



NOME DO PROJETO  
FACULDADE DE MEDICINA DA USP

## Controle do Documento

### Histórico de revisões

Data	Autor	Versão	Resumo da atividade
12/08/2022	Vitor, Jonas, Giovane, Renato, Theo, Kaique	1.1	Criação da parte de negócios (análise do mercado, SWOT, Matriz de Riscos, Canvas Value Proposition), persona e jornada do usuário, descrição da solução e descrição dos dados.

# Sumário

<b>1. Introdução</b>	<b>4</b>
<b>2. Objetivos e Justificativa</b>	<b>5</b>
2.1. Objetivos	5
2.2. Proposta de Solução	5
2.3. Justificativa	5
<b>3. Metodologia</b>	<b>6</b>
3.1. CRISP-DM	6
3.2. Ferramentas	6
3.3. Principais técnicas empregadas	6
<b>4. Desenvolvimento e Resultados</b>	<b>7</b>
4.1. Compreensão do Problema	7
4.1.1. Contexto da indústria	7
4.1.2. Análise SWOT	7
4.1.3. Planejamento Geral da Solução	7
4.1.4. Value Proposition Canvas	7
4.1.5. Matriz de Riscos	7
4.1.6. Personas	8
4.1.7. Jornadas do Usuário	8
4.2. Compreensão dos Dados	9
4.3. Preparação dos Dados	10
4.4. Modelagem	11
4.5. Avaliação	12
<b>5. Conclusões e Recomendações</b>	<b>13</b>
<b>6. Referências</b>	<b>14</b>
<b>Anexos</b>	<b>15</b>

# 1. Introdução

Inaugurado em 2008, o Instituto do Câncer do Estado de São Paulo (ICESP), unidade do Hospital das Clínicas da Faculdade de Medicina da USP (HCFMUSP), é um dos maiores centros de oncologia da América Latina, com pilares em assistência, ensino e pesquisa, e atendimento 100% pela rede pública de saúde. Ao longo dos anos consolidou um padrão de gestão acolhedora ao prestar uma assistência com atendimento humanizado e de qualidade. Possui índices de 96,2% de satisfação dos usuários e foi eleito duas vezes como melhor hospital do estado pelos pacientes (2010 e 2014). Em pouco mais de uma década, já atendeu mais de 110 mil pessoas, foram realizados mais de 24 milhões de exames de análises clínicas e 1,8 milhão de exames de imagem. Diariamente circulam pelas instalações cerca de 10 mil pessoas. Com alta densidade tecnológica, o Instituto conta com um dos maiores parques radioterápicos da América Latina com aceleradores lineares para radioterapia, equipamento de braquiterapia e tomógrafo para simulação de procedimentos.

O problema a ser resolvido é a falta de padrões e a grande variabilidade da evolução do câncer de mama e sua resposta a tratamentos convencionais. Pacientes que possuem o mesmo subtipo de câncer de mama ou estão em uma mesma faixa de risco apresentam respostas diferentes a tratamentos iguais, algumas vivem mais do que o esperado e recebem alta e outras vão a óbito precocemente.

## 2. Objetivos e Justificativa

### 2.1. Objetivos

O objetivo principal é a criação de um modelo preditivo a partir de coortes de pacientes acompanhadas em projetos de pesquisa do Instituto do Câncer do Estado de São Paulo/Faculdade de Medicina da Universidade de São Paulo. O modelo deve mostrar um score de risco (esquema de cores) que ajudará na decisão da abordagem e do tratamento mais adequados, além de aumentar a objetividade no acompanhamento com dados mais precisos.

### 2.2. Proposta de Solução

A solução proposta do projeto - Modelo preditivo a partir de variáveis clínico-laboratoriais de pacientes com câncer de mama - refere-se a dois pontos : a) Tempo de sobrevida do paciente (número de dias da sobrevida) e b) Resposta ao tratamento do paciente

A evolução do câncer de mama e sua resposta a tratamentos convencionais é muito variável. O ICESP quer identificar padrões preditivos dessa variabilidade a partir de dados clínicos e do seguimento desses pacientes, dessa maneira temos o objetivo da criação de modelo preditivo a

partir de dados de pacientes acompanhadas em projetos de pesquisa do Instituto do Câncer do Estado de São Paulo/Faculdade de Medicina da Universidade de São Paulo.

O projeto vai fornecer uma *classificação* de acordo com o *grau de prioridade e urgência* da paciente com câncer de mama. Por exemplo, um sistema de cores sendo verde para pacientes com pouco risco e que não precisam de um acompanhamento tão recorrente, amarelo para pacientes com um risco médio e vermelho para pacientes com alto risco que devem ser acompanhados de perto.

## 2.3. Justificativa

Nossa proposta de solução trará uma ferramenta a mais para auxiliar médicos a tomar decisões, tornando o acompanhamento e tratamento mais eficientes e precisos ao encontrar padrões e tendências ocultas. Potencialmente, será uma inteligência na qual aprende ao longo do tempo, quanto mais fazer uso da I.A mais ela se torna eficiente e gera resultados mais sólidos e isso torna uma exponencial de aprendizado e eficiência em relação ao tempo de uso. Além disso, é um produto o qual terá escalabilidade para outros tratamentos e tipos de câncer. Desse modo, nosso produto vai gerar grande valor para a sociedade melhorando a área de saúde e tecnologia, trazendo valores diretos para médicos, hospitais e pacientes. Em relação ao diferencial podemos citar: Interfaces e estrutura de interpretação de dados de modo moderno, mais eficiente, simples e menos confuso. A maneira atual na qual as telas expõem os dados retidos e trabalhos ainda é confusa, chata e trabalhosa. No ano de 1840 teve o surgimento de gráficos de linhas, colunas e de pizza, e até os dias de hoje fazemos uso dessa estrutura para interpretação de dados devido ao fato de que nos acostumamos com isso, mas podemos mudar e fazer diferente.

## 3. Metodologia

Descreva as etapas metodológicas que foram utilizadas para o desenvolvimento, citando o referencial teórico. Você deve apenas enunciar os métodos, sem dizer ainda como ele foi aplicado e quais resultados obtidos.

### 3.1. CRISP-DM

Descreva brevemente a metodologia CRISP-DM e suas etapas de processo

### 3.2. Ferramentas

Descreva brevemente as ferramentas utilizadas e seus papéis (Google Colaboratory)

### 3.3. Principais técnicas empregadas

Descreva brevemente as principais técnicas empregadas, algoritmos e seus benefícios

## 4. Desenvolvimento e Resultados

De maneira geral, você deve descrever nesta seção a aplicação dos métodos aprendidos e os resultados obtidos por seu grupo em seu projeto

### 4.1. Compreensão do Problema

#### 4.1.1. Contexto da indústria

As 5 Forças de Porter são um framework de análise setorial que permite entender o nível de competitividade de um mercado. O modelo apresenta os atores envolvidos (concorrentes, fornecedores, compradores, novos entrantes e substitutos), como eles se relacionam e como influenciam o sucesso dos negócios.

#### Rivalidade entre os concorrentes:

##### Principais players:

1) Fleury

O Grupo Fleury é uma empresa de saúde brasileira fundada em 1926, cuja principal atividade é a prestação de serviços médicos e medicina diagnóstica.

2) DASA

A Dasa é a maior rede de saúde integrada do Brasil e tem a posição de líder em medicina diagnóstica no Brasil e na América Latina e é a 5ª maior do setor no mundo

3) Centro de Oncologia e Hematologia do Hospital Israelita Albert Einstein

É um centro de excelência no atendimento ao paciente oncológico, reunindo toda cadeia de atendimento: da prevenção e diagnóstico às diversas modalidades de tratamento, além de oferecer uma ampla gama de serviços para ajudar no enfrentamento de doenças, com o objetivo de minimizar seus impactos e maximizar a qualidade de vida, aliado com um atendimento integral e humanizado.

4) Centro de Oncologia e Hematologia da Beneficência Portuguesa

O Centro de Oncologia e Hematologia da BP oferece cuidado integrado de saúde para pacientes com câncer e doenças hematológicas, permeando todos os serviços oferecidos pela instituição, desde prevenção, diagnóstico e tratamentos.

#### 5) A.C. Camargo Cancer Center

Referência internacional, o A.C. Camargo Cancer Center é um centro integrado de diagnóstico, tratamento, ensino e pesquisa do câncer, além de ser modelo sustentável de atuação social. Há 65 anos, é a principal instituição de ensino em oncologia no país, responsável pela formação de mais de mil especialistas e residentes, além de mais de 600 mestres e doutores.

**Conclusão:** Esse setor possui diversos players que rivalizam na concorrência. Em um setor concorrido, o diferencial será o investimento em novas tecnologias que atraiam mais pacientes em busca de diagnósticos e prognósticos mais precisos e de melhores atendimentos.

#### **Poder de barganha dos fornecedores:**

Pode-se inferir que o Instituto do Câncer do Estado de São Paulo, por ser um hospital público, depende da verba do governo. Dessa forma, o poder de barganha dos fornecedores é alto, já que o hospital depende quase que 100% de uma única fonte de verba, não tendo escolha de quanto será investido nem por qual instituição.

#### **Poder de barganha dos clientes:**

Por se tratar de um serviço público o poder de barganha dos clientes pode ser considerado pequeno, já que muitos pacientes não têm condições de ir a um hospital particular e acabam dependendo de hospitais públicos. Dessa forma, elas dependem do que é oferecido pelos hospitais, mesmo que não sejam de qualidade.

#### **Ameaça de novos entrantes:**

Agentes que trazem ameaças : Hospitais particulares , públicos e Startups.

Hoje, a tecnologia de Inteligência Artificial utilizada para fins de diagnóstico, detecção de câncer e classificação de grau de prioridade de tratamento está em alta e muitos falam sobre isso, a tendência do mercado é que cada vez mais Startups e Hospitais privados e públicos invistam nessa tecnologia, se tornando ameaças para o setor e possíveis futuros concorrentes.

**Barreiras de entrada:** falta de profissionais técnicos na área de tecnologia e banco de dados robustos( imagens e informações ) nos quais poucos têm acesso.

Características de empresas que podem trazer ameaça: Boa imagem da marca (autoridade e reconhecimento ), forte Know-how e acesso a matérias primas (dados, profissionais...)

### **Ameaça de produtos substitutos:**

Linda lifetech

função : Detectar o câncer de mama

Proposta : consiste em um sensor infravermelho que captura a imagem da mama da paciente e envia para um servidor na nuvem. De lá, com ajuda de inteligência artificial, a imagem é comparada com um banco de dados, que conta com mais de 5 milhões de informações, em busca de indícios de lesões cancerígenas.Quanto mais a ferramenta for utilizada, mais dados ela terá e o diagnóstico será cada vez mais aprimorado

Um potencial serviço substituto seria a Linda Lifetech, uma startup que oferece serviços de diagnóstico de câncer de mama utilizando Inteligência Artificial de baixo custo, que poderia substituir uma grande parte dos pacientes do Hospital Universitário, pois os diagnósticos são mais precisos e eficientes, pois são compatíveis com telemedicina, com o custo sendo um grande ponto que ainda não o torna como solução número 1º sobre diagnóstico de câncer de mama.

### **Modelo de negócios:**

Por ser uma instituição pública, o ICESP recebe verba do Estado de São Paulo, proveniente de dinheiro público, sendo pago por ciclos de atendimentos, já que por existir diferentes tipos de neoplasias e cada tumor não possuir um medicamento específico, o SUS aloca valores definidos para cada tratamento de cada paciente, e cabe aos médicos escolherem os medicamentos e protocolos que utilizarão no tratamento com base em dados científicos e padrões da instituição em que atendem.

### **Tendências do setor de saúde:**

1. Interoperabilidade em nuvem: O armazenamento de informações coletadas dos pacientes em plataformas nativas em nuvem ajudam o médico que está prestando o atendimento a fazê-lo de uma forma mais eficiente, humanizada e assertiva, já que ele terá disponível todo prontuário médico da pessoa mesmo numa primeira consulta. Além disso, diminuem custos com a criação e manutenção de bancos de dados locais e dificultam vazamento de dados.
2. Inteligência artificial cada vez mais presente: A inteligência artificial (IA) possibilita novos estudos e pesquisas, o aprimoramento de tratamentos e ajuda na detecção de diagnósticos. A análise preditiva, aliada a IA, traz a possibilidade de identificação de um

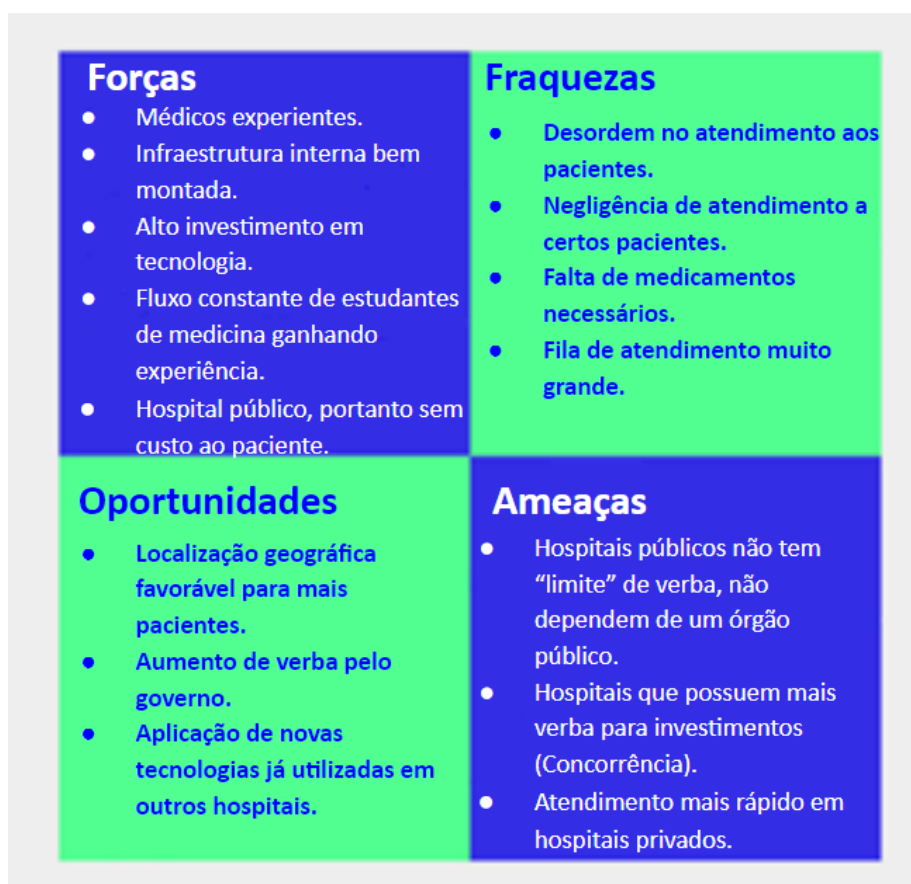


problema de forma mais rápida a partir da antecipação dos riscos pela identificação de dados históricos e padrões.

3. Saúde Mental: Por conta da pandemia muitas pessoas sofreram com problemas relacionados com a saúde mental e emocional. Dessa forma, os tratamentos e tecnologias especializadas nesse tipo de tratamento devem ganhar um destaque ainda maior.
4. Medicina robótica, o uso de robôs para auxiliar nas cirurgias: Com um crescimento estimado em 30% ao ano, segundo a Sociedade Brasileira de Cirurgia Minimamente Invasiva e Robótica (SOBRACIL), o procedimento garante uma abordagem menos invasiva e consideravelmente mais precisa em comparação a outras técnicas. Por isso, é esperado que esse tipo de cirurgia ganhe mais espaço no Brasil e que o avanço dessa tecnologia possa trazer novas melhorias para o setor.
5. Atendimento remoto: O distanciamento social imposto pela pandemia fez com que a telemedicina ganhasse um enorme espaço no setor, possibilitando atendimentos à distância sem que o médico esteja presencialmente com o paciente. Tecnologias como o 5G e Internet das Coisas devem fortalecer o atendimento remoto ainda mais. Vale ressaltar ainda a sua importância social ao garantir ainda um avanço democrático da saúde no país ao possibilitar que pacientes e médicos fora dos grandes centros possam usufruir deste meio.
6. Crescimento das Health Techs: Com tantas transformações digitais demandadas pelas mudanças atuais, a força do mercado de health tech contribui significativamente para otimizar estratégias, por meio de soluções focadas no atendimento mais acessível e seguro das redes de saúde. Esse crescimento está atrelado à procura de hospitais, clínicas e laboratórios que recorrem às empresas de tecnologia para criar soluções que atuam na prevenção, detecção e tratamento de doenças e otimizar suas estratégias de gestão.

### 4.1.2. Análise SWOT

A análise SWOT tem como objetivo ter uma visão externa e interna do negócio. Desse modo, a matriz é organizada em quatro quadrantes que levam em conta o ambiente externo e interno, que são divididos em fatores internos controláveis e fatores externos incontrolláveis. Essa forma de abordagem contribui para o fortalecimento dos pontos fortes e no amadurecimento dos pontos fracos, além de prevenir possíveis danos.



### 4.1.3. Planejamento Geral da Solução

**a) quais os dados disponíveis (fonte e conteúdo - exemplo: dados da área de Compras da empresa descrevendo seus fornecedores)**

Dados dos prontuários de pacientes com câncer de mama do ICESP. Esses dados descrevem informações pessoais das pacientes e informações a respeito do câncer.

**b) Qual a solução proposta (pode ser um resumo do texto da seção 2.2)**

O objetivo principal é a criação de um modelo preditivo a partir de coortes de pacientes acompanhadas em projetos de pesquisa do Instituto do Câncer do Estado de São Paulo/Faculdade de Medicina da Universidade de São Paulo.

A solução proposta do projeto refere-se a passar o risco do paciente de acordo com o tempo de sobrevivência. O modelo irá fornecer uma *classificação* de acordo com o *grau de prioridade e urgência* da paciente com câncer de mama a partir de um sistema de cores que identificam risco alto, médio ou baixo.

**c) qual o tipo de tarefa (regressão ou classificação)**

O tipo de tarefa será de classificação, ou seja, prever qual o risco do paciente e colocá-lo em uma classificação por grupos (risco alto, médio e baixo).

**d) como a solução proposta deverá ser utilizada**

A solução deverá ser utilizada como uma ferramenta de pesquisa para auxiliar o médico a tomar a melhor decisão de tratamento e acompanhamento. A partir da cor, o médico saberá a melhor abordagem para cada paciente.

**e) quais os benefícios trazidos pela solução proposta**

O benefício principal será a construção de uma plataforma de análise de coortes de pacientes assistidos no Instituto do Câncer como ferramenta para pesquisa. A solução fará com que os prognósticos sejam mais objetivos e assertivos de acordo com o risco de cada paciente.

**f) qual será o critério de sucesso e qual medida será utilizada para o avaliar**

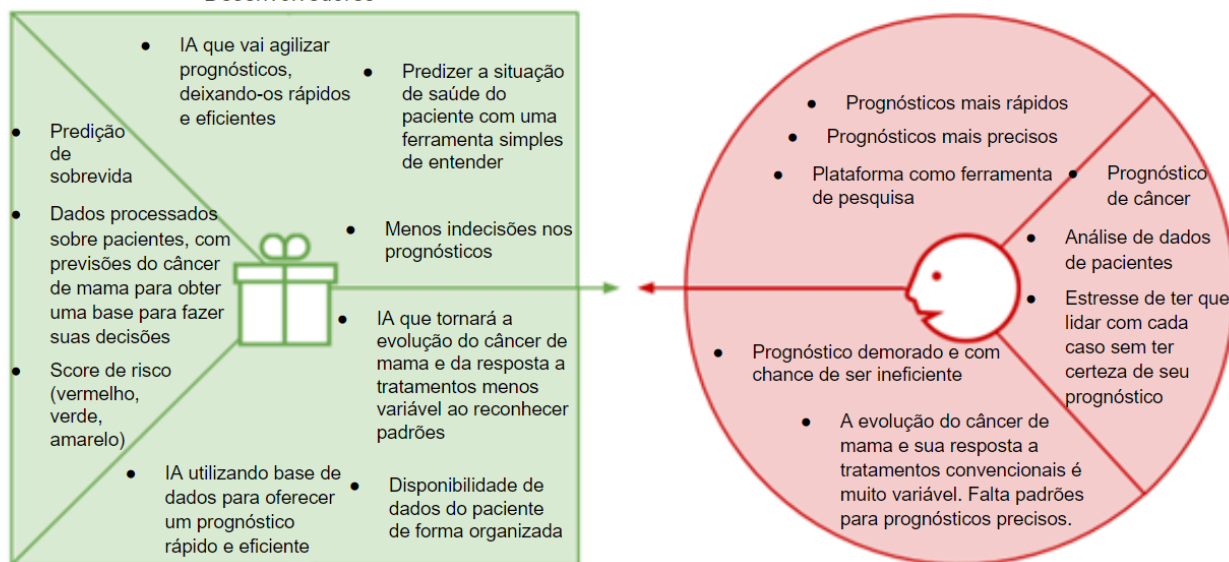
O critério de sucesso será uma assertividade de 80% e para avaliá-la serão utilizadas pacientes já do banco de dados, já que sabemos o resultado previamente e poderá ser comparado com o resultado do modelo preditivo.

#### **4.1.4. Value Proposition Canvas**

O Canvas da proposta de valor serve para ajudar a criar e organizar produtos/serviços que se alinhem com o que seu cliente realmente valoriza e precisa.

## Perfil do Cliente

Faculdade de Medicina da USP / ICESP



A matriz de risco é uma forma usual de se avaliar o risco. Uma matriz de risco é uma representação da combinação da probabilidade de ocorrer um evento associando a esta probabilidade a consequência caso o evento ocorra.

		Ameaças				
Probabilidade	90%					
	70%					Pouco tempo de desenvolvimento
	50%				Falta de conhecimento técnico	
	30%				Desalinhamento/desentendimento da equipe	IA possuir viés negativo
	10%				Recusa dos médicos em utilizarem a IA	Dados errados/escasso
		Muito Baixo	Baixo	Moderado	Alto	Muito Alto
		Impacto				

		Oportunidades				
Probabilidade	90%		Oportunidade de aprender novas tecnologias como I.A			
	70%		Conhecer uma nova área para ter interesse em seguir como carreira			
	50%	IA ser espalhada para outros hospitais públicos	Diagnósticos rápidos			
	30%	Referência para criação de outras IA			Criar empregos de manutenção da AI	
	10%	Diminuir a taxa de mortalidade do câncer de mama				
		Muito Alto	Alto	Moderado	Baixo	Muito Baixo
		Impacto				

#### 4.1.6. Personas



NOME: João Dias Pinto Filho

IDADE: 43 anos

GÊNERO: Masculino

OCUPAÇÃO: Oncologista

“O melhor médico é aquele que mais esperança inspira”

João é um oncologista do Instituto do Câncer do Estado de São Paulo especializado em câncer de mama. Calmo e atencioso, quer sempre ajudar o máximo de pacientes possível. Contudo, em todos os seus anos como oncologista presenciou muitas de suas pacientes vindo a óbito por causa do câncer, já que sua evolução e respostas a tratamentos são muito variáveis. Assim, João precisa de um modelo preditivo que seja fácil de utilizar e que de alguma forma encontre padrões nessa variabilidade para ajudá-lo a decidir a melhor abordagem no tratamento de suas pacientes de acordo com o risco de cada uma. No fim, ele terá seu trabalho facilitado e poderá atender ainda mais pacientes de forma precisa e eficaz.



NOME: Ana Maria Gabriela de Jesus

IDADE: 33 anos

GÊNERO: Feminino

OCUPAÇÃO: Marceneira

"Não deixe que as pessoas te façam desistir daquilo que você mais quer na vida. Acredite. Lute. Conquiste. E acima de tudo, seja feliz!"

Ana Maria vive na cidade de São Paulo e trabalha como marceneira. Sua renda não é tão elevada, então ela acaba dependendo do atendimento público quando tem algum problema de saúde. Recentemente, ela descobriu que tinha câncer de mama, mesmo mantendo hábitos saudáveis, e por não confiar 100% no sistema público ela tem grande receio em como será seu tratamento e em relação ao verdadeiro risco de seu câncer. Para tranquilizar Ana Maria, o modelo preditivo irá, através do médico, garantir um prognóstico e tratamento eficientes além de mostrar o possível risco do câncer dela. Assim, ela terá um atendimento de qualidade e personalizado, e saberá explicitamente o risco da doença.

## 4.1.7. Jornadas do Usuário

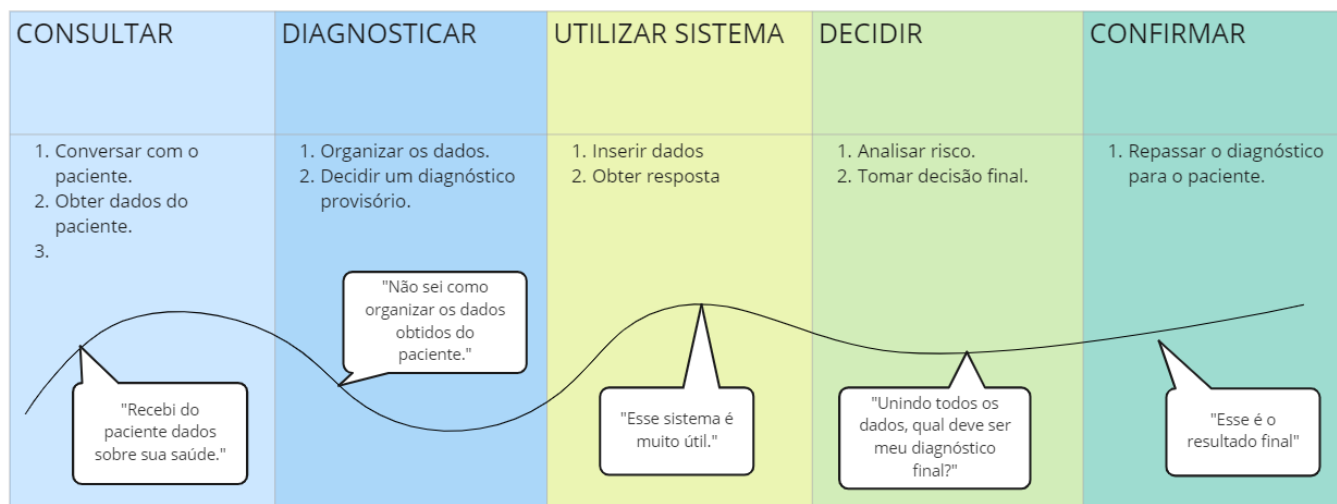


João Dias Pinto

**Cenário:** João precisa fazer a consulta e decidir o melhor diagnóstico. Ele quer uma solução que facilite a análise de dados e mostre um score de risco.

### Expectativas

Atender o máximo de pessoas possíveis com um diagnóstico rápido e preciso.



### Oportunidades

- Score de risco para definir um diagnóstico mais preciso
- Processamento de dados com maior velocidade

### Responsabilidades

Time de desenvolvimento: treinar o modelo com dados suficientes e completos para aumentar a eficiência.

miro

## 4.2. Compreensão dos Dados

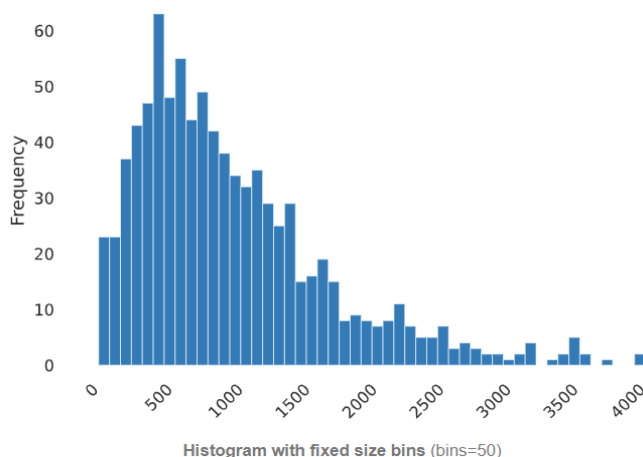
1. **Descreva os dados a serem utilizados (disponibilizados pelo cliente e outros se tiverem sido incluídos), detalhando a fonte, o formato (CSV, XLSX, banco de dados, etc.), o conteúdo e o tamanho.**

Os dados, no formato CSV, são dos prontuários de pacientes com câncer de mama do ICESP. Esses dados descrevem informações pessoais das pacientes (como escolaridade, IMC, raça, câncer na família, etc.) e informações a respeito do câncer (como subtipo, recidiva, regime de tratamento, etc.). A tabela possui aproximadamente 63500 linhas, com 4132 pacientes (linhas únicas). As pacientes são identificadas pela primeira coluna (record\_id), que nesse caso foi alterada para proteger a identificação. Contudo, muitos registros estão vazios, assim, mesmo tendo uma grande diversidade de informações, muitas acabam se tornando inutilizáveis pela falta de informação. Dessa forma, uma análise aprofundada dos dados é necessária para saber quais estão completos e quais são relevantes para, então, dividi-los em subconjuntos e retirar dados nulos.

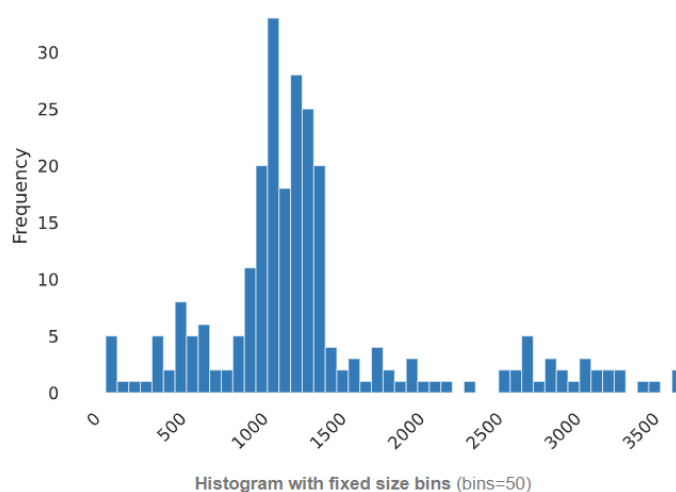
## 2. Descrição estatística básica dos dados, principalmente dos atributos de interesse, com inclusão de visualizações gráficas e como essas análises embasam suas hipóteses.

No primeiro momento, foram escolhidos três atributos para análise: o tempo decorrido entre o diagnóstico e a última vez que houve contato com o paciente (follow\_up\_days), tipos de tratamento (treatment) e última informação que se tem em relação a óbito ou alta (utlinfo). A primeira análise foi feita com a relação entre follow\_up\_days e utlinfo, abaixo seguem os gráficos:

### follow\_up\_days de pessoas que vieram a óbito, por câncer (Média: 1045 dias)

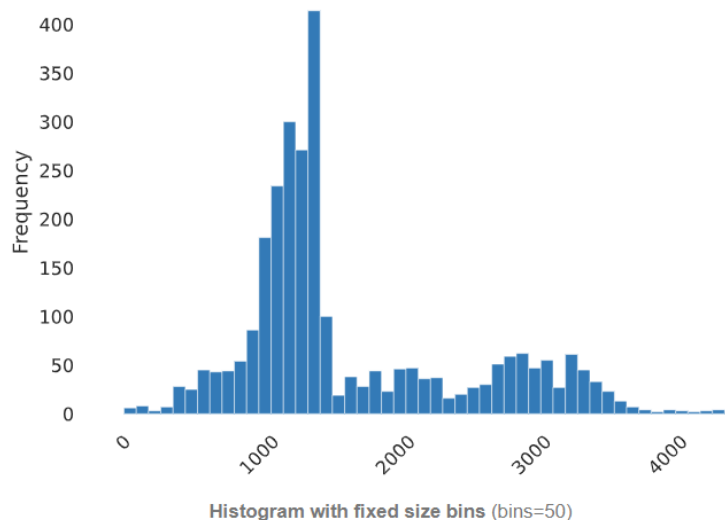


### follow\_up\_days de pessoas que vivas, com câncer (Média: 1387 dias)





## follow\_up\_days de pessoas vivas, sem outras especificações (Média: 1668 dias)



Essa análise embasa a hipótese de que algumas pacientes que vieram a óbito não terminaram ou começaram um tratamento e, por isso, a média de dias é menor do que pacientes que tiveram alta (1045 dias contra 1387 dias para vivas com câncer e 1668 dias para vivas sem câncer).

A outra análise feita foi a relação entre o utlinfo e o treatment. Abaixo estão os gráficos e uma legenda para os valores:

Legenda:

0, Não fez quimioterapia

1, Terapia Adjuvante (realizada depois de uma cirurgia ou radioterapia)

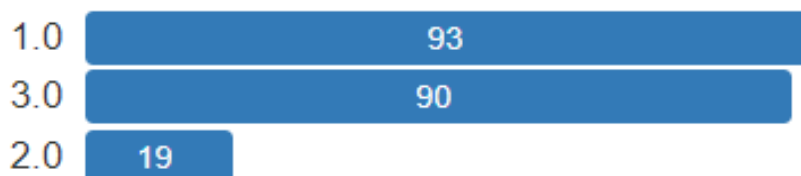
2, Terapia Neoadjuvante (realizada antes de uma cirurgia ou radioterapia)

3, Paliativo (realizada em pacientes em um estado mais grave)

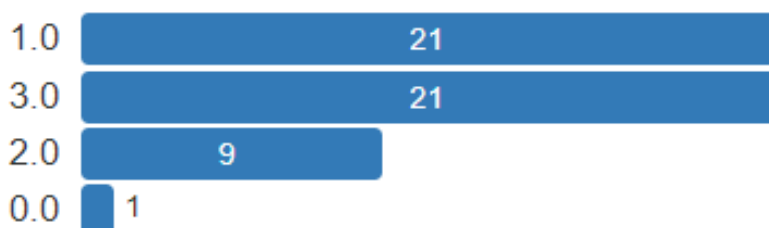
## treatment de pessoas vivas, sem outras especificações (23.3% respondido)



treatment de pessoas que vieram a óbito, por câncer (23.2% respondido)



treatment de pessoas vivas, com câncer (20.7% respondido)



Pelos gráficos é possível notar que pacientes que vivem sem câncer acabam realizando mais a terapia adjuvante enquanto que os outros grupos possuem duas terapias com praticamente o mesmo número de casos.

**3. Descrição da predição desejada (“target”), identificando sua natureza (binária, contínua, etc.)**

Deseja-se uma predição multiclasse (três tipos de classificação) em que a saída do modelo apresente corretamente uma das três classificações (baixa, média e alta) a partir de um esquema de cores. Se o risco da nova paciente for baixo, mostrar como output a cor verde, se o risco for médio, mostrar a cor amarela e se o risco for alto, mostrar a cor vermelha. Caso esses outputs sejam corretos em 80% dos casos, a predição terá sido como desejada.

## 4.3. Preparação dos Dados

Descreva as etapas realizadas para definir os dados e os atributos descritivos dos dados (“features”) a serem utilizados. Essa descrição deve ser feita de modo a garantir uma futura reprodução do processo por outras pessoas, e deve conter:

- a) Descrição de quaisquer manipulações necessárias nos registros e suas respectivas features.
- b) Se aplicável, como deve ser feita a agregação de registros e/ou derivação de novos atributos.
- c) Se aplicável, como devem ser removidos ou substituídos valores ausentes/em branco.
- d) Identificação das features selecionadas, com descrição dos motivos de seleção.

Não deixe de usar tabelas e gráficos de visualização de dados para melhor ilustrar suas descrições.

**IMPORTANTE:** Crie tópicos utilizando a formatação “Heading 3” (ou menor) para que o Google Docs identifique e atualize o Sumário (é necessário apertar o botão Refresh no Sumário para ele coletar as atualizações)

## 4.4. Modelagem

Para a Sprint 3, você deve descrever aqui os experimentos realizados com os modelos (treinamentos e testes) até o momento. Não deixe de usar equações, tabelas e gráficos de visualização de dados para melhor ilustrar seus experimentos e resultados.

Para a Sprint 4, você deve realizar a descrição final dos experimentos realizados (treinamentos e testes), comparando modelos. Não deixe de usar equações, tabelas e gráficos de visualização de dados para melhor ilustrar seus experimentos e resultados.

## 4.5. Avaliação

Nesta seção, descreva a solução final de modelo preditivo, e justifique a escolha. Alinhe sua justificativa com a seção 4.1, resgatando o entendimento do negócio e explicando de que formas seu modelo atende os requisitos. Não deixe de usar equações, tabelas e gráficos de visualização de dados para melhor ilustrar seus argumentos.

## 5. Conclusões e Recomendações

Escreva, de forma resumida, sobre os principais resultados do seu projeto e faça recomendações formais ao seu parceiro de negócios em relação ao uso desse modelo. Você pode aproveitar este espaço para comentar sobre possíveis materiais extras, como um manual de usuário mais detalhado na seção “Anexos”.

Não se esqueça também das pessoas que serão potencialmente afetadas pelas decisões do modelo preditivo, e elabore recomendações que ajudem seu parceiro a tratá-las de maneira estratégica e ética.

## 6. Referências

Nesta seção você deve incluir as principais referências de seu projeto, para que seu parceiro possa consultar caso ele se interessar em aprofundar.

Utilize a norma ABNT NBR 6023 para regras específicas de referências. Um exemplo de referência de livro:

LUCK, Heloisa. **Liderança em gestão escolar**. 4. ed. Petrópolis: Vozes, 2010.

SOBRENOME, Nome. **Título do livro**: subtítulo do livro. Edição. Cidade de publicação: Nome da editora, Ano de publicação.

## Anexos

Utilize esta seção para anexar materiais como manuais de usuário, documentos complementares que ficaram grandes e não couberam no corpo do texto etc.