

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO SOCIOECONÔMICO
DEPARTAMENTO DE CIÊNCIAS DA ADMINISTRAÇÃO**

Tiago Lamêu da Silva

**Gestão de dados em startups: percepções de gestores acerca de sua importância em seu
sucesso**

Florianópolis

2022

Tiago Lamêu da Silva

Gestão de dados em startups: percepções de gestores acerca de sua importância em seu sucesso

Trabalho de Curso apresentado à disciplina CAD 7305 como requisito parcial para a obtenção do grau de Bacharel em Administração pela Universidade Federal de Santa Catarina.

Enfoque: Monográfico – Artigo

Área de concentração: Administração da tecnologia da informação

Orientador(a): Prof. Dr. Alexandre Marino Costa

Florianópolis

2022

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

Lamêu da Silva, Tiago

Gestão de dados em startups: percepções de gestores
acerca de sua importância em seu sucesso / Tiago Lamêu da
Silva ; orientador, Alexandre Marino Costa, 2022.
26 p.

Trabalho de Conclusão de Curso (graduação) -
Universidade Federal de Santa Catarina, Centro Sócio
Econômico, Graduação em Administração, Florianópolis, 2022.

Inclui referências.

1. Administração. 2. Análise de dados. 3. Startups. I.
Marino Costa, Alexandre. II. Universidade Federal de Santa
Catarina. Graduação em Administração. III. Título.

Tiago Lamêu da Silva

Gestão de dados em startups: percepções de gestores acerca de sua importância em seu sucesso

Este Trabalho de Curso foi julgado adequado e aprovado na sua forma final pela Coordenadoria Trabalho de Curso do Departamento de Ciências da Administração da Universidade Federal de Santa Catarina.

Florianópolis, 25 de Julho de 2022.

Prof. Ana Luiza Paraboni
Coordenador de Trabalho de Curso

Avaliadores:

Prof. Alexandre Marino Costa, Dr.
Orientador
Universidade Federal de Santa Catarina

Prof^a. Andressa Sasaki Vasques Pacheco, Dra.
Avaliadora
Universidade Federal de Santa Catarina

Prof^a. Gabriela Mattei de Souza, Dra.
Avaliadora
Universidade Federal de Santa Catarina

2022
RESUMO

A análise de dados é um processo que visa entender o comportamento dos dados, seja pelo histórico ou por tendências. É um processo considerado importante para a identificação de problemas e o crescimento de empresas. Ao trazer o tema para empreendimentos, nota-se uma possibilidade de implementação considerável a fim de reduzir uma possível falência precoce, já que é um tipo de empresa muito relevante no Brasil. Dentre essas empresas, encontram-se as *startups*, que, por serem de base tecnológica, tendem a ser mais habituadas a se preocuparem com dados. A partir disso, o presente trabalho tem como objetivo elencar dados para medir o sucesso de *startups* na Associação Catarinense de Tecnologia (ACATE). Ressaltam-se como limitações o período, já que foi feito no segundo semestre de 2021 ao fim do primeiro semestre de 2022, a região, pois foram entrevistados gestores da ACATE em Florianópolis e os entrevistados, pois a amostragem foi feita a partir da disponibilidade dos entrevistados. As entrevistas foram feitas de forma remota e semiestruturadas, e, com as respostas, foi possível definir o que é sucesso de uma *startup*, especificar os dados necessários e propor uma estrutura de dados para medir esse feito.

Palavras-chave: *Startup*, Análise de dados, Tecnologia

1 INTRODUÇÃO

Segundo o Ministério da Economia, em 2020, os empreendimentos representam 99% dos negócios brasileiros, têm participação de 30% no PIB e são responsáveis por 55% dos empregos gerados no país. Apesar da grande relevância no cenário nacional, as micro e pequenas empresas sofrem desafios para não chegarem a seu período de mortalidade precoce. De acordo com Chiavenato (2008, p. 15), existem muitos riscos e perigos que levam os novos negócios a falência prematura, e isso é refletido, no Brasil, ao levar em conta que 23,4% das empresas decretam falência antes de completar dois anos de existência, segundo o relatório de sobrevivência das empresas no Brasil, feito pelo Sebrae.

Dentro do ramo de empreendedorismo, têm-se as chamadas *startups*, que são, segundo a Associação Brasileira de Startups (Abstartups), empresas que nascem a partir de um modelo de negócio ágil e enxuto, capaz de gerar valor para seu cliente resolvendo um problema real, do mundo real; oferecem uma solução escalável para o mercado e, para isso, usa tecnologia como ferramenta principal. Segundo os dados do Google for Startups no Brasil, em agosto de 2021, o país passou de cinco mil para mais de treze mil startups, de 2016 a 2021. Por se tratar de um modelo crescente e que é aplicado a diversas áreas da sociedade, é importante identificar os fatores remetentes a esse tipo de empresa para que, dessa forma, possa ser identificada tendência de crescimento ou não.

Com base nessas informações, são feitas entrevistas para coleta de dados, respondidas por quatro gestores de *startups* da Associação Catarinense de Tecnologia (ACATE), a fim de analisar como eles enxergam a importância e utilidade dos dados para o desenvolvimento de suas empresas.

A partir disso, ao levar em conta respostas levantadas por gestores de startups da ACATE, questiona-se: qual a importância dos dados para que uma startup tenha sucesso?

1.1 Objetivos

Tendo em vista os pontos destacados de forma introdutória, faz-se necessário estabelecer os objetivos com o presente Trabalho de Conclusão de Curso.

1.1.1 Objetivos gerais

Preliminarmente, à título de objetivo de caráter geral, ou seja, inerente à todas as etapas do projeto, têm-se a tentativa de definir se há importância de dados para o sucesso de uma

startup e, caso positivo, elencar os principais dados para fazer a medição.

1.1.2 Objetivos específicos

Após, ao tomar como base o constatado a partir do objetivo geral, será feita uma contextualização acerca dos objetivos específicos do projeto. Inicialmente, houve a formulação de uma definição clara de sucesso e indicadores imprescindíveis para esse feito de *startups*. Depois, o raciocínio foi a partir de uma análise desses fatores, a fim de obter uma estrutura capaz de detectar correlações e tendências deles. Por fim, é proposta essa estrutura de *dataframe* a ser utilizada para *startups* medirem seu desempenho.

1.2 Justificativa

Estabelecidos os objetivos do presente projeto, cabe demonstração a respeito de seus motivos justificadores.

De início, cabe discorrer acerca da importância das startups no contexto brasileiro. Observa-se que estas possuem grande capacidade de "girar" a economia, conquistando o horizonte do empreendedorismo digital não apenas brasileiro, mas, inclusive, mundial. Isso é possível graças aos "hubs de inovação", os quais podem ser definidos como um local em que as startups possuem possibilidades de desenvolvimento de ideias (ARCENALES, SENA e ARAUJO, 2021). O Brasil é reconhecido por ser possuidor de diversos hubs de inovação, o que faz com que o ecossistema das startups represente grande papel na economia interna e externa do país.

Todavia, não se pode prestar atenção apenas no cenário econômico: os empreendedores possuem o compromisso de melhorar a educação do país para que a sociedade como um todo possa também ser melhorada (ARCENALES, SENA e ARAUJO, 2021). As startups, desse modo, devem sua importância a sua inovadora capacidade de interferirem, de modo positivo, no contexto educacional, a partir do incentivo a investimentos no ramo, com vistas à educar uma sociedade mais justa e igualitária.

Nesse contexto promissor, percebe-se, na sociedade pós-moderna, uma crescente importância atribuída aos dados utilizados na tomada de decisão pelos empreendedores. Isso porque é a partir dos dados e informações coletados previamente que torna-se possível aos decisores identificar o problema em questão a ser solucionado, tendo em vista a tomada de

decisão ser primordial para estabelecer potenciais alternativas de solução, a partir de uma ordem de preferência (BUENO, AZEVEDO, 2011).

Nota-se, neste giro, que a ausência de uma análise feita de modo embasado a respeito do ambiente de mercado e, portanto, voltada aos clientes e à possível concorrência, é capaz de levar pequenas organizações a um fim precoce (ROCHA, CARVALHO e FILHO, 2020).

Outro fator justificativo da alta mortalidade das startups diz respeito aos aspectos do ambiente no qual estas estão inseridas: a estrutura das empresas, determinada no momento de sua concepção, a falta de escolaridade adequada, de exemplos no cenário do empreendedorismo e de conhecimentos específicos na área de gestão, são fatores capazes de encurtar significativamente sua expectativa de vida (NOGUEIRA, ARRUDA, 2014). Nogueira e Arruda ainda destacam que, para que tal realidade possua menos chances de se concretizar, o empreendedor de startup possui o dever de se atentar a três aspectos essenciais: ao número de sócios envolvidos, ao volume de capital investido, e ao local onde a empresa será instalada (2014).

Diante dessa realidade, startups capazes de gerar novas oportunidades, a partir do contexto mercadológico, são as mais procuradas pelos fundos de investimentos brasileiros, já que estes buscam apoiar o trabalho inicial de validação e consequente desenvolvimento de negócio (GONÇALVES, 2021). Para estes fundos, é imprescindível a existência de grandes evidências de segurança nos possíveis investimentos a serem realizados nas startups, o que só pode ser devidamente garantido mediante a verificação da capacidade de geração de novas oportunidades.

Por fim, destaca-se que a efetivação de abordagens acerca da análise de dados como etapa do processo de investigação científica no campo das organizações denota grande relevância (TEIXEIRA, 2003). Isso porque, ao discriminar especificidades, pode-se gerar melhorias nas condições de desenvolvimento de novos estudos, tomando como base um entendimento mais efetivo e conceitual do processo.

Nesse viés, o presente trabalho possui como intuito a identificação dos fatores considerados como indicadores de sucesso de uma startup, a ser feito a partir da visão de gestores dessa modalidade de empresa. Este intuito será realizado para que torne-se possível a aplicação de um modelo de base de dados a fim de medir o progresso das startups com base em uma análise estatística e com possibilidades de análises preditivas.

2. FUNDAMENTAÇÃO TEÓRICA

De início, deve-se realizar conceitualização teórica acerca de diversos conceitos, os quais se mostram relevantes para que se possa verificar a importância dos dados para o sucesso de uma startup.

2.1 Empreendedorismo

O empreendedorismo, mais comumente definido como a “*arte de fazer acontecer com criatividade e motivação*” (BAGGIO, 2015), representa o prazer advindo de realizar projetos (pessoais ou organizacionais) por meio de ideias inovadoras. Permite-se assim, o desafio aos fatores de oportunidade e riscos, tornando possível que indivíduos tenham um comportamento proativo, tomando a dianteira em questões que precisam de resolução.

Baggio, QUEM É, ainda, define o empreendedorismo como o “*despertar*” do indivíduo, fazendo com que o aproveitamento total de suas capacidades racionais e intuitivas seja exequível. As atividades empreendedoras, neste giro, representam a busca pelo autoconhecimento, efetivada a partir de um processo permanente de aprendizado (BAGGIO, 2015).

A partir disso, entende-se que o empreendedor de uma *startup* adota como principal objetivo a capacidade de originar um novo negócio. Para tanto, ele efetiva uma análise concisa do cenário, o que faz com que não deixe passar oportunidades significativas, apresentando um novo empreendimento. O empreendedor de uma startup, contudo, apresenta certos desafios (apud de Pessoa, 2005, em BAGGIO, 2015): buscar e apresentar diferenciais competitivos em um mercado já existente; vencer a concorrência; conquistar clientes; e alcançar a lucratividade e a produtividade necessárias à manutenção do empreendimento.

Nota-se, mediante a conceitualização apresentada, que atividade empreendedora possui extremo valor no contexto das startups. Isso porque, a partir de ideias inovadoras e atitudes desafiadoras frente a oportunidades e riscos, o empreendedor torna possível o crescimento exponencial de uma startup, contornando os desafios que mostram-se ao longo do caminho.

2.1.1 Startups

Startup, nesse contexto, deve ser compreendida como uma instituição destinada a criar novos produtos e/ou serviços em um cenário de incertezas (RIES, 2012, *tradução livre*). Ainda, são empresas que se apresentam em fase embrionária, geralmente efetivando o processo de

implementação e organização de suas operações. Apesar de não haver obrigatoriedade quanto ao início da comercialização de seus produtos e/ou serviços, para que seja considerada uma *startup*, a empresa embrionária deve estar em funcionamento (TABORDA , 2006).

De acordo com dados fornecidos pelo SEBRAE, as *startups* possuem, ainda, outra peculiaridade: optam pela busca de novos modelos de negócios, refletindo e representando a velocidade das mudanças em um mundo globalizado, o que é capaz de influenciar de modo positivo a construção de novos conhecimentos e o seu desenvolvimento econômico (SEBRAE, 2012, p. 5).

Cabe ressaltar que os conceitos colacionados acima são resgatados da conceitualização proposta pela Abstartups, previamente apresentada na introdução do presente trabalho. Nesse viés, as startups são definidas como empresas que surgem de um modelo de negócio ágil e enxuto. Tal modelo é capaz de gerar valor para os clientes, o que torna possível a resolução de problemas do mundo real.

2.2 Dados

Em relação aos dados, estes devem ser entendidos como todo objeto que tenha sido criado em formato digital (*digital born*, no inglês) ou que tenha sido posteriormente convertido para formato digital (digitalizado). Para ser considerado como “dado”, o objeto deve permitir seu uso para a geração dos chamados “*insights de conhecimento*” (SEMELLER et al, 2017).

De modo análogo, os dados podem ser compreendidos como fatos passíveis de serem gravados, possuidores de significado implícito (ELMASRI & NAVATHE, 2011). Podem ser divididos em dois tipos principais: quantitativos e qualitativos. O primeiro grupo representa os dados que podem ser reproduzidos por valores numéricos, além de poderem ser discretos e contínuos. Os dados qualitativos, por outro lado, são os que contêm valores nominais e ordinais, ou seja, categóricos (CAMILO, 2009).

Neste giro, uma coleção de dados relacionados é denominada de banco de dados, que possui certas propriedades implícitas: representa determinados aspectos do mundo real, podendo ser chamado de “universo de discurso” (UoD), cujas mudanças podem ser refletidas em um banco de dados; é formado por uma coleção lógica e coerente de dados que possuem certo significado inerente; é projetado, construído e povoado por dados, de modo a atender uma proposta específica e possuindo grupo de usuários definido e certas aplicações preconcebidas

(ELMASRI & NAVATHE, 2011).

Por fim, outra conceitualização importante diz respeito ao denominado “*dataframe*”. Na prática, observa-se que os bancos de dados podem ser armazenados em um *dataframe*, que nada mais é do que um objeto de duas dimensões, onde cada coluna pode possuir um tipo primário diferente (SITE R).

2.3 Análise de dados

De início, cabe destacar que a Análise Envoltória de Dados (DEA) pode ser compreendida como uma técnica que, por meio do uso de programação matemática, é capaz de construir fronteiras no contexto da produção de unidades produtivas (DMUs), as quais utilizam processos de características tecnológicas para transformar diversos insumos em produtos (CASADO, 2007).

2.3.1 Análise descritiva

“A descrição dos dados tem como objetivo básico resumir uma série de valores de mesma natureza através de um conjunto de ferramentas e técnicas: tabelas, gráficos, medidas (estatísticas) de variabilidade e de tendência central que ajudam na produção de uma visão global dos dados” (MANCUSO, 2018)

Em relação à análise descritiva, ressalta-se que as ferramentas descritivas dizem respeito aos variados tipos de gráficos, tabelas e medidas de síntese, tais como a porcentagem, os índices e as médias (REIS, 2002). Vale ressaltar que, no processo de condensamento de dados, algumas informações podem ser perdidas, devido ao fato de as observações originais não serem mantidas. Entretanto, se comparado com os ganhos que podem ser obtidos graças a clareza de interpretação, a perda de informação é pequena.

O processo de descrição de dados possui como objetivo, também, discriminar anomalias que podem, inclusive, advir do registro incorreto de valores e dados mal inseridos (que não seguem a tendência do restante do conjunto). Logo, utiliza-se medidas de tendência central, capazes de representar o valor típico de determinada variável, medidas de variabilidade, que tratam das variações de valores que ocorrem em uma mesma variável, medidas de posição, as quais representam a posição em relação ao total, e o *boxplot*, nome dado aos gráficos capazes de detectar valores que são considerados discrepantes (REIS, 2002).

2.3.2 Análise preditiva

Sobre a análise preditiva, esta visa descobrir um valor posterior de uma determinada variável. As predições podem, nesse contexto, serem numéricas ou de variáveis discretas. As primeiras servem para a previsão de valores de variáveis contínuas, enquanto que, para as segundas, pode-se aplicar as técnicas de classificação já apresentadas.

Ainda a respeito da análise preditiva, destaca-se que as técnicas de regressão são as responsáveis por realizar a modelagem do contato de variáveis preditoras (ou independentes), que dizem respeito aos atributos de determinados registros, com variáveis resposta (ou dependentes), que se referem ao que se pretende prever (CAMILO, 2009).

2.3.3 Métodos de análise preditiva

As análises preditivas comportam dois importantes métodos: as regressões lineares e as não-lineares. As regressões lineares são aquelas em que a relação entre variáveis preditoras e variáveis resposta seguem um comportamento denominado de linear, o que torna possível a criação de modelos onde “y” representa uma função linear de “x”. As regressões não-lineares, por sua vez, possui um comportamento não-linear, em que a relação entre as variáveis pode ser determinada a partir de uma função polinomial, por exemplo, as quais podem ser convertidas para regressões lineares (CAMILO, 2009).

Nesse sentido, cumpre destacar estudos feitos sobre predição de sucesso de startups (RODRIGUES et al, 2021):

Imagem 1 - Estudos citados por Rodrigues

Tabela 2 - Predição do sucesso de *startups*: síntese de estudos relacionados

Referência	Objetivo	Amostra	Variável de Interesse	Técnicas de <i>Machine Learning</i>
Liang e Daphne Yuan (2013)	Investigar o papel das relações sociais entre investidores e empresas para a predição do comportamento de investimento.	11.916 <i>startups</i> , 12.127 pessoas, 1.122 organizações financeiras	Ocorrência de investimento	<i>SVM</i> , <i>Árvore de Decisão</i> e <i>Naive Bayes</i>
Shan, Cao e Lin (2014)	Prever se um investidor investirá em uma <i>startup</i> específica com base em sinais textuais, topológicos e específicos do domínio	214.290 <i>startups</i> , 286.659 pessoas, 31.942 investimentos	Investidor efetiva investimento em <i>startup</i>	Regressão Logística
Bento (2017)	Desenvolver um modelo preditivo para classificar uma <i>startup</i> como bem-sucedida, ou não (classificação binária)	86588 <i>startups</i> (estados americanos)	<i>IPO</i> ou Aquisição da <i>startup</i>	Floresta Aleatória , <i>SVM</i> e Regressão Logística
Pan, Gao e Luo (2018)	Predizer o sucesso de <i>startups</i> , definido como um evento que dá uma grande quantia aos fundadores e investidores	+60.000 <i>startups</i>	Processo de <i>M&A</i> ou <i>IPO</i>	KNN , Floresta Aleatória e Regressão Logística
Arroyo et al. (2019)	Desenvolvimento e avaliação de uma abordagem orientada por dados que usa <i>machine learning</i> para ajudar os investidores de <i>VC</i> a explorar e selecionar as melhores empresas para apoiar.	120.507 <i>startups</i> , 34.180 <i>funding rounds</i>	Aquisição, rodada de financiamento, <i>IPO</i> , fechamento ou nenhum evento	Gradient Tree Boosting , <i>Árvore de Decisão</i> , Floresta Aleatória e <i>SVM</i> .
Gastaud, Carniel e Dalle (2019)	Predizer o sucesso de <i>startups</i> na arrecadação de investimentos em diferentes estágios (<i>early</i> -, <i>growth</i> - e <i>late stage</i>)	65.957 <i>startups</i>	Obtenção de <i>funding</i> em diferentes estágios (<i>seed</i> , <i>series A</i> e <i>B</i>)	Floresta Aleatória, <i>Graph Neural Networks</i>

Fonte: Elaborado pelos autores (2021).

Fonte: RODRIGUES et al, 2021

2.3.4 Aplicação

Nesse viés, cabe ressaltar que as análises descritivas e preditivas podem ser feitas pelo software “R”. Este pode ser definido como uma linguagem utilizada para computação de dados estatísticos e gráficos.

O *software R* é capaz de disponibilizar uma ampla variedade de técnicas estatísticas (modelos lineares e não-lineares, testes estatísticos clássicos, classificações e etc) e gráficas, além de ser altamente adaptável (SITE R). Ele torna possível que se realizem análises por meio de *machine learning*, considerado um dos principais ramos atuais da Inteligência Artificial. Isso porque, graças aos algoritmos de *machine learning*, tem-se uma retenção de grandes volumes de dados, os quais podem ou não ser estruturados, e, conseqüentemente, uma identificação de padrões para regressão ou classificação (RODRIGUES et al, 2021).

2.4 Informação

Pode-se definir informação como um conjunto de representações de características mentais e/ou emocionais. Estas representações são, frequentemente, codificadas por meio de

sinais e símbolos, o que permite que sejam modeladas pela interação social. Neste viés, a informação pode ser registrada em qualquer formato, seja ele físico ou digital, podendo, por consequência, ser comunicada de modo assíncrono e multidirecional interativo (SILVA, RIBEIRO, 2012).

Por fim, a informação emerge sempre que determinados fatos são inseridos em um certo contexto, onde são combinados com uma estrutura. É a partir da interpretação que se dá significado para a informação, o que faz com que ela se transforme em conhecimento (TUOMI, 1999). Ainda segundo Tuomi (1999), para que a informação seja criada, é necessário que primeiro haja conhecimento acerca de determinado tema: os dados surgem como um subproduto de artefatos cognitivos, que são capazes de presumir a existência de práticas, compartilhadas socialmente, de uso de tais artefatos.

3. PROCEDIMENTOS METODOLÓGICOS

Nesta etapa, serão apresentados os procedimentos metodológicos utilizados para o alcance dos objetivos específicos. Este capítulo é composto por pontos que definem os meios responsáveis pela execução da pesquisa, e serão levados em conta a natureza da pesquisa, delimitação da pesquisa, técnica de coleta de dados, análise e tratamento de dados e resultados esperados.

3.1 Natureza da pesquisa

A fim de caracterização da pesquisa, tem-se o conceito de pesquisa aplicada que, segundo Silveira e Córdova (2009), se trata de uma pesquisa cujo objetivo é gerar conhecimentos para aplicação prática, dirigidos à solução de problemas específicos, e que envolve verdades e interesses locais. A partir desse conceito, pode-se afirmar que a presente pesquisa compartilha dessa natureza, uma vez que seus objetivos remetem a construção de um modelo aplicável a um contexto específico.

Acerca dos objetivos gerais e específicos apresentados nesta pesquisa, podemos classificá-la como exploratória. Segundo Gil (2007), a pesquisa exploratória se dá por aquela que objetiva trazer maior familiaridade com um problema para, assim, torná-lo mais explícito ou proporcionar a criação de hipóteses. O presente trabalho se encaixa nessa definição, pois tem o intuito de explicitar o ponto de vista dos gestores acerca dos dados coletados em suas empresas e, dessa forma, é criada proximidade a respeito do tema.

De acordo com Rodrigues (2007), a pesquisa descritiva se caracteriza pela observação, registro, análise, classificação e interpretação de dados, de forma que o pesquisador não interfira nesse processo. Além disso, são utilizadas técnicas padronizadas de coleta de dados. A partir dessa definição, essa pesquisa é considerada descritiva, uma vez que será aplicado um questionário padrão sem a intervenção do pesquisador.

Ademais, a pesquisa terá um teor qualitativo, uma vez que os dados coletados serão elementos de cunho conceitual, que caracterizariam uma pesquisa qualitativa.

3.2 Delimitação da Pesquisa

A fim de trazer clareza ao presente trabalho, deve-se estabelecer os limites da pesquisa a ser realizada. Desse modo, a respeito desta, será feita a definição acerca de seu grupo componente, sujeitos e a região em que será aplicada.

Quanto ao grupo de pesquisa, este será definido por empresas que impactam no setor de tecnologia, enquanto os sujeitos de pesquisa serão colaboradores que exercem funções estratégicas nas empresas em questão. O estudo será focado na região da ACATE, em Florianópolis.

O modelo de análise da presente pesquisa foi pautado na descrição das respostas, obtidas a partir de uma amostragem por acessibilidade, ou seja, por quatro gestores que estavam disponíveis para responder às perguntas. Além disso, foi feita uma filtragem para considerar apenas empresas que se encaixam no conceito de *startups*, ou que passaram por esse modelo.

3.3 Técnica de Coleta e Análise de Dados

Para a coleta de dados, foram feitas entrevistas de forma remota, com a intenção de obter uma grande especificidade de respostas cujos dados primários serão coletados a partir dela. As perguntas em questão foram separadas em três blocos: identificação do perfil profissional, identificação do perfil de empresa e informações sobre relevância e tipos de dados a serem analisados.

As perguntas do primeiro bloco foram:

- Qual o seu cargo na empresa onde faz parte?

- Há quanto tempo está na empresa?

As perguntas do segundo bloco foram:

- Qual o ramo da empresa que você faz parte?
- Qual a idade da empresa?

As perguntas do terceiro bloco foram:

- Considera que o processo de análise de dados é essencial para o sucesso de uma startup?
- O que é considerado para uma startup ter sucesso, na sua opinião?
- Levando em conta que parâmetros são itens que compõem indicadores, por exemplo receitas e despesas, que compõem lucro, quais parâmetros você acha importante medir para o sucesso de uma startup?
- A partir disso, quais os indicadores você acha importante medir para o sucesso de uma startup?

A partir disso, resgata-se os objetivos específicos da seguinte forma:

A - Formulação de uma definição clara de sucesso e indicadores imprescindíveis para crescimento de startups

B - Obtenção de uma estrutura capaz de detectar correlações e tendências entre os indicadores

C - Proposição de estrutura de *dataframe* a ser utilizada para *startups* medirem seu desempenho frente ao mercado

Dessa forma, relaciona-se os blocos com os objetivos conforme mostrado na tabela abaixo:

Quadro 1: Relação entre bloco de perguntas e objetivos

Bloco	Objetivo
1	A
2	A
3	B e C

Fonte: elaborada pelo autor

As perguntas que compõem os blocos 1 e 2 visam identificar o perfil dos gestores e empresas presentes na pesquisa. Já as perguntas do bloco 3 têm como objetivo identificar, conceitualmente, o que é sucesso no contexto de uma startup e os parâmetros e indicadores para medir esse feito, de forma que possam ser organizados como um dataframe passível de análises descritivas e preditivas.

3.5 Limitações da Pesquisa

A pesquisa foi feita a partir da seleção de gestores disponíveis, levando em conta apenas empresas presentes na ACATE de Florianópolis, na região do Saco Grande, feito em sua totalidade entre junho de 2021 a julho de 2022, denso as entrevistas feitas entre maio de 2022 e julho de 2022.

4. ANÁLISE DE RESULTADOS

A partir das respostas dadas pelos entrevistados, são apresentados quatro tipos de perfis, que variam de hierarquia dos respondentes, tempo presente nas respectivas empresas, assim como idade e ramo das organizações. Além disso, também são enumerados parâmetros que compõem indicadores e indicadores considerados importantes a serem medidos para definir o sucesso de uma *startup*.

4.1 Perfil dos entrevistados e empresas

A tabela abaixo mostra os perfis dos entrevistados, obtidos a partir do primeiro bloco da entrevista:

Tabela 1: Perfil dos entrevistados e das empresas

	Cargo	Ramo da empresa	Cargo na empresa	Idade da empresa
Entrevistado 1	Diretor comercial	Tecnologia	4 anos	4 anos
Entrevistado 2	CEO	Comunicação	6 anos	6 anos
Entrevistado 3	Analista financeiro	Comunicação	2 anos e meio	6 anos
Entrevistado 4	Analista financeiro	<i>Fintech</i>	6 meses	9 meses

Fonte: Elaborada pelo autor

Nota-se que foram selecionados gestores que estão presentes há pelo menos 40% do tempo de empresa, sendo o entrevistado 4 o que possui menos vivência no contexto de sua empresa atual. Vale ressaltar que, além de diretor comercial, o entrevistado 1 é, também, sócio fundador da organização, o que colabora para um ponto de vista empreendedor, assim como o Entrevistado 2.

Referente às empresas, a amostra conta com duas ocorrências de empresas de comunicação, uma de tecnologia e uma de *fintech*, esta que, segundo o entrevistado, é um ramo financeiro atrelado à tecnologia. Seus produtos são, respectivamente, soluções para comunicação empresarial, soluções de inteligência de dados e gestão de fundos na *blockchain*.

4.2 Importância e descrição de dados

Após serem perguntados sobre a importância de dados para o crescimento de uma *startup*, todos os entrevistados afirmaram que é um processo imprescindível. Apesar disso, o entrevistado 1 acrescentou que é importante para a agilidade na tomada de decisão, mas que, no início, deve-se atentar à qualidade das informações, pois podem estar enviesadas pela oscilação do início da trajetória da organização. O entrevistado 3 também acrescentou que a análise de dados traz um panorama explícito geral da empresa, capaz de identificar, de forma

mais precisa, quais aspectos mudar para corrigir determinados gargalos, conforme o contexto.

Para o entrevistado 1, a definição de sucesso para uma *startup* é fazer com que seu crescimento e relevância sejam responsáveis para que uma empresa maior a compre. Para o entrevistado 2, o sucesso se dá pela sua consolidação no mercado e otimização em resolver os problemas os quais ela se propôs, corroborado pelo entrevistado 3, que acrescentou a importância de ter uma base tecnológica. Já o entrevistado 4 resumiu o sucesso de uma *startup* em três pilares: faturamento, adoção e inovação. Em suma, todas as respostas consideram o crescimento de autoridade como sucesso de uma *startup* até o ponto que ela deixe de ser esse tipo de empresa.

Quanto aos parâmetros indicados para a medição de sucesso de uma *startup*, foram coletadas as seguintes respostas:

Tabela 2: Parâmetros e indicadores citados pelo entrevistado 1

Entrevistado 1	
Parâmetros	Indicadores
Receita	Novas receitas mensais (NMRR)
Despesa	Resultado acumulado (geração de caixa)
Número de clientes	Churn (cancelamento de clientes) (%)

Fonte: elaborado pelo autor

Tabela 3: Parâmetros e indicadores citados pelo entrevistado 2

Entrevistado 2	
Parâmetro	Indicador
Receita	Novas receitas mensais (NMRR)
Despesa	Resultado acumulado (geração de caixa)
	Margem de lucro
Número de clientes	Ticket médio
	Churn (cancelamento de clientes) (%)

Fonte: elaborado pelo autor

Tabela 4: Parâmetros e indicadores citados pelo entrevistado 3

Entrevistado 3	
Parâmetro	Indicador
Receita operacional	Lucro operacional
Custo operacional	
Investimento em vendas	Custo de aquisição de clientes (CAC)
Número de novos clientes	
Número de clientes	Churn (cancelamento de clientes) (%)
Receita	Novas receitas mensais (NMRR)

Fonte: elaborado pelo autor

Tabela 5: Parâmetros e indicadores citados pelo entrevistado 4

Entrevistado 4	
Parâmetro	Indicador
Receita	Variação de receitas (%)
Despesa	Resultado acumulado (geração de caixa)
Número de clientes	Ticket médio
	Variação de clientes (%)

Fonte: elaborado pelo autor

Nota-se que a receita é um parâmetro unânime entre os respondentes. Além disso, os parâmetros e indicadores mostraram uma preocupação no ponto de vista geral da empresa, ao observar dados majoritariamente financeiros.

Ao compararmos os parâmetros e indicadores, nota-se coerência nas respostas, uma vez que os parâmetros realmente fazem parte dos indicadores. Isso traz a possibilidade de, caso não seja ideal projetar estes diretamente, projetar os que o compõem para trazer dados secundários e, mesmo que com menos confiabilidade, explicitar uma situação projetada.

Com relação às fórmulas, conforme os entrevistados:

- NMRR: não se trata de uma fórmula em si, mas é um indicador que deve ser preenchido

conforme a chegada de novas receitas

- Resultado acumulado: $Resultado\ anterior + (Receita\ atual - despesa\ atual)$
- Churn (%) = $Churn / Clientes\ atuais$
- Margem de lucro = $(Receita\ atual - Despesa\ atual) / Receita\ atual$
- Ticket médio = $Receita\ atual / Número\ de\ clientes$
- Lucro operacional = $Receita\ operacional - Despesa\ operacional$
- Custo de aquisição de clientes = $Investimento\ em\ vendas / Número\ de\ novos\ clientes$
- Variação de receitas (%) = $(Receita\ anterior / Receita\ atual) - 1$
- Variação de clientes (%) = $(Clientes\ anteriores / Clientes\ atuais) - 1$

4.3 Proposta de Dataframe

A partir das informações coletadas, propõe-se um *dataframe* capaz de ser analisado no *software* R. Sua construção se dá por meio de uma planilha no *Google Sheets*, pois é capaz de exportá-la como arquivo .csv, que, por sua vez, pode ser aberto para análise no programa proposto. Sua estrutura será composta pelo seguinte cabeçalho, que, na prática, será composto apenas por uma linha:

Tabela 6: Primeira parte da proposta de *dataframe*

Mês de referência	Receita	Variação de receita (%)	Receita operacional	NMR R	Despesa	Despesa operacional	Investimento em vendas
-------------------	---------	-------------------------	---------------------	-------	---------	---------------------	------------------------

Fonte: elaborado pelo autor

Tabela 7: Segunda parte da proposta de *dataframe*

Lucro operacional	Número de clientes	Número de novos clientes	Ticket médio	CA C	Variação de clientes (%)	Resultado acumulado	Margem de lucro (%)	Churn	Churn (%)
-------------------	--------------------	--------------------------	--------------	------	--------------------------	---------------------	---------------------	-------	-----------

Fonte: elaborada pelo autor

Devido a indicadores que marcam recorrência, como o NMRR, foi escolhido medi-los mensalmente, a fim de obter resultados alinhados a todos os parâmetros e indicadores. Dessa forma, é possível utilizar o R para identificar tendências centrais e visualizar, por meio de análise descritiva, a situação atual da empresa ao exibir indicadores. Além disso, torna-se viável codificar uma tabela de correlação entre esses fatores, tanto quanto fazer testes capazes de identificar tendências passíveis de analisar resultados futuros de determinados parâmetros em função dos demais, ou seja, fazer uma análise preditiva. Conforme o resultado dos testes, é possível desenvolver modelos de *machine learning* para explicitar projeções estatisticamente embasadas.

5. CONSIDERAÇÕES FINAIS

A fim de tentar definir se há importância de dados para o sucesso de uma *startup*, este trabalho aborda o tema análise de dados sob a perspectiva gerencial de empresas da ACATE, o que trouxe esclarecimento acerca de dados a serem medidos para identificar o sucesso desse tipo de empresa.

Dito isso, o objetivo geral é considerado contemplado, uma vez que, no ponto de vista dos gestores consultados, a análise de dados se mostrou importante e os dados fundamentais para uma *startup* ser considerada bem sucedida foram devidamente elencados.

Quanto aos objetivos específicos, consideram-se atendidos. O primeiro se trata da formulação de definição clara de sucesso e indicadores essenciais para o crescimento de startups, o que foi contemplado de forma que, em resumo, foram considerados relevância no mercado e transição de startup para uma grande empresa os fatores de sucesso, assim como foram elencados indicadores, majoritariamente financeiros, para o prosseguimento da pesquisa.

Para o segundo objetivo específico, após a análise dos indicadores propostos, foi definida uma estrutura para análise mensal, uma vez citado o indicador de novas receitas mensais. Por isso, a estrutura segue de forma a alinhar todos os indicadores e ser possível fazer análises comparativas entre eles, assim como projeções temporais.

Referente ao terceiro e último objetivo específico, foi proposto, então, um dataframe

construído por meio de planilha eletrônica, capaz de ser importado ao software R, para que possam ser feitos cálculos de correlações e tendências com o uso de machine learning.

A partir do presente trabalho, sugere-se pesquisas posteriores, propostas com o enfoque nos algoritmos de *machine learning*, de forma a trazer os aspectos técnicos para o campo da administração, além de possibilitar uma abordagem mais detalhada e personalizada dentro das empresas. Além destas, recomenda-se pesquisas de campo, com o intuito de entender como a inteligência artificial é implementada em negócios, seja na parte de análise de dados interna ou como produto, a fim de compreender a estrutura dessa tecnologia pelo ponto de vista da administração.

Referências

ARCENALES, John Jairo G. SENA, Priscila M.. DE ARAUJO, Nelma C. **O papel das aceleradoras para o desenvolvimento das startups e do empreendedorismo no Brasil.** AtoZ: novas práticas em informação e conhecimento, v. 10, n. 3, p. 1-10, 2021.

ARRUDA, C. NOGUEIRA, V. Causas da mortalidade de startups brasileiras. 2014. Núcleo de Inovação e Empreendedorismo. **Fundação Dom Cabral.** Disponível em: <<http://acervo.ci.fdc.org.br/AcervoDigital/Artigos%20FDC/Artigos%20DOM%2025/Causas%20da%20mortalidade%20das%20startups%20brasileiras.pdf>> Acesso em 10/dez./2021> .

BAGGIO, Adelar Francisco; BAGGIO, Daniel Knebel. **Empreendedorismo: Conceitos e definições.** Revista de empreendedorismo, inovação e tecnologia, v. 1, n. 1, p. 25-38, 2015.

BARRENCE, André. Cinco anos de Google for Brasil: celebrando a evolução do ecossistema e a reinvenção do cotidiano. **Blog do Google Brasil**, 19 ago. 2021. Disponível em: <<https://brasil.googleblog.com/2021/08/cinco-anos-de-google-for-startups-no.html>>. Acesso em: 19 nov. 2021.

BEDÊ, Marco Aurélio. **Sobrevivência das empresas no Brasil.** Brasília: Sebrae, 2016.

BUENO, R. L. P.; AZEVEDO, M. C. **Produção científica sobre racionalidade na tomada de decisão organizacional estratégica.** Anais do Encontro da Associação Nacional de Pós-Graduação e Pesquisa em Administração-EnANPAD, v. 35, 2011.

CAMILO, Cássio Oliveira; SILVA, João Carlos da. **Mineração de dados: Conceitos, tarefas, métodos e ferramentas.** Universidade Federal de Goiás (UFG), v. 1, n. 1, p. 1-29, 2009.

CASADO, Frank Leonardo. **Análise envoltória de dados: conceitos, metodologia e estudo da arte na educação superior.** Revista Sociais e Humanas, v. 20, n. 1, p. 59-71, 2007.

CHIAVENATO, Idalberto. **Empreendedorismo: dando asas ao espírito empreendedor.** 2. ed. São Paulo: Saraiva, 2008.

SILVEIRA, Denise Tolfo; CÓRDOVA, Fernanda Peixoto. **A pesquisa científica. Métodos de pesquisa.** Editora da Universidade Federal do Rio Grande do Sul. Porto Alegre, 2009.

DA ROCHA, Rodrigo E.; DE CARVALHO, Rodrigo B.; DE AGUIAR FILHO, Armando

S. Inteligência Competitiva na Tomada de Decisão: Estudo de Caso em Startups de Belo Horizonte. Revista Inteligência Competitiva, v. 10, n. 3, p. 18-21, 2020.

DA SILVA, Emanoella S. et al. **Originação de startups: um estudo sobre a tomada de decisão de fundos de investimento em startups no Brasil.** 2021.

ELMASRI, Ramez; NAVATHE, Shamkant B. **Sistemas de Banco de Dados: Fundamentos e Aplicações.** Pearson. São Paulo, 2011.

GIL, A. C. Como elaborar projetos de pesquisa. 4. ed. São Paulo: Atlas, 2007. GOVERNO destaca papel da Micro e Pequena Empresa para a economia do país. **Gov**, 05 out. 2020. Disponível em: <https://www.gov.br/economia/pt-br/assuntos/noticias/2020/outubro/governo-destaca-papel-da-micro-e-pequena-empresa-para-a-economia-do-pais#:~:text=Governo%20destaca%20papel20da%20Micro%20e%20Pequena%20Empresa%20para%20a%20economia%20do%20pa%C3%ADs,-empreendimentos%20representam%2099&text=Juntas%2C%20elas%20representam%2099%25%20dos,dos%20empregos%20gerados%20no%20Brasil>. Acesso em: 19 de nov. de 2021.

MANCUSO, Aline Castello Branco et al. **Estatística descritiva: perguntas que você sempre quis fazer, mas nunca teve coragem.** Clinical & Biomedical Research, v. 38, n. 4, 2018.

R Foundation. The R Project for Statistical Computing. Vienna: **R Foundation**. Disponível em: www.r-project.org Acesso em: 13 de jan. de 2022.

REIS, Edna Afonso; REIS, Ilka Afonso. **Análise descritiva de dados.** Relatório Técnico do Departamento de Estatística da UFMG, v. 1, 2002.

RIES, Eric, “The Lean Startup Methodology, Develop an MVP”. 2012. Disponível em: <http://theleanstartup.com/principles>>. Acesso em: 29 jan.2022.

RODRIGUES, Fabiano; RODRIGUES, Francisco A.; RODRIGUES, Thelma V.. **Modelos de machine learning para predição do sucesso de startups.** Revista de Gestão e Projetos, v. 12, n. 2, p. 28-55, 2021.

RODRIGUES, William Costa et al. Metodologia científica. Faetec/IST. Paracambi, p. 2-20, 2007.

SEBRAE. Termo de Referência: Atuação do Sistema SEBRAE na Economia Digital. Disponível em: <https://www.sebrae.com.br/Sebrae/Portal%20Sebrae/Anexos/tr_economia_digital_2012.pdf>. Acesso em: 29 jan.2022.

SEMELER, Alexandre Ribas et al. Ciência da informação em contextos de e-science: bibliotecários de dados em tempos de Data Science. 2017.

SILVA, A.; RIBEIRO, F. Information Science and Philosophy of Information: Approaches and Differences. In: DEMIR, H. (Ed.). Luciano Floridi's Philosophy of Technology: Critical Reflections., Heidelberg: Springer Science+Business Media Dordrecht, 2012. (Philosophy of Engineering and Technology 8).

TABORDA, Ana. O que é uma start up? 2006. Disponível em: <http://www.gesentrepreneur.com/pdf/o_que_e_uma_start_up.pdf//> Acesso em: 29 jan.2022.

TEIXEIRA, Enise Barth. **A análise de dados na pesquisa científica: importância e desafios em estudos organizacionais.** Desenvolvimento em questão, v. 1, n. 2, p. 177-201, 2003.

TUDO o que você precisa saber sobre startups. **Abstartups**, 05 jul. 2017. Disponível em: <<https://abstartups.com.br/definicao-startups/>>. Acesso em: 19 nov. 2021.

TUOMI, I. **Data is more than knowledge: implications of the reversed knowledge hierarchy for knowledge management and organization memory.** Journal of Management Information Systems, v. 16, n. 3, p. 103-117, Winter 1999.