

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO TECNOLÓGICO  
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO E SISTEMAS  
ENGENHARIA DE PRODUÇÃO MECÂNICA

Pedro Baldissera Trajano

**Análise dos fatores de sucesso de uma rede de franquias do setor fitness pela Regressão  
Logística**

Florianópolis

2022

Pedro Baldissera Trajano

# **Análise dos fatores de sucesso de uma rede de franquias do setor fitness pela Regressão Logística**

Trabalho de Conclusão de Curso submetido ao curso de Engenharia de Produção Mecânica do Centro Tecnológico da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Bacharel em Engenharia Mecânica com habilitação em Engenharia de Produção.

Orientador: Prof. Rogério Feroldi Miorando, Dr.

Florianópolis

2022

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Trajano, Pedro Baldissera

Análise dos fatores de sucesso de uma rede de franquias do setor fitness pela regressão logística / Pedro Baldissera Trajano ; orientador, Rogério Feroldi Miorando, 2022.

85 p.

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Santa Catarina, Centro Tecnológico, Graduação em Engenharia de Produção Mecânica, Florianópolis, 2022.

Inclui referências.

1. Engenharia de Produção Mecânica. 2. Regressão logística. 3. Fatores de sucesso. 4. Franquias. 5. Setor fitness. I. Miorando, Rogério Feroldi. II. Universidade Federal de Santa Catarina. Graduação em Engenharia de Produção Mecânica. III. Título.

Pedro Baldissera Trajano

**Análise dos fatores de sucesso de uma rede de franquias do setor fitness pela Regressão  
Logística**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do título de Engenheiro Mecânico com habilitação em Produção e aprovado em sua forma final pelo Curso de Engenharia de Produção Mecânica.

Florianópolis, 08 de dezembro de 2022.

---

Mônica Maria Mendes Luna  
Coordenadora do curso

**Banca examinadora**

---

Prof. Rogério Feroldi Miorando, Dr.  
Orientador(a)

---

Prof. Lynceo Falavigna Braghirolli, Dr.(a)  
UFSC

---

Prof. Marco Antonio de Oliveira Vieira Goulart, Dr.(a)  
UFSC

Florianópolis, 2022.

## AGRADECIMENTOS

Aos meus queridos pais, Cristina e José Eduardo, que nunca mediram esforços para me incentivar e fornecer todo o suporte para que eu construísse meu próprio caminho. À minha amada irmã Deborah que, mesmo distante, consegue se fazer presente e ser uma excelente amiga.

Agradeço em especial à minha companheira, Daniela Barbim Gonçalves, que esteve sempre presente, me apoiou incondicionalmente durante os dias mais difíceis dessa jornada e meu deu forças para prosseguir.

Aos amigos, responsáveis por uma infinidade de momentos de felicidade compartilhados e pela parceria que é impossível descrever.

Ao meu orientador, Rogério Feroldi Miorando, por me guiar na aplicação deste trabalho e fornecer os esclarecimentos para sua finalização.

À UFSC e ao Departamento de Engenharia de Produção, pelo ensino de qualidade e portas abertas que ampliaram meu horizonte de conhecimento.

À toda equipe da EFX7, Equal e parceiros, onde encontrei pessoas incríveis que me proporcionaram um desenvolvimento pessoal e profissional que jamais poderia esperar.

Por fim, agradeço a todos aqueles que de alguma forma contribuíram para essa pesquisa ou para minha formação.

Muito grato!

O aumento do conhecimento é como uma esfera dilatando-se no espaço: quanto maior a nossa compreensão, maior o nosso contato com o desconhecido.

(Blaise Pascal)

## RESUMO

A inauguração e expansão de novos modelos de negócios é marcada por incertezas que influenciam o desempenho das organizações. No setor de varejo e de serviço, já é conhecida a importância dos fatores de localização na performance das operações, e também o desafio que é encontrar o ponto ideal para uma unidade. O desafio se estende para uma ampla pluralidade de fatores - muitas das vezes despercebidos - que estão ligados com a entrega do serviço ou produto, ou com a percepção de valor pelo cliente. O presente trabalho busca investigar quais os principais fatores de sucesso das operações de uma rede de franquia do setor fitness, e com isso, auxiliar no processo de avaliação de pontos para novas unidades de franquia. Foi realizado um amplo entendimento do modelo de negócio proposto pela franqueadora seguido por um mapeamento e definição das principais variáveis relacionadas ao sucesso de suas franquias. Após uma priorização, foi estabelecido que o sucesso de uma operação pode ser impactado principalmente por variáveis nas categorias de Gestão, Imóvel e Estrutura, e Sociodemográfica. Através de um modelo de regressão logística, foi possível avaliar quais destes fatores exercem maior influência no sucesso de uma operação da franquia em questão, e modelar a probabilidade de sucesso por estas variáveis. Os resultados obtidos demonstraram que franquias com boa capacidade de gestão, localizadas em regiões com renda média elevada e boa concentração de trabalhadores apresentam maior probabilidade de obterem sucesso ao longo de sua operação. Assim, o processo de abertura de uma nova unidade da franquia pode ser auxiliado pela estimativa de probabilidade de sucesso da futura operação.

**Palavras-chave:** Regressão logística; Fatores de sucesso; Localização; Franquias; Setor *fitness*.

## ABSTRACT

The implementation and the expansion of a new business model are characterized with uncertainties which influences the organizations' performance. In the retail and services sector, the relevance of the localization factors on the performance of the operations is already well known, and so is the challenge of locating the ideal spot for a unit. The challenge extends to a plurality of other factors - frequently unnoticed - that are linked to the product or service quality, or to the customer's perception of value. This study aims to investigate the main success factors in the operations of a franchise network in the fitness sector, and thus to assist in the evaluation process for new franchise units. An extensive analysis was conducted to understand the franchisee's proposed business model, followed by a mapping and a definition of the main variables associated with the successful operation of their franchises. Once prioritized, it was settled that the success of an operation can be impacted mainly by variables in the following categories: Management, Property and Structure, and Sociodemographic. Through a logistic regression model, it became possible to assess which of these factors have the most influence on a successful franchise operation and model the probability of success by these variables. The obtained results revealed that the franchises with adequate management capacity, located in regions with high average income and a good concentration of workers, are more likely to be successful throughout their operation. Therefore, the process of opening a new franchise unit can be assisted by the estimate of the probability of success of the future operation.

**Keywords:** Logistic regression; Success factors; Localization; Franchising; Fitness industry.

## LISTA DE FIGURAS

Figura 1 - Evolução do faturamento de franquias no Brasil.	17
Figura 2 – Reta de regressão linear vs.curva logística.	20
Figura 3 – Enquadramento metodológico do trabalho.	24
Figura 4 – Etapas e sub etapas metodológicas do trabalho.	25
Figura 5 – Exemplo de fatores e subfatores mapeados e priorizados.	28
Figura 6 – Subfatores que compõem o fator “Desempenho Operacional”.	35
Figura 7 – Subfatores que compõem o fator “Imóvel e Estrutura”.	36
Figura 8 – Elementos de “Personalidade” em um imóvel da franquia.	37
Figura 9 – Subfatores que compõem o fator “Gestor”.	38
Figura 10 – Subfatores que compõem o fator “Sociodemografia”.	40
Figura 11 – Mapa de fatores e subfatores de sucesso.	41
Figura 12 – Mapa de fatores e subfatores de sucesso priorizado	43
Figura 13 – Localização das unidades de franquia	47
Figura 14 – Conjunto de studios e alunos ativos.	48
Figura 15 – Áreas de estudo de São Paulo.	49
Figura 16 – Perfil etário dos alunos da rede de franquias (quantidade).	50
Figura 17 – Alunos ativos de unidades em São Paulo/SP.	52
Figura 18 – Alunos ativos de unidades em Caxias do Sul/RS.	52
Figura 19 – Estatística de colinearidade das variáveis predictoras.	55
Figura 20 – Correlações das variáveis do modelo	56

## LISTA DE QUADROS

Quadro 1 – Pesos dos subfatores de Imóvel e Estrutura.	45
Quadro 2 – Pesos dos subfatores de Imóvel e Estrutura.	45

## LISTA DE TABELAS

Tabela 1 – Matriz de confusão.	58
Tabela 2 – Resultados da regressão logística.	58
Tabela 3 – Resultados da regressão logística.	59

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	<b>12</b>
1.1 OBJETIVOS	13
<b>1.1.1 Objetivo Geral</b>	<b>13</b>
<b>1.1.2 Objetivos Específicos</b>	<b>13</b>
1.2 DELIMITAÇÕES E LIMITAÇÕES DO TRABALHO	14
<b>2 FUNDAMENTAÇÃO TEÓRICA</b>	<b>15</b>
2.1 <i>BUSINESS FRANCHISE MODEL</i> (MODELO DE FRANQUIA DE NEGÓCIOS)	15
2.2 SEGMENTAÇÃO DE CENTROS DE ATIVIDADE FÍSICA	17
2.3 REGRESSÃO LOGÍSTICA (RL)	18
<b>2.3.1 Testes de ajustes e qualidade</b>	<b>21</b>
<b>3 METODOLOGIA</b>	<b>24</b>
3.1 CARACTERIZAÇÃO DE PESQUISA	24
3.2 ETAPAS METODOLÓGICAS	24
<b>3.2.1 Identificação de variáveis</b>	<b>26</b>
3.2.1.1 <i>Definição de fatores de sucesso de operação</i>	26
3.2.1.1.1 Entendimento do modelo de negócio	26
3.2.1.1.2 Análise de indicadores chave de performance	26
3.2.1.1.3 Entrevistas	27
3.2.1.1.4 Priorização	27
<b>3.2.2 Coleta e tratamento de dados</b>	<b>28</b>
3.2.2.1 <i>Dados de inteligência de negócio</i>	28
3.2.2.2 <i>Dados socioeconômicos e demográficos</i>	29
3.2.2.3 <i>Tratamento dos dados</i>	30
<b>3.2.3 Modelo de regressão logística</b>	<b>31</b>
3.2.3.1 <i>Preparação do modelo de regressão logística</i>	31
3.2.3.2 <i>Análise dos resultados</i>	31
<b>4 APLICAÇÃO DO MÉTODO</b>	<b>33</b>
4.1 A ORGANIZAÇÃO	33
4.2 IDENTIFICAÇÃO E FORMULAÇÃO DE VARIÁVEIS DE SUCESSO	33
<b>4.2.1 Entendimento do modelo de negócio e definição das variáveis</b>	<b>34</b>
4.2.1.1 <i>Fator Desempenho Operacional</i>	34
4.2.1.2 <i>Fator Imóvel e Estrutura</i>	35

4.2.1.3 Fator Gestor	38
4.2.1.4 Fator Sociodemografia	39
4.2.1.5 Priorização	40
<b>4.2.2 Coleta de dados e informações</b>	<b>44</b>
4.2.2.1 Sucesso das franquias	44
4.2.2.2 Qualidade dos imóveis e estrutura das franquias	45
4.2.2.3 Qualidade da gestão das franquias	46
4.2.2.4 Dados sociodemográficos	47
4.3 CONSTRUÇÃO DO MODELO	51
<b>5 ANÁLISE E DISCUSSÃO DOS RESULTADOS</b>	<b>54</b>
<b>6 CONCLUSÃO E CONSIDERAÇÕES FINAIS</b>	<b>61</b>
6.1 SUGESTÕES PARA TRABALHOS FUTUROS	62
<b>REFERÊNCIAS</b>	<b>63</b>
<b>GLOSSÁRIO</b>	<b>66</b>
<b>APÊNDICE A – VALORES CONSOLIDADOS DAS VARIÁVEIS</b>	<b>68</b>
Apêndice A1 - Valores para modelo R1000m	68
Apêndice A2 - Valores para modelo R2000m	72
Apêndice A3 - Valores para modelo misto R1000m e R2000m	76
<b>APÊNDICE B – NOTAS CONSOLIDADAS DE IMÓVEIS</b>	<b>80</b>

## 1 INTRODUÇÃO

Em março de 2020, a Organização Mundial da Saúde (OMS) declarou a consolidação de uma pandemia, utilizando o isolamento social como primeira estratégia para evitar a contaminação por COVID-19 (CUCINOTTA, VANELLI 2020). Dessa maneira, surge a imediata necessidade da readaptação de hábitos da população, principalmente pela restrição de circulação e pelo isolamento compulsório.

O cenário pandêmico resultou em redução da prática de atividade física, logo, constata-se que o setor fitness (negócios voltados à saúde e bem-estar de modo geral) foi um dos setores mais impactados financeiramente em função da proibição de aglomeração de pessoas e da dependência do espaço físico para exercer atividades. No entanto, este mercado performou uma rápida recuperação, pois 88% das assinaturas dos desistentes foram reativadas logo que as academias reabriram, em agosto de 2020 (JÚNIOR E PAIANO, 2020; FIA, 2021; BERNUZZI, 2022).

Em números, conforme os resultados apresentados no relatório de 2017 da *International Health Racquet and Sportsclub Association* (IHRSA, 2017), o Brasil ficou em segundo lugar na quantidade de academias, cerca de 34,5 mil estabelecimentos, apenas atrás dos Estados Unidos. No ranking de clientes, o país ocupava a quarta posição, com 9,6 milhões de pessoas, e a décima colocação em faturamento, tendo movimentado aproximadamente 2 bilhões de dólares naquele ano.

Dada tamanha dimensão do setor fitness e a constante expansão mercadológica, a oferta de produtos e serviços consolidados, a inovação e a existência de concorrentes relevantes, buscam-se recursos que conduzam uma organização a uma condição diferenciada em relação aos competidores. Os fatores que constituem estas vantagens competitivas são temas centrais nas discussões sobre estratégias organizacionais (PORTER, 1991).

Em um cenário competitivo, os destaques podem ser traduzidos em fatores de sucesso que se consolidam ao longo de todo o caminho de uma operação de negócios. No setor fitness, as particularidades em relação à demanda e ao comportamento de consumo de serviços e produtos podem ser um desafio. A localização geográfica das unidades é um grande recurso competitivo, tendo um impacto considerável na projeção comercial da unidade.

Decidir a localização de um empreendimento é considerada uma decisão estratégica que requer embasamento em estudos criteriosos e bem fundamentados. O uso dos recursos adequados que auxiliem a tomada de decisão mostra-se crucial para tal definição (CARNASCIALI, DELAZARI, 2011).

As ferramentas de informação geográfica podem auxiliar na escolha do lugar mais adequado para uma nova franquia, no mapeamento dos locais existentes de negócios similares, no entender da demografia da região e para identificar áreas potenciais de crescimento. Além disso, também podem ser utilizadas para auxiliar na gestão de um negócio em funcionamento, pois contribuem em informações atualizadas sobre o perfil de consumo e presença de novos concorrentes e polos de atração de clientes.

Este trabalho busca identificar os principais fatores de sucesso na operação de uma rede de franquias no setor fitness, assim como analisar a importância dessas variáveis no sucesso do empreendimento por modelos estatísticos de regressão. Dessa forma é possível promover, no processo de abertura de uma franquia, uma ponderação dos fatores de sucesso e seus respectivos pesos para obter uma probabilidade de sucesso de uma futura operação por meio de uma regressão logística.

## 1.1 OBJETIVOS

Nas seções abaixo estão descritos o objetivo geral e os objetivos específicos deste trabalho.

### 1.1.1 Objetivo Geral

Definir os principais fatores e sua influência no sucesso da operação de franquias do setor fitness, de maneira que se auxilie na tomada de decisão para aprovação de novas unidades.

### 1.1.2 Objetivos Específicos

O alcance do objetivo geral descrito acima é sustentado pelo sucesso dos objetivos específicos descritos a seguir:

- a) Identificar os fatores e subfatores que compõem o sucesso operacional da franquia;
- b) Analisar a importância dos fatores no sucesso da operação;
- c) Construir um modelo preditivo do sucesso operacional baseado nos fatores levantados;

## 1.2 DELIMITAÇÕES E LIMITAÇÕES DO TRABALHO

As delimitações do trabalho decorrem principalmente do caráter prático da aplicação em uma rede de franquias do setor fitness. As análises são delimitadas por fatores alinhados com a realidade das franquias em questão e do setor que estas se encontram.

As limitações do trabalho se concentram principalmente na coleta e qualidade dos dados que fundamentam a construção do modelo e dos indicadores. As fontes dos dados podem apresentar erros e incertezas justamente por se tratar de plataformas intermediárias de censos, outros levantamentos paralelos e estudos datados. Outro fator que contribui para a não assertividade dos resultados é o preenchimento com informações incorretas por parte dos franqueados, depois coletadas pelo sistema de *business intelligence* da franqueadora e usadas para metrificar desempenhos. As limitações também se estendem para a aferição de alguns dados, que podem ter sua coleta limitada por fatores de custo ou dificuldades processuais.

## 1.3 ESTRUTURA DO TRABALHO

O presente trabalho está dividido em 6 capítulos: (i) Introdução, (ii) Fundamentação Teórica, (iii) Metodologia, (iv) Aplicação do Método, (v) Análise e Discussão dos Resultados e (vi) Conclusão e Considerações Finais.

**I. Introdução:** O capítulo contextualiza o tema abordado e a importância do trabalho, apresenta os objetivos e destaca as limitações e delimitações da pesquisa.

**II. Referencial Teórico:** Explicam-se, de forma suficiente, os principais conceitos que fornecem um pleno entendimento do trabalho e outras revisões bibliográficas pertinentes.

**III. Metodologia:** Detalham-se, de forma generalizada, as etapas metodológicas que guiaram o desenvolvimento do estudo, e fornece roteiro metodológico para eventuais replicações ou adaptações do trabalho.

**IV. Aplicação do Método:** São desenvolvidos os detalhes da aplicação das etapas metodológicas na empresa em questão, por meio da definição das categorias e das variáveis de sucesso, priorização e coleta dos respectivos dados. Também é descrita a preparação do modelo estatístico de regressão logística em seus parâmetros e configurações utilizadas.

**V. Análise e Discussão dos Resultados:** São descritas as principais observações referentes aos resultados do modelo de regressão logística, seus testes de aderência e qualidade e conclusão sobre as variáveis preditivas. A seção também visa trazer uma discussão destes resultados quando comparados ao contexto atual da empresa em questão.

**VI. Conclusão e Considerações Finais:** São feitas considerações sobre os objetivos alcançados, o processo de desenvolvimento do trabalho e resultados obtidos. Conclui-se com sugestões para trabalhos futuros.

## 2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta uma revisão da literatura sobre os principais temas abordados neste trabalho. A revisão dos temas, que fornece a base teórica para o capítulo 3 - Metodologia, está estruturada da seguinte forma: (i) *Business Franchise Model*, (ii) Segmentação de Centros de Atividades Físicas e (iii) Regressão Logística.

### 2.1 BUSINESS FRANCHISE MODEL (MODELO DE FRANQUIA DE NEGÓCIOS)

O mercado fitness funciona principalmente em três modelos de negócio: novas empresas, licenciamento e franquia. Conforme a *International Franchise Association* (Associação Internacional de Franquias) (IFA)<sup>1</sup>, o modelo de franquia, ou *franchising*, é um método de distribuição de produtos e serviços. Ele envolve duas partes: a concessão de uma licença por um franqueador (que possui a marca e o modelo de negócio) e o franqueado, que, sob o pagamento de royalties e taxas, obtém o direito de reproduzir o nome e modelo de negócio conforme os termos acordados.

Em um contrato de franquia, o franqueado e franqueador se articulam sob a motivação de ganhos bilaterais no uso compartilhado da marca franqueada no que se concerne, por exemplo, na ampliação de uma tecnologia, de uma linha de produtos ou de processos administrativos, ou ainda no desenvolvimento do volume de compras e de distribuição. O *franchising* também deve beneficiar o aporte de capital à rede, possibilitando uma expansão

---

<sup>1</sup> Disponível em: <https://www.franchise.org/> e <https://www.franchise.org/about-us>

rápida com custos comparativamente inferiores a uma expansão verticalizada (estruturada completamente em lojas próprias) (SILVA, 2004).

Segundo a Lei Nº 13.966, de 26 de dezembro de 2019, que dispõe sobre o sistema de franquia empresarial e revoga a Lei nº 8.955, de 15 de dezembro de 1994 (Lei de Franquia), entende-se que:

(...) um franqueador autoriza por meio de contrato um franqueado a usar marcas e outros objetos de propriedade intelectual, sempre associados ao direito de produção ou distribuição exclusiva ou não exclusiva de produtos ou serviços e também ao direito de uso de métodos e sistemas de implantação e administração de negócio ou sistema operacional desenvolvido ou detido pelo franqueador, mediante remuneração direta ou indireta, sem caracterizar relação de consumo ou vínculo empregatício em relação ao franqueado, ou a seus empregados, ainda que durante o período de treinamento.

Além disso, a lei da Franquia estabelece que seja oferecido pelo franqueador ao franqueado: a) suporte; b) supervisão de rede; c) serviços; d) incorporação de inovações tecnológicas às franquias; e) treinamento do franqueado e de seus funcionários, especificando duração, conteúdo e custos; f) manuais de franquia; g) auxílio na análise e na escolha do ponto onde será instalada a franquia; e, h) layout e padrões arquitetônicos das instalações do franqueado, incluindo arranjo físico de equipamentos e instrumentos, memorial descritivo, composição e croqui.

Entre as principais vantagens do modelo de franquias, além de ser um conceito previamente testado e aprovado, pode-se destacar o notável potencial para escalabilidade acelerada, a intensa cooperação entre as partes envolvidas e o rateamento de custos. Comparado com outras estratégias de negócios, o modelo de franquia é uma estratégia mais eficaz e eficiente para entrar em novos mercados e expandir as operações (GUETTA, 2013; LEE *et al.*, 2016).

No Brasil, a Pesquisa de Desempenho do Setor de Franchising, realizada pela Associação Brasileira de Franchising (ABF), foi realizada entre os dias 4 de julho e 2 de agosto de 2022. O crescimento do faturamento, como indicado na Figura 1, especialmente no segundo trimestre de 2021, tem relação com o maior fluxo em lojas físicas e com a retomada de segmentos com forte impacto da pandemia.

Figura 1 - Evolução do faturamento de franquias no Brasil.



Fonte: ABF (2022)<sup>2</sup>.

Os resultados da pesquisa de desempenho mostram que a variação no 2º trimestre de 2022 representou um acréscimo de 4 mil operações de franquia no país, totalizando 178 mil operações.

## 2.2 SEGMENTAÇÃO DE CENTROS DE ATIVIDADE FÍSICA

Os centros de atividade física (ou *Fitness*) integram o segmento de serviços, estabelecendo conexões com três outros setores: estética, entretenimento e saúde (SCALDAFERRI, 2007). Segundo os Descritores em Ciência da Saúde (DeCS, 2022), centros de atividade física são “instalações com programas que pretendem promover e manter um estado de bem-estar físico para ótimo desempenho e saúde”. Toscano (2001) define:

As academias de ginástica são os centros de atividades físicas onde se presta um serviço de avaliação, prescrição e orientação de exercícios físicos, sob supervisão direta de profissionais de educação física.

Blocken *et al.* (2020) descrevem os centros esportivos *indoor* como espaços que dispõem de equipamentos e serviços designados à prática de exercícios físicos, podendo abranger variadas atividades: equipamentos para exercícios aeróbicos/cardio, equipamentos de treino isodinâmicos (máquinas) e pesos livres, aulas coletivas como ciclismo/*spinning*, yoga, alongamento, pilates, e os espaços alternativos como as pistas de corrida cobertas, aulas de natação, quadras de basquete, artes marciais etc. Logo, nesta classificação, estão inclusos os estúdios de pilates e academias em geral. Blocken *et al.* (2020) acrescentam que esses

<sup>2</sup> Disponível em: [https://www.abf.com.br/wp-content/uploads/2022/10/Relatorio\\_Desempenho\\_Franchising\\_2TRI\\_2022\\_diagramada.pdf](https://www.abf.com.br/wp-content/uploads/2022/10/Relatorio_Desempenho_Franchising_2TRI_2022_diagramada.pdf)

locais têm potencial para impulsionar a saúde primária, considerando que o bem-estar e a qualidade de vida está diretamente relacionada à aptidão física individual.

Em razão de uma nova perspectiva conscienciosa da sociedade em vista da atividade física, em que os adeptos, em qualquer idade e forma, buscam entender melhor o próprio corpo, os malefícios que estão suscetíveis, e os meios para desenvolver uma vida mais saudável (TOSCANO, 2001). Posto isso, atualmente o mercado *fitness* é caracterizado pela diversidade de alternativas e pela alta demanda por equipamentos modernos e tecnológicos, pelo excelente atendimento aos clientes, e pela especificidade de nichos, como o público da terceira idade, atletas de alto rendimento, reabilitação física, entre outros.

Em um contexto mundial, o mercado *fitness* é dominado pelos EUA, alcançando um faturamento de US\$ 30 bilhões por ano até 2019. Os 10 maiores mercados no setor somam 71% da receita total da área *fitness*, que corresponde a US\$ 64,9 bilhões, segundo o levantamento da *International Health, Racquet & Sportsclub Association* (IHRSA). Quanto ao faturamento no Brasil, estima-se o valor de US\$ 2,1 bilhões em 2019, alcançando a terceira posição das Américas, atrás somente dos EUA e do Canadá (IHRSA, 2021).

Apesar de o Brasil ser a segunda nação com mais academias do mundo (35 mil unidades oficiais), – a primeira é novamente os Estados Unidos, com 40 mil unidades, quando o mercado é avaliado sob a perspectiva do número de clientes, o Brasil aparece em quarto lugar, com 9,6 milhões de clientes contabilizados, depois de EUA (62 milhões), Alemanha (11 milhões) e Reino Unido (9,9 milhões) (FIA, 2021).

A perspectiva para 2023 é o crescimento do setor *fitness* ao nível mundial, contudo, o setor deverá se basear nas experiências de 2020 (restrições devido ao coronavírus), a exemplo da digitalização, atenção às necessidades do consumidor e planejamento para resistir às crises.

### 2.3 REGRESSÃO LOGÍSTICA (RL)

Uma análise de regressão consiste em um método estatístico para investigar a relação entre variáveis. A definição do modelo a ser utilizado depende da quantidade de variáveis em análise e as suposições que possam ser feitas sobre a distribuição dos dados. Dentre os modelos de regressão, tem-se a Regressão Logística (RL). Field (2009) esclarece:

(...) a regressão logística é uma regressão múltipla, mas com uma variável de saída categórica dicotômica e variáveis previsoras contínuas ou categóricas. Simplificando, isso quer dizer que podemos prever a qual de duas categoriais é provável que uma pessoa pertença dado certas informações.

Em análises preditivas, a variável dependente é binária (0 e 1), categórica ordenada (há uma hierarquia determinada entre as variáveis resposta) ou mesmo categórica desordenada (não possui nenhuma relação de ordem entre elas) (CHATTERJEE e HADI, 2006; SMOLSKI, 2017; NHU *et al.*, 2020).

Desenvolver modelos precisos e robustos a partir de muitos dados é um grande desafio devido à sua multidimensionalidade, ou seja, os dados são de natureza diversa e vêm de fontes variadas, como pesquisas de campo, fotos aéreas, imagens de satélite, registros históricos, entre outros (BUI *et al.*, 2016).

A vantagem da RL é haver uma facilidade para lidar com variáveis independentes categóricas; fornece o resultado em termos de probabilidade; facilita a classificação de indivíduos em categorias; requer pequeno número de suposições; e possui alto grau de confiabilidade (BAI *et al.*, 2010; BUI *et al.*, 2016; SHIRZADI *et al.*, 2017; PHAM *et al.*, 2019).

A RL determina a relação da ocorrência de um evento e possíveis fatores causadores, logo a probabilidade de ocorrer tal evento é denotado pela Equação 1, no qual  $\log(p/p1-p)$  é previsto a partir de uma combinação de cada variável previsora multiplicada pelo seu respectivo coeficiente de regressão, e examinar os resíduos na regressão logística é relevante para isolar pontos em que o modelo tem pouca aderência e isolar pontos que exercem uma influência indevida no modelo. (FIELD, 2019; NHU *et al.*, 2020):

$$\log \frac{p}{p1-p} = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n + \epsilon \quad (1)$$

Onde:

$\log \frac{p}{p1-p}$  = probabilidade de um evento ocorrer em relação ao evento não ocorrer;

$b_0, b_1, b_2, \dots, b_n$  = coeficientes expressos em logaritmos;

$x_1, x_2, \dots, x_n$  = variáveis independentes;

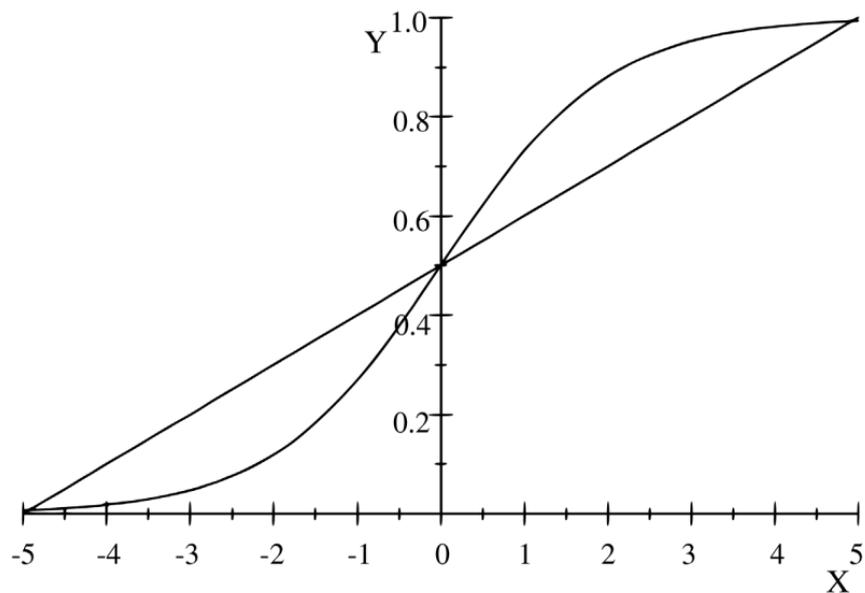
$\epsilon$  = erro ou resíduo da regressão.

Pode-se observar na Figura 2 que a variável dependente (Y) no modelo de RL assume apenas dois valores (0 ou 1), a probabilidade predita pelo modelo também deve se

limitar ao referido intervalo. Quando  $X$  (variável independente) assume valores mais baixos, a probabilidade se aproxima de zero. No outro oposto, enquanto  $X$  aumenta, a probabilidade se aproxima de um. A curva apresenta um aumento acelerado, se assemelhando de uma relação linear e volta a se estabilizar lentamente conforme se aproxima de 1, nunca ultrapassando esse valor.

Além disso, na Figura 2 é possível visualizar que a função logística usada para descrever a relação entre a probabilidade e as variáveis independentes, é representada pela curva *sigmoide*, que se assemelha à forma de um "S".

Figura 2 – Reta de regressão linear vs. curva logística.



Fonte: Fernandes (2021).

Considerando que a natureza binária da variável dependente viola alguns pressupostos do modelo linear, a utilização do modelo RL para analisar variáveis binárias pode gerar coeficientes mais eficientes (menor variância), e menos viesados (tende a não superestimar ou subestimar o valor do parâmetro populacional).

Quanto a interpretação dos coeficientes, observa-se que o impacto de cada coeficiente é sobre a própria razão de chances, e não mais sobre a quantidade de *logit*; identifica-se o impacto do coeficiente da variável independente da razão de chances;

determina-se o efeito que os coeficientes exercem sobre a chance de um evento ocorrer; enfatiza-se que um coeficiente positivo aumenta a probabilidade e um negativo diminui a probabilidade.

Para a aplicação do modelo de RL, os autores Fernandes *et al.* (2020) descrevem um planejamento em cinco estágios:

- 1) Identificar que variável dependente seja originalmente dicotômica;
- 2) Observar os requisitos técnicos: a RL é mais flexível do que outras técnicas estatísticas, mas é sensível, por exemplo, os problemas de altos níveis de correlação entre as variáveis independentes (multicolinearidade), ou quando as amostras muito pequenas tendem a produzir estimativas inconsistentes e amostras excessivamente grandes aumentam o poder dos testes estatísticos de tal maneira que qualquer efeito tende a ser estatisticamente significativo, independentemente da magnitude;
- 3) Estimar e ajustar o modelo. Para isso, deve-se a) reportar o software e b) compartilhar os materiais de replicação, que incluem os dados originais, os dados tratados e os scripts computacionais; o próximo passo é avaliar a qualidade do ajuste.
- 4) Interpretar os resultados, por exemplo: entender os coeficientes e discutir substantivamente como os resultados se relacionam com a hipótese de pesquisa; analisar o impacto das variáveis independentes sobre a chance (odds); estimar o aumento percentual na chance de ocorrência de Y.
- 5) Validar os resultados observados com uma subamostra de sua base de dados original. Isso confere maior confiabilidade aos resultados de pesquisa, principalmente quando se trabalha com amostras pequenas.

### **2.3.1 Testes de ajustes e qualidade**

#### **Multicolinearidade**

A multicolinearidade é um problema dos dados. Altos níveis de correlação entre variáveis independentes produzem efeitos adversos sobre a consistência dos coeficientes. O principal efeito de altos níveis de correlação entre as variáveis explicativas é prejudicar a capacidade preditiva do modelo de regressão (LONG, 1997).

Para identificar a multicolinearidade, a maneira mais simples é estimar uma matriz de correlação entre as variáveis independentes. Quanto maior a magnitude dos coeficientes, maiores os eventuais problemas. Em síntese, a literatura indica duas medidas para diagnosticar problemas de colinearidade: (1) Tolerância e (2) Fator de Inflação da Variância (*Variance Inflation Factor* - VIF).

Figueiredo Filho *et al.* (2015) apresentam quatro procedimentos para superar os problemas multicolinearidade: (1) verificar a codificação e a transformação das variáveis; (2) aumentar o tamanho da amostra; (3) utilizar alguma técnica de redução de dados (análise fatorial ou análise de componentes principais); e (4) consultar a literatura específica sobre o tema.

### **R<sup>2</sup> de Nagelkerke**

Em uma regressão linear, o coeficiente de correlação múltiplo R e o correspondente R<sup>2</sup> são medidas úteis para avaliar como o modelo se ajusta aos dados. Sendo assim, é possível uma versão de coeficiente de correlação para a RL, conhecido como pseudo R<sup>2</sup> ou estatística-R. Esse valor representa a correlação parcial entre a variável de saída e cada uma das variáveis preditoras, variando de -1 a 1. O cálculo de R<sup>2</sup> é determinado pela equação 2 (SMOLSKI, 2017).

$$R^2 = \frac{-2LL_{nulo} - 2LL_{modelo}}{-2LL_{nulo}} \quad (2)$$

Onde:

$-2LL_{nulo}$  = -2 vezes o logaritmo do valor da verossimilhança

O valor R<sup>2</sup>, quando positivo, indica que a variável previsoras aumenta, também aumenta a probabilidade da ocorrência do evento. Se o valor é negativo, indica que se a variável previsoras aumenta, a probabilidade do valor de saída ocorrer diminui. Se uma variável tem um valor pequeno de R, ela contribui pouco para o modelo.

**Hosmer e Lemeshow (2000)**

Esse teste é utilizado para avaliar a qualidade do ajuste do modelo, ou seja, se o modelo proposto pode explicar bem o que se observa. Neste teste, os dados são agrupados conforme o seu valor, do maior para o menor, e divididos em grupos de igual tamanho. A seguir, para certificar-se de que o modelo não difere significativamente dos dados observados, para cada grupo são estimados os eventos e não eventos observados. Isso indica se o modelo escolhido está ou não ajustado aos dados (FIELD, 2009).

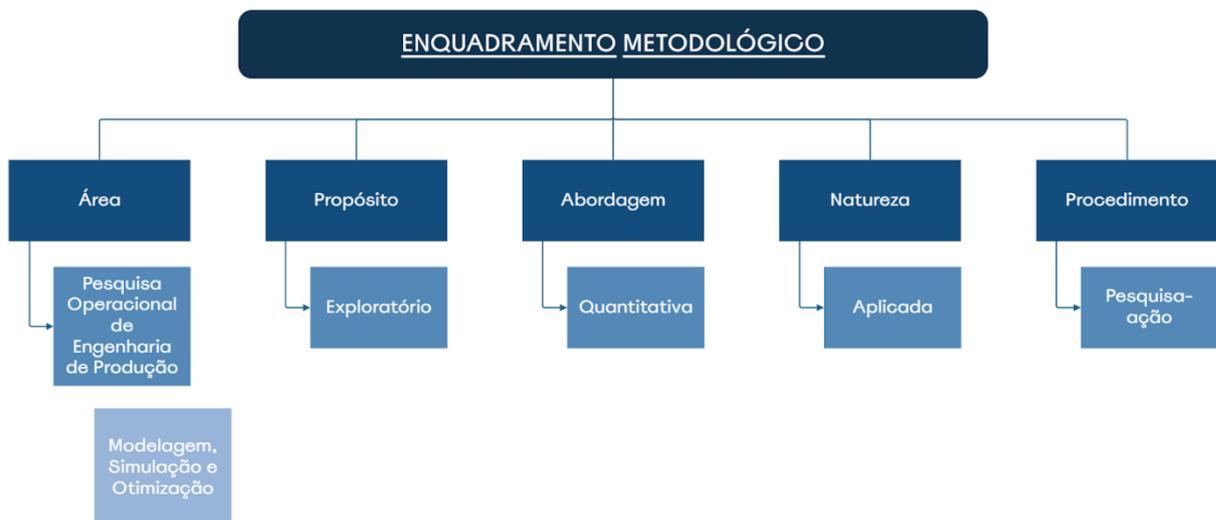
O valor de eventos é igual à soma das probabilidades estimadas de todos os indivíduos do grupo, como a estimativa do número de não eventos equivale ao tamanho do grupo menos o número de eventos estimados. O modelo é rejeitado para valores abaixo do nível de significância (FIELD, 2009).

### 3 METODOLOGIA

#### 3.1 CARACTERIZAÇÃO DE PESQUISA

Este trabalho se enquadra na área de Pesquisa Operacional de Engenharia de Produção e na subárea de Modelagem, Simulação e Otimização. A pesquisa se enquadra como exploratória, de abordagem quantitativa e de natureza aplicada. O procedimento técnico se caracteriza como uma pesquisa-ação. O resumo do enquadramento metodológico está representado pela Figura 3.

Figura 3 – Enquadramento metodológico do trabalho.



Fonte: elaboração própria (2022).

#### 3.2 ETAPAS METODOLÓGICAS

As etapas metodológicas que guiam este trabalho podem ser separadas em dois pilares que sustentam os objetivos apresentados no Capítulo 1. Suas etapas e respectivas sub etapas são elencadas a seguir:

- I. Identificação de variáveis:
  - a. Definição de fatores de sucessos da operação;
  - b. Coleta e tratamento de dados;

- II. Construção do modelo de regressão logística:
  - a. Preparação do modelo de regressão logística
  - b. Análise dos resultados

Essa divisão de etapas e sub etapas é apresentada na Figura 4.

Figura 4 – Etapas e sub etapas metodológicas do trabalho.



Fonte: Elaboração própria (2022).

### **3.2.1 Identificação de variáveis**

Essa etapa tem o objetivo de compreender e definir os fatores candidatos que podem caracterizar o sucesso de uma operação. Também, busca-se quantificar e tratar os dados das métricas definidas.

#### *3.2.1.1 Definição de fatores de sucesso de operação*

Os fatores de sucesso que configuram uma operação com excelência de resultados podem ser estabelecidos pelo uso dos seguintes recursos: (i) Entendimento do modelo de negócio; (ii) Análise de indicadores chave de performance; (iii) Entrevistas; (iv) Priorização.

##### 3.2.1.1.1 Entendimento do modelo de negócio

Para compreender as principais variáveis que constituem o sucesso de uma franquia, é necessário mapear os fatores e eventuais subfatores que contribuem para tal resultado. Inicialmente, apresenta-se de forma abrangente os principais pontos idealizados para um entendimento do modelo de negócio. Deve-se abranger todas as categorias que têm potencial de impactar os resultados e que envolvem o funcionamento de uma nova unidade, desde sua implantação até a operação do negócio.

A organização dos fatores e subfatores deve ser de tal forma que fique explícito a relação de causalidade com o sucesso do negócio. Sugere-se que as dependências e interdependências das variáveis sejam evidentes e representadas de forma visual para melhor entendimento e organização do problema.

##### 3.2.1.1.2 Análise de indicadores chave de performance

A definição dos fatores de sucesso pode ser fundamentada por uma avaliação dos principais indicadores de performance de uma operação, como taxas de conversão, retenção e satisfação de clientes. Estes indicadores já são normalmente utilizados para representar métricas de desempenho de unidades e servem como suporte para gestão comercial e estratégica das franquias. Por estarem diretamente ligados aos resultados operacionais, tais fatores terão grande contribuição para o mapeamento dos fatores de sucesso.

É possível que alguns indicadores de performance sejam resultantes de uma composição de um ou mais indicadores, portanto, mostra-se necessário realizar uma investigação e mapeamento da ramificação destes elementos. Por exemplo, uma métrica de desempenho comum em serviços é a retenção de clientes. Esta, por vez, é frequentemente resultante ou acompanhada de indicadores de renovação e cancelamento do serviço.

#### 3.2.1.1.3 Entrevistas

Para essa etapa, é utilizada a própria experiência e conhecimento dos colaboradores e de seus clientes para o levantamento dos pontos-chave para o sucesso do negócio. As entrevistas têm como objetivo recolher informações qualitativas e quantitativas das principais fontes percebidas por quem de fato vive o dia a dia do empreendimento - às vezes distante ou complementar ao que é planejado idealmente como modelo de negócio.

É recomendado envolver pessoas de diferentes setores da empresa, assim como gestores e clientes de diversificadas conjunturas de resultados de operação. Dessa forma, evita-se que depoimentos viesados por fatores dependentes de algum setor da empresa ou por aspectos de operações similares afetem a amostragem das declarações.

Recomenda-se que a estrutura das entrevistas inicie pelo objetivo da conversa e que haja um alinhamento do que e de como deve ser o resultado esperado. As perguntas devem ser elaboradas de tal forma que se investigue os fatores de sucesso mais pertinentes ao problema e a coleta de dados, mas também que se dê liberdade para levantar outros fatores que possam ser significativos.

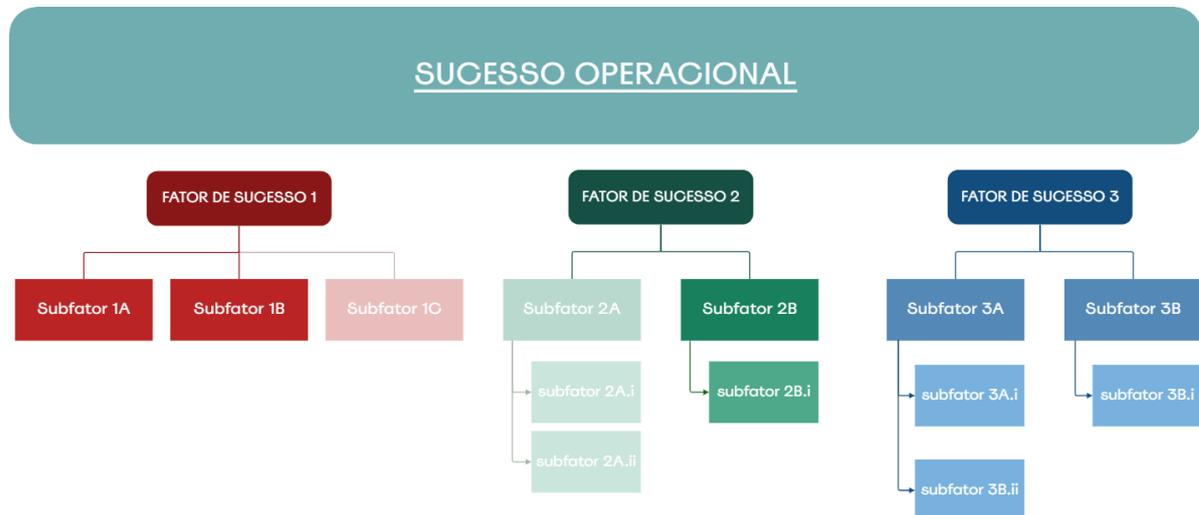
#### 3.2.1.1.4 Priorização

O desenvolvimento das etapas anteriores pode resultar em um mapeamento complexo, composto por diversos fatores e subfatores que reflitam o sucesso de uma operação. Entretanto, para a elaboração de um modelo estatístico baseado em tais variáveis, é necessário considerar a capacidade de mensuração destes fatores. Dessa forma, uma priorização é recomendável, uma vez que há possibilidade de alguns aspectos serem de difícil aferição.

Os fatores priorizados devem possuir métodos claros e fundamentados de avaliação, sejam eles qualitativos ou quantitativos, e não limitados por notas ou valores próprios, podendo haver agregações e cálculos entre fatores. Ainda, o conjunto de fatores priorizados

não pode deixar de caracterizar uma operação de sucesso. A Figura 5 representa um conjunto de fatores e subfatores mapeados, destacando-se aqueles sendo priorizados.

Figura 5 – Exemplo de fatores e subfatores mapeados e priorizados.



Fonte: Elaboração própria (2022).

### 3.2.2 Coleta e tratamento de dados

Nessa etapa, concentram-se os processos de aquisição e tratamento dos dados que constituirão o modelo a ser desenvolvido.

#### 3.2.2.1 Dados de inteligência de negócio

O levantamento de dados referentes ao negócio tem por finalidade o agrupamento e organização de informações relevantes para a caracterização dos empreendimentos que compõem a base do trabalho. Essas informações devem ser suficientes para respaldar diretamente os fatores de sucesso levantados nas etapas anteriores, ou ao menos fornecer elementos para eventuais procedimentos que fundamentam a construção de outras métricas de interesse. Por exemplo, o fator de sucesso “Conversão de Clientes” de um serviço pode ser diretamente representado pela própria taxa de conversão em um funil de vendas realizada por uma unidade. Em contrapartida, os fatores sociodemográficos podem ser adquiridos ou

formulados a partir de informações de Logradouro, Número, CEP, Cidade e Estado das respectivas unidades.

Dada a possibilidade de se obter fatores qualitativos decorrentes dos recursos de definição dos fatores de sucesso, a quantificação destas métricas torna-se essencial para incluí-las nos modelos estatísticos que decorrem no trabalho. O processo de quantificação pode ser auxiliado por uma definição de unidades de referência na escala de valores, de forma que se tenham exemplos que facilitem a classificação do restante dos elementos. Por exemplo, o fator qualitativo da qualidade da fachada de um imóvel pode ter sua quantificação auxiliada caso haja imóveis que representem uma nota 10 ou uma nota 5. Dessa forma, a comparação da unidade avaliada com as unidades de referência é útil para facilitar a aproximação da nota escolhida.

### 3.2.2.2 *Dados socioeconômicos e demográficos*

Retomando o conhecido desafio de avaliação da localização para abertura de novos negócios e a premissa de que fatores de região podem influenciar no desempenho de uma operação, faz-se necessário coletar e aferir estes indicadores. O cumprimento da etapa anterior deve fornecer informações suficientes para a completa caracterização da geolocalização das unidades que serão estudadas.

Com as geolocalizações conhecidas, os dados sociodemográficos e econômicos das áreas podem ser adquiridos gratuitamente por pesquisas e levantamentos feitos pelo IBGE - Instituto Brasileiro de Geografia e Estatística, assim como outros órgãos ou instituições de pesquisa. Outra alternativa são empresas especializadas em inteligência geográfica que fornecem plataformas e serviços que podem se enquadrar com as necessidades de diversos estudos, muitas delas baseada em dados do próprio IBGE e complementadas com levantamentos das próprias empresas.

A obtenção dos dados mencionados, sobretudo, sucede uma fundamentada definição das dimensões de áreas de estudo e áreas de influência dos empreendimentos. As áreas de estudo refletem uma área delimitada de onde as informações demográficas serão retiradas. Já as áreas de influência estão ligadas diretamente com a amplitude do impacto que o negócio pode gerar na região. É importante basear esses dados com informações adquiridas na etapa anterior de *dados de inteligência de negócios*, e também no mapeamento gerado no *entendimento do modelo de negócio*. Ambas as áreas estão fortemente relacionadas, justamente pela área de estudo ser apoiada pela área de influência.

A definição da dimensão destas áreas poderá seguir algumas possibilidades conforme a necessidade de cada estudo, da quantidade e qualidade das informações adquiridas dos empreendimentos e das funcionalidades das plataformas de inteligência geográfica escolhida. O comum será escolher entre opções que variam em sua abrangência territorial, partindo de dados sociodemográficos nacionais, estaduais, municipais, de bairros até setores censitários. Ainda, em algumas plataformas é possível gerar áreas por deslocamento a partir do endereço das unidades - criadas a partir da distância ou tempo percorridos e até considerar seus meios de locomoção: a pé ou de carro.

### 3.2.2.3 *Tratamento dos dados*

A conclusão das etapas anteriores deve resultar em uma base de dados robusta que relacione as diversas unidades de um negócio com seus respectivos fatores de sucesso de forma quantificada. Dado o contexto do presente trabalho, tem-se um objetivo de realizar procedimentos estatísticos a partir destas informações, portanto, é aconselhado tratar e adaptar a organização destes dados para tal finalidade. Como mencionado na etapa *Dados de inteligência de negócio*, a formulação de alguns fatores de sucesso pode requerer cálculos utilizando os dados adquiridos. Dessa forma, parte desta etapa se destina justamente a execução destes procedimentos matemáticos.

Ainda, cabe a ponderação de um processo filtragem de alguns indicadores que não estejam ligados tão diretamente com o respectivo fator de sucesso. Por exemplo, deseja-se avaliar o impacto da variável “Quantidade de Pessoas” de uma área no sucesso da operação. Entretanto, é conhecida a faixa etária alvo do público consumidor do negócio, logo, utiliza-se a “Quantidade de Pessoas” somente dentro da faixa etária de interesse. Isso assegura que os dados estejam mais alinhados com o modelo de negócio estipulado.

Por fim, recomenda-se uma revisão geral da formatação dos dados e da estrutura do banco de dados, a fim de evitar complicações na leitura estatística dos elementos. Alguns aspectos importantes para a aferição são a padronização de escrita dos valores, a parametrização dos fatores e a tabulação dos dados.

### 3.2.3 Modelo de regressão logística

Essa etapa busca, inicialmente, configurar os parâmetros estatísticos necessários para a execução da regressão logística sobre os dados consolidados na etapa anterior. Finaliza-se com a interpretação dos *outputs* gerados pelo modelo de regressão logística e a avaliação e conclusão dos objetivos propostos no início do trabalho.

#### 3.2.3.1 Preparação do modelo de regressão logística

Após a conclusão das etapas anteriores, os dados do problema deverão estar preparados para serem incluídos no modelo de regressão logística. A aplicação do modelo estatístico de regressão poderá ser realizada por diversas opções, os mais comuns *softwares* e pacotes estatísticos como o SPSS (IBM®), ou linguagens de programação como R e Python complementadas das respectivas bibliotecas estatísticas.

A configuração do modelo de regressão logística, independente da plataforma utilizada, inicia-se pela definição da variável dependente - representada pelo interesse do estudo. As variáveis independentes são provenientes do processo realizado para o entendimento e priorização dos fatores dos quais queremos mensurar o impacto na variável dependente.

Outra parametrização importante para o modelo é a definição do nível de significância adotado para o problema, os mais comuns 0,01, 0,05 ou 0,1, conforme o que está sendo avaliado. Esse valor auxiliará na análise dos resultados para a rejeição - ou não - da hipótese nula.

#### 3.2.3.2 Análise dos resultados

A avaliação dos resultados obtidos concentra-se na verificação dos *outputs* retornados pelo modelo de regressão logística no sistema escolhido e na interpretação destes resultados. É importante verificar principalmente as métricas de aderência do modelo, da significância geral e das variáveis inseridas. A caracterização das variáveis predictoras pode ser feita pela significância estatística de contribuição para o modelo, pelo coeficiente do predictor e sua razão de chance. Os principais testes de aderência e qualidade do modelo são os

de Colinearidade,  $R^2$  de Nagelkerke, Hosmer Lemeshow Test e a Estatística de Verossimilhança-Log.

Ressalta-se o importante caráter iterativo deste procedimento de análise com a etapa anterior de preparação do modelo. Frequentemente, será preciso reavaliar as variáveis independentes selecionadas para a composição do modelo a fim de aprimorar seus resultados.

## 4 APLICAÇÃO DO MÉTODO

Este capítulo apresenta a aplicação da metodologia descrita no Capítulo 3 em uma rede de franquias do setor *fitness*. É apresentado o desenvolvimento das etapas realizadas para a aplicação do modelo de regressão logística, assim como uma contextualização da empresa alvo do estudo.

### 4.1 A ORGANIZAÇÃO

A organização na qual o presente trabalho tem seu foco é uma franqueadora de *studios fitness* com matriz em Florianópolis. A empresa tem seu nome associado a uma marca de esportes reconhecida nacionalmente e direciona seus esforços para treinamento e suporte de franqueados, assim como preparação e controle do serviço oferecido. Em novembro de 2022 a franqueadora conta com 26 colaboradores, 56 unidades ativas e mais 23 em implantação.

Os *studios fitness* são caracterizados pela metodologia de treinamento exclusiva e propriamente desenvolvida, que engloba técnicas de Pilates, treinamento funcional, treinamentos cardiorrespiratórios e da metodologia *Five Konzept*. A entrega principal do serviço são treinos customizados para a individualidade dos alunos de uma turma de até 6 pessoas. O modelo de negócio da franquia acaba por ser caracterizado - mesmo que informalmente - como uma academia de boutique.

O público-alvo dos *studios* são alunos da faixa etária de 30 a 60 anos e com perfil socioeconômico alto. Isso é resultado das características do serviço oferecido - que vão ao encontro dos objetivos dos alunos dessa faixa etária - e do ticket médio praticado pelas unidades.

### 4.2 IDENTIFICAÇÃO E FORMULAÇÃO DE VARIÁVEIS DE SUCESSO

A respectiva seção é dedicada ao aprofundamento das etapas iniciais da aplicação do método. As seguintes ações executadas foram voltadas para a identificação das principais variáveis de sucesso, para coleta dos dados necessários e de preparação para o modelo de regressão logística.

#### 4.2.1 Entendimento do modelo de negócio e definição das variáveis

O passo inicial para a formulação do problema foi o entendimento do modelo de negócio. Buscou-se compreender quais são e qual o comportamento das variáveis que influenciam o sucesso de uma unidade de franquia com base no que já é realizado pela franqueadora.

O objetivo da aplicação focou no fator sucesso das operações. A definição de sucesso para esta aplicação reflete o êxito operacional de uma unidade da franquia. Para isso, leva-se em conta o tempo de operação, resultados de métricas projetadas *versus* realizadas do plano de negócio e o potencial da região onde a unidade é localizada. O fator Sucesso é denominado variável dependente do modelo de regressão logística. Como explicado no Capítulo 2, a variável dependente necessita ser qualitativa e dicotômica, além de ser mutuamente exclusiva - o sucesso das operações é representado por 1 e o insucesso por 0.

##### 4.2.1.1 Fator Desempenho Operacional

Com estas definições, foi possível dar início ao processo de identificação das variáveis ligadas com o sucesso de uma unidade de franquia. O bom desempenho de uma operação pode ser bem fundamentado por seus principais indicadores de performance, comuns ao próprio gerenciamento operacional da franquia no estudo em questão. Em conjunto com o setor de Operação da franqueadora, foi definido que o sucesso da franquia seria impactado pelo desempenho operacional, que por vez tem seus respectivos subfatores detalhados na Figura 6.

Figura 6 – Subfatores que compõem o fator “Desempenho Operacional”.



Fonte: Elaboração própria (2022).

Os subfatores de “Captação” e “Conversão” representam as principais taxas que formam o funil de aquisição de alunos matriculados - fonte principal de receita de um *studio*. Já o indicador de “Retenção” é composto pelas taxas de “Cancelamento” e de “Renovação”, que configuram quão bem um *studio* consegue preservar sua carteira de clientes. Por último, a métrica de “Alunos ativos” representa o número de clientes matriculados na unidade de franquia. Essa métrica também auxilia a franqueadora na mensuração da maturidade de uma unidade, pois o número de alunos matriculados reflete, muitas das vezes, o êxito em alcançar os objetivos e projeções estabelecidas no plano de negócio e possíveis problemas comuns relacionados à magnitude desse número.

A seleção destes subfatores foi feita por uma avaliação das métricas mais utilizadas para o gerenciamento das unidades de franquia. O estudo foi facilitado pela plataforma de inteligência de negócios adotada pela empresa, que permitiu a verificação destes indicadores na totalidade da rede de franquias. As métricas levantadas também foram validadas junto ao setor de Operação da empresa e consideradas pertinentes ao problema.

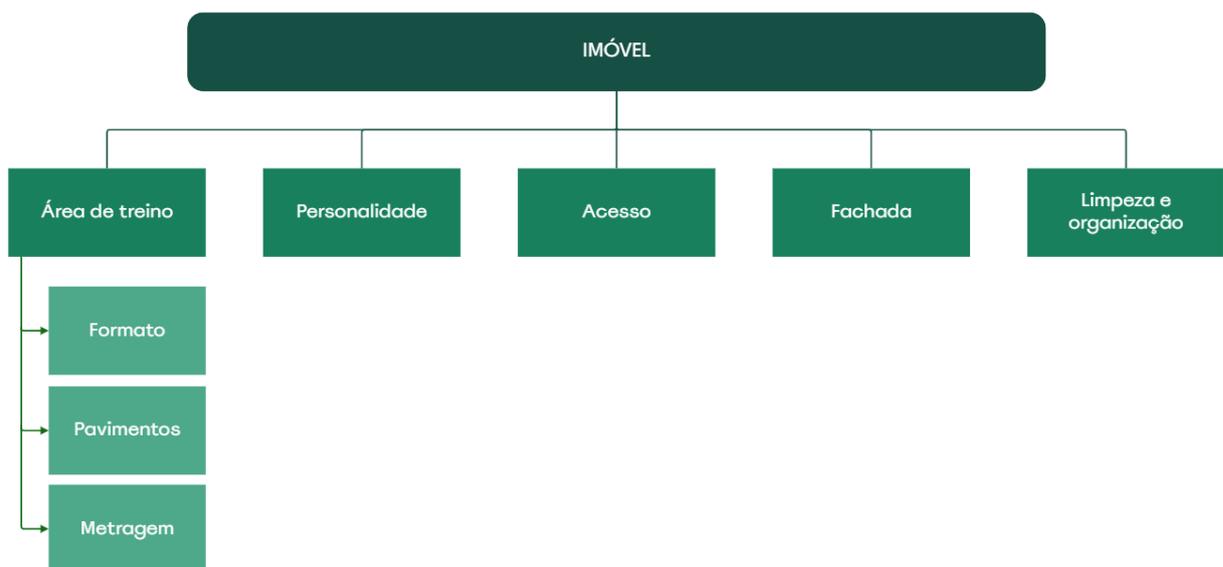
#### 4.2.1.2 Fator Imóvel e Estrutura

No entanto, desejou-se expandir a abrangência do entendimento para além dos fatores de desempenho operacional das unidades. Isso foi motivado pela percepção da

franqueadora, que enxerga que o sucesso geral da franquia pode depender de fatores relacionados à localização, ao imóvel, à gestão praticada pelos franqueados, dentre outros.

Os fatores relacionados ao imóvel e estrutura foram contemplados como variáveis de sucesso por serem um ponto-chave de relevância na experiência do serviço oferecido. Além de possuir requisitos mínimos para a viabilidade da franquia - como área mínima e recomendada, acessibilidade e outros - o imóvel e estrutura apresentam importante papel na diferenciação e destaque da entrega de valor do serviço. Diferenciação esta, que corrobora com o propósito da franqueadora de se distinguir dos serviços *fitness* difundidos no mercado e que pode servir como um amplificador dos resultados operacionais. Observa-se os subfatores detalhados na Figura 7.

Figura 7 – Subfatores que compõem o fator “Imóvel e Estrutura”.



Fonte: Elaboração própria (2022).

O subfator “Área de Treino” corresponde ao espaço do imóvel destinado à entrega do serviço - treinos para turmas de até 6 alunos. Existe uma eficiência associada à distribuição do espaço de treino, influenciada por sua metragem, divisão em pavimentos e formato. A franqueadora exige uma metragem mínima para o setor de treino, mas é recomendado áreas mais amplas possíveis, pois se trata de um diferencial na experiência do aluno.

A variável “Acesso” engloba pontos de acessibilidade para pessoas com necessidades especiais e de facilidade de acesso ao *studio*. Aqui são considerados pontos relacionados ao número de pavimentos e escadas, estacionamentos e vias de acesso.

O item “Fachada” diz respeito à comunicação do *studio* com o ambiente externo. Concentra elementos da marca, destaca o propósito do serviço oferecido e serve de vitrine para novos clientes. É considerado um importante atrativo para captação de possíveis novos alunos.

O fator “Personalidade” é definido como o conjunto de diferenciais estruturais e arquitetônicos que o imóvel possui. São características próprias que tornam o imóvel único e de difícil replicação para o restante da rede de franquia. Normalmente são elementos que tornam o ambiente mais agradável e, de alguma forma, transformam a experiência do cliente. Por exemplo, uma área externa com plantas e um local para descanso e mesa de café, ou um espaço para acomodar alguma arte ou pintura de parede. A Figura 8 exemplifica esses elementos em uma unidade da franquia.

Figura 8 – Elementos de “Personalidade” em um imóvel da franquia.



Fonte: Elaboração própria (2022).

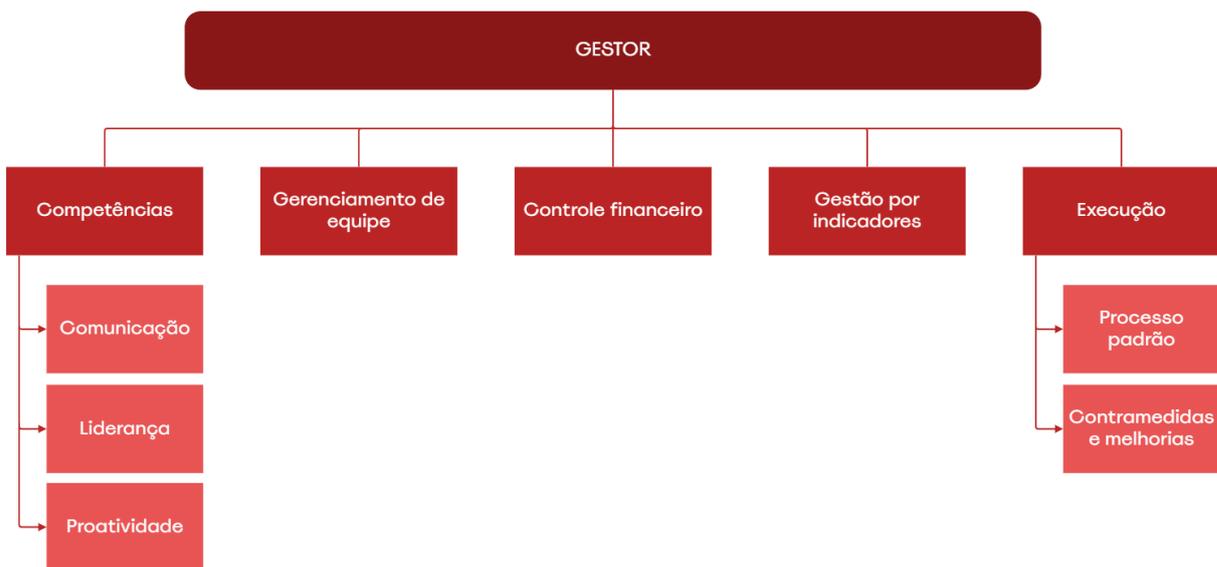
O elemento “Limpeza e Organização” é associado a rotina de manutenção, organização e higiene dos *studios*. Tal ponto também tem grande importância em fornecer uma experiência agradável para o aluno.

A definição destes subfatores foi feita por entrevistas com o setor de Arquitetura da franqueadora - responsável por parte da implantação de novas unidades e da manutenção do padrão estrutural dos *studios* em operação. As entrevistas consistiram em perguntas diretas sobre quais os principais pontos de atenção ao avaliar a qualidade de um imóvel, seja para uma nova unidade ou de uma já em funcionamento, e na análise dos resultados de uma pesquisa de satisfação aplicada anualmente, onde uma seção aborda perguntas sobre a estrutura dos *studios*.

#### 4.2.1.3 Fator Gestor

Um dos pontos que se julga de extrema contribuição para o sucesso de uma operação é o gerenciamento da unidade de franquia. Nas franquias da empresa em questão, o gerenciamento é feito pelo gestor da unidade - o próprio franqueado ou algum colaborador que exerce esta função. O gestor é responsável por zelar pelo controle e aplicação dos processos impostos pela franqueadora, pelo gerenciamento do time, pelo controle financeiro, além de guiar as ações da unidade através do acompanhamento dos indicadores de desempenho. O fator Gestão, portanto, pode ser resumido em um conjunto de comportamento e competências com sua hierarquia detalhada pela Figura 9.

Figura 9 – Subfatores que compõem o fator “Gestor”.



Fonte: Elaboração própria (2022).

O fator “Competências” está relacionado às capacidades do gestor voltadas à comunicação, liderança e proatividade. Tais capacidades são necessárias para a execução de uma gestão eficiente das operações.

O fator “Gerenciamento de Equipe” tem relação com a direção da equipe que compõe o *studio*. As equipes são compostas por pessoas responsáveis pela parte comercial e de atendimento ao cliente, e fundamentalmente por professores que fazem a entrega das aulas.

O “Controle Financeiro” é voltado à gestão monetária das franquias. Possui importante papel no sucesso do negócio, pois engloba a saúde financeira da unidade. Tal subfator exige eficiência do uso de recursos disponíveis em cada franquia.

A “Gestão por indicadores” retrata o gerenciamento diário, guiado pelos principais dados operacionais do negócio. É o acompanhamento das informações seguido da tomada de decisão pelo gestor.

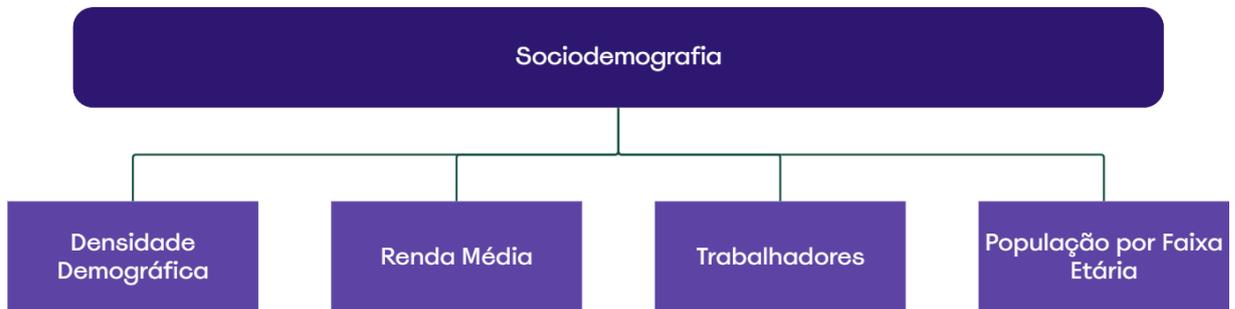
Já o fator “Execução” compreende o cumprimento dos processos padrão determinados pela franqueadora, assim como constante aplicação de contramedidas e melhorias para oportunidades. Tal fator está diretamente ligado com a entrega de valor proposto pelo modelo de negócio.

A apuração dos elementos que caracterizam o fator Gestor foi realizada por meio da seleção de competências já mapeadas pela franqueadora, em formulários de avaliação de candidatos à possíveis franqueados. Outras condutas pertinentes para um gestor foram levantadas por entrevistas com o time do setor de Operação - responsável por acompanhar, guiar e auxiliar as franquias após a inauguração. Tentou-se agrupar métricas que pudessem refletir pontos observados em toda a rede de franquias.

#### *4.2.1.4 Fator Sociodemografia*

Os fatores de Sociodemografia retratam os aspectos sociais, econômicos e demográficos de uma determinada região. O perfil sociodemográfico visa caracterizar uma parcela populacional em sua condição de idade, gênero, residência, nível de escolaridade, ocupação e renda, principalmente. Tais fatores possuem influência no sucesso operacional das franquias por serem um retrato de um conjunto de pessoas com potencial de se tornarem clientes do negócio. Portanto, é de interesse que as operações das franquias estejam localizadas em regiões com um perfil populacional alinhado com a proposta do negócio e com concentração do seu público-alvo. O fator e seus respectivos componentes são representados pela Figura 10.

Figura 10 – Subfatores que compõem o fator “Sociodemografia”.



Fonte: Elaboração própria (2022).

O subfator “Densidade Demográfica” corresponde à medida expressa pela relação entre a população sobre uma determinada área. Tal fator exerce influência no sucesso, pois diz respeito à concentração de clientes potenciais em uma região.

A “Renda Média” é referente aos rendimentos médios mensais por domicílio em uma dada região. O subfator foi considerado importante por se relacionar com o poder aquisitivo dos clientes.

A variável “Trabalhadores” confere a quantidade de trabalhadores de uma região. O subfator compõe o sucesso por quantificar o número de pessoas com uma fonte de renda constante e previsível.

Já a “População por Faixa Etária” é a distribuição da quantidade de pessoas por faixa de idade. O subfator alinha-se com o sucesso, pois se relaciona com o público-alvo dos *studios*.

A definição dos aspectos sociodemográficos relevantes para o sucesso da franquia foi feita em conjunto com o time de Implantação, setor responsável pela etapa de aprovação de região para novas unidades. O processo de avaliação de pontos realizado pela franqueadora já engloba a análise de algumas métricas sociodemográficas, como a Densidade Demográfica e Renda Média, principalmente. O presente trabalho foi aproveitado para investigar outros fatores pressupostos de exercer influência nos resultados das franquias, como concentração de Trabalhadores e a População por Faixa Etária.

#### 4.2.1.5 Priorização

A execução do mapeamento das métricas de sucesso da franquia resultou em um robusto diagrama com os fatores e seus respectivos subfatores de importância para esse resultado. A Figura 11 elucida o mapeamento realizado e sua hierarquia destes fatores.

Figura 11 – Mapa de fatores e subfatores de sucesso.



Fonte: Elaboração própria (2022).

Embora o Desempenho Operacional possa ter possibilidade relevante de influenciar o sucesso da franquia, a variável foi desprezada no presente estudo. Considerando que as métricas de desempenho operacional são posteriores a abertura da unidade, não seria possível trabalhar com dados de operações ainda em implantação ou expansão, as quais são o objetivo desse estudo. Outro fator relevante para a não priorização é a falta de confiabilidade dos dados disponíveis para análise.

Os dados de desempenho são produtos dos processos executados na ponta operacional das franquias. Por conta da recente migração de sistema gerencial das unidades, muitas unidades encontram-se com falta de dados inseridos na nova plataforma. Somado a isso, o processo de entendimento desse fator explicitou a não conformidade do preenchimento de informações com o padrão imposto pela franqueadora em diversas unidades. Em razão disso, os fatores e subfatores da métrica de desempenho operacional foram ignorados nesta aplicação.

O fator Imóvel também foi afetado por uma dificuldade de aferição de dados. O subfator de Limpeza e Organização foi desconsiderado por ter sua aferição impossibilitada na totalidade das unidades da rede de franquias, por conta de deslocamentos e custos envolvidos.

A métrica de Sociodemografia teve seus fatores selecionados de acordo com algumas variáveis de interesse da franqueadora. Dentre os fatores disponíveis na plataforma, foram selecionadas as variáveis condizentes com o que já é conhecido e acompanhado no atual processo de avaliação de pontos de franquia e métricas que são comuns a outros setores, como as utilizadas pelo Marketing na aplicação de campanhas digitais. Dessa forma, priorizaram-se os fatores de Densidade Demográfica, Renda Média, Trabalhadores e População por Faixa Etária.

Com a priorização, o mesmo mapa de fatores e subfatores de sucesso agora é representado pela Figura 12. Estes fatores guiarão o estudo e procedimentos nas etapas seguintes do desenvolvimento desta aplicação.

Figura 12 – Mapa de fatores e subfatores de sucesso priorizado



Fonte: Elaboração própria (2022).

## 4.2.2 Coleta de dados e informações

O trabalho de coleta dos dados que constituem as métricas definidas acima iniciou-se pelos dados gerais das unidades da franquia. Através da plataforma de inteligência de negócios utilizada pela franqueadora - Microsoft Power BI - foi possível coletar as informações de endereço, meses de operação, relação de alunos e ex-alunos e seus respectivos dados cadastrais. Com isso, foi formado uma grande base de dados relacional dos *studios* e seus respectivos alunos e ex-alunos.

### 4.2.2.1 Sucesso das franquias

A classificação em sucesso ou insucesso das franquias foi atribuída por consenso entre gestores e diretores da franqueadora. Foi realizada uma dinâmica de discussão objetiva sobre cada uma das 61 unidades de franquia consideradas, onde se atentou aos resultados operacionais conquistados até o momento do encontro. Fatores como o projetado *versus* realizado do plano de negócio, o tempo de operação e o aproveitamento de potencial da cidade também foram considerados.

O resultado foi uma tabela com que relaciona cada unidade com a variável categórica de sucesso, atribuído o valor 1 em casos positivos e 0 para casos negativos. Das 61 franquias analisadas, 35 foram classificadas como sucesso pelos diretores da franquia. A lista de classificação das franquias encontra-se disponível no **Apêndice A**.

A dinâmica de atribuição de notas gerou uma reflexão no grupo de diretores sobre a quantidade de unidades classificadas como bem-sucedidas. Discutiu-se a relatividade e dificuldade de uma atribuição desse tipo, mesmo baseada em fatores estruturados. A conversa teve resultados positivos, pois também contribuiu para a identificação dos fatores que impediam algumas unidades de serem classificadas como sucesso, gerando um direcionamento para tomada de decisão e acompanhamento mais próximo por parte da franqueadora.

#### 4.2.2.2 Qualidade dos imóveis e estrutura das franquias

Além da definição de quais subfatores influenciam o fator Imóvel e Estrutura, a entrevista realizada com o time de Arquitetura foi importante para estabelecer os pesos e notas para cada um desses componentes das respectivas unidades. O intuito foi formar um *score* resultante das notas e pesos atribuídos para ser possível quantificar de forma mais direta a qualidade dos imóveis das unidades de franquia. Os pesos definidos para os subfatores estão representados pela Quadro 1.

Quadro 1 – Pesos dos subfatores de Imóvel e Estrutura.

<b>Subfator</b>	<b>Peso</b>
Fachada	30%
Personalidade	15%
Acesso	15%
Área treino	40%

Fonte: Elaboração própria (2022).

As notas de “Fachada”, “Personalidade” e “Acesso” foram atribuídas diretamente pelo próprio time de Arquitetura e baseadas nos projetos arquitetônicos e nos relatórios fotográficos das unidades. Já a nota de “Área de Treino” foi calculada utilizando também uma média ponderada, mas desta vez pelos fatores de metragem do espaço de treino, divisão de pavimentos e do formato da área por suas respectivas notas. Para estas notas, os valores também derivaram do detalhamento dos projetos arquitetônicos. Os pesos definidos para os componentes da Área de Treino estão representados no Quadro 2.

Quadro 2 – Pesos dos componentes do fator área de treino.

<b>Subfator</b>	<b>Peso</b>
Metragem	50%
Pavimentos	25%
Formato	25%

Fonte: Elaboração própria (2022).

Todas as notas estão dentro do intervalo de 0 a 10 para facilitar a atribuição e interpretação dos resultados. O cálculo do *Score de Imóvel* foi feito através de uma média ponderada das notas com seus respectivos pesos, e os resultados das notas e *scores* das unidades são consolidados no Apêndice B.

O processo de coleta dos dados dos imóveis com a equipe de Arquitetura se mostrou benéfico para o processo de avaliação de pontos já realizado pela franqueadora. Esse processo não contava com métricas com seus respectivos pesos definidos, muito menos uma quantificação de *scores* para a tomada de decisão. A aplicação das entrevistas foi favorável para uma fundamentação quantitativa das informações, obtendo melhor direcionamento para os pontos críticos e positivos dos imóveis pendentes de aprovação.

#### 4.2.2.3 *Qualidade da gestão das franquias*

A quantificação do fator Gestor se deu de forma similar ao já realizado para o Sucesso e Imóvel. Por meio de entrevistas com diretores e gestores da franqueadora, foi possível atribuir uma nota de 0 a 10 referente a capacidade de gestão atual de cada *studio*. Essas notas foram baseadas na gestão realizada até o momento da avaliação - novembro de 2022. Os valores definidos são reflexo do desempenho do gestor em Competências, Gerenciamento de Time, Controle Financeiro, Gestão por Indicadores e Execução de Processos, principalmente. A relação de notas para cada unidade da franquia é detalhada no Apêndice A.

O processo de atribuição de notas para os gestores teve um papel importante em evidenciar quantitativamente a percepção da franqueadora sobre a gestão de algumas unidades. As discussões surgidas nas entrevistas foram benéficas para um direcionamento de planos de ação sobre os principais pontos de atenção voltados ao gerenciamento de certas operações.

#### 4.2.2.4 Dados sociodemográficos

Para a coleta de dados de Sociodemografia, foi aproveitado o acesso disponível à plataforma de inteligência geográfica utilizada pela franqueadora - OnMaps da Geofusion. A base de dados relacionais construída anteriormente permitiu a definição da geolocalização dos *studios* e seus respectivos alunos no território nacional. A distribuição das unidades de franquia é representada pela Figura 13.

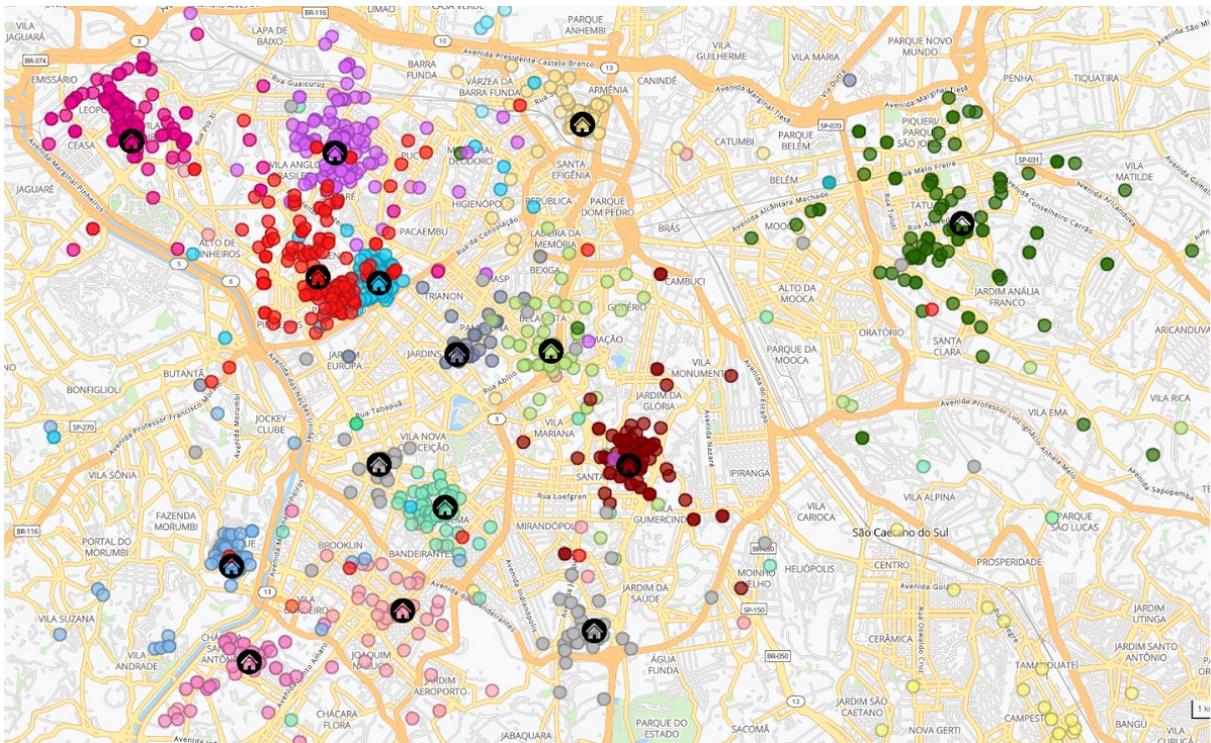
Figura 13 – Localização das unidades de franquia



Fonte: Elaboração própria (2022).

Com as coordenadas das franquias conhecidas, foi possível definir as áreas de estudo de interesse para a coleta dos respectivos dados sociodemográficos. Levantamentos realizados pela franqueadora previamente ao presente trabalho indicam que, em média, cerca de 80% dos alunos estão concentrados em um raio de 3 quilômetros a partir das unidades da franquia. Tal informação caracteriza uma área de influência consideravelmente próxima aos *studios*. A Figura 14 exemplifica a distribuição dos alunos ativos para um conjunto de *studios*, onde cada cor representa alunos de uma determinada unidade, sendo estas destacadas pelos ícones pretos.

Figura 14 – Conjunto de studios e alunos ativos.



Fonte: Elaboração própria (2022).

As áreas de estudo da aplicação, portanto, foram dimensionadas segundo o conhecido comportamento dessas áreas de influência. Foi decidido criar áreas considerando raios de 1 e 2 quilômetros a partir do ponto das franquias. Dessa forma, ambas as áreas geradas fornecem informações sociodemográficas que refletem o entorno do *studio* e de uma região com boa concentração de clientes. A Figura 15 corresponde às áreas de estudo delimitadas em algumas unidades das franquias de São Paulo.

Figura 15 – Áreas de estudo de São Paulo.



Fonte: Elaboração própria (2022).

Definida as áreas de estudo para as unidades, fez-se o *download* dos respectivos dados sociodemográficos. Para alguns destes dados, foi realizado um tratamento das informações adquiridas para que estas fossem ainda mais condizentes com o contexto e objetivo do estudo. Foram tratados os dados das variáveis de Renda Média, Trabalhadores e População por Faixa Etária, de forma que estas representassem melhor o público-alvo dos *studios* conforme parâmetros já conhecidos pela franqueadora.

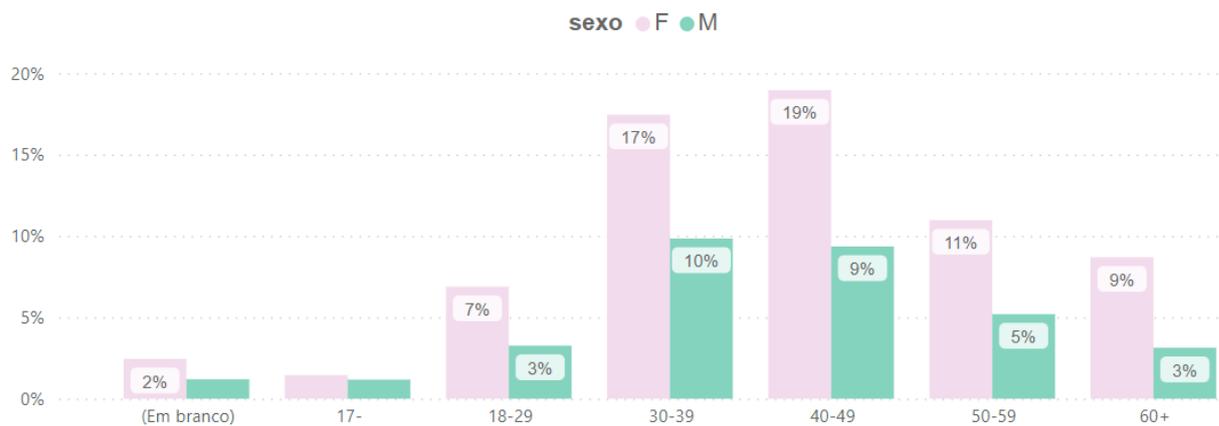
Em relação à Renda Média, optou-se por calcular a Renda Média de Interesse, calculada pela média ponderada da “renda média domiciliar por faixa de renda” como valores, e “Domicílios por faixa de renda domiciliar mensal” como pesos. As faixas de renda consideradas foram A++ (mais de R\$ 29.176,01), A+ (de R\$ 16.768,01 a R\$ 29.176,00), B1 (de R\$ 8.255,01 a R\$ 16.768,00) e B2 (de R\$ 4.458,01 a R\$ 8.255,00). O resultado representa

melhor o público de interesse dos *studios* por se tratar de uma parcela populacional com renda compatível com o *ticket* médio mais elevado, característico da franquia.

Analogamente, a variável Trabalhadores foi refinada para representar o novo fator chamado Trabalhadores de Interesse. O novo valor é calculado por uma soma simples da quantidade de trabalhadores com renda de R\$ 3.718,08 a +R\$ 24.000 nas respectivas áreas de estudo. Assim, o fator se torna mais relevante por selecionar a fração populacional que teria condições financeiras de se tornar um possível aluno do *studio*.

A última modificação na base de dados foi realizada na variável de População por Faixa Etária. Foi realizada uma soma simples da quantidade de pessoas dentro da faixa etária de 25 a 59 anos, o que resultou na chamada População por Faixa Etária de Interesse. A faixa selecionada é condizente com o perfil de aluno mais comum em toda a rede da franquia, como mostra a Figura 16. Essa faixa etária também é associada ao modelo de aulas proposto e a estabilidade financeira que provavelmente a acompanha.

Figura 16 – Perfil etário dos alunos da rede de franquias (quantidade).



Fonte: Elaboração própria (2022).

O resultado dessa etapa, portanto, foi uma base com dados de Densidade Demográfica, Renda Média de Interesse, Trabalhadores de Interesse e População por Faixa Etária de Interesse para as duas áreas de estudo estipuladas para cada uma das 61 unidades analisadas. Tal relação encontra-se detalhada no Apêndice A.

A coleta dos dados sociodemográficos teve sua importância estendida para outras frentes da franquadora. Foi possível compilar outros indicadores que permitiram um estudo melhor do comportamento de regiões e cidades, sendo de grande utilidade para o setor de Expansão e Implantação. Deu-se início em trabalhos internos para investigar padrões na presença de concorrentes, polos atrativos e outros aspectos sociodemográficos.

### 4.3 CONSTRUÇÃO DO MODELO

Para a aplicação da regressão logística nos dados compilados até a presente etapa, foi escolhido o *software* estatístico SPSS da IBM. O programa é considerado de fácil usabilidade, além de contemplar todas as análises necessárias para a execução da aplicação sem exigir uma curva de aprendizado muito grande.

A parametrização de alguns fatores foi importante para a construção do modelo. Definiu-se o nível de significância em 10%, dada a subjetividade de alguns dados, como a avaliação dos fatores de imóvel e de gestão das franquias. Dessa forma, a rejeição - ou não - da hipótese nula considera estas incertezas.

O modelo de regressão logística prevê dois tipos de variáveis: a dependente e as independentes. Nesta aplicação, a variável dependente é definida pelo Sucesso de uma unidade de franquia. Já para as variáveis independentes - também chamadas de previsores - são considerados os 7 fatores explicados e priorizados nas etapas anteriores: Gestor, Imóvel, Densidade Demográfica, Renda Média de Interesse, Trabalhadores de Interesse e População da Faixa Etária de Interesse.

O método escolhido para realizar o modelo de regressão no SPSS foi o de “passo a passo” e “para trás”, que inicia a regressão com todas as variáveis independentes e as retira seguindo um critério de remoção por cada etapa. Os métodos passo a passo são apropriados quando há interesse em testar hipóteses e encontrar um modelo para ajustar os dados, como sugere Menard (1995). O método “para trás” também é recomendado, pois limita o efeito supressor, que ocorre quando uma variável independente só tem efeito significativo quando outra variável é constante (FIELD, 2009).

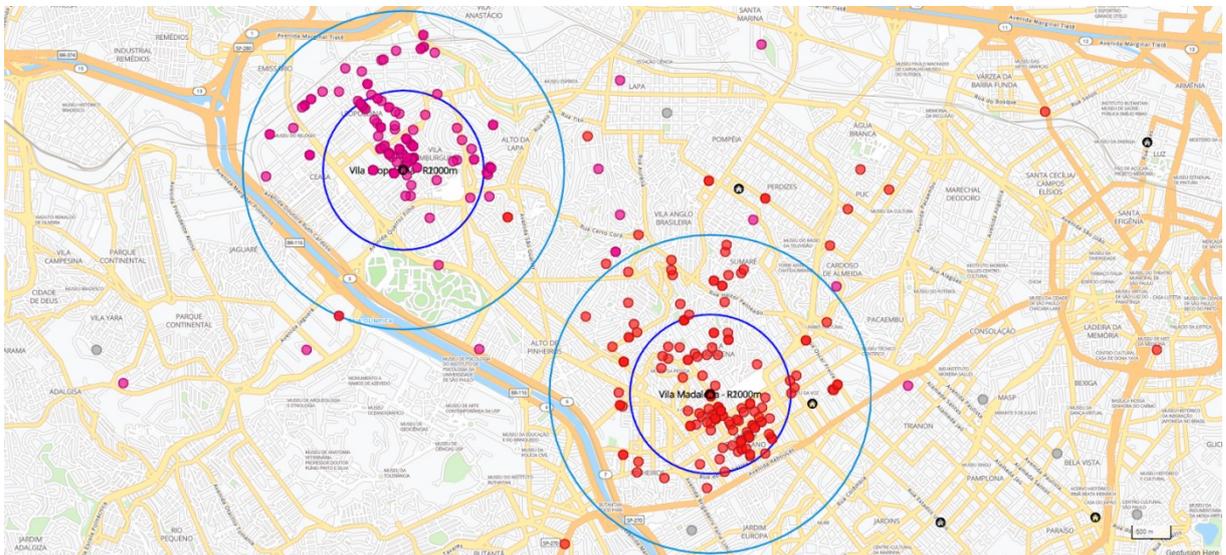
A criação de áreas de estudo com raios diferentes contribuiu para a formação de 3 combinações que serviram de base para a aplicação do modelo de regressão logística. O primeiro modelo tem seus fatores de sociodemografia decorrentes de áreas formadas por um raio de 1 quilômetro a partir do endereço das unidades. O segundo modelo considera um raio de 2 quilômetros para os dados desses fatores. Já para o terceiro modelo, foi feita uma combinação das áreas geradas por raios de 1 e 2 quilômetros, sendo o critério de escolha o tamanho da cidade onde a unidade de franquia está localizada. Foi definido que cidades com mais de 1.000.000 de habitantes teriam as áreas de estudo definidas por raios de 1 quilômetro.

Para os 3 modelos definidos, os dados de Sucesso, Imóvel e Gestor permaneceram os mesmos. Portanto, analisou-se somente a diferenciação do comportamento sociodemográfico conforme as variações das áreas. As combinações tiveram o objetivo de abranger as incertezas

associadas à variabilidade das características demográficas de cada cidade, além de fatores associados à locomoção e perfil de consumo.

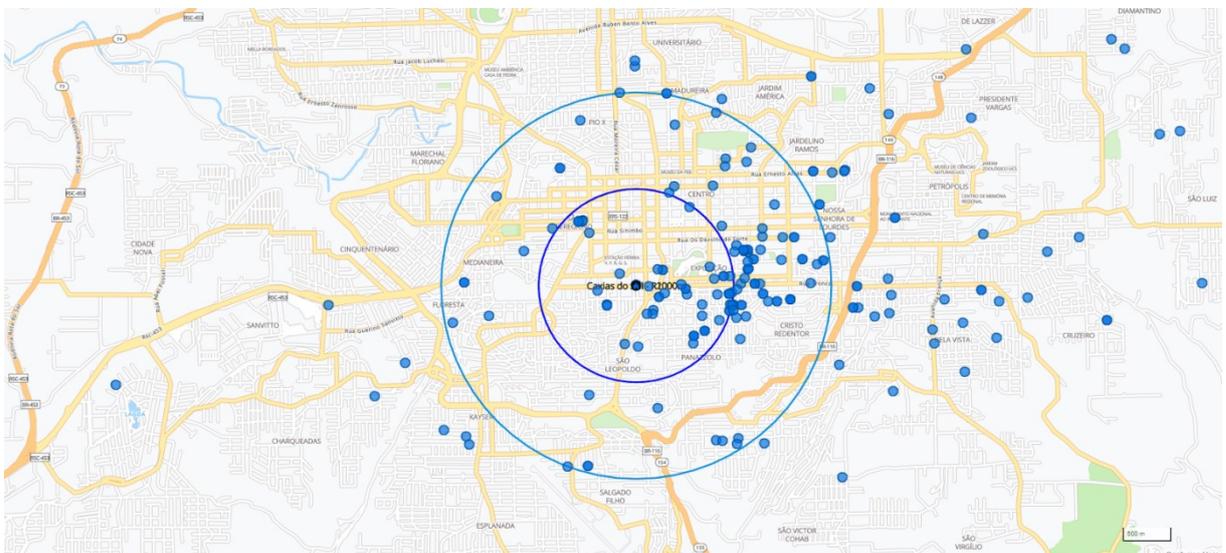
A partir da plotagem dos alunos ativos das unidades, foi notável a diferença da área de influência nos grandes centros como São Paulo/SP, quando comparada a cidades menores como Caxias do Sul/RS, representadas pelas Figuras 17 e 18 respectivamente. Nos grandes centros observa-se a tendência de alunos ativos residirem mais próximos aos studios, concentrados em um raio de até 1 quilômetro. Em contrapartida, cidades menores mostram uma maior dispersão dos alunos em relação ao endereço do *studio*.

Figura 17 – Alunos ativos de unidades em São Paulo/SP.



Fonte: Elaboração própria (2022).

Figura 18 – Alunos ativos de unidades em Caxias do Sul/RS.



Fonte: Elaboração própria (2022).

Portanto, nesta etapa definiu-se o software de aplicação do modelo, a separação das variáveis dependentes e independentes, bem como os parâmetros de configuração. Consolidou-se, por vez, os dados coletados em 3 modelos distintos preparados para compor o leque de cenários de análise. Com isso definido e devidamente configurado, é possível prosseguir para as análises e discussões do resultado.

## 5 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Este capítulo apresenta as análises dos resultados obtidos pela aplicação no Capítulo 4. São descritas as principais observações referentes aos produtos do modelo de regressão logística, seus testes de aderência e qualidade e conclusão sobre as variáveis preditivas. A seção também visa trazer uma discussão destes resultados quando comparados ao contexto atual da empresa em questão.

Embora o mesmo modelo tenha sido aplicado para os 3 cenários - R1000m, R2000m e Misto (R1000m e R2000m) - a discussão será sobre o modelo com resultados de maior significância para o trabalho. O modelo que obteve os melhores resultados foi o que considerou áreas geradas pelo raio de 1 quilômetro.

As análises iniciaram pelo diagnóstico de multicolinearidade das variáveis previsoras. A presença de multicolinearidade resultaria em um modelo viciado, de forma que gerasse uma tendência de aumento dos erros padrões dos coeficientes dos previsores, que por vez afetariam as significâncias estatísticas dessas variáveis. Para o modelo escolhido, foi constatado que as variáveis Densidade Demográfica e População por Faixa Etária de Interesse apresentam colinearidade entre si.

A Figura 19 apresenta as estatísticas que indicam a presença de colinearidade nestas variáveis. Os valores sugeridos por Menard (1995) para a Tolerância e por Myers (1990) para VIF, menor que 0,1 e maior que 10 respectivamente, sugerem que as variáveis apresentem colinearidade. Essa colinearidade foi depois confirmada pela análise dos autovalores e das proporções de variância.

Figura 19 – Estatística de colinearidade das variáveis predictoras.

**Coefficientes<sup>a</sup>**

Modelo		Estatísticas de colinearidade	
		Tolerância	VIF
1	Imóvel	,851	1,176
	Gestor	,749	1,336
	Densidade_demográfica	▷ ,002	▷ 405,198
	Renda_média_interesse	,609	1,643
	Trabalhadores	,522	1,916
	Pop_faixa_etária_interesse	▷ ,002	▷ 418,654

a. Variável dependente: Sucesso

Fonte: Elaboração própria (2022).

Essa característica também é fundamentada pela correlação dessas mesmas duas variáveis, como mostra a Figura 20. Observa-se uma correlação positiva e quase perfeita entre os fatores de Densidade Demográfica e População por Faixa Etária de Interesse ( $r = +0,999$ ). A interpretação disso é que à medida que a concentração de pessoas presentes em uma área aumenta, a quantidade de pessoas da faixa etária escolhida também aumenta na mesma proporção.

Figura 20 – Correlações das variáveis do modelo

**Correlações**

		Sucesso	Imóvel	Gestor	Densidade_d emográfica	Renda_médi a_interesse	Trabalhadore s	Pop_faixa_etá ria_interesse
Sucesso	Correlação de Pearson	1	,314*	,643**	,138	,117	-,091	,139
	Sig. (2 extremidades)		,014	,000	,288	,370	,484	,286
	N	61	61	61	61	61	61	61
Imóvel	Correlação de Pearson	,314*	1	,365**	,004	-,005	-,033	,005
	Sig. (2 extremidades)	,014		,004	,974	,967	,801	,971
	N	61	61	61	61	61	61	61
Gestor	Correlação de Pearson	,643**	,365**	1	,323*	,140	,165	,328**
	Sig. (2 extremidades)	,000	,004		,011	,281	,205	,010
	N	61	61	61	61	61	61	61
Densidade_demográfica	Correlação de Pearson	,138	,004	,323*	1	,542**	,622**	,999**
	Sig. (2 extremidades)	,288	,974	,011		,000	,000	,000
	N	61	61	61	61	61	61	61
Renda_média_interesse	Correlação de Pearson	,117	-,005	,140	,542**	1	,554**	,552**
	Sig. (2 extremidades)	,370	,967	,281	,000		,000	,000
	N	61	61	61	61	61	61	61
Trabalhadores	Correlação de Pearson	-,091	-,033	,165	,622**	,554**	1	,632**
	Sig. (2 extremidades)	,484	,801	,205	,000	,000		,000
	N	61	61	61	61	61	61	61
Pop_faixa_etária_interesse	Correlação de Pearson	,139	,005	,328**	▷ ,999**	,552**	,632**	1
	Sig. (2 extremidades)	,286	,971	,010	,000	,000	,000	
	N	61	61	61	61	61	61	61

\*. A correlação é significativa no nível 0,05 (2 extremidades).

\*\*.. A correlação é significativa no nível 0,01 (2 extremidades).

Fonte: Elaboração própria (2022).

A constatação da multicolinearidade teve como consequência a retirada de uma das variáveis preditoras no modelo. Retirou-se a variável de População por Faixa Etária de Interesse pois, pensando no processo de avaliação de pontos, a métrica de Densidade Demográfica já é utilizada pela franqueadora e não há necessidade de qualquer tratamento dos dados. Ainda, a variável de faixa etária acaba por excluir uma parcela da população, que mesmo não sendo o público que mais consome o serviço, ainda apresenta potencial de serem convertidos em clientes.

Com isso, foi feita a aplicação do modelo de regressão logística nas variáveis restantes. Os resultados obtidos demonstram que o modelo que inclui as variáveis preditoras é significativamente melhor que o modelo nulo, que considera somente a constante. Isso é demonstrado pelo valor de qui-quadrado de 46,046, no qual valores positivos e maiores que 0 indicam uma melhora no modelo e serve de comparação entre as diferentes etapas. Portanto, fez sentido avançar com o restante das análises do modelo. O grau de aderência do modelo é confirmado pelo valor de verossimilhança-log -2 e  $R^2$  de Nagelkerke, que indicam quanto do modelo não é explicado pelas variáveis consideradas. Obteve-se os valores de 37,182 e 0,721, para verossimilhança-log -2 e  $R^2$  de Nagelkerke, respectivamente. Estes valores indicam uma melhora significativa em relação ao modelo nulo.

Outro fator que permite avaliar a qualidade do ajuste do modelo é o teste de Hosmer e Lemeshow. Os resultados do modelo para o respectivo teste apontam um valor de qui-quadrado de 3,148 e *p-value* de 0,925. Os valores refletem um bom ajuste do modelo - lembrando que o modelo é rejeitado para valores abaixo do nível de significância (FIEL, 2009).

O modelo criado conseguiu classificar as unidades corretamente em 85,2% dos casos. A Tabela 1 representa a matriz de confusão que demonstra a distribuição das classificações reais e previstas pelo modelo. Os resultados foram satisfatórios para previsões corretas de unidades que obteriam sucesso e para unidades que não obteriam.

Tabela 1 – Matriz de confusão

		PREVISTO		
		Não Sucesso (1)	Sucesso (1)	Porcentagem Correta
REAL	Não Sucesso (0)	22	4	84,6%
	Sucesso (1)	5	30	85,7%
	Porcentagem global			<b>85,2%</b>

Fonte: Elaboração própria (2022).

Na Tabela 2 estão representados os resultados das variáveis finais selecionadas pelo método de passo a passo do SPSS. São detalhados os coeficientes de regressão logística  $\beta$ , o desvio padrão e o *p-value* obtido para cada um dos fatores.

Tabela 2 – Resultados da regressão logística.

Fator	Coefficiente ( $\beta$ )	Desvio padrão	p-value
Gestor (score)	2,318	0,76	0,002
Renda_média_interesse	0,00012	0,00008	0,047
Trabalhadores	-0,00012	0,00005	0,02
Constante	-12,825	4,053	0,002

Fonte: Elaboração própria (2022).

Adotado o nível de significância em 10%, as variáveis preditoras mais significativas foram Gestor, Renda Média de Interesse e Trabalhadores. O fator Gestor teve o *p-value* em 0,02, a Renda Média de Interesse em 0,047 e Trabalhadores em 0,020. As outras variáveis foram descartadas pelo modelo por não atenderem a métrica de significância.

Pode-se associar o valor positivo coeficiente de regressão  $\beta$  a maiores chances de classificar como Sucesso ao incrementar em 1 unidade no valor de Gestão e Renda Média de Interesse. Já Trabalhadores tem valor negativo e possui interpretação oposta, onde o aumento do valor da variável implica em menores chances de classificar uma unidade como Sucesso.

Na Tabela 3 estão representados para as mesmas variáveis os resultados de  $\text{Exp}(\beta)$  e seus limites inferiores e superiores para um intervalo de confiança de 90%.

Tabela 3 – Resultados da regressão logística

Fator	Exp( $\beta$ )	90% C.I para Exp( $\beta$ )	
		Inferior	Superior
Gestor (score)	10,153000	2,910000	35,419000
Renda_média_interesse	1,000120	0,999996	1,000244
Trabalhadores	0,999875	0,999786	0,999964

Fonte: Elaboração própria (2022).

O valor de  $\text{Exp}(\beta)$ , também conhecido como Razão de Chances, pode ser interpretado como a variação da chance de pertencimento à categoria Sucesso ao incremento de 1 unidade no valor da variável preditora. Portanto, temos que o acréscimo de 1 unidade na nota de Gestão aumenta a chances da unidade ser classificada com Sucesso em 10 vezes. De forma similar, o acréscimo de R\$ 1,00 no valor da Renda Média de Interesse implica em um aumento de chances em 1,00012 vezes. Já aumento de 1 Trabalhador diminui a chance de classificação em sucesso. A equação resultante do modelo de regressão que descreve essa relação é representada pela Equação 3.

$$\log p/(p1 - p) = -12,825 + 2,318 * (\text{nota gestor}) + 0,00012 * (\text{Renda média interesse}) - 0,00012 * (\text{Trabalhadores}) \quad (3)$$

Em resumo, o modelo de regressão logística foi aplicado para compreender os efeitos do imóvel, gestão, densidade demográfica, renda média de interesse e número de trabalhadores de interesse no sucesso de uma unidade de franquia do setor *fitness*. O modelo resultante foi estatisticamente significante quando comparado ao modelo nulo, ( $X^2(3) = 46,046$   $p < 0,01$ ), explicou 72,10% do sucesso (Nagelkerke  $R^2$ ) e previu corretamente 85,20% dos casos. Gestor ( $p < 0,003$ ), Renda Média de Interesse ( $p < 0,048$ ) e Trabalhadores ( $p < 0,03$ ) foram estatisticamente significantes, enquanto Imóvel ( $p = 0,892$ ) e Densidade Demográfica ( $p = 0,575$ ) não foram.

É interessante analisar os dados obtidos pelo modelo de regressão pela perspectiva da realidade vivenciada pelas franquias e acompanhada pela franqueadora. De fato, observa-se a importância do papel exercido pelo Gestor nas operações. O modelo de negócio é desenhado com a figura de um gestor muito presente e executor, de forma que seja um dos diferenciais

da franquia por resultar em um acompanhamento e atendimento muito próximo aos alunos. Já são constatados resultados interessantes em unidades localizadas em regiões com menor potencial sociodemográfico, mas que apresentam uma boa gestão. As variáveis sociodemográficas também são de interesse da franqueadora. Embora na prática não se interprete a quantidade de Trabalhadores como um fator de potencial negativo, consideram-se interessantes os resultados obtidos sugeridos pelo modelo.

## 6 CONCLUSÃO E CONSIDERAÇÕES FINAIS

O presente trabalho tinha o objetivo geral de definir os principais fatores que influenciam no sucesso da operação das franquias de *studios fitness*, de maneira que se auxilie na tomada de decisão para aprovação de novas unidades. Para se tornar possível o alcance do objetivo geral, foram elencados quatro objetivos específicos.

O primeiro objetivo específico era voltado para a identificação dos fatores e subfatores que compõem o sucesso operacional da franquia. Esse objetivo foi atingido através de um mapeamento das variáveis e subvariáveis que impactam no sucesso da franquia. Foi realizado um amplo processo de entendimento do modelo de negócio e utilizou-se de entrevistas com o time da franqueadora e análise de dados para a construção do mapa de fatores e subfatores. Por vez, estes fatores foram priorizados, quantificados e consolidados.

O segundo objetivo visava analisar qual a importância de cada fator na composição do sucesso. O propósito foi alcançado mediante a aplicação de uma regressão logística nos dados da franquia, pela qual foi possível verificar o comportamento dos fatores mapeados como influenciadores no sucesso. Das 6 variáveis preditoras priorizadas no modelo, apenas 3 tiveram significância estatística e peso considerável de influência no sucesso.

O mesmo procedimento de regressão logística foi importante para o alcance do terceiro objetivo específico, que visava construir um modelo preditivo do sucesso operacional das franquias em questão que fosse baseado nos fatores determinados. Com os resultados da regressão logística, os coeficientes resultantes forneceram suporte para a elaboração do modelo de previsão. O modelo construído obteve 85,20% da previsão de sucesso correta.

Infere-se, pois, que o desenvolvimento e a aplicação do trabalho atingiu seu objetivo geral de forma satisfatória. Foi possível elaborar ao menos uma versão inicial de um modelo que forneça um suporte quantitativo para a expansão das franquias e aprovação de novas unidades. Pode-se analisar os resultados obtidos por uma perspectiva da realidade das franquias e da franqueadora, confirmando assim a direção das informações apontadas pelo modelo.

Este trabalho permitiu a expansão e o esclarecimento de alguns fatores, acompanhados de forma constante no dia a dia da franqueadora, mas que por vezes são tratados de forma distante. Julga-se muito benéfica à participação das equipes especializadas da franqueadora na aplicação do trabalho, onde a colaboração foi crucial para o entendimento do problema e das variáveis envolvidas. As dinâmicas de mapeamento e quantificação dos

fatores realizadas com estas equipes foram responsáveis por fomentar discussões benéficas e encaminhamentos de novos projetos para a empresa em questão.

## 6.1 SUGESTÕES PARA TRABALHOS FUTUROS

Como sugestão para trabalhos futuros, recomenda-se realizar um trabalho interno da franqueadora para, de alguma maneira, quantificar competências e variáveis que neste trabalho não foram possíveis aferir. A sugestão se estende principalmente para os subfatores de algumas variáveis, onde mesmo tendo as suas características bem definidas, sua quantificação é difícil e muito relativizada pela falta de processos que a facilitem. Tal sugestão não se concentra apenas na melhoria do modelo, mas tem importância também na própria tomada de ação ao acompanhar esses dados.

Dado o contexto prático do trabalho e da utilização de dados reais da franqueadora, é de interesse que o modelo construído seja aproveitado e adaptado à rotina de avaliação de novos pontos comerciais da franqueadora. Dessa forma, finaliza-se a recomendação de próximos passos para a devida aplicação do modelo na rotina e de sua padronização como processo dentro da franqueadora na medida do possível. Tudo isso, claro, triado por um rigoroso processo de melhoria do modelo elaborado.

## REFERÊNCIAS

- AGUIAR, H. de S.; LUPPE, M. R.; NASCIMENTO, P. T. de S. Modelo de Fatores Influentes na Operação de Franquias. *BBR. Brazilian Business Review*, v. 18, p. 45-61, 2021.
- ALVES, M. P.; CHAVES, A. P. A. Formação profissional em Educação Física e gestão fitness: possíveis olhares a partir da pesquisa documental. *Professional training in Physical Education and fitness management: Possible views from documentary research. Brazilian Journal of Development*, v. 7, n. 12, p. 112879-112897, 2021.
- BAI, S. *et al.* GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China. *Geomorphology*, v. 115, n. 1-2, p. 23-31, 2010.
- BERNUZZI, G. M. Segmentação de mercado em academias esportivas: uma revisão da literatura. 2022.
- BLOCKEN, B; VAN DRUENEN, T; VAN HOOFF, T; VERSTAPPEN, P. A; MARCHAL, T; MARR, L. C. Can indoor sports centers be allowed to re-open during the COVID-19 pandemic based on a certificate of equivalence? *Build Environ.* 2020; 180: 107022.  
doi:10.1016/j.buildenv.2020.107022
- BUI, D.T.; TUAN, T.A.; KLEMPE, H.; PRADHAN, B.; REVHAUG, I. Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* 2016, 13, 361–378.
- BURROUGH, P. A. *et al.* Principles of geographical. Information systems for land resource assessment. Clarendon Press, Oxford, 1986.
- CÂMARA, G.; DAVIS, C.; MONTEIRO, A. Miguel V. Introdução à ciência da geoinformação. 2001.
- CARNASCIALI, A. M. dos S.; DELAZARI, L. S. A localização geográfica como recurso organizacional: utilização de sistemas especialistas para subsidiar a tomada de decisão locacional do setor bancário. *Revista de Administração Contemporânea*, v. 15, p. 103-125, 2011.
- CHATTERJEE, S.; HADI A. S. *Regression Analysis By Example*. 4. Ed. Mexico: Wiley Interscience, 2006.
- COWEN, D. J. GIS versus CAD versus DBMS: what are the differences? *Photogrammetric Engineering and Remote Sensing*, 54(1), 1551-1554.1988.

CUCINOTTA, D.; VANELLI, M. WHO declares COVID-19 a pandemic. *Acta Bio Medica: Atenei Parmensis*, v. 91, n. 1, p. 157, 2020.

FERNANDES, A. A. T. *et al.* Leia este artigo se você quiser aprender regressão logística. *Revista de Sociologia e Política*, v. 28, 2021.

FIA BUSINESS SCHOOL. Mercado fitness: evolução, desafios e tendências para 2023. São Paulo. 2021. Disponível em: <https://fia.com.br/blog/mercado-fitness/>

FIELD, A. *Descobrimo a estatística usando o SPSS*. 2. ed. Porto Alegre. 2009.

FIGUEIREDO FILHO, D; SILVA, L; DOMINGOS, A. O Que é e como Superar a Multicolinariedade? Um Guia Para Ciência Política. *Conexão Política*, v. 4, n. 2, p. 95-104, 2015.

GUETTA, A. *et al.* *Franchising: Aprenda com os especialistas*. Prime Books Brasil. Rio de Janeiro, 2013.

HODGE, C.; OPPEWAL, H.; TERAWATANAVONG, C. Determinants of franchise conversion: a franchisee perspective. *European Journal of Marketing*, 2013.

HOSMER JR, D. W.; LEMESHOW, S; STURDIVANT, R. X. *Applied logistic regression*. John Wiley & Sons, 2013.

IHRSA [International Health, Racquet & Sportsclub Association]. Findings from the IHRSA Global Report 2017. *Health Club Management*, v. 7, 2017.

JÚNIOR, P. G. F; PAIANO, R; DOS SANTOS, A.C. Isolamento social: consequências físicas e mentais da inatividade física em crianças e adolescentes. *Revista Brasileira de Atividade Física & Saúde*, v. 25, p. 1-2, 2020.

LEE, C. K. H. *et al.* A cloud-based responsive replenishment system in a franchise business model using a fuzzy logic approach. *Expert systems*, v. 33, n. 1, p. 14-29, 2016.

MENARD, S. An introduction to logistic regression diagnosis. *Applied Logistic Regression Analysis: Sage University Series on Quantitative Applications in the Social Sciences*. 1995.

MYERS, R. H. *Classical and modern regression with applications*. Belmont, CA: Duxbury press, 1990.

NHU, V. *et al.* Shallow landslide susceptibility mapping: A comparison between logistic model tree, logistic regression, naïve bayes tree, artificial neural network, and support vector machine algorithms. *International journal of environmental research and public health*, v. 17, n. 8, p. 2749, 2020.

PHAM, B.T.; PRAKASH, I.; SINGH, S.K.; SHIZARDI, A.; SHAHABI, H.; BUI, D.T. Landslide susceptibility modeling using reduced error pruning trees and different ensemble techniques: Hybrid machine learning approaches. *Catena*. 2019.

PORTER, M. E. Towards a dynamic theory of strategy. *Strategic management journal*, v. 12, n. S2, p. 95-117, 1991.

PRIHANDONO, D.; WIJAYANTO, A.; CAHYANINGDYAH, D. Franchise business sustainability model: Role of conflict risk management in Indonesian franchise businesses. *Problems and Perspectives in Management*, v. 19, n. 3, p. 383, 2021.

SCALDAFERRI, D. B. D.; MATOS, V. de A. A competitividade do setor de academias de ginástica em Salvador. *Seminário Estudantil de Produção Acadêmica*, v. 11, n. 1, 2007.

SHIRZADI, A.; BUI, D.T.; PHAM, B.T.; SOLAIMANI, K.; CHAPI, K.; KAVIAN, A.; SHAHABI, H.; REVHAUG, I. Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. *Environ. Earth Sci.* 2017, 76, 60.

SILVA, V. L. S. Ambiente institucional e organização e rede de franquias: uma comparação entre Brasil e França. Tese. Programa de pós-graduação em engenharia de produção – Centro de ciências exatas e de tecnologia, Universidade Federal de São Carlos, São Carlos. 2004.

SMOLSKI, F. M. da S. *et al.* Análise do perfil da produção científica da Revista de Administração Pública (RAP) no período 2003-16. *Revista de Administração Pública*, v. 51, p. 1139-1163, 2017.

TOSCANO, J. J. de O. Academia de ginástica: um serviço de saúde latente. *Rev. bras. ciênc. mov*, p. 40-2, 2001.

WARBURTON, D. ER.; BREDIN, S. SD. Health benefits of physical activity: a systematic review of current systematic reviews. *Current opinion in cardiology*, v. 32, n. 5, p. 541-556, 2017.

## GLOSSÁRIO

**Área de influência:** é o território ao redor do estabelecimento que concentra a maioria dos clientes ou das vendas daquele ponto.

**Delimitação de área de influência:** a distância que o consumidor está disposto a percorrer - ou seu tempo de deslocamento - para consumir tal produto ou serviço oferecido por uma unidade.

**Densidade demográfica:** é o total de pessoas dividido por quilômetro quadrado (habitantes/km<sup>2</sup>), ou seja, é a concentração de moradores de uma determinada região.

**Faixas de renda:** define as classes de renda da população brasileira, utilizando como base as informações contidas no Cadastro Nacional de Informações Sociais - CNIS.

**Faixas de renda Geofusion:** A++ (mais de 26.164,01); A+ (15.071,01 – 26.164,00); B1 (7.053,01 – 15.071,00); B2 (3.779,01 – 7.053,00); C1 (2.165,01 – 3.779,00); C2 (1.197,01 – 2.165); D (880,01 – 1.197); e E (até 880).

**Faixas Etárias Geofusion:** 0 a 4 anos; 5 a 14 anos; 15 a 19 anos; 20 a 24 anos; 25 a 34 anos; 35 a 44 anos; 45 a 49 anos; 50 a 59 anos; 60 anos ou mais.

**Fatores sociodemográficos:** elementos que contribuem ou influenciam nas condições, causas e resultados relacionados com aspectos sociais e demográficos, especialmente usado para se referir a um lugar e aos grupos sociais que dele fazem parte.

**Geofusion:** plataforma de dados dinâmicos que identifica o perfil do consumidor, entende os seus hábitos e como, onde e por que ele consome tal marca.

**Mercado de trabalho:** é um conceito de trocas entre pessoas que buscam emprego e as empresas que oferecem vínculo empregatício.

**Mercado de trabalho Geofusion:** definido pela base de empresas do mercado formal; pirâmide salarial por setor de atividade (por município); taxas de informalidade por setor de atividade e faixas salariais (por região metropolitana e estados).

**OnMaps:** Software de inteligência geográfica que trabalha com base em dados assertivos para auxiliar a escolha do melhor local para negócios em geral. O OnMaps tem dados como idade, renda, faixa etária e nível de instrução da população, além do cruzamento de idade e renda.

**Renda média:** a média dos rendimentos dos domicílios de determinada região, obtidas a partir da soma das rendas de todas as famílias residentes, dividida pelo número total de domicílios no local.

**Renda média domiciliar:** a soma das receitas individuais dos moradores de um domicílio.

**Plataforma de dados sociodemográficos:** fonte dos dados da renda domiciliar para as análises e aplicações mercadológicas.

**População por faixa etária:** classificação da população de acordo com sua idade em anos.

**Trabalhador:** quem oferece a sua força de trabalho e recebe uma remuneração em troca desta especialidade.

**APÊNDICE A – VALORES CONSOLIDADOS DAS VARIÁVEIS**  
**Apêndice A1 - Valores para modelo R1000m**

R1000m							
# Unidade	Sucesso	Imóvel (score)	Gestor (score)	Densidade demográfica	Renda média de dom. da faixa de interesse	Trabalhadores de interesse	População por faixa etária de interesse
#1	0	7,68	3	17.499,36	21.809	18.022	42.383
#2	1	7,45	5	14.699,38	19.417	2.668	33.321
#3	0	7,3	4	8.864,08	23.130	3.858	20.399
#4	1	7,05	4	7.380,83	31.752	7.725	18.548
#5	1	8,95	10	11.960,38	18.871	15.814	30.034
#6	0	7,93	4	16.377,25	20.684	32.091	40.934
#7	1	9,18	5	13.145,57	15.128	27.054	28.326
#8	0	7,45	2	4.543,48	12.870	1.196	10.574
#9	1	7,7	7	3.252,12	13.093	908	7.347
#10	0	7,55	4	3.317,33	23.578	31.646	8.059
#11	1	8,95	7	16.071,57	19.527	14.294	40.056
#12	0	9,05	5	16.596,01	27.645	14.573	41.034
#13	1	8,58	6	7.777,92	13.263	3.261	18.595
#14	1	7,35	8	11.480,92	12.535	9.543	28.464
#15	1	8,25	7	17.333,52	27.248	12.578	40.798

R1000m							
# Unidade	Sucesso	Imóvel (score)	Gestor (score)	Densidade demográfica	Renda média de dom. da faixa de interesse	Trabalhadores de interesse	População por faixa etária de interesse
#16	0	8,6	6	7.936,17	25.826	47.454	19.513
#17	0	7,35	5	7.717,19	12.071	5.963	17.315
#18	0	6,6	5	22.176,11	28.634	63.808	54.919
#19	1	6,6	4	4.675,31	20.862	6.130	10.425
#20	1	9,5	7	5.579,93	10.534	2.510	13.097
#21	0	6,6	5	13.168,69	26.896	32.405	32.827
#22	1	8,6	5	1.982,76	9.750	404	4.501
#23	0	8,45	4	4.508,72	18.082	11.881	10.912
#24	1	7,65	10	44.299,02	21.964	70.511	106.914
#25	0	5,05	4	2.484,85	8.090	1.192	4.818
#26	1	8,9	9	17.707,18	32.346	76.620	44.816
#27	0	8,35	5	5.628,13	12.265	7.774	13.485
#28	1	8,65	5	18.570,34	30.320	29.863	46.750
#29	1	8,55	9	4.142,96	18.336	3.585	9.405
#30	1	8,35	5	10.479,09	16.252	9.131	24.715
#31	0	8,58	4	4.403,93	11.061	3.560	9.162

R1000m							
# Unidade	Sucesso	Imóvel (score)	Gestor (score)	Densidade demográfica	Renda média de dom. da faixa de interesse	Trabalhadores de interesse	População por faixa etária de interesse
#32	1	7,95	5	17.463,12	30.298	18.414	43.223
#33	1	7,9	5	24.896,97	31.604	21.259	61.600
#34	1	5,6	5	6.557,23	19.105	1.911	15.167
#35	1	9	10	25.520,20	22.595	5.220	63.536
#36	0	7,45	5	12.294,17	11.893	9.569	28.547
#37	0	7,55	4	28.767,20	27.516	66.890	71.055
#38	0	8,95	5	21.667,93	28.893	23.302	51.889
#39	1	8,45	8	22.072,34	26.362	33.104	55.571
#40	0	8,15	3	3.808,07	15.190	4.517	9.480
#41	1	7,93	8	5.887,90	20.404	1.947	14.258
#42	1	7,65	9	30.600,67	27.247	15.506	73.657
#43	0	7,98	5	6.877,69	35.482	18.122	14.598
#44	1	8,35	8	17.804,92	20.317	6.266	42.043
#45	1	8,8	8	5.678,47	12.164	1.354	12.520
#46	1	9,48	9	13.552,46	21.617	11.649	33.784
#47	0	4,55	5	13.894,73	23.175	22.870	33.631

R1000m							
# Unidade	Sucesso	Imóvel (score)	Gestor (score)	Densidade demográfica	Renda média de dom. da faixa de interesse	Trabalhadores de interesse	População por faixa etária de interesse
#48	1	9,25	6	9.159,96	22.224	7.048	21.594
#49	0	7,45	4	22.649,54	16.957	11.323	56.769
#50	0	7,78	5	2.494,55	11.542	859	5.650
#51	1	9,35	6	14.512,95	20.124	15.908	34.664
#52	0	6,9	3	20.864,63	20.779	11.418	48.927
#53	0	9,85	5	4.331,91	10.795	3.945	8.045
#54	1	8,15	6	584,28182	18.497	569	1.177
#55	0	7,38	4	4.351,15	11.426	789	10.462
#56	1	8,53	5	14.150,22	25.234	15.638	33.698
#57	1	8,35	5	3.419,20	13.165	1.978	7.637
#58	1	8,9	9	11.547,91	26.172	11.076	26.817
#59	1	8	9	19.575,23	26.422	25.016	48.898
#60	1	7,95	6	18.428,88	27.384	13.530	44.018
#61	0	8,25	4	12.433,76	23.374	6.521	29.808

## Apêndice A2 - Valores para modelo R2000m

R2000m							
# Unidade	Sucesso	Imóvel (score)	Gestor (score)	Densidade demográfica	Renda média de domicílios da faixa de interesse	Trabalhadores de interesse	População por faixa etária de interesse
#1	0	7,68	3	11.046,52	20.986	44.348	107.291
#2	1	7,45	5	7.525,48	19.449	7.441	67.592
#3	0	7,3	4	7.918,24	19.463	12.917	72.932
#4	1	7,05	4	5.978,43	31.458	18.239	59.774
#5	1	8,95	10	11.343,27	19.156	41.378	112.075
#6	0	7,93	4	17.096,21	19.770	114.959	166.532
#7	1	9,18	5	14.446,62	15.510	172.244	130.607
#8	0	7,45	2	3.546,34	11.555	5.294	32.743
#9	1	7,7	7	3.019,96	12.175	4.921	27.480
#10	0	7,55	4	3.367,91	21.980	81.556	32.297
#11	1	8,95	7	10.905,76	16.321	48.771	108.841
#12	0	9,05	5	14.358,86	28.147	54.446	141.450
#13	1	8,58	6	6.075,26	11.736	8.705	55.161
#14	1	7,35	8	9.587,82	11.684	23.188	92.386
#15	1	8,25	7	18.325,56	24.292	73.035	176.208
#16	0	8,6	6	7.416,39	25.323	98.839	71.246

R2000m							
# Unidade	Sucesso	Imóvel (score)	Gestor (score)	Densidade demográfica	Renda média de domicílios da faixa de interesse	Trabalhadores de interesse	População por faixa etária de interesse
#17	0	7,35	5	5.828,42	10.927	15.857	50.938
#18	0	6,6	5	15.215,60	30.838	173.495	151.396
#19	1	6,6	4	3.560,23	18.762	12.980	31.513
#20	1	9,5	7	4.196,20	9.799	8.474	38.375
#21	0	6,6	5	5.759,81	24.329	57.563	55.246
#22	1	8,6	5	1.121,68	12.187	928	10.121
#23	0	8,45	4	8.134,73	18.394	30.282	75.919
#24	1	7,65	10	30.019,40	21.683	276.182	285.575
#25	0	5,05	4	2.822,09	8.150	4.106	21.892
#26	1	8,9	9	11.707,29	31.426	176.686	115.873
#27	0	8,35	5	5.088,71	10.719	15.543	46.877
#28	1	8,65	5	18.036,50	30.190	209.300	181.168
#29	1	8,55	9	4.064,94	17.818	13.745	37.490
#30	1	8,35	5	8.279,49	14.698	27.858	77.160
#31	0	8,58	4	4.156,74	11.506	6.689	34.733
#32	1	7,95	5	15.591,71	29.169	60.169	154.975

R2000m							
# Unidade	Sucesso	Imóvel (score)	Gestor (score)	Densidade demográfica	Renda média de domicílios da faixa de interesse	Trabalhadores de interesse	População por faixa etária de interesse
#33	1	7,9	5	15.524,04	29.921	85.951	153.626
#34	1	5,6	5	6.038,17	14.032	7.244	52.545
#35	1	9	10	13.948,56	22.013	25.039	138.388
#36	0	7,45	5	6.519,46	11.090	17.351	59.203
#37	0	7,55	4	24.644,17	25.865	182.887	240.393
#38	0	8,95	5	20.501,04	26.868	92.394	197.304
#39	1	8,45	8	14.975,78	28.417	144.500	149.904
#40	0	8,15	3	4.306,43	14.717	16.734	42.269
#41	1	7,93	8	4.088,91	17.996	5.711	38.033
#42	1	7,65	9	18.400,03	26.799	66.195	177.745
#43	0	7,98	5	8.234,85	27.190	111.463	69.411
#44	1	8,35	8	16.720,86	18.926	19.176	155.412
#45	1	8,8	8	4.014,19	16.538	6.739	36.346
#46	1	9,48	9	11.896,29	19.501	45.600	115.578
#47	0	4,55	5	11.905,04	22.621	54.173	111.741
#48	1	9,25	6	9.613,15	19.151	32.307	90.432

R2000m							
# Unidade	Sucesso	Imóvel (score)	Gestor (score)	Densidade demográfica	Renda média de domicílios da faixa de interesse	Trabalhadores de interesse	População por faixa etária de interesse
#49	0	7,45	4	16.315,03	16.661	26.570	159.687
#50	0	7,78	5	1.659,44	10.695	2.043	14.753
#51	1	9,35	6	10.865,70	16.419	42.267	100.861
#52	0	6,9	3	16.356,32	20.836	67.578	154.019
#53	0	9,85	5	3.347,31	18.962	29.692	26.936
#54	1	8,15	6	678,53348	17.944	2.686	5.488
#55	0	7,38	4	2.732,39	10.374	1.317	24.371
#56	1	8,53	5	12.609,64	21.830	55.570	117.223
#57	1	8,35	5	4.034,10	11.925	5.465	34.448
#58	1	8,9	9	8.439,05	25.096	44.254	78.475
#59	1	8	9	13.747,09	26.766	92.546	137.021
#60	1	7,95	6	18.605,93	25.239	83.172	179.946
#61	0	8,25	4	8.908,45	22.377	25.633	81.680

**Apêndice A3 - Valores para modelo misto R1000m e R2000m**

R1000m e R2000m							
# Unidade	Sucesso	Imóvel (score)	Gestor (score)	Densidade demográfica	Renda média de domicílios da faixa de interesse	Trabalhadores de interesse	População por faixa etária de interesse
#1	0	7,68	3	11.046,52	20.986	44.348	107.291
#2	1	7,45	5	7.525,48	19.449	7.441	67.592
#3	0	7,3	4	8.864,08	23.130	3.858	20.399
#4	1	7,05	4	5.978,43	31.458	18.239	59.774
#5	1	8,95	10	11.343,27	19.156	41.378	112.075
#6	0	7,93	4	17.096,21	19.770	114.959	166.532
#7	1	9,18	5	14.446,62	15.510	172.244	130.607
#8	0	7,45	2	4.543,48	12.870	1.196	10.574
#9	1	7,7	7	3.252,12	13.093	908	7.347
#10	0	7,55	4	3.367,91	21.980	81.556	32.297
#11	1	8,95	7	10.905,76	16.321	48.771	108.841
#12	0	9,05	5	14.358,86	28.147	54.446	141.450
#13	1	8,58	6	7.777,92	13.263	3.261	18.595
#14	1	7,35	8	11.480,92	12.535	9.543	28.464
#15	1	8,25	7	18.325,56	24.292	73.035	176.208

R1000m e R2000m							
# Unidade	Sucesso	Imóvel (score)	Gestor (score)	Densidade demográfica	Renda média de domicílios da faixa de interesse	Trabalhadores de interesse	População por faixa etária de interesse
#16	0	8,6	6	7.416,39	25.323	98.839	71.246
#17	0	7,35	5	7.717,19	12.071	5.963	17.315
#18	0	6,6	5	15.215,60	30.838	173.495	151.396
#19	1	6,6	4	3.560,23	18.762	12.980	31.513
#20	1	9,5	7	5.579,93	10.534	2.510	13.097
#21	0	6,6	5	13.168,69	26.896	32.405	32.827
#22	1	8,6	5	1.982,76	9.750	404	4.501
#23	0	8,45	4	8.134,73	18.394	30.282	75.919
#24	1	7,65	10	30.019,40	21.683	276.182	285.575
#25	0	5,05	4	2.484,85	8.090	1.192	4.818
#26	1	8,9	9	11.707,29	31.426	176.686	115.873
#27	0	8,35	5	5.628,13	12.265	7.774	13.485
#28	1	8,65	5	18.036,50	30.190	209.300	181.168
#29	1	8,55	9	4.142,96	18.336	3.585	9.405
#30	1	8,35	5	10.479,09	16.252	9.131	24.715
#31	0	8,58	4	4.403,93	11.061	3.560	9.162

R1000m e R2000m							
# Unidade	Sucesso	Imóvel (score)	Gestor (score)	Densidade demográfica	Renda média de domicílios da faixa de interesse	Trabalhadores de interesse	População por faixa etária de interesse
#32	1	7,95	5	15.591,71	29.169	60.169	154.975
#33	1	7,9	5	15.524,04	29.921	85.951	153.626
#34	1	5,6	5	6.557,23	19.105	1.911	15.167
#35	1	9	10	25.520,20	22.595	5.220	63.536
#36	0	7,45	5	12.294,17	11.893	9.569	28.547
#37	0	7,55	4	24.644,17	25.865	182.887	240.393
#38	0	8,95	5	20.501,04	26.868	92.394	197.304
#39	1	8,45	8	14.975,78	28.417	144.500	149.904
#40	0	8,15	3	3.808,07	15.190	4.517	9.480
#41	1	7,93	8	4.088,91	17.996	5.711	38.033
#42	1	7,65	9	18.400,03	26.799	66.195	177.745
#43	0	7,98	5	8.234,85	27.190	111.463	69.411
#44	1	8,35	8	16.720,86	18.926	19.176	155.412
#45	1	8,8	8	5.678,47	12.164	1.354	12.520
#46	1	9,48	9	11.896,29	19.501	45.600	115.578
#47	0	4,55	5	11.905,04	22.621	54.173	111.741

R1000m e R2000m							
# Unidade	Sucesso	Imóvel (score)	Gestor (score)	Densidade demográfica	Renda média de domicílios da faixa de interesse	Trabalhadores de interesse	População por faixa etária de interesse
#48	1	9,25	6	9.159,96	22.224	7.048	21.594
#49	0	7,45	4	22.649,54	16.957	11.323	56.769
#50	0	7,78	5	2.494,55	11.542	859	5.650
#51	1	9,35	6	14.512,95	20.124	15.908	34.664
#52	0	6,9	3	16.356,32	20.836	67.578	154.019
#53	0	9,85	5	3.347,31	18.962	29.692	26.936
#54	1	8,15	6	678,53348	17.944	2.686	5.488
#55	0	7,38	4	4.351,15	11.426	789	10.462
#56	1	8,53	5	12.609,64	21.830	55.570	117.223
#57	1	8,35	5	3.419,20	13.165	1.978	7.637
#58	1	8,9	9	8.439,05	25.096	44.254	78.475
#59	1	8	9	13.747,09	26.766	92.546	137.021
#60	1	7,95	6	18.605,93	25.239	83.172	179.946
#61	0	8,25	4	12.433,76	23.374	6.521	29.808

**APÊNDICE B – NOTAS CONSOLIDADAS DE IMÓVEIS**

# Unidade	Fachada	Personalidade	Acesso	Área treino	Formato	Pavimento	Metragem	Score ponderado
#1	6	7	5,5	10	10	10	10	7,68
#2	4	6	9	10	10	10	10	7,45
#3	7	7	7	7,75	7	10	7	7,3
#4	7	7	6	7,5	5	7	9	7,05
#5	8	8	9	10	10	10	10	8,95
#6	7,5	7,5	9	8	5	7	10	7,93
#7	8,5	7,5	10	10	10	10	10	9,18
#8	7	5	10	7,75	5	10	8	7,45
#9	8	8	10	6,5	5	7	7	7,7
#10	7	7	6	8,75	5	10	10	7,55
#11	10	7	10	8,5	10	10	7	8,95
#12	9	9	10	8,75	10	7	9	9,05
#13	10	8,5	10	7	7	7	7	8,58
#14	6	7	8	8,25	10	7	8	7,35

# Unidade	Fachada	Personalidade	Acesso	Área treino	Formato	Pavimento	Metragem	Score ponderado
#15	9	10	7	7,5	5	7	9	8,25
#16	8	6	10	9,5	10	10	9	8,6
#17	6	6	9	8,25	5	10	9	7,35
#18	5	6	6	8,25	5	10	9	6,6
#19	7	5	9	6	7	7	5	6,6
#20	10	8	10	9,5	10	10	9	9,5
#21	5	5	5	9	10	10	8	6,6
#22	8	10	8	8,75	5	10	10	8,6
#23	8	10	9	8	7	7	9	8,45
#24	7	8	7	8,25	5	10	9	7,65
#25	0	6	5	8,5	10	10	7	5,05
#26	10	9	7	8,75	10	7	9	8,9
#27	6,5	6	10	10	10	10	10	8,35
#28	10	9	8	7,75	10	7	7	8,65
#29	9	9	8	8,25	10	5	9	8,55
#30	9	7	8	8,5	10	10	7	8,35

# Unidade	Fachada	Personalidade	Acesso	Área treino	Formato	Pavimento	Metragem	Score ponderado
#31	7	7	9,5	10	10	10	10	8,58
#32	8	8	9	7,5	5	7	9	7,95
#33	8	6	10	7,75	10	7	7	7,9
#34	3	5	7	7,25	5	10	7	5,6
#35	10	10	10	7,5	5	7	9	9
#36	7	7	8	7,75	5	10	8	7,45
#37	8,5	6	10	6,5	5	5	8	7,55
#38	10	9	10	7,75	10	7	7	8,95
#39	8	8	9	8,75	10	7	9	8,45
#40	6,5	8	10	8,75	10	7	9	8,15
#41	7	8,5	7	8,75	5	10	10	7,93
#42	7	7	8	8,25	5	10	9	7,65
#43	7	7,5	9	8,5	7	7	10	7,98
#44	6	10	7	10	10	10	10	8,35
#45	10	8	10	7,75	10	7	7	8,8
#46	9	8,5	10	10	10	10	10	9,48

# Unidade	Fachada	Personalidade	Acesso	Área treino	Formato	Pavimento	Metragem	Score ponderado
#47	1	5	0	8,75	5	10	10	4,55
#48	9,5	9	9	9,25	10	7	10	9,25
#49	6	5	10	8,5	10	10	7	7,45
#50	6	8,5	8	8,75	7	10	9	7,78
#51	10	7	10	9,5	10	10	9	9,35
#52	5	7	9	7,5	10	10	5	6,9
#53	10	9	10	10	10	10	10	9,85
#54	8	8	7	8,75	10	7	9	8,15
#55	7	6	8,5	7,75	10	7	7	7,38
#56	8	7,5	10	8,75	10	7	9	8,53
#57	9	9	10	7	7	7	7	8,35
#58	9,5	9	8	8,75	5	10	10	8,9
#59	8	9	7	8	5	7	10	8
#60	8	7	8	8,25	5	10	9	7,95
#61	8,5	8	10	7,5	5	7	9	8,25

