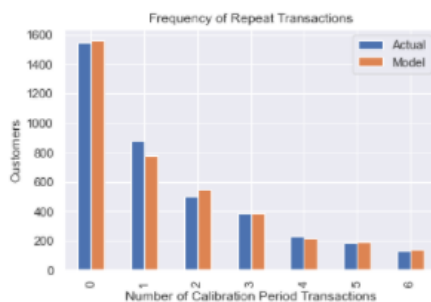


Figura 31: *Jupyter Notebook*. Treinamento do modelo X transações atuais que demonstra que o aplicativo está equilibrado. Imagem da autora.

Portanto, o modelo do aplicativo programado até então tem previsões equilibradas com os dados reais de transações e atende às análises preditivas de forma satisfatória.

Por meio do método *plot_calibration_purchases_vs_holdout_purchases*, da *lifetime*, testa-se se o modelo atende satisfatoriamente à previsão da frequência de compras em comparativo com a frequência de compras reais no período. Conforme observado no gráfico 04:

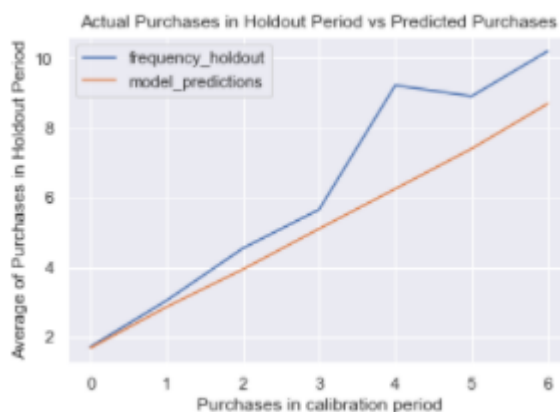


Gráfico 04: Comparativo da frequência de compras do modelo preditivo X compras reais no período. Imagem da autora.

O modelo é coerente com a frequência de compras reais. Percebe-se que acima de três compras, o modelo se destoa do real pois estes eventos têm uma probabilidade menor de ocorrer (um cliente comprar mais de três vezes dentro do período de 240 dias), mas podem ocorrer de forma pontual, como observado.

4.1.3.2 Modelo BG/NBD aplicado à base de dados em integridade

Após testes e calibração do modelo em número limitado de dias, amplia-se a aplicação em toda a base de dados. Foi aplicada a função *summary_data_from_transaction_data*, já apresentada no tópico anterior.

A figura 32 possibilita uma visão geral das principais estatísticas descritivas, portanto, trabalhou-se com a totalidade da base de dados estudada.

```
In [123]: df_rft.describe()
```

```
Out[123]:
```

	frequency	recency	T	monetary_value
count	4,339	4,339	4,339	4,339
mean	3	131	223	307
std	6	132	118	2,612
min	0	0	0	0
25%	0	0	112	0
50%	1	93	248	177
75%	3	252	326	360
max	131	373	373	168,470

Figura 32: *Jupyter Notebook*. Visão geral de estatísticas descritivas. Imagem da autora.

Assim como na seção anterior, o modelo ajusta a base de dados à distribuição para derivação de seus parâmetros por meio da função *BetaGeoFitter* e testa-se a frequência de transações previstas *versus* transações reais por meio do método *plot_period_transactions*, o resultado é dado pelo gráfico 05 e demonstra equilíbrio do modelo preditivo (em laranja) em comparação com o histórico real de transações (em azul).

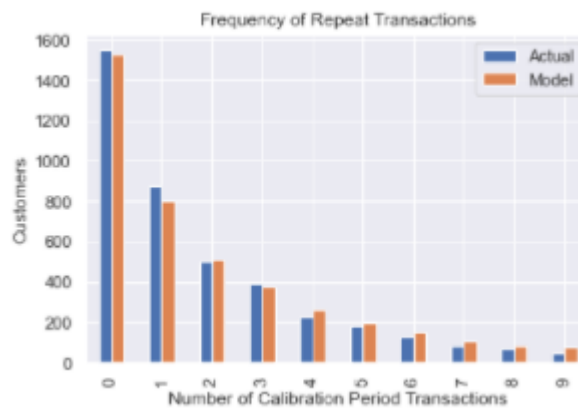


Gráfico 05: *Frequência de transações*. Treinamento do modelo X transações atuais que demonstra que o aplicativo está equilibrado. Imagem da autora.

Por último, testa-se o aplicativo para verificar a previsão de compras previstas *versus* reais no período analisado, o que pode ser observado no gráfico 06:

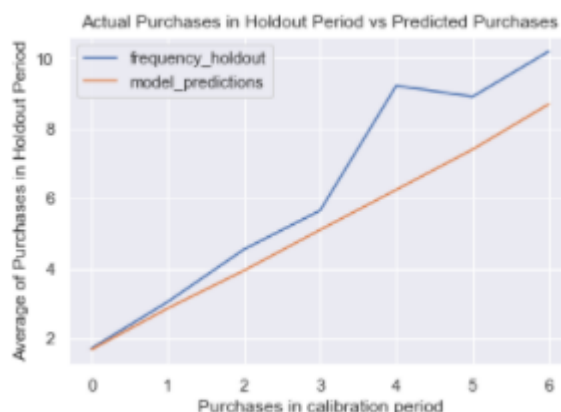


Gráfico 06: *Compras previstas X Compras reais no período*. Treinamento do modelo X compras atuais que demonstra que o aplicativo está equilibrado. Imagem da autora.

Por meio do gráfico 06 é perceptível que o aplicativo prevê suficientemente as transações de compras no período. Um pico não previsto em um intervalo entre três e quatro compras (em azul) se destoa do modelo (em laranja) pois, para esta base de clientes, uma frequência acima de três compras no período é pouco provável (ver figura 18), contudo pode acontecer de forma pontual, assim como é observado neste gráfico.

Na próxima seção, serão apresentados os resultados preditivos de compras, probabilidade de *churn* e, por último, do *lifetime value*, possibilitados pela metodologia desenvolvida até aqui com aplicação do modelo *BG/NBD* e *Gamma-Gamma*, que faculta a estimativa para valores monetários (faturamento).

5 APRESENTAÇÃO E ANÁLISE DE DADOS DA PESQUISA

5.1 Compras previstas

Para encontrar a previsão de compras futuras pela base de clientes analisada, utiliza-se uma matriz de frequência X recência. Este dois indicadores, por meio do método *plot_frequency_recency_matrix* da *lifetime*, predizem as compras esperadas para um período.

Nesse sentido, considera-se dois períodos comuns entre empresas de serviços recorrentes no Brasil: 90 e 180 dias. O resultado pode ser observado nos gráficos 07 e 08, respectivamente.

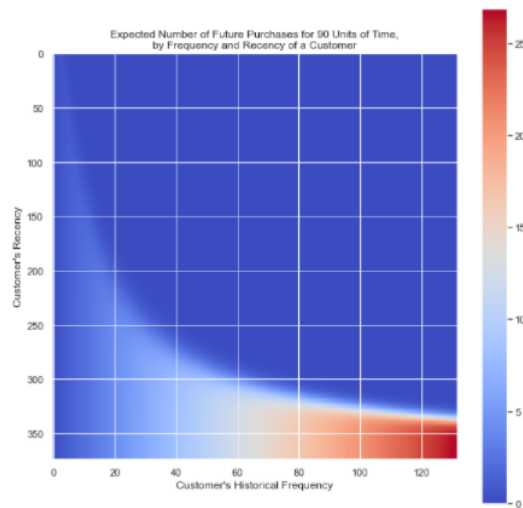


Gráfico 07: *Compras previstas nos próximos 90 dias.* Imagem da autora.

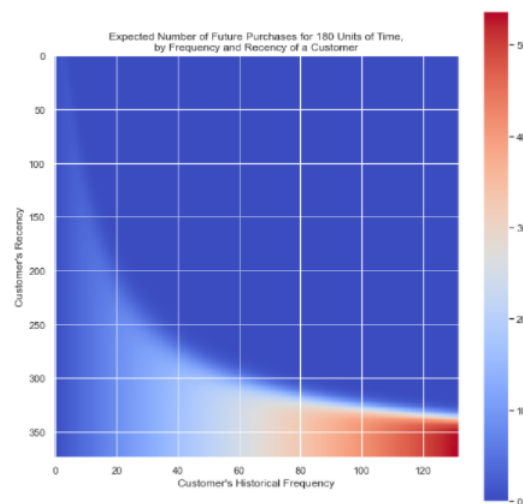


Gráfico 08: *Compras previstas nos próximos 180 dias.* Imagem da autora.

Ambos os gráficos apontam, por meio da zona vermelha no canto inferior direito, que clientes com alta frequência, a partir de 70 transações e longa recência, acima de 300 dias, têm maior propensão a realizar compras futuras, o que representa cerca de 15 a 25 compras no gráfico 07 e cerca de 30 a 50 compras no gráfico 08.

Ainda pode-se identificar e consultar previsões para clientes específicos. Por meio da função *bgf.predict*, seleciona-se um cliente aleatório (aqui considerou-se o 12500) e retorna a estimativa de 4.4 compras para os próximos 180 dias (ver figura 33).

```
In [76]: # previsão de compras de cliente selecionado em período "T"
t = 180
custID = 12500

df_rft_C = df_rft.loc[custID,:]
predC = bgf.predict(
    t,
    df_rft_C["frequency"],
    df_rft_C["recency"],
    df_rft_C["T"])
print("Cliente", custID, ": número esperado de compras em", t, "dias =", f'{predC:.1f}')

Cliente 12500 : número esperado de compras em 180 dias = 4.4
```

Figura 33: *Jupyter Notebook*. Função *bgf.predict* para previsões de clientes específicos. Imagem da autora.

Por meio da função *predict_purch*, aplica-se a análise preditiva para toda a base de clientes, conforme visualização obtida pela figura 34. A partir daqui, é possível a exportação da tabela para o formato *.CSV*, por exemplo, que possibilita manipulação posterior em *Excel* ou *Power B.I, softwares* esses que são mais comumente utilizados em empresas.

ID Cliente	frequency	recency	T	monetary_value	predict_purch_90	prob_alive	prob_vivo	predict_purch_180	predict_purch_270
12346	0.0	0.0	325.0	0.0	0.2	1.0	1.0	0.4	0.6
12347	6.0	365.0	367.0	599.7	1.4	1.0	1.0	2.8	4.2
12348	3.0	283.0	358.0	301.5	0.8	1.0	1.0	1.6	2.4
12349	0.0	0.0	18.0	0.0	0.9	1.0	1.0	1.7	2.6
12350	0.0	0.0	310.0	0.0	0.2	1.0	1.0	0.4	0.6
...
18280	0.0	0.0	277.0	0.0	0.2	1.0	1.0	0.4	0.6
18281	0.0	0.0	180.0	0.0	0.3	1.0	1.0	0.6	0.9
18282	1.0	119.0	126.0	77.8	0.8	1.0	1.0	1.7	2.5
18283	13.0	334.0	337.0	152.8	3.1	1.0	1.0	6.1	9.2
18287	2.0	159.0	201.0	536.0	0.9	1.0	1.0	1.9	2.8

4339 rows x 12 columns

Figura 34: *Jupyter Notebook*. Função *predict_purch* para previsões de toda a base de clientes. Imagem da autora.

É possível, ainda, uma visualização por ordem decrescente dos principais compradores (top 10), para facilitar ações de *marketing* específicas para o sucesso do cliente, por exemplo (ver figura 35).

ID Cliente	frequency	recency	T	monetary_value	predict_purch_10	predict_purch_30	predict_purch_60	predict_purch_90
14911	131.0	372.0	373.0	1,093.7	3.0	8.9	17.9	28.8
12748	113.0	373.0	373.0	298.4	2.6	7.7	15.5	23.2
17841	111.0	372.0	373.0	384.5	2.5	7.6	15.2	22.8
15311	89.0	373.0	373.0	677.7	2.0	6.1	12.2	18.3
14606	88.0	372.0	373.0	135.9	2.0	6.0	12.1	18.1
12971	70.0	369.0	372.0	159.2	1.6	4.8	9.6	14.5
13089	65.0	367.0	369.0	893.7	1.5	4.5	9.0	13.5
14527	53.0	367.0	369.0	155.0	1.2	3.7	7.4	11.1
13798	52.0	371.0	372.0	706.7	1.2	3.6	7.2	10.8
16422	47.0	352.0	369.0	702.5	1.1	3.3	6.6	9.8

Figura 35: *Jupyter Notebook*. Top 10 clientes pela previsão de compras nos próximos 90 dias. Imagem da autora.

5.2 Probabilidade de *churn*

A previsão sobre a evasão de um cliente (ou sua fidelidade) é obtida por meio do método `plot_probability_alive_matrix`, da *lifetime*. Este método considera a frequência e recência de um cliente ao longo do tempo para gerar, assim como o método anterior, uma matriz.

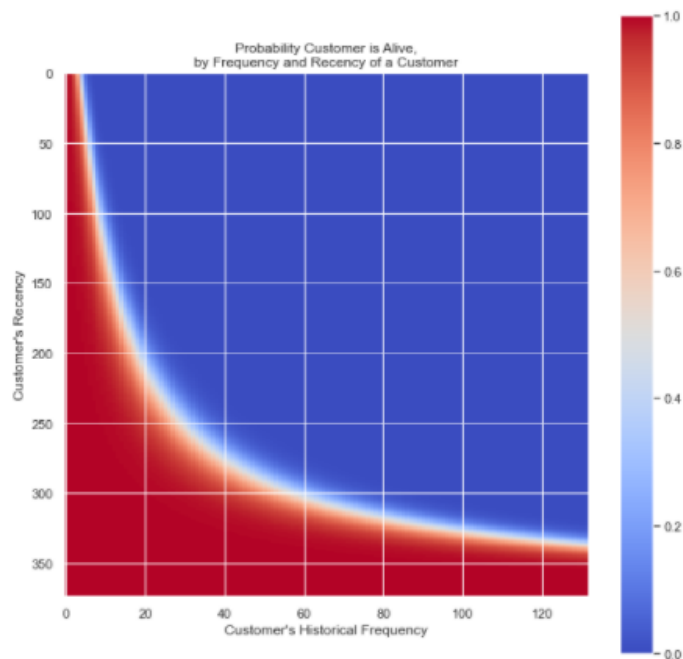


Gráfico 09: *Jupyter Notebook*. Probabilidade de *churn*. Imagem da autora.

No gráfico 09, a zona vermelha representa a probabilidade de o cliente estar fidelizado, orientado por altos indicadores de recência e alto número de transações.

Nesse sentido, há a possibilidade de aplicação da função *conditional_probability_alive*, que gera o *dataframe* ilustrado na figura 36:

	frequency	recency	T	monetary_value	predict_purch_90	predict_purch_180	predict_purch_270	prob_vivo
count	4,339.00	4,339.00	4,339.00	4,339.00	4,339.00	4,339.00	4,339.00	4,339.00
mean	2.86	130.74	222.78	306.94	1.08	2.15	3.23	1.00
std	5.95	132.21	117.88	2,612.45	1.27	2.54	3.81	0.01
min	0.00	0.00	0.00	0.00	0.17	0.34	0.50	0.75
25%	0.00	0.00	112.00	0.00	0.45	0.90	1.35	1.00
50%	1.00	93.00	248.00	176.94	0.77	1.53	2.30	1.00
75%	3.00	252.00	326.00	360.16	1.27	2.54	3.81	1.00
max	131.00	373.00	373.00	168,469.60	26.84	53.68	80.50	1.00

Figura 36: *Jupyter Notebook*. Função *conditional_probability_alive*. Imagem da autora.

Na coluna inserida e nomeada “*prob_alive*” (probabilidade de estar “vivo”), percebe-se que 75% da base de clientes, no mínimo, estará ativa no período máximo do histórico.

Pode-se, ainda, identificar os clientes mais sensíveis a se evadir da empresa, aqueles a quem se deve traçar planos (comerciais ou de relacionamento) com objetivo de retê-lo. Na figura 37, segmentou-se a base de clientes com probabilidade menor de 90% de recompra, ou seja, maior probabilidade de inativação, que resulta em quatro clientes:

	frequency	recency	T	monetary_value	predict_purch_90	predict_purch_180	predict_purch_270	prob_vivo
ID Cliente								
13093	6.00	98.00	373.00	1,172.18	1.18	2.36	3.54	0.85
15107	5.00	61.00	372.00	40.70	0.89	1.79	2.68	0.75
15235	9.00	156.00	373.00	240.89	1.77	3.54	5.32	0.89
16725	5.00	69.00	371.00	131.77	0.97	1.94	2.91	0.81

Figura 37: *Jupyter Notebook*. Clientes com probabilidade menor de 90% de recompra. Imagem da autora.

O método *plot_history_alive* analisa, ainda, o histórico de um cliente específico e como se desenvolveu durante todo o período de tempo. Este método é útil, por exemplo, para compreender o histórico de comportamento de compra e experiência dos principais clientes, assim como dos clientes evasores. Aqui, considerou-se o cliente 13093 (ID Cliente), o primeiro da lista com menor probabilidade de recompra para visualizar seu histórico junto ao negócio (ver figura 38):

Cliente 13093 : probabilidade de estar vivo ao longo do tempo



Cliente 13093 probabilidade de estar vivo, desenvolvimento ao longo do tempo

```
count    374.00
mean     0.97
std      0.04
min      0.85
25%     0.97
50%     1.00
75%     1.00
max      1.00
dtype: float64
```

Figura 38: *Jupyter Notebook*. Probabilidade de *churn* de cliente específico (ID Cliente 13093). Imagem da autora.

O modelo ainda considera uma probabilidade mínima de 85% para que este cliente retorne a comprar com a empresa. Isso se dá devido ao curto período de histórico (para esta base, dois anos que compreendem entre 2010 e 2011) e grande percentual de clientes de média frequência. Desse modo, caso se analisasse a partir de uma base de alta frequência com clientes que comprem mensalmente, o modelo tenderia a diminuir o percentual mínimo de probabilidade de recompra.

Na próxima seção, apresenta-se as modelagens preditivas do CLV, incluindo valores previstos de faturamento geral e por cliente.

5.3 Previsão monetária do *customer lifetime value* (CLV)

A previsão do *customer lifetime value* é possível por meio da integração entre o modelo BG/NBG e Gamma-Gamma. A longevidade do cliente considera todo seu histórico disponível de compra e a probabilidade de tempo de vida leva em conta parâmetros que devem ser informados à máquina de aprendizado para aplicação da função *customer_lifetime_value*.

Esses parâmetros são:

- O período esperado que deve ser aplicado ao modelo probabilístico (em meses);
- A taxa de desconto do retorno líquido que a empresa espera receber neste prazo, ou seja, a taxa de dividendo entre acionistas (em percentual);

Para esse estudo, foi definido um período de 48 meses, para alcançar a previsão de indicadores pelos próximos quatro anos, que é o período proposto como objetivo específico para esta monografia; e uma taxa de desconto anual de 10%, a exemplo de uma taxa de dividendo definida entre os acionistas para o período, o que pode variar entres as empresas.

A *função utiliza* as métricas obtidas no modelo BG/NBD por meio da função *summary_data_from_transaction_data* (recência, frequência, longevidade “T” e valor monetário) e retorna uma matriz de valores de CLV, conforme figura 39:

	count	mean	std	min	25%	50%	75%	max
CLV_previsto	2,790.000	9,723.082	35,132.922	535.168	2,587.213	4,639.829	8,684.076	999,257.238
CLV	2,790.000	9,723.082	35,132.922	535.168	2,587.213	4,639.829	8,684.076	999,257.238
LTV	2,790.000	9,723.082	35,132.922	535.168	2,587.213	4,639.829	8,684.076	999,257.238
frequency	2,790.000	4.454	6.930	1.000	1.000	3.000	5.000	131.000
recency	2,790.000	203.329	111.460	1.000	106.000	211.000	301.750	373.000
T	2,790.000	260.121	101.480	9.000	197.250	284.000	358.000	373.000
monetary_value	2,790.000	477.359	3,245.622	1.000	187.165	309.473	458.371	168,469.600
predict_purch_90	2,790.000	1.428	1.459	0.365	0.672	1.087	1.660	26.843
predict_purch_180	2,790.000	2.856	2.918	0.731	1.345	2.133	3.319	53.677
predict_purch_270	2,790.000	4.283	4.376	1.096	2.017	3.200	4.977	80.504
prob_vivo	2,790.000	0.998	0.007	0.752	0.999	0.999	1.000	1.000
faturamento_esperado	2,790.000	469.989	2,437.796	35.325	228.001	329.162	455.733	125,940.166
media_esperada	2,790.000	477.359	3,245.622	1.000	187.165	309.473	458.371	168,469.600
margem_erro	2,790.000	-7.370	814.660	-42,529.434	-1.025	11.492	33.887	111.687

Figura 39: *Jupyter Notebook*. Descrição de previsões de toda a base de clientes. Fonte: imagem da autora.

A partir do *dataframe*, conclui-se que em quatro anos, a média prevista do CLV para esta base de dados é de €9.723,08, com uma probabilidade média de 99,8% da base de clientes estar viva.

É possível ainda, ampliar a visualização por cliente. Na figura 40, percebe-se, por exemplo, que o cliente 16.446 é o cliente mais valioso, com um CLV esperado de quase €1 milhão e faturamento esperado, nos próximos 48 meses, de €125.940,16.

ID Cliente	CLV_previsto	CLV	prob_vivo	faturamento_esperado	media_esperada	margem_erro
16446	999,257.238	999,257.238	0.999	125,940.166	168,489.600	-42,529.434
14646	797,753.273	797,753.273	1.000	6,321.425	6,366.706	-45.281
18102	651,125.338	651,125.338	1.000	9,230.351	9,349.477	-119.126
17450	535,978.639	535,978.639	1.000	7,315.109	7,404.690	-89.581
14096	470,650.978	470,650.978	1.000	3,996.161	4,071.434	-75.273
...
12755	624.393	624.393	0.996	125.048	17.550	107.498
14865	607.251	607.251	0.999	123.144	15.000	108.144
17194	603.290	603.290	0.996	119.409	10.000	109.409
18037	566.816	566.816	0.997	35.325	7.744	27.581
17816	535.168	535.168	0.999	50.983	6.783	44.199

Figura 40: *Jupyter Notebook*. *Dataframe* de visualização de indicadores por cliente. Fonte: Imagem da autora.

5.4 Visualização de dados

Por fim, a base de dados é exportada do ambiente de programação e importada para a ferramenta *Data Studio*, da Google. A ferramenta de visualização de dados possibilita o compartilhamento do *dashboard* preditivo para gerenciamento proativo da carteira de clientes (ver figura 41).



Figura 41: Visualização de dados preditivos em *dashboard* hospedado no *Data Studio*. Fonte: Imagem da autora.

O *dashboard* é interativo, gratuito e compartilhável, levando a equipe de *marketing* à tomadas de decisões orientadas a dados preditivos. Foram selecionadas métricas relevantes para acompanhamento a longo prazo, como a frequência média de compras, CLV, faturamento esperado, margem de erro e *churn*, por exemplo.

Os dados são podem ser atualizados em tempo real, a partir do ambiente de programação, sendo atualizados à medida do processamento de novas transações ou pela falta delas, como abandonos, recompras ou entrada de novos clientes.

5.5 Ações estratégicas

A partir dos recursos fundamentados a respeito da experiência do cliente, gestão de relacionamento e *marketing* preditivo, todos em prol do sucesso do cliente, este trabalho resulta em uma aplicação que baseia-se em *machine learning*, esta que é melhorada à medida que recebe mais histórico de transações, para prever resultados monetários pelos próximos quatro anos - período esse que foi limitado apenas para fins metodológicos e que pode ser expandido.

Nesse sentido, ações estratégicas para gerenciamento desses resultados de forma proativa, gerenciando o *onboarding*, fidelidade e recuperação são propostas e aplicáveis para modelos de negócios de receita recorrente, correspondendo ao objetivo geral desta monografia.

Com intuito de induzir ao faturamento a curto, médio e longo prazo, são propostas as seguintes ações:

- Divisão dos clientes em *clusters* (agrupamento de clientes de características específicas) baseados em seu CLV, para tratamento de interesses, distribuição de benefícios, suporte de sucesso, acompanhamento da jornada e indicador de satisfação (o NPS), o que visa reter e fidelizar os principais clientes que compõem a carteira que é o ponto-chave para a sustentabilidade financeira do negócio, como apresentado na seção 2.3.2.1 desta monografia;
- Definição de um grupo de trabalho específico para tratamento do grupo de clientes com probabilidade de evasão acima de 80%, por exemplo - esta taxa deverá ser definida pela empresa, embasada nos resultados estatísticos obtidos e fundamentados na seção 2.3.2.2 desta monografia, conforme defende a seção 3.1 que, apesar da estratégia aliada a dados, a intervenção humana é essencial para o sucesso do plano tático;
- Pesquisas com o cliente interno que caracterizem qualitativamente cada *cluster* e análises quantitativas exploratórias, demarcando padrões estatísticos sobre produtos e serviços que são consumidos e quais têm potencial para crescer a participação de clientes na carteira, gerenciando estratégias de *up-sell* e *cross-sell* orientadas à análise relacional entre os demais participantes do *cluster*;
- Revisão do mix de produtos/serviços e divisão dos quais atendem cada *cluster*, formando comitês de clientes que opinam ativamente e guiam a produção de produtos ou melhoria de serviços, como fundamentado no tópico 2.3.2 desta monografia que argumenta acerca do valor do cliente para a estratégia da empresa, gerando a oportunidade de participação em decisões;

5.5.1 Mapeamento da jornada do cliente

O mapeamento da jornada do cliente para empresas B2B precisa considerar as etapas do ciclo de vida do cliente, prever os pontos de contato e plano de canais de comunicação que considerem a hierarquia de valor dos clientes (ver figura 10).

As figuras 42, 43 e 44 contemplam o modelo desenvolvido pela autora para empresas de receita recorrente, baseado no ciclo de vida do cliente, fundamentado no tópico 2.4.1 e na evolução do funil de *marketing*, apresentada na seção 2.3 deste estudo.

Fase da Jornada	Etapa 01: Captação de clientes	Etapa 02: Onboarding
Ações O que o cliente vai fazer?	Engajar, conhecer, investigar, avaliar, compreender e converter	Receber promessas feitas durante a captação. Obter acesso a materiais de implementação do serviço, treinamento do produto.
Touchpoint: Que parte do serviço o cliente vai interagir?	Canais de comunicação <ul style="list-style-type: none"> • High touch: prospecção <i>hunter/ativa por consultores especializados e SDRs</i>; • Low touch: <i>Canais de publicidade físicos ou digitais</i> • High tech: <i>Canais de publicidade físicos ou digitais</i> 	Canais de comunicação <ul style="list-style-type: none"> • High touch: reuniões presenciais de treinamento e implementação com especialistas; • Low touch: reuniões periódicas de treinamento/implementação e canal digital de comunicação; • High tech: Treinamento/Implementação através de portal de relacionamento, webinars, canais digitais;
Pensamento do cliente: o que o cliente está pensando?	A empresa tem uma boa solução/produto para meu negócio/para mim.	Terei sucesso com esse produto/serviço.
Sentimento do cliente: o que o cliente está sentindo?	Neutro/Positivo	Neutro/Positivo
Propriedade do processo: quem é o líder nisso?	Equipe de <i>Marketing</i>	<i>Equipe de Customer Success</i>
Oportunidades	Estratégias de captação online, offline e ativa (<i>hunter</i>).	Participação do canal de SAC; Premiações e benefícios exclusivos; Ofertas segmentadas; Brindes;
Indicadores de performance	Taxa de ativação de novos clientes (captação X conversão)	Compras previstas, frequência prevista X Real, NPS

miro

Figura 42: Mapeamento da jornada do cliente: captação e *onboarding*. Fonte: Imagem da autora.

Fase da Jornada	Etapa 03: Energização	Etapa 04: Proteção/fidelização
Ações O que o cliente vai fazer?	Aumentar o ticket médio, comprando mais produtos e experimentando outras soluções/serviços/produtos.	Contribuir ativamente com o negócio, promover positivamente a empresa, beneficiar-se com as soluções adquiridas, aumentar ou manter seu ticket médio.
Touchpoint: Que parte do serviço o cliente vai interagir?	Canais de comunicação <ul style="list-style-type: none"> • High touch: Consultores especializados; • Low touch: Consultores especializados, loja física ou atendimento <i>online</i>; • Tech touch: E-commerce, canais de vendas digitais, e-mail ou loja física 	Canais de comunicação <ul style="list-style-type: none"> • High touch: Consultores especializados; • Low touch: Consultores especializados, portal de relacionamento e canais de publicidade digitais; • Tech touch: Portal de relacionamento e canais de publicidade digitais;
Pensamento do cliente: o que o cliente está pensando?	Preciso de mais soluções/produtos ou soluções/produtos complementares para obter mais sucesso.	Tenho sucesso com essa solução/produto.
Sentimento do cliente: o que o cliente está sentindo?	Neutro/Positivo	Positivo
Propriedade do processo: quem é o líder nisso?	Equipe de Marketing Digital/Comercial	Equipe de Marketing/Comercial/Customer Success
Oportunidades	Participação do canal de SAC; Premiações e benefícios exclusivos; Ofertas segmentadas; Brindes;	
Indicadores de performance	Compras previstas, frequência prevista X Real, NPS	Compras previstas, frequência prevista X Real, CSAT

míro

Figura 43: Mapeamento da jornada do cliente:energização, proteção e fidelização. Fonte: Imagem da autora.

Fase da Jornada	Etapa 05: Reconquista/Recuperação
Ações O que o cliente vai fazer?	Solucionar empecilhos para seu sucesso, retomar o negócio.
Touchpoint: Que parte do serviço o cliente vai interagir?	Canais de comunicação <ul style="list-style-type: none"> • High touch: Consultores especializados; • Low touch: Consultores especializados e canais de publicidade digitais; • Tech touch: Canais de publicidade digitais;
Pensamento do cliente: o que o cliente está pensando?	Não obteve sucesso com essa solução/produto.
Sentimento do cliente: o que o cliente está sentindo?	Negativo
Propriedade do processo: quem é o líder nisso?	Equipe de Marketing/Customer Success
Oportunidades	Participação do canal de SAC; Premiações e benefícios exclusivos; Ofertas segmentadas; Brindes;
Indicadores de performance	Churn rate

Figura 44: Mapeamento da jornada do cliente: retenção e recuperação. Fonte: Imagem da autora.

5.5.2 Canais de comunicação

O planejamento dos canais de comunicação segue a metodologia de hierarquia do valor do cliente (ver seção 2.3.2), prevendo diferentes experiências e pontos de contato para consumidores de alto, intermediário e baixo *ticket* médio, com base nas fundamentações apresentadas no tópico 2.3 do estudo.

Para consumidores de alto *ticket* médio, são propostos canais de comunicação *high touch*, com maior intervenção e contato de consultores especializados, que contemplam reuniões mensais, por meios físicos ou digitais, além de profissionais capacitados para atuar dentro do negócio do cliente como facilitadores.

Para consumidores de *ticket* médio intermediário, propõe-se canais de comunicação *low touch*, com intervenções de consultores especializados apenas quando necessário e/ou em

período pré-definido, com a possibilidade de reuniões trimestrais, por exemplo. Nesse sentido, também são ofertados ao cliente como canais de comunicação o e-mail, portal de relacionamento, lives, *What'sApp* exclusivo de atendimento, SAC, loja física, entre outros.

Para clientes de baixo *ticket* médio, planeja-se a comunicação digital, por meio de portais de relacionamento, e-mail, chat automatizado, webinars, lives, loja física, central de relacionamento, SAC, entre outros.

Na fase de captação, propõe-se, ainda, a atuação de *hunters* (caçadores, em inglês). *Hunters* são agentes de captação de clientes que podem atuar em meios físicos, como captação em eventos relevantes, ou *online*, atuando em redes sociais como o LinkedIn ao abordar empresários, investidores, empreendedores ou qualquer outro perfil que seja relevante para o negócio.

5.5.3 Plano tático

Baseado na análise de indicadores preditivos resultantes da implementação da aplicação desenvolvida neste estudo, o relacionamento do cliente, principalmente nas fases de energização, proteção e reconquista pode ser monitorado e controlado com base nas faixas de probabilidade de *churn* e CLV.

Na seção “Probabilidade de *churn*”, pode-se, ainda, segmentar a base de clientes por faixas de probabilidade de evasão, levando a equipe a agir proativamente ao evitar a inativação do cliente e a diminuição do potencial de compra. Algumas ações podem ser propostas, a exemplo:

- Clientes com probabilidade de *churn* abaixo de 60% (quanto menor o percentual, mais provável é a evasão) e CLV acima de €50.000 são contatados e recebem consultorias do serviço/produto pelo próximo mês;
- Clientes com probabilidade de *churn* abaixo de 80% (quanto menor o percentual, mais provável é a evasão) e CLV até €50.000 são contatados, recebem treinamento gravado e ofertas específicas;
- Clientes que declinaram na previsão CLV no último mês recebem ofertas customizadas;

Portanto, o plano de ação de *marketing* e vendas deve ser guiado pelos *insights* de valor do cliente, em prol de seu sucesso. Existe um horizonte de ações que podem ser adotadas a partir da aplicação desenvolvida nesta monografia, possibilitando inclusive a existência de relacionamentos *one to one*, que consideram benefícios, ofertas e oportunidades personalizadas por cliente, dentro do universo de toda a base de contatos. Esta é uma das contribuições em que a ciência de dados é aplicada a favor do *marketing*, operacionalizando análises de *big data* e resultando em estratégias assertivas e otimização de orçamento.

6 CONCLUSÃO E TRABALHOS FUTUROS

O presente estudo debruçou-se a aplicar as soluções da ciência de dados a favor de estratégias preditivas de *marketing*, voltadas ao sucesso do cliente e à sustentabilidade financeira do negócio de receita recorrente (B2B).

Atualmente, empresas têm adotado a disciplina *Customer Oriented*, na qual estruturam-se internamente para entregar o melhor produto/serviço ao cliente, solucionando quaisquer atritos durante o relacionamento. São vários os benefícios conquistados ao adotar essa cultura: clientes satisfeitos, comprando mais, indicando o negócio para possíveis novos consumidores e com participação ativa no desenvolvimento de novos produtos, campanhas, entre outros.

A área de CRM, que rotineiramente administra canais de comunicação e ofertas para a base de clientes, pode ser aliada à estratégias de *Customer Success* e soluções da ciência de dados para gerenciar operações com grandes quantidades de dados que resultam em ações guiadas por previsões estatísticas e matemáticas.

O setor de *Customer Success* tem o objetivo de entregar as promessas de valor e sucesso durante todas as fases da jornada do cliente com o negócio por meio de uma postura proativa e observação de indicadores de sucesso, evitando evasões e fidelizando consumidores.

Observou-se ao longo do estudo que as metodologias relacionadas à matemática e estatística possibilitam o gerenciamento do relacionamento com toda a base de clientes de maneira otimizada, orientando a equipe de *marketing* por meio de dados. Conforme demonstrado, os modelos RFM, BG/NBD e Gamma-Gamma, aplicados por meio da linguagem de programação *Python*, tornaram possível a previsão por máquina de aprendizado pelos próximos quatro anos da frequência de compras (geral e por cliente), probabilidade de

evasão e valor de vida do cliente (CLV) por meio de operações em larga escala de dados históricos (*big data*).

Dessa forma, ações estratégicas que otimizam investimentos e melhoram o relacionamento com o consumidor são propostas a partir dos resultados observados na seção de visualização e fundamentado por técnicas da experiência do cliente, em que se mapeia a jornada do cliente, prevê a operação dos canais de comunicação e/ou pontos de contato e ações proativas de energização, proteção e recuperação do cliente quando propício.

Como próximos passos de desenvolvimento, que visam a continuidade deste estudo, é possível o cruzamento de dados de faturamento segmentando esforços por praça, resultados integrados com ações que acontecem em redes sociais e, ainda mais relevante, dados de faturamento cruzados com o tempo climático, podendo, por exemplo, compreender qual a correlação da performance de vendas em dias chuvosos ou ensolarados.

Estes *insights* poderão entregar profundas análises de oportunidades para ações estratégicas de *marketing*, fazendo com que a equipe realize ações com maiores resultados e o mínimo de investimento e que se tornam cada vez mais assertivos, tendo em vista que a máquina de aprendizado desenvolvida na aplicação tende a se tornar cada vez mais confiável com a atualização diária de sua base de dados.

REFERÊNCIAS

BORGES, Luiz Eduardo. **Python para desenvolvedores: aborda Python 3.3**. Novatec Editora, 2014.

CRUZ, Ana Carolina Narciso Ferreira da. **Customer success: os benefícios de uma cultura com foco no sucesso do cliente em uma indústria de alimentos do interior do Rio Grande do Sul**. Tese (Bacharelado em Administração) - Centro de Ciências da Administração, Universidade de Caxias do Sul. Rio Grande do Sul. 2021.

DEMO, Gisela et al. **Marketing de relacionamento (CRM): estado da arte, revisão bibliográfica da produção nacional de primeira linha, institucionalização da pesquisa no Brasil e agenda de pesquisa**. RAM. Revista de Administração Mackenzie, v. 16, p. 127-160, 2015.

FADER, Peter S.; HARDIE, Bruce GS. **The Gamma-Gamma model of monetary value**. February, v. 2, p. 1-9, 2013.

GABRIEL, Martha. **Marketing na era digital. Conceitos, plataformas e estratégias**. Novatec Editora, v.1, São Paulo. 2010.

GANTI, Akhil. **A Validation of and Extension to a Non-Parametric Approach to Buy-'Til-You-Die Models**. 2019.

HOCHSTEIN, Bryan et al. **An industry/academic perspective on customer success management**. Journal of Service Research, v. 23, n. 1, p. 3-7, 2020.

KOTLER, Philip; ARMSTRONG, Gary. **Princípios de marketing**. Editora Pearson - Prentice Hall, 9. ed, São Paulo. 2003.

KOTLER, Philip et al. **Marketing 5.0. Tecnologia para a humanidade**. Sextante, v.1, 2021.

MADRUGA, Roberto. **Gestão de relacionamento e customer experience. A revolução na experiência do cliente**. Editora Atlas, 2. ed, São Paulo. 2021.

Newman, D. J., Hettich, S., Blake, C., & Merz, C. J. **UCI repository of machine learning databases [Machine-readable data repository]**. University of California, Department of Information and Computer Science, Irvine, CA. 2006.

SCHMITTLEIN, David C.; MORRISON, Donald G.; COLOMBO, Richard. **Counting your customers: Who-are they and what will they do next?.** Management science, v. 33, n. 1, p. 1-24, 1987.

SEGAL, Troy. **Recency, Frequency, Monetary Value (RFM).** Investopedia. 23/08/2021. Disponível em: <https://www.investopedia.com/terms/r/rfm-recency-frequency-monetary-value.asp>. Acesso em 01/05/2022.

STEINMAN, Dan et al. **Customer Success. Como as empresas inovadoras descobriram que a melhor forma de aumentar a receita é garantir o sucesso dos clientes.** Autêntica Business, 1. ed, 2020.

WEI, Jo-Ting; LIN, Shih-Yen; WU, Hsin-Hung. **A review of the application of RFM model.** African Journal of Business Management, v. 4, n. 19, p. 4199-4206, 2010.

ZEITHAML, Valarie A. et al. **A theories-in-use approach to building marketing theory.** Journal of Marketing, v. 84, n. 1, p. 32-51, 2020.

ZENDESK. **Customer Experience Trends.** 2022. Disponível em: <https://www.zendesk.com.br/customer-experience-trends/>. Acesso em 01/05/2022.

APÊNDICE A – CÓDIGO DE PROGRAMAÇÃO

In [2]: `pip install lifetimes`

```
Requirement already satisfied: lifetimes in c:\users\gabri\anaconda3\lib\site-packages (0.11.3)Note: you may need to restart the kernel to use updated packages.

Requirement already satisfied: pandas>=0.24.0 in c:\users\gabri\anaconda3\lib\site-packages (from lifetimes) (1.3.4)
Requirement already satisfied: dill>=0.2.6 in c:\users\gabri\anaconda3\lib\site-packages (from lifetimes) (0.3.4)
Requirement already satisfied: numpy>=1.10.0 in c:\users\gabri\anaconda3\lib\site-packages (from lifetimes) (1.20.3)
Requirement already satisfied: scipy>=1.0.0 in c:\users\gabri\anaconda3\lib\site-packages (from lifetimes) (1.7.1)
Requirement already satisfied: autograd>=1.2.0 in c:\users\gabri\anaconda3\lib\site-packages (from lifetimes) (1.3)
Requirement already satisfied: future>=0.15.2 in c:\users\gabri\anaconda3\lib\site-packages (from autograd>=1.2.0->lifetimes) (0.18.2)
Requirement already satisfied: pytz>=2017.3 in c:\users\gabri\anaconda3\lib\site-packages (from pandas>=0.24.0->lifetimes) (2021.3)
Requirement already satisfied: python-dateutil>=2.7.3 in c:\users\gabri\anaconda3\lib\site-packages (from pandas>=0.24.0->lifetimes) (2.8.2)
Requirement already satisfied: six>=1.5 in c:\users\gabri\anaconda3\lib\site-packages (from python-dateutil>=2.7.3->pandas>=0.24.0->lifetimes) (1.16.0)
```

In [3]: *#Importação das bibliotecas que serão utilizadas para a análise da base de dados que está em Excel*

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from datetime import timedelta
import seaborn as sns

from sklearn.metrics import mean_absolute_percentage_error

from lifetimes import BetaGeoFitter, GammaGammaFitter
from lifetimes.utils import \
    calibration_and_holdout_data, \
    summary_data_from_transaction_data, \
    calculate_alive_path
from lifetimes.plotting import \
    plot_frequency_recency_matrix, \
    plot_probability_alive_matrix, \
    plot_period_transactions, \
    plot_history_alive, \
    plot_cumulative_transactions, \
    plot_calibration_purchases_vs_holdout_purchases, \
    plot_transaction_rate_heterogeneity, \
    plot_dropout_rate_heterogeneity

import warnings
warnings.filterwarnings("ignore")

%matplotlib inline
sns.set(rc={'image.cmap': 'coolwarm'})

pd.set_option("display.precision",2)
np.set_printoptions(precision=2, suppress=True)
pd.options.display.float_format = '{:,.0f}'.format
```

```
In [4]: #Importação (abertura) da base de dados
df0 = pd.read_excel(r"Users\gabri\OneDrive\Documents\Online Retail.xlsx")
```

```
In [5]: #Conferência se o campo "ID do cliente" está vazio, afinal, precisamos trabalhar apenas com os clientes identificáveis
# para compreender de forma real os indicadores sugeridos
df0[df0.isnull().any(axis=1)]
```

```
Out[5]:
```

	Transação	Cód Estoque	Descrição	Quantidade	Data Transação	Preço unitário	ID Cliente	País
	622	538414	22139	NaN	58	2010-12-01 11:52:00	0	NaN United Kingdom
	1443	538544	21773	DECORATIVE ROSE BATHROOM BOTTLE	1	2010-12-01 14:32:00	3	NaN United Kingdom
	1444	538544	21774	DECORATIVE CATS BATHROOM BOTTLE	2	2010-12-01 14:32:00	3	NaN United Kingdom
	1445	538544	21788	POLKADOT RAIN HAT	4	2010-12-01 14:32:00	1	NaN United Kingdom
	1446	538544	21787	RAIN PONCHO RETROSPOT	2	2010-12-01 14:32:00	2	NaN United Kingdom
...
	541536	581498	85099B	JUMBO BAG RED RETROSPOT	5	2011-12-09 10:28:00	4	NaN United Kingdom
	541537	581498	85099C	JUMBO BAG BAROQUE BLACK WHITE	4	2011-12-09 10:28:00	4	NaN United Kingdom
	541538	581498	85150	LADIES & GENTLEMEN METAL SIGN	1	2011-12-09 10:28:00	5	NaN United Kingdom
	541539	581498	85174	S/4 CACTI CANDLES	1	2011-12-09 10:28:00	11	NaN United Kingdom
	541540	581498	DOT	DOTCOM POSTAGE	1	2011-12-09 10:28:00	1,714	NaN United Kingdom

135080 rows x 8 columns

```
In [6]: # Deleta-se as linhas que não se pode identificar o cliente
# Copia-se a base de dados por segurança
df1 = df0.copy()
df1 = df1[pd.notnull(df1["ID Cliente"])]

# Ainda existem valores faltando? Se sim, excluí-los
boolMiss = df1.isnull().values.any()
if boolMiss:
    _ = [print(k,":",v) for k,v in df1.isnull().sum().items() if v!=0] # valores faltando
```

```
In [7]: # Revisão dos numerais numéricos para conferência
df1.describe
```

```
Out[7]:
```

<bound method NDFrame.describe of	Transação	Cód Estoque	Descrição	Quantidade	\
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	
1	536365	71053	WHITE METAL LANTERN	6	
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	
...
541904	581587	22613	PACK OF 20 SPACEBOY NAPKINS	12	
541905	581587	22899	CHILDREN'S APRON DOLLY GIRL	6	
541906	581587	23254	CHILDRENS CUTLERY DOLLY GIRL	4	
541907	581587	23255	CHILDRENS CUTLERY CIRCUS PARADE	4	
541908	581587	22138	BAKING SET 9 PIECE RETROSPOT	3	

	Data Transação	Preço unitário	ID Cliente	País
0	2010-12-01 08:26:00	3	17,850	United Kingdom
1	2010-12-01 08:26:00	3	17,850	United Kingdom
2	2010-12-01 08:26:00	3	17,850	United Kingdom
3	2010-12-01 08:26:00	3	17,850	United Kingdom
4	2010-12-01 08:26:00	3	17,850	United Kingdom
...
541904	2011-12-09 12:50:00	1	12,680	France
541905	2011-12-09 12:50:00	2	12,680	France
541906	2011-12-09 12:50:00	4	12,680	France
541907	2011-12-09 12:50:00	4	12,680	France
541908	2011-12-09 12:50:00	5	12,680	France

[406829 rows x 8 columns]>

```
In [8]: #Para prevenir que compras devolvidas de clientes contabilizem, exclui-se os valores numéricos que estejam negativos
df1[df1["Quantidade"] <= 0]
```

```
Out[8]:
```

	Transação	Cód Estoque	Descrição	Quantidade	Data Transação	Preço unitário	ID Cliente	País
141	C538379	D	Discount	-1	2010-12-01 09:41:00	28	14,527	United Kingdom
154	C538383	36004C	SET OF 3 COLOURED FLYING DUCKS	-1	2010-12-01 09:49:00	5	15,311	United Kingdom
235	C538391	22558	PLASTERS IN TIN CIRCUS PARADE	-12	2010-12-01 10:24:00	2	17,548	United Kingdom
236	C538391	21984	PACK OF 12 PINK PAISLEY TISSUES	-24	2010-12-01 10:24:00	0	17,548	United Kingdom
237	C538391	21983	PACK OF 12 BLUE PAISLEY TISSUES	-24	2010-12-01 10:24:00	0	17,548	United Kingdom
...
540449	C581490	23144	ZINC T-LIGHT HOLDER STARS SMALL	-11	2011-12-09 09:57:00	1	14,397	United Kingdom
541541	C581499	M	Manual	-1	2011-12-09 10:28:00	225	15,498	United Kingdom
541715	C581588	21258	VICTORIAN SEWING BOX LARGE	-5	2011-12-09 11:57:00	11	15,311	United Kingdom
541716	C581589	84978	HANGING HEART JAR T-LIGHT HOLDER	-1	2011-12-09 11:58:00	1	17,315	United Kingdom
541717	C581589	20979	38 PENCILS TUBE RED RETROSPOT	-5	2011-12-09 11:58:00	1	17,315	United Kingdom

8905 rows x 8 columns

```
In [9]: # Restringir análise apenas à transações com quantidade positiva
df1 = df1[df1["Quantidade"] > 0]

# Na coluna "Data Transação" temos o dia e hora que o cliente comprou. Como apenas a data será utilizada na análise, remove-se
# os Algarismos relativos à hora, pois não serão necessários.
df1["Data Transação"] = pd.to_datetime(df1["Data Transação"]).dt.date #normalize()
#df1.set_index("InvoiceDate", inplace=False)

# A biblioteca Pandas interpreta o "ID Cliente" como uma variável numérica. Será necessário redefinir
# como um tipo de objeto/string e não como número.
df1["ID Cliente"] = df1["ID Cliente"].astype(np.int64).astype(object)

# revisão das variáveis categóricas
df1.describe(include='object').T
```

```
Out[9]:
```

	count	unique	top	freq
Transação	397924	18538	576339	542
Cód Estoque	397924	3885	85123A	2035
Descrição	397924	3877	WHITE HANGING HEART T-LIGHT HOLDER	2028
Data Transação	397924	305	2011-11-08	3423
ID Cliente	397924	4339	17841	7847
País	397924	37	United Kingdom	354345

```
In [10]: # O conjunto de dados compreende 4.339 IDs de clientes exclusivos e 305 datas de compra dentro do período analisado
# O país (Localidade) com mais transações é o Reino Unido.

# Os números das faturas, códigos de estoque e descrições não fornecem informações significativas para a análise,
# portanto, exclui-se.

# A coluna "país" apresenta dados relevantes para a segmentação de cliente, contudo, por trabalhar-se com uma quantidade
# de mais de 400.000 dados será removida, para facilitar a manipulação. Os clientes serão analisados de uma perspectiva global.

# Deleta-se colunas que não são úteis
try:
    d1 = df1.drop(["Transação", "Cód Estoque", "Descrição", "País"], axis=1, inplace=True)
except:
    pass
df1.describe(include='object').T
```

```
Out[10]:
```

	count	unique	top	freq
Data Transação	397924	305	2011-11-08	3423
ID Cliente	397924	4339	17841	7847

```
In [11]: # Para determinar os valores monetários das transações - e posteriormente também os LTV de todos os clientes -
# insere-se uma nova coluna "Faturamento" multiplicando as quantidades pelos preços unitários correspondentes.

df1["Faturamento"] = df1["Quantidade"] * df1["Preço unitário"]
df1
```

```
Out[11]:
```

	Quantidade	Data Transação	Preço unitário	ID Cliente	Faturamento
0	6	2010-12-01	3	17850	15
1	6	2010-12-01	3	17850	20
2	8	2010-12-01	3	17850	22
3	6	2010-12-01	3	17850	20
4	6	2010-12-01	3	17850	20
...
541904	12	2011-12-09	1	12680	10
541905	6	2011-12-09	2	12680	13
541906	4	2011-12-09	4	12680	17
541907	4	2011-12-09	4	12680	17
541908	3	2011-12-09	5	12680	15

397924 rows x 5 columns

```
In [12]: # Revisão dos valores numéricos
df1.describe()
```

```
Out[12]:
```

	Quantidade	Preço unitário	Faturamento
count	397.924	397.924	397.924
mean	13	3	22
std	180	22	309
min	1	0	0
25%	2	1	5
50%	6	2	12
75%	12	4	20
max	80.995	8.143	168.470

```
In [13]: #Estrutura de data
date_grp = df1.groupby(df1["Data Transação"]).mean()
date_grp
```

```
Out[13]:
```

Data Transação	Quantidade	Preço unitário	Faturamento
2010-12-01	12	3	24
2010-12-02	16	3	24
2010-12-03	11	3	22
2010-12-05	6	3	12
2010-12-06	8	3	18
...
2011-12-05	11	3	17
2011-12-06	11	2	18
2011-12-07	19	3	32
2011-12-08	11	3	20
2011-12-09	147	2	299

305 rows x 3 columns

```
In [14]: # Recência, Frequência e Idade "T"
dfx = df1[df1["ID Cliente"] == 14527]
xmax_date = dfx["Data Transação"].max()
xmin_date = dfx["Data Transação"].min()

# Recência:
print("Data da primeira compra:", xmin_date)
print("Data da última compra:", xmax_date)
xrec = (xmax_date - xmin_date).days
print("Recência:", xrec) # recência é o tempo entre a primeira e última compra

# Idade "T":
xmaxall_date = df1["Data Transação"].max()
print("Data máxima do histórico:", xmaxall_date)
xage = (xmaxall_date - xmin_date).days # Idade "T"
print("T:", xage)

# Frequência:
xfreq = len(dfx[dfx["Quantidade"] > 0]).groupby("Data Transação").-1 # Frequência: períodos entre compras
print("Frequência:", xfreq)
```

```
Data da primeira compra: 2010-12-05
Data da última compra: 2011-12-07
Recência: 367
Data máxima do histórico: 2011-12-09
T: 369
Frequência: 53
```



```
In [15]: # Etapa de Treinamento/Teste (calibration/holdout)
t_holdout = 240 # dias reservados para o período de espera

max_date = df1["Data Transação"].max() # data final de observação
print("Data fim da observação:", max_date)

max_cal_date = max_date - timedelta(days=t_holdout) # data final escolhida para o período de calibração
print("Data fim do período de calibração:", max_cal_date)

df_ch = calibration_and_holdout_data(
    transactions = df1,
    customer_id_col = "ID Cliente",
    datetime_col = "Data Transação",
    monetary_value_col = "Faturamento",
    calibration_period_end = max_cal_date,
    observation_period_end = max_date,
    freq = "D")

print("Comportamento do cliente no período de calibração e espera")
pd.options.display.float_format = '{:,.0f}'.format
df_ch
```

Data fim da observação: 2011-12-09
 Data fim do período de calibração: 2011-04-13
 Comportamento do cliente no período de calibração e espera

```
Out[15]:
```

	frequency_cal	reency_cal	T_cal	monetary_value_cal	frequency_holdout	monetary_value_holdout	duration_holdout
ID Cliente							
12346	0	0	85	0	0	0	240
12347	2	121	127	556	4	25	240
12348	2	110	118	297	1	103	240
12350	0	0	70	0	0	0	240
12352	3	34	56	422	3	20	240
...
18270	0	0	26	0	1	29	240
18272	0	0	6	0	5	18	240
18273	0	0	17	0	2	76	240
18280	0	0	37	0	0	0	240
18283	2	53	97	105	11	3	240

2299 rows x 7 columns

```
In [16]: df_ch.describe()
```

```
Out[16]:
```

	frequency_cal	reency_cal	T_cal	monetary_value_cal	frequency_holdout	monetary_value_holdout	duration_holdout
count	2,299	2,299	2,299	2,299	2,299	2,299	2,299
mean	1	31	79	176	3	28	240
std	2	44	43	438	5	113	0
min	0	0	0	0	0	0	240
25%	0	0	42	0	1	3	240
50%	0	0	82	0	2	16	240
75%	1	65	124	252	4	23	240
max	35	133	133	8,030	96	4,307	240

```
In [17]: # treinamento: frequência
pd.options.display.float_format = '{:,.3f}'.format
x = df_ch["frequency_cal"].value_counts(normalize=True)
x = x.nlargest(15)
print("Frequência:")
x.sort_index(ascending=True)
```

```
Out[17]:
```

0.000	0.591
1.000	0.184
2.000	0.092
3.000	0.051
4.000	0.027
5.000	0.017
6.000	0.009
7.000	0.008
8.000	0.004
9.000	0.004
10.000	0.003
11.000	0.001
13.000	0.001
14.000	0.001
17.000	0.001

Name: frequency_cal, dtype: float64

```

In [18]: # training: axis length
max_freq = df_ch["frequency_cal"].quantile(0.98)
max_rec = df_ch["recency_cal"].max()
max_T = df_ch["T_cal"].max()

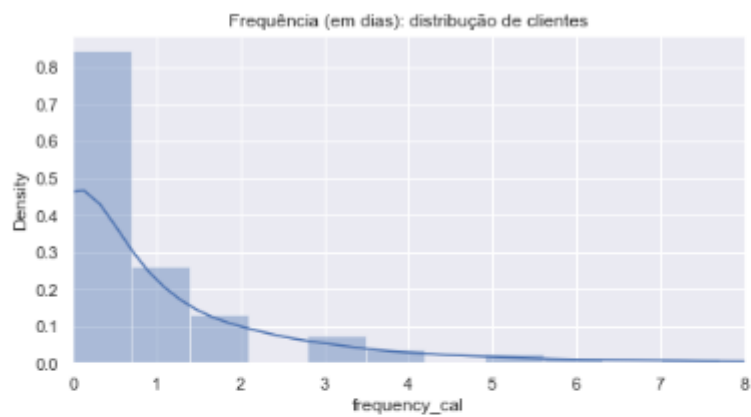
# treinamento
fig = plt.figure(figsize=(8, 4))
ax = sns.distplot(df_ch["frequency_cal"])
ax.set_xlim(0, max_freq)
ax.set_title("Frequência (em dias): distribuição de clientes");

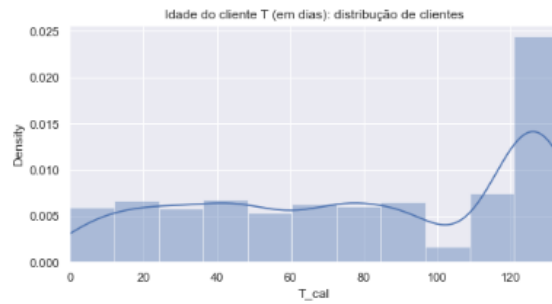
# treinamento
fig = plt.figure(figsize=(8, 4))
ax = sns.distplot(df_ch["recency_cal"])
ax.set_xlim(0, max_rec)
ax.set_title("Recência (em dias): distribuição de clientes")

# treinamento
fig = plt.figure(figsize=(8, 4))
ax = sns.distplot(df_ch["T_cal"])
ax.set_xlim(0, max_T)
ax.set_title("Idade do cliente T (em dias): distribuição de clientes")

```

Out[18]: Text(0.5, 1.0, 'Idade do cliente T (em dias): distribuição de clientes')





In [19]: # determinar recência, frequência, idade do cliente T, faturamento de cada cliente

```
df_rft = summary_data_from_transaction_data(
    transactions = df1,
    customer_id_col = "ID Cliente",
    datetime_col = "Data Transação",
    monetary_value_col = "Faturamento",
    observation_period_end = max_date,
    freq = "D")

pd.options.display.float_format = '{:,.0f}'.format
df_rft
```

In [19]: # determinar recência, frequência, idade do cliente T, faturamento de cada cliente

```
df_rft = summary_data_from_transaction_data(
    transactions = df1,
    customer_id_col = "ID Cliente",
    datetime_col = "Data Transação",
    monetary_value_col = "Faturamento",
    observation_period_end = max_date,
    freq = "D")

pd.options.display.float_format = '{:,.0f}'.format
df_rft
```

Out[19]:

ID Cliente	frequency	recency	T	monetary_value
12346	0	0	325	0
12347	6	365	367	600
12348	3	283	358	301
12349	0	0	18	0
12350	0	0	310	0
...
18280	0	0	277	0
18281	0	0	180	0
18282	1	119	126	78
18283	13	334	337	153
18287	2	159	201	538

4339 rows × 4 columns

Aplicação do modelo beta-geometric/NBD

```
In [20]: # Modelo BG/NBD
bgf = BetaGeoFitter(penalizer_coef=1e-06)
bgf.fit(
    frequency = df_rft["frequency"],
    recency = df_rft["recency"],
    T = df_rft["T"],
    weights = None,
    verbose = True,
    tol = 1e-06)
pd.options.display.float_format = '{:,.3f}'.format
bgf.summary
```

```
Optimization terminated successfully.
Current function value: -2.709794
Iterations: 38
Function evaluations: 39
Gradient evaluations: 39
```

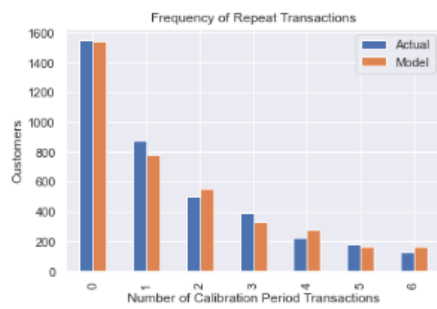
```
Out[20]:
```

	coef	se(coef)	lower 95% bound	upper 95% bound
r	0.827	0.027	0.774	0.879
alpha	88.910	2.618	83.779	74.041
a	0.002	0.006	-0.010	0.014
b	2.967	6.267	-9.317	15.250

```
In [21]: # treinamento: o modelo reflete dados reais de forma suficientemente próxima?
```

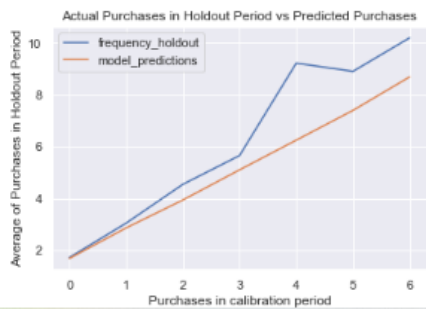
```
# frequência de transações repetidas: previsão X real
fig = plt.figure(figsize=(20, 20))
plot_period_transactions(bgf);
```

```
<Figure size 1440x1440 with 0 Axes>
```



```
In [22]: # Teste de compras previstas X reais no período de holdout
fig = plt.figure(figsize=(40, 40))
plot_calibration_purchases_vs_holdout_purchases(bgf, df_ch);
```

<Figure size 2880x2880 with 0 Axes>



```
In [23]: # Determina-se recência, frequência, faturamento para toda a base de clientes
```

```
df_rft = summary_data_from_transaction_data(
    transactions = df1,
    customer_id_col = "ID Cliente",
    datetime_col = "Data Transação",
    monetary_value_col = "Faturamento",
    observation_period_end = max_date,
    freq = "D")

pd.options.display.float_format = '{:,.0f}'.format
df_rft
```

```
Out[23]:
```

ID Cliente	frequency	recency	T	monetary_value
12346	0	0	325	0
12347	6	365	367	600
12348	3	283	358	301
12349	0	0	18	0
12350	0	0	310	0
...
18280	0	0	277	0
18281	0	0	180	0
18282	1	119	126	78
18283	13	334	337	153
18287	2	159	201	536

4339 rows x 4 columns

```
In [24]: df_rft.describe()
```

```
Out[24]:
```

```
Out[24]:
```

	frequency	recency	T	monetary_value
count	4,339	4,339	4,339	4,339
mean	3	131	223	307
std	6	132	118	2,812
min	0	0	0	0
25%	0	0	112	0
50%	1	93	248	177
75%	3	262	326	360
max	131	373	373	168,470

```
In [25]: # Modelo BG/NBD
bgf = BetaGeoFitter(penalizer_coef=1e-06)
bgf.fit(
    frequency = df_rft["frequency"],
    recency = df_rft["recency"],
    T = df_rft["T"],
    weights = None,
    verbose = True,
    tol = 1e-06)
pd.options.display.float_format = '{:,.3f}'.format
bgf.summary
```

```
Optimization terminated successfully.
Current function value: -2.709794
Iterations: 38
Function evaluations: 39
Gradient evaluations: 39
```

```
Out[25]:
```

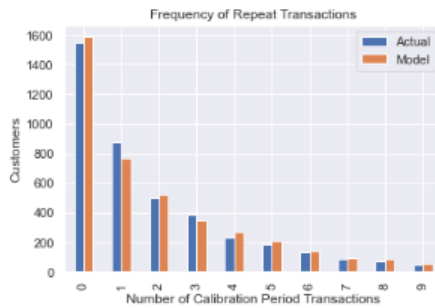
	coef	se(coef)	lower 95% bound	upper 95% bound
r	0.827	0.027	0.774	0.879
alpha	68.910	2.618	63.779	74.041
a	0.002	0.006	-0.010	0.014
b	2.967	6.267	-9.317	15.250

```
In [26]: #axis Length
max_freq = int(df_rft["frequency"].max())
max_T = int(df_rft["T"].max())
max_rec = int(df_rft["recency"].max())
print(max_freq)

# frequência de transações repetidas: prevista X real
fig = plt.figure(figsize=(20, 20))
plot_period_transactions(
    model = bgf,
    max_frequency = 10);
```

131

<Figure size 1440x1440 with 0 Axes>



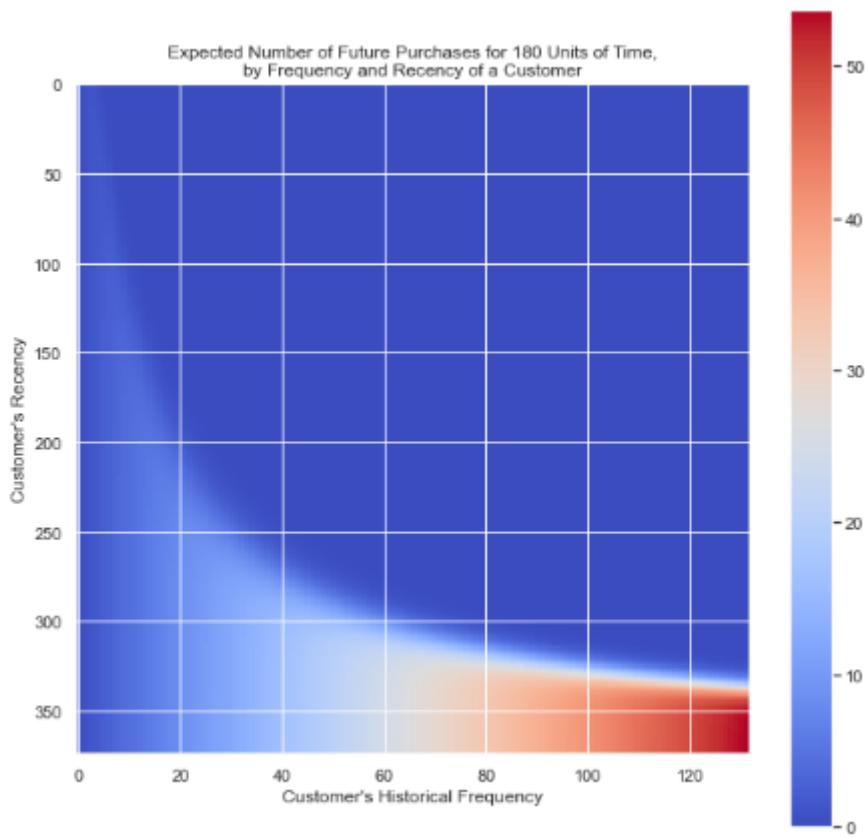
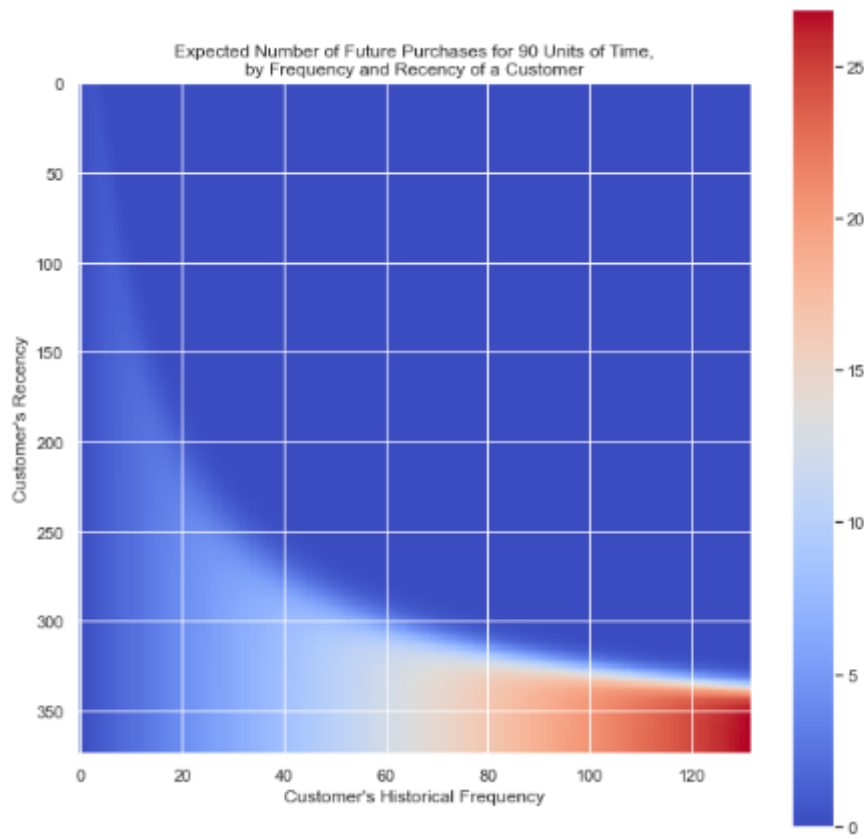
Compras previstas

```
In [27]: #Compras previstas em até 90 dias

# função auxiliar: plota o número esperado de compras futuras que um cliente com
# frequência f e recência r farão no período de previsão

def plot_freq_rec(t):
    fig = plt.figure(figsize=(10, 10))
    plot_frequency_recency_matrix(
        model = bgf,
        T = t,
        max_frequency = max_freq,
        max_recency = max_rec);

# Chamado da função plotter entre 90 e 180 dias
tFC = [90, 180]
_ = [plot_freq_rec(t) for t in tFC]
```




```
In [28]: # previsão de compras de cliente selecionado em período "T"
t = 180
custID = 12500

df_rft_C = df_rft.loc[custID,:]
predC = bgf.predict(
    t,
    df_rft_C["frequency"],
    df_rft_C["recency"],
    df_rft_C["T"])
print("Cliente", custID, ": número esperado de compras em", t, "dias =", f'{predC:.1f}')
```

Cliente 12500 : número esperado de compras em 180 dias = 4.4

```
In [29]: # função auxiliar "predict_purch": previsão de compras de todos os clientes em período "T"
def predict_purch(df, t):
    df["predict_purch_" + str(t)] = \
        bgf.predict(
            t,
            df["frequency"],
            df["recency"],
            df["T"])

## função auxiliar "predict_purch": previsão de compras de todos os clientes em múltiplos períodos "T"
t_FC = [90,180,270]
_ = [predict_purch(df_rft, t) for t in t_FC]
pd.options.display.float_format = '{:,.1f}'.format
print("Previsão do número de compras de cada cliente nos próximos 'T' dias:")
df_rft
```

Previsão do número de compras de cada cliente nos próximos 'T' dias:

```
Out[29]:
```

	frequency	recency	T	monetary_value	predict_purch_90	predict_purch_180	predict_purch_270
ID Cliente							
12346	0.0	0.0	325.0	0.0	0.2	0.4	0.6
12347	6.0	365.0	367.0	599.7	1.4	2.8	4.2
12348	3.0	283.0	358.0	301.5	0.8	1.6	2.4
12349	0.0	0.0	18.0	0.0	0.9	1.7	2.6
12350	0.0	0.0	310.0	0.0	0.2	0.4	0.6
...
18280	0.0	0.0	277.0	0.0	0.2	0.4	0.6
18281	0.0	0.0	180.0	0.0	0.3	0.6	0.9
18282	1.0	119.0	126.0	77.8	0.8	1.7	2.6
18283	13.0	334.0	337.0	152.8	3.1	6.1	9.2
18287	2.0	159.0	201.0	536.0	0.9	1.9	2.8

4339 rows x 7 columns

```
In [30]: # top 10 clientes nos próximos 30 dias
print("top 10 clientes pela previsão de compras nos próximos 90 dias")
df_rft.sort_values(by="predict_purch_90", ascending=False).head(10)
```

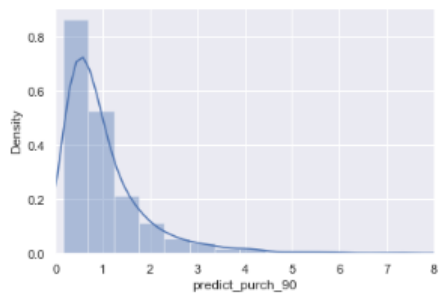
top 10 clientes pela previsão de compras nos próximos 90 dias

```
Out[30]:
```

	frequency	recency	T	monetary_value	predict_purch_90	predict_purch_180	predict_purch_270
ID Cliente							
14911	131.0	372.0	373.0	1,093.7	28.8	53.7	80.5
12748	113.0	373.0	373.0	298.4	23.2	46.3	69.5
17841	111.0	372.0	373.0	384.5	22.8	45.5	68.3
15311	89.0	373.0	373.0	677.7	18.3	36.6	54.9
14606	88.0	372.0	373.0	135.9	18.1	36.2	54.2
12971	70.0	369.0	372.0	159.2	14.5	28.9	43.3
13089	65.0	367.0	369.0	893.7	13.5	27.0	40.6
14527	53.0	367.0	369.0	155.0	11.1	22.1	33.2
13798	52.0	371.0	372.0	706.7	10.8	21.6	32.3
16422	47.0	352.0	369.0	702.5	9.8	19.6	29.5

```
In [31]: #distplot da seaborn
print("Previsão do número de compras nos próximos 90 dias:")
ax = sns.distplot(df_rft["predict_purch_90"]).set_xlim(0,8)
```

Previsão do número de compras nos próximos 90 dias:



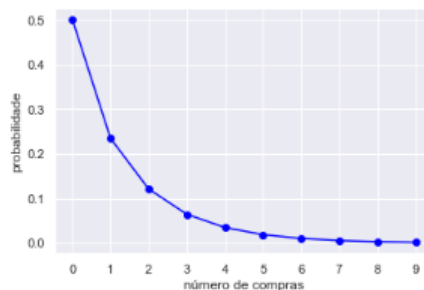
```
In [32]: # O método Probabilidades de tempo de vida probabilidade_of_n_purchases_up_to_time
# calcula o número médio de transações por cliente em um período de tempo.

# função auxiliar: probabilidade 'n' compras repetidas em um período 't'
def prob_purch(t, n):
    p = bgf.probability_of_n_purchases_up_to_time(t, n)
    return p

# Chamada da função auxiliar: probabilidade 'n' compras repetidas em um período 't'
t = 90 # período analisado
purch = range(0,10,1) # número de compras repetidas
probs = [prob_purch(t,n) for n in purch]

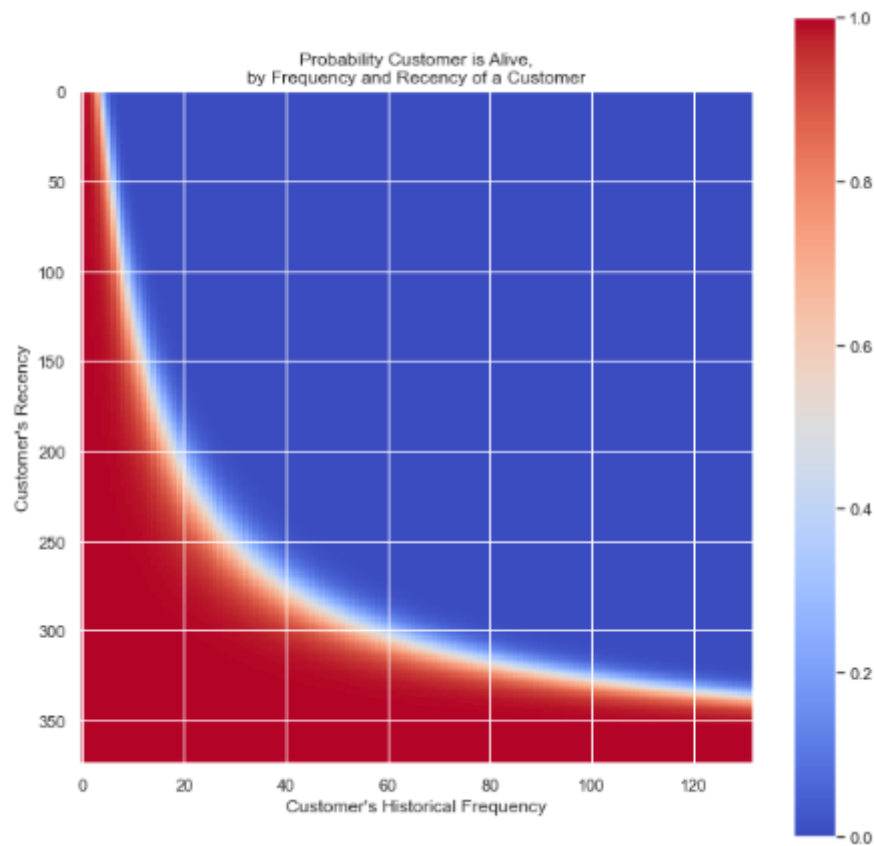
dict_probs = dict(zip(purch, probs))
#_ = [print(k,":",f'{v:.2f}') for k,v in dict_probs.items()]

plt.plot(purch, probs, '-o', color='blue')
plt.xlabel("número de compras")
plt.ylabel("probabilidade")
plt.xticks(purch)
plt.show()
```



Probabilidade de churn dos clientes

```
In [33]: # probabilidade de um cliente não se evadir (continuar vivo)
# baseado em clientes com recência r e frequência f específicas
fig = plt.figure(figsize=(10, 10))
plot_probability_alive_matrix(
    model = bgf,
    max_frequency = max_freq,
    max_recency = max_rec);
```



```
In [34]: # probabilidade de cada cliente estar 'vivo' em dataframe
prob_alive = bgf.conditional_probability_alive(
    frequency = df_rft["frequency"],
    recency = df_rft["recency"],
    T = df_rft["T"])

df_rft["prob_vivo"] = prob_alive
pd.options.display.float_format = '{:,.2f}'.format
df_rft.describe()
```

```
Out[34]:
```

	frequency	recency	T	monetary_value	predict_purch_90	predict_purch_180	predict_purch_270	prob_vivo
count	4,339.00	4,339.00	4,339.00	4,339.00	4,339.00	4,339.00	4,339.00	4,339.00
mean	2.86	130.74	222.78	306.94	1.08	2.15	3.23	1.00
std	5.95	132.21	117.88	2,612.45	1.27	2.54	3.81	0.01
min	0.00	0.00	0.00	0.00	0.17	0.34	0.50	0.75
25%	0.00	0.00	112.00	0.00	0.45	0.90	1.35	1.00
50%	1.00	93.00	248.00	176.94	0.77	1.53	2.30	1.00
75%	3.00	252.00	326.00	360.16	1.27	2.54	3.81	1.00
max	131.00	373.00	373.00	166,469.60	28.84	53.68	80.50	1.00

```
In [37]: # Mostrar clientes com probabilidade menor de 90% de estarem 'vivos'
df_rft[df_rft["prob_vivo"] < 0.9]
```

```
Out[37]:
```

ID Cliente	frequency	recency	T	monetary_value	predict_purch_90	predict_purch_180	predict_purch_270	prob_vivo
13093	6.00	98.00	373.00	1,172.18	1.18	2.36	3.54	0.85
15107	5.00	61.00	372.00	40.70	0.89	1.79	2.68	0.75
15235	9.00	166.00	373.00	240.89	1.77	3.54	5.32	0.89
16725	5.00	69.00	371.00	131.77	0.97	1.94	2.91	0.81

Método plot_history_alive

Mostra o histórico de transações do cliente durante todo o período disponível

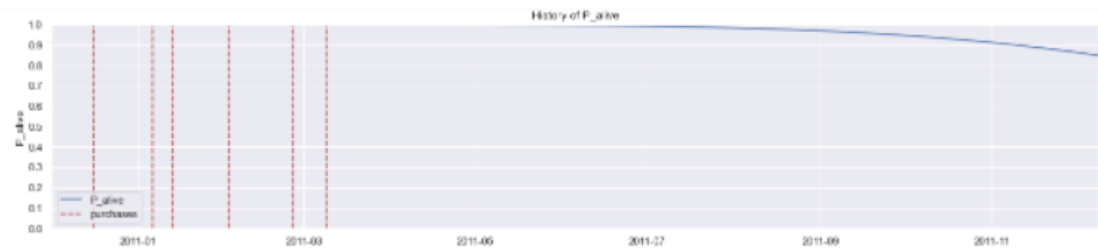
```
In [38]: # Selecionar um cliente
cliente_id = 13093

df1C = df1[df1["ID Cliente"] == cliente_id]
df1C.describe()

#Definição
max_date = df1["Data Transação"].max()
min_date = df1["Data Transação"].min()
período_dias = (max_date - min_date).days
print(período_dias)

# histórico do cliente selecionado: probabilidade ao longo do tempo de estar 'vivo'
fig = plt.figure(figsize=(20,4))
print("Cliente",cliente_id,": probabilidade de estar vivo ao longo do tempo")
plot_history_alive(
    model = bgf,
    t = período_dias,
    transactions = df1C,
    datetime_col = "Data Transação");
```

373
Cliente 13093 : probabilidade de estar vivo ao longo do tempo



```
In [39]: # probabilidade de um cliente selecionado estar 'vivo': desenvolvimento ao longo do tempo
path_alive = calculate_alive_path(
    model = bgf,
    t = período_dias,
    transactions = df1C,
    datetime_col = "Data Transação");
print("Cliente",cliente_id,"probabilidade de estar vivo, desenvolvimento ao longo do tempo")
path_alive.explode().astype(np.float64).describe()
```

Cliente 13093 probabilidade de estar vivo, desenvolvimento ao longo do tempo

```
Out[39]: count    374.00
mean      0.97
std       0.04
min       0.85
25%      0.97
50%      1.00
75%      1.00
max       1.00
dtype: float64
```

Estimando valor monetário com modelo Gamma-Gamma

```
In [40]: # selecionar apenas clientes com valor de compra > 0
df_rftv = df_rft[df_rft["monetary_value"] > 0]
pd.options.display.float_format = '{:,.2f}'.format
df_rftv.describe()
```

```
Out[40]:
```

	frequency	recency	T	monetary_value	predict_purch_90	predict_purch_180	predict_purch_270	prob_vivo
count	2,790.00	2,790.00	2,790.00	2,790.00	2,790.00	2,790.00	2,790.00	2,790.00
mean	4.45	203.33	260.12	477.36	1.43	2.86	4.28	1.00
std	6.93	111.48	101.48	3,245.82	1.46	2.92	4.38	0.01
min	1.00	1.00	0.00	1.00	0.37	0.73	1.10	0.75
25%	1.00	106.00	197.25	187.17	0.87	1.34	2.02	1.00
50%	3.00	211.00	284.00	309.47	1.07	2.13	3.20	1.00
75%	5.00	301.75	358.00	458.37	1.66	3.32	4.98	1.00
max	131.00	373.00	373.00	168,469.80	26.84	53.88	80.50	1.00

```
In [41]: # Gamma-Gamma requer uma correlação de Pearson próxima de 0
# entre frequência de compra e valor da transação
corr_matrix = df_rftv[["monetary_value", "frequency"]].corr()
corr = corr_matrix.iloc[1,0]
print("Correlação de Pearson: {:.3f} % corr)
```

Correlação de Pearson: 0.016

```
In [42]: # Modelo Gamma-Gamma
ggf = GammaGammaFitter(penalizer_coef = 1e-06)
ggf.fit(
    frequency = df_rftv["frequency"],
    monetary_value = df_rftv["monetary_value"],
    weights = None,
    verbose = True,
    tol = 1e-06,
    q_constraint = True)
pd.options.display.float_format = '{:,.3f}'.format
ggf.summary
```

```
Out[42]:
```

	coef	se(coef)	lower 95% bound	upper 95% bound
p	4.176	0.188	3.811	4.541
q	2.415	0.070	2.279	2.552
v	149.875	8.304	133.600	166.151

```
In [51]: # Previsão do faturamento com cada cliente, baseado na frequência e valores de todas as últimas compras
faturamento_esperado = ggf.conditional_expected_average_profit(
    df_rftv["frequency"],
    df_rftv["monetary_value"])

df_rftv["faturamento_esperado"] = faturamento_esperado
df_rftv["media_esperada"] = df_rftv["monetary_value"]
df_rftv["margem_erro"] = df_rftv["faturamento_esperado"] - df_rftv["media_esperada"]

mape = mean_absolute_percentage_error(faturamento_esperado, df_rftv["monetary_value"])
print("Mapa de faturamento esperado:", f'{mape:.2f}')

pd.options.display.float_format = '{:,.3f}'.format
df_rftv.head()
```

Mapa de faturamento esperado: 0.12

Out[51]:

	CLV	LTV	frequency	recency	T	monetary_value	predict_purch_90	predict_purch_180	predict_purch_270	prob_vivo	faturamento_e
ID Cliente											
12347	11,024.213	11,024.213	6.000	365.000	367.000	599.702	1.409	2.817	4.228	1.000	
12348	3.368.058	3.368.058	3.000	283.000	358.000	301.480	0.806	1.612	2.417	0.999	
12352	8,286.615	8,286.615	6.000	280.000	296.000	368.257	1.683	3.365	5.046	1.000	
12356	2,518.454	2,518.454	2.000	303.000	325.000	269.905	0.845	1.291	1.936	0.999	
12358	6,175.758	6,175.758	1.000	149.000	150.000	683.200	0.750	1.500	2.250	0.999	

LTV

```
In [64]: # Calculo do CLV
desconto_anual = 0.10 # média de desconto anual
VIDA = 48 # tempo de vida esperado para os clientes em meses

discount_m = (1 + desconto_anual)**(1/12) - 1 # distribui o desconto anual por mês

LTV = ggf.customer_lifetime_value(
    transaction_prediction_model = bgf,
    frequency = df_rftv["frequency"],
    recency = df_rftv["recency"],
    T = df_rftv["T"],
    monetary_value = df_rftv["monetary_value"],
    time = VIDA,
    freq = "D",
    discount_rate = discount_m)

df_rftv.insert(0, "CLV_previsto2", LTV) # valor CLV esperado
df_rftv.describe().T
```

Out[64]:

	count	mean	std	min	25%	50%	75%	max
CLV_previsto2	2,790,000	9,723.082	35,132.922	535.168	2,587.213	4,639.829	8,684.078	999,257.238
CLV_previsto1	2,790,000	9,723.082	35,132.922	535.168	2,587.213	4,639.829	8,684.078	999,257.238
CLV_previsto	2,790,000	9,723.082	35,132.922	535.168	2,587.213	4,639.829	8,684.078	999,257.238
CLV	2,790,000	9,723.082	35,132.922	535.168	2,587.213	4,639.829	8,684.078	999,257.238
LTV	2,790,000	9,723.082	35,132.922	535.168	2,587.213	4,639.829	8,684.078	999,257.238
frequency	2,790,000	4.454	8.930	1.000	1.000	3.000	5.000	131.000
recency	2,790,000	203.329	111.460	1.000	106.000	211.000	301.750	373.000
T	2,790,000	260.121	101.480	9.000	197.250	284.000	358.000	373.000
monetary_value	2,790,000	477.359	3,245.622	1.000	187.165	309.473	458.371	168,469.600
predict_purch_90	2,790,000	1.428	1.459	0.365	0.872	1.087	1.660	28.843
predict_purch_180	2,790,000	2.856	2.918	0.731	1.345	2.133	3.319	63.677
predict_purch_270	2,790,000	4.283	4.376	1.096	2.017	3.200	4.977	80.504
prob_vivo	2,790,000	0.998	0.007	0.752	0.999	0.999	1.000	1.000
faturamento_esperado	2,790,000	469.989	2,437.796	35.325	228.001	329.162	455.733	125,940.166
media_esperada	2,790,000	477.359	3,245.622	1.000	187.165	309.473	458.371	168,469.600
margem_erro	2,790,000	-7.370	814.660	-42,529.434	-1.025	11.492	33.887	111.687

```
In [65]: # Visão CLV por cliente em dataframe
df_rftv.sort_values(by="LTV", ascending=False)
```

Out[65]:

	CLV_previsto2	CLV_previsto1	CLV_previsto	CLV	LTV	frequency	recency	T	monetary_value	predict_purch_90	predict_purch_180
ID Cliente											
16446	999,257.238	999,257.238	999,257.238	999,257.238	999,257.238	1.000	205.000	205.000	168,469.600	0.600	1.16
14646	797,753.273	797,753.273	797,753.273	797,753.273	797,753.273	44.000	363.000	364.000	6,366.706	9.537	19.01
18102	651,125.338	651,125.338	651,125.338	651,125.338	651,125.338	25.000	367.000	367.000	9,349.477	5.331	10.66
17450	535,978.639	535,978.639	535,978.639	535,978.639	535,978.639	26.000	359.000	367.000	7,404.690	5.537	11.01
14096	470,650.978	470,650.978	470,650.978	470,650.978	470,650.978	16.000	97.000	101.000	4,071.434	8.908	17.81
...
12755	624.393	624.393	624.393	624.393	624.393	1.000	85.000	366.000	17.650	0.377	0.76
14865	607.251	607.251	607.251	607.251	607.251	1.000	365.000	372.000	15.000	0.373	0.74
17194	603.290	603.290	603.290	603.290	603.290	1.000	87.000	360.000	10.000	0.382	0.76
18037	566.816	566.816	566.816	566.816	566.816	5.000	208.000	362.000	7.744	1.213	2.42
17816	535.168	535.168	535.168	535.168	535.168	3.000	329.000	366.000	6.783	0.793	1.56

2790 rows x 16 columns