



Integration

BIG DATA

INTEGRAÇÃO, GERENCIAMENTO E ANÁLISE

INSTITUTO DE TECNOLOGIA E LIDERANÇA – INTELI

INTEGRAÇÃO, GERENCIAMENTO E ANÁLISE DE BIG DATA

INTEGRATION

Autores:

Gabriel Santos

Henri Harari

Rafael Moritz

Patrick Miranda

Raduan Muarrek

Vitor Moura

Data de criação: 24 de Outubro de 2023

SÃO PAULO – SP

Controle de Documento

Histórico de Revisões

Data	Autor	Versão	Resumo da atividade
26/10/2023	Patrick	1.0	Adição dos artefatos da sprint 1: Artefatos de análise de negócio, Análise de Experiência do usuário e arquitetura da solução.
09/11/2023	Henri e Vitor	1.1	Adição da documentação do wireframe.
12/11/2023	Patrick	2.0	Revisão da formatação do arquivo para entrega

Table 1: Controle de documento

Sumário

Controle de Documento	3
Sumário	3
1. Introdução	5
1.1 Parceiro de Negócios	5
1.2 Problema	5
1.2.1 Definição do Problema	5
2. Objetivos	6
2.1 Objetivos Gerais	6
2.2 Objetivos Específicos	6
2.3 Justificativa	6
3. Análise de Negócios	7
3.1 Proposta de Valor	7
3.2 Matriz de Risco	8

3.3 TAM SAM SOM	9
4. Análise de Experiência do Usuário	12
4.1 Personas	12
4.2 Jornada do Usuário	13
4.3 User Stories	14
5. Análise Exploratória	18
5.1 Introdução	18
5.2 Método	22
6. Arquitetura Macro	23
6.1 Identificação dos dados:	23
6.1.1 Dados públicos:	23
6.1.2 Dados do cliente:	25
6.2 Gestão de dados:	25
6.2.1 Dados públicos:	25
6.2.2 Dados privados:	25
6.3 Seleção dos serviços AWS:	26
6.4 Fluxo dos dados:	27
6.5 Segurança:	28
6.6 Monitoramento e gerenciamento:	28
7. Data lake	29
7. Wireframe	29
7.1 Objetivo	29
7.2 Informações Requeridas	29
7.3 Justificativa das Escolhas de Design	30
7.4 Técnicas Avançadas	31
8. Conclusões	33
9. Referências	34
10. Anexos	35

1. Introdução

1.1 Parceiro de Negócios

1.2 Problema

1.2.1 Definição do Problema

2. Objetivos

2.1 Objetivos Gerais

2.2 Objetivos Específicos

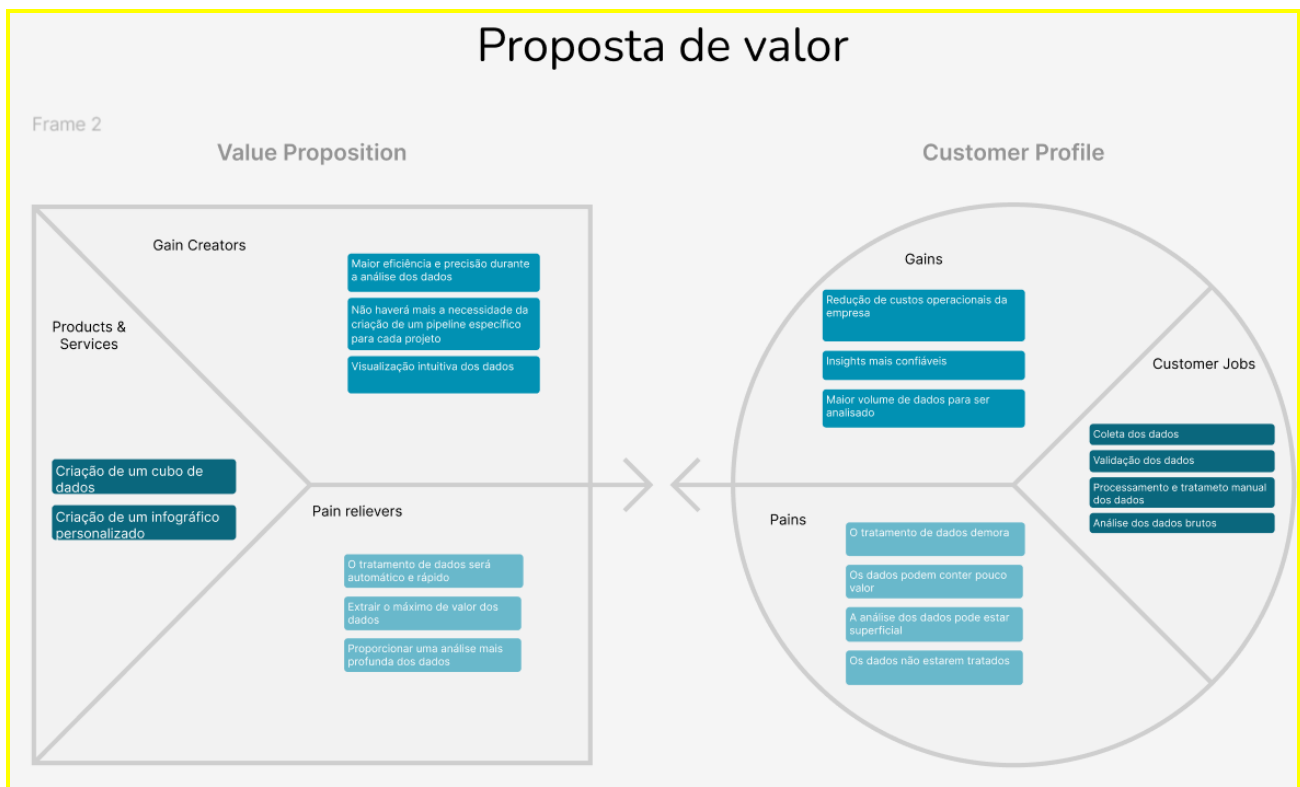
2.3 Justificativa

3. Análise de Negócios

3.1 Proposta de Valor

Ao usar o Value Proposition Canvas, as empresas podem alinhar suas ofertas com as necessidades do cliente, melhorar suas mensagens de marketing e se diferenciar de seus concorrentes.

imagem 1: Canvas proposta de valor



Fonte: Dados dos autores (2023)

Pains (Dores) e Pain Relievers (Aliviadores de Dor):

As "Dores" representam os desafios específicos que consultorias enfrentam ao tentar entender o mercado na indústria alimentícia, como a demora na coleta e análise de dados, a falta de insights relevantes e a dificuldade em interpretar grandes volumes de dados brutos. Os "Aliviadores de Dor" demonstram como o pipeline de Big Data pode solucionar esses desafios. Isso inclui um tratamento de dados mais ágil, extração de

insights valiosos de vastos conjuntos de dados e a capacidade de realizar análises mais profundas e precisas.

Gains (Ganhos) e Gain Creators (Criadores de Ganho):

Os "Ganhos" expressam os resultados positivos e benefícios que as consultorias desejam alcançar, como otimização de estratégias de "go to market", maior eficiência operacional e geração de insights inovadores para seus clientes na indústria alimentícia. Os "Criadores de Ganho" elucidam como o pipeline de Big Data facilita esses ganhos. Isso pode incluir uma análise mais precisa dos hábitos do consumidor, tendências emergentes no mercado alimentício e identificação de oportunidades inexploradas. Products & Services (Produtos e Serviços) e Customer Jobs (Trabalhos do Cliente):

Conclusão

A seção "Produtos e Serviços" destaca as soluções específicas proporcionadas pelo pipeline, como ferramentas de visualização de dados, análise preditiva e segmentação avançada do mercado. Os "Trabalhos do Cliente" representam as tarefas ou atividades que as consultorias precisam realizar, como entender padrões de consumo, identificar novos nichos de mercado e formular estratégias de penetração de mercado eficazes. Em síntese, o canva "Proposta de Valor" para este projeto de pipeline de Big Data busca direcionar e articular o valor tangível oferecido às consultorias voltadas para a indústria alimentícia. Ele ilustra como a solução aborda dores específicas do mercado, potencializa

3.2 Matriz de Risco

É uma das principais ferramentas na análise de negócios, utilizada para o gerenciamento de riscos operacionais existentes na empresa. A Figura 2, ilustra a construção da matriz de risco para o projeto.

imagem 1: matriz de risco

Matriz de Risco										
Probabilidade	Riscos					Oportunidades				
	Muito Alta	1					Reconhecimento no mercado pela solução inovadora			Entrega do MVP antes do prazo previsto
	Alta	2		Erros na análise estatística	Falha na integração entre serviços		Integração mais profunda com outras ferramentas	Otimização dos custos	Melhoria contínua através do feedback dos usuários finais	
	Média	3		Limitações com ferramentas open-source	Falha na ingestão de dados	Custo inesperado		Descoberta de novas fontes de dados relevantes		
	Baixa	4		Problemas na visualização	Dados incompletos ou corrompidos	Falha no processamento	Exceder o tempo estipulado para entrega do MVP	Expansão da solução para outros distribuidores e mercados	Adoção de novas ferramentas open-source	
	Muito Baixa	5				Falhas de segurança nos dados				
			1	2	3	4	5	5	4	3
Impacto										
			Muito Baixa	Baixa	Média	Alta	Muito Alta	Muito Alta	Alta	Média
									2	1
									Baixa	Muito Baixa

Fonte: Dados dos autores (2023)

Enquanto a matriz de risco proporciona uma visão clara das possíveis contingências e desafios que podem surgir, abaixo apresenta-se o plano de ação onde representa a resposta estratégica para enfrentar tais eventualidades.

imagem 2: plano de ação

Plano de Ação					
Matriz	Descrição do Impacto	Responsável	Ação	Descrição da Ação	Previsão
2-3	Erros na análise estatística.	Rafael	Mitigar	Revisar métodos e validar com especialistas.	Sprint 4
2-4	Falha na integração entre serviços.	Patrick	Mitigar	Testes de integração frequentes.	Sprint 3
3-3	Limitações com ferramentas open-source.	Patrick	Prevenir	Pesquisa prévia e testes de adequação.	Sprint 3
3-4	Falha na ingestão de dados.	Henri	Mitigar	Monitorar processos de ingestão e ter backups.	Sprint 2
3-5	Custo inesperado.	Raduan	Mitigar	Monitorar e controlar gastos.	Sprint 3
4-3	Problemas na visualização.	Vitor	Mitigar	Treinamento na ferramenta e testes frequentes.	Sprint 5
4-4	Dados incompletos ou corrompidos.	Gabriel	Prevenir	Validação dos dados antes da ingestão.	Sprint 2
4-4	Falhas no processamento.	Henri	Mitigar	Monitorar a performance e otimizá-la.	Sprint 4
4-5	Exceder o tempo estipulado para entrega do MVP.	Gabriel	Prevenir	Gestão eficaz do projeto e revisões frequentes.	Sprint 5
5-5	Falhas de segurança nos dados.	Raduan	Prevenir	Implementar protocolos de segurança rigorosos.	Sprint 3

Fonte: Dados dos autores (2023)

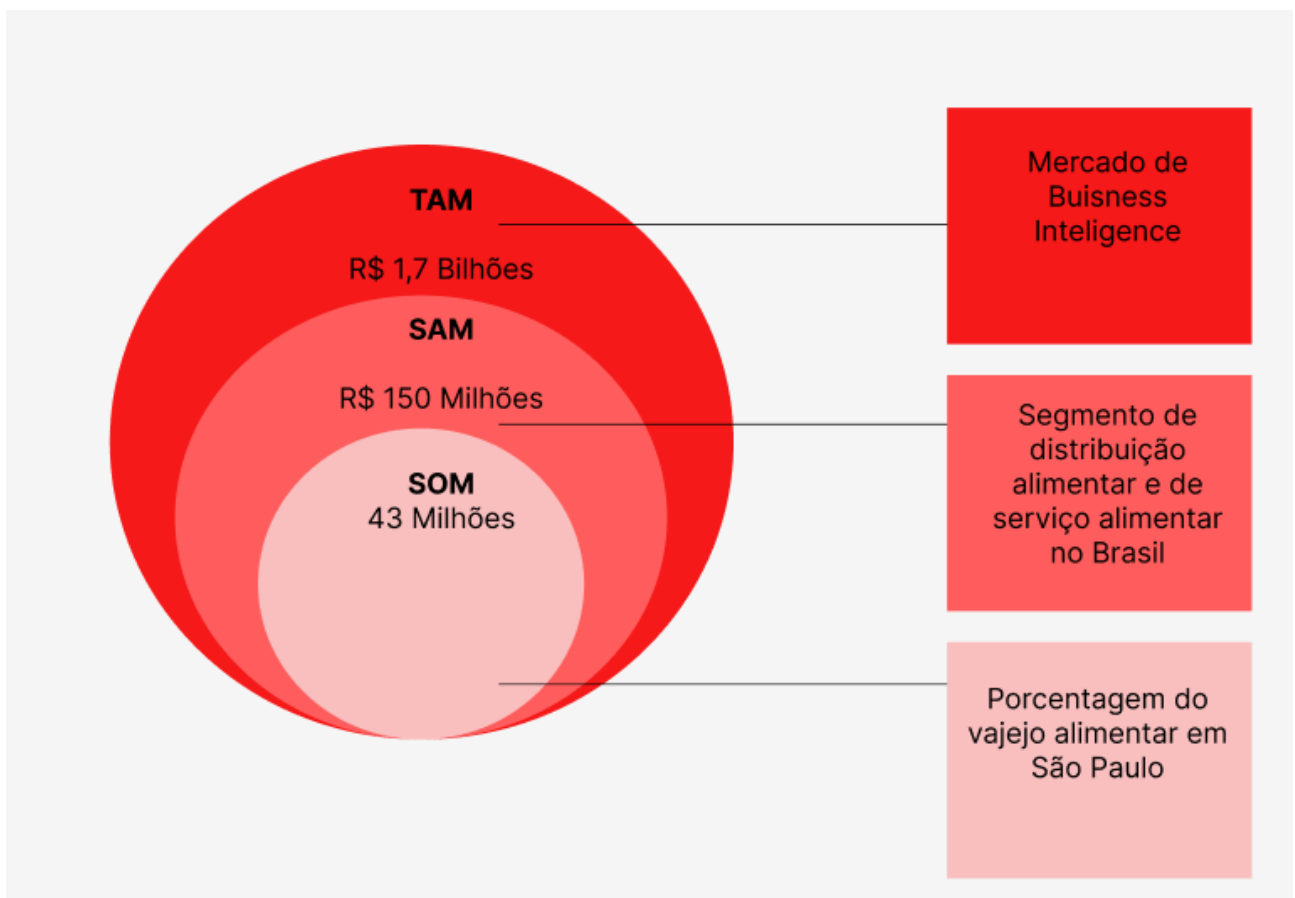
3.3 TAM SAM SOM

TAM (Total Addressable Market): Representa o mercado total que poderia se beneficiar ou necessitar do seu serviço ou produto. É o valor total caso 100% do mercado adotasse seu produto.

SAM (Serviceable Addressable Market): É a parcela do TAM que realmente pode ser alcançada por seu produto ou serviço, levando em consideração as limitações geográficas, de distribuição, capacidade e outras.

SOM (Serviceable Obtainable Market): Representa a porção do SAM que se espera atingir em um determinado período de tempo, considerando fatores como concorrência, barreiras de entrada e estratégia de implementação.

Imagem 1: TAM SAM SOM



TAM

Descrição: Total estimado de receita no mercado de Business Intelligence relacionado à indústria alimentícia.

Premissa: Existe uma grande demanda por insights e análises na indústria alimentícia. O TAM engloba todas as consultorias, empresas e indivíduos que poderiam potencialmente se beneficiar do uso de ferramentas de Business Intelligence na indústria alimentícia.

Valor: R\$ 1,7 bilhões.

SAM

Descrição: Total estimado de receita que pode ser direcionada pelas consultorias que servem o segmento de distribuição alimentar e serviço alimentar no Brasil utilizando insights de big data.

Premissa: Dentro do amplo mercado de Business Intelligence para a indústria alimentícia, há um segmento específico focado na distribuição e no serviço alimentar que pode ser diretamente beneficiado por insights de big data. Este segmento tem necessidades mais específicas e pode ser atendido de forma mais direcionada pelo seu serviço.

Valor: R\$ 150 milhões.

SOM

Descrição: Receita potencial estimada que pode ser capturada pelas consultorias em um determinado período de tempo, focando especificamente no varejo alimentar de São Paulo.

Premissa: São Paulo, sendo um grande hub comercial, possui um segmento significativo de varejo alimentar. Focar neste segmento proporciona uma oportunidade tangível e mensurável. A premissa é que, ao focar em um mercado específico e conhecido, como o varejo alimentar de São Paulo, você pode oferecer soluções mais personalizadas e alcançar uma maior penetração de mercado.

Valor: R\$ 43 milhões.

4. Análise de Experiência do Usuário

4.1 Personas

A persona é uma representação humanizada do público-alvo ideal e é usada para ajudar a equipe de desenvolvimento a compreender melhor suas necessidades, desejos e comportamentos. No projeto atual, foram identificadas duas personas, o tech lead, responsável pela construção do cubo de dados e o consultor de marketing, responsável pela análise do cubo.

Imagem 1: Persona, Moisés Aragão



Moisés Aragão
Tech Lead - Integration

- 36 anos
- Dedicado ao trabalho
- Atenção aos detalhes
- Golfe aos domingos
- Casado
- "Para ter reconhecimento, você deve demonstrar o esforço necessário"

01 KPIs.

- Veracidade dos dados
- Volume de dados
- Velocidade de processamento

02 Desejos

- Desenvolvimento Profissional
- Otimização das tarefas

03 Necessidades:

- Solução adaptável pelos interesses do consultor
- Projetar um cubo de dados

04 Dores:

- Dados descentralizados
- Processos muito demorados

05 Objetivos (Momento que usará o sistema):

- Processar grande volume de dados
- Fornecer dados confiáveis

Fonte: Dados dos autores (2023)

Imagem 2: Persona, Enzo Ananias



Enzo Ananias

Consultor de Marketing - Integration

- 25 anos
- Análise de tendências
- Inovador
- Beach tennis
- Solteiro
- "Não há tempo para perder"



01

KPIs.

- Lucro real da empresa consultada
- Qualidade dos insights
- Eficiência das análises

02

Desejos

- Eficiência Operacional
- Acompanhamento de Tendências

03

Necessidades:

- Praticidade para analisar os dados
- Insights de melhor qualidade

04

Dores:

- Grande volume de dados
- Repetição de tarefas analíticas

05

Objetivos (Momento que usará o sistema):

- Analisar grande volume de dados
- Fornecer insights confiáveis

Fonte: Dados dos autores (2023)

4.2 Jornada do Usuário

A jornada do usuário é uma representação visual ou narrativa do percurso que um indivíduo realiza ao interagir com um produto, serviço ou sistema, desde o primeiro contato até a conclusão de um objetivo específico, levando em consideração suas emoções, experiências e desafios ao longo do caminho. Ela ajuda a compreender as necessidades, motivações e pontos de atrito do usuário, facilitando a criação de experiências mais eficientes e satisfatórias.

Imagem 1: jornada, Enzo Ananias



Fonte: Dados dos autores (2023)

Imagem 2: jornada, Moisés Aragão



Fonte: Dados dos autores (2023)

4.3 User Stories

Abaixo seguem quatro user stories realizados no padrão INVEST, para garantia do padrão de qualidade. Duas referentes ao teach lead e duas referentes ao consultor de marketing.

Número	US01
Título	Atualizar dados Públicos do IBGE
Personas	Moisés Aragão
História	Eu como teach lead, quero atualizar os dados do IBGE, de forma a garantir que os dados utilizados para análise estejam condizentes com o último censo.
Critérios de aceitação	<ol style="list-style-type: none">1. Validar formato da planilha .csv2. Validar valores da planilha .csv3. Efetuar atualização dos dados em .csv para o S3.
Testes de aceitação	<p>Critério 1:</p> <ul style="list-style-type: none">- A planilha está no formato .csv<ul style="list-style-type: none">- Aceitou: Correto, começar o próximo critério.- Recusou: Errado, analisar o processamento da planilha enviada.- O nome da planilha corresponde à lista de planilhas esperadas pelo pipeline<ul style="list-style-type: none">- Aceitou: Correto, começar o próximo critério.- Recusou: Errado, analisar nome e fonte da planilha enviada. <p>Critério 2:</p> <ul style="list-style-type: none">- Dados duplicados ou corrompidos foram removidos.<ul style="list-style-type: none">- Aceitou: Correto, começar o próximo critério- Recusou: Errado, tratar os dados <p>Critério 3:</p> <ul style="list-style-type: none">- Os novos dados governamentais foram enviados para atualização do S3.<ul style="list-style-type: none">- Aceitou: Correto, começar o próximo critério- Recusou: Errado, analisar formato e origem dos dados- As colunas das planilhas no S3 não estão traduzidas:<ul style="list-style-type: none">- Aceitou: Errado, revisar tratamento antes do consumo- Recusou: Correto, dados atualizados com sucesso.

Número	US02
Título	Atualizar dados Públicos de CNPJ

Personas	Moisés Aragão
História	Eu como teach lead, quero atualizar os dados de CNPJ, de forma a garantir que os dados utilizados para análise estejam condizentes com o último censo.
Critérios de aceitação	<ol style="list-style-type: none"> 1. Validar formato da planilha .csv 2. Validar valores da planilha .csv 3. Efetuar atualização dos dados em .csv para o S3.
Testes de aceitação	<p>Critério 1:</p> <ul style="list-style-type: none"> - A planilha está no formato .csv <ul style="list-style-type: none"> - Aceitou: Correto, começar o próximo critério. - Recusou: Errado, analisar o processamento da planilha enviada. - O nome da planilha corresponde à lista de planilhas esperadas pelo pipeline <ul style="list-style-type: none"> - Aceitou: Correto, começar o próximo critério. - Recusou: Errado, analisar nome e fonte da planilha enviada. <p>Critério 2:</p> <ul style="list-style-type: none"> - Dados duplicados ou corrompidos foram removidos. <ul style="list-style-type: none"> - Aceitou: Correto, começar o próximo critério - Recusou: Errado, tratar os dados <p>Critério 3:</p> <ul style="list-style-type: none"> - Os novos dados governamentais foram enviados para atualização do S3. <ul style="list-style-type: none"> - Aceitou: Correto, começar o próximo critério - Recusou: Errado, analisar formato e origem dos dados - As colunas das planilhas no S3 não estão traduzidas: <ul style="list-style-type: none"> - Aceitou: Errado, revisar tratamento antes do consumo - Recusou: Correto, dados atualizados com sucesso.

Número	US03
Título	Disponibilizar visualização do infográfico
Personas	Enzo Ananias
História	Eu, como consultor de marketing, quero visualizar o infográfico gerado com base no cubo de dados, para que eu possa ter auxílio na tomada de decisões .

Critérios de aceitação	1. O usuário deve ser capaz de acessar a página do infográfico. 2. O usuário deve ser capaz de exportar o infográfico.
Testes de aceitação	Critério 1: <ul style="list-style-type: none"> - Usuário iniciou o acesso a página do infográfico. <ul style="list-style-type: none"> - Aceitou: Correto, começar o próximo critério - Recusou: Errado, analisar origem do erro e resolvê-lo Critério 2: <ul style="list-style-type: none"> - Usuário efetuou a exportação do infográfico <ul style="list-style-type: none"> - Aceitou: Correto, infográfico exportado com sucesso - Recusou: Errado, analisar origem do erro e resolvê-lo

Número	US04
Título	Atualizar dados do cliente
Personas	Moisés Aragão
História	Eu como teach lead, quero que os dados de vendas do cliente sejam semanalmente atualizados, garantindo que os dados analisados do cliente estejam sempre atualizados.
Critérios de aceitação	1. O usuário consegue visualizar os dados do cliente . 2. Verificar se os dados da API do parceiro estão atualizados 3. Verificar se o consumo ocorreu automaticamente no horário definido.
Testes de aceitação	Critério 1: <ul style="list-style-type: none"> - A requisição foi aceita pela API do parceiro. <ul style="list-style-type: none"> - Aceitou: Correto, começar o próximo critério. - Recusou: Errado, entrar em contato com o parceiro para discutir a disponibilidade da api. Critério 2: <ul style="list-style-type: none"> - Os dados presentes no cubo correspondem com o registro mais recente. <ul style="list-style-type: none"> - Aceitou: Correto, começar o próximo critério - Recusou: Errado, entrar em contato com o parceiro para solicitar revisão dos dados fornecidos. Critério 3: <ul style="list-style-type: none"> - A atualização semanal foi acionada automaticamente.

	<ul style="list-style-type: none"> - Aceitou: Correto, dados foram atualizados com sucesso. - Recusou: Revisar código de automatização
--	--

Número	US5
Título	Atualizar dados consumidos pelo infográfico.
Personas	Enzo Ananias
História	Eu como consultor de marketing, quero que o infográfico reflita a realidade do cubo de dados, de forma a garantir que os dados visualizados correspondam aos últimos dados do cliente.
Crterios de aceitação	<ol style="list-style-type: none"> 1. O usuário deve ter acesso aos dados pelo infográfico de forma correspondente ao cubo. 2. Os dados consumidos devem ser os mais recentes gerados.
Testes de aceitação	<p>Critério 1:</p> <ul style="list-style-type: none"> - Verificar se as informações no cubo e nos infográficos correspondem. <ul style="list-style-type: none"> - consistentes: Correto, começar o próximo critério - inconsistentes: Errado, revisar código de consumo dos dados para geração dos infográficos <p>Critério 2:</p> <ul style="list-style-type: none"> - Verificar se os infográficos foram devidamente atualizados junto ao cubo de dados. <ul style="list-style-type: none"> - Dados recentes: Correto, cubo de dados atualizados - Dados anteriores: Errado, revisar código de consumo dos dados para geração dos infográficos

5. Análise Exploratória e fonte dos dados

5.1 Introdução

Bucket base dos dados

<https://basedosdados.org/dataset/49ace9c8-ae2d-454b-bed9-9b9492a3a642?table=3880670f-eceb-47ec-802b-4579ee62ae3>

Importância dos dados: Os *dados geográficos* brasileiros foram selecionados para entender as demandas regionais.

Formato: CSV

Tamanho: base de dados não operante

Frequência de atualização: trimestralmente (2000-2021)

<https://basedosdados.org/dataset/9fa532fb-5681-4903-b99d-01dc45fd527a?table=4b025d5a-5af0-4fa8-bd04-59de13b378ae>

Importância dos dados: A *Pesquisa Nacional* por Amostra de Domicílios Contínua foi escolhida para entender a situação social das famílias.

Formato: CSV

Tamanho: 14 conjuntos de dados

Frequência de atualização: trimestralmente (2012-2020)

<https://basedosdados.org/dataset/a1b6d2b6-4aa6-47e7-a517-8a21b28b7254?table=7b880731-ffa2-4bde-a290-ae058b3acf51>

Importância dos dados: A *Pesquisa de Orçamentos Familiares* foi escolhida para, complementando os dados geográficos e a PNAD, viabilizar um perfil socioeconômico das famílias brasileiras.

Formato: CSV

Tamanho: 8 conjuntos de dados

Frequência de atualização: trimestralmente (2017-2018)

Bucket dados abertos ibge

<https://dados.gov.br/dados/conjuntos-dados/io-produto-interno-bruto-dos-municipios>

Importância dos dados: O PIB de cada município foi selecionado para adicionar informações sobre concentração de renda.

Formato: HTML; JSON; ODS; XML;

Tamanho: 422 conjuntos de dados

Frequência de atualização: anualmente

<https://dados.gov.br/dados/conjuntos-dados/pc-indice-nacional-de-precos-ao-consumidor-inpc>

Importância dos dados: Índice Nacional de Preços ao Consumidor para demonstrar a variância dos preços.

Formato: HTML; JSON; ODS; XML;

Tamanho: 422 conjuntos de dados

Frequência de atualização: anualmente

Bucket POF IBGE

<https://www.ibge.gov.br/estatisticas/sociais/saude/24786-pesquisa-de-orcamentos-familiares-2.html?=&t=downloads>

Importância dos dados: Pesquisa de Orçamentos Familiares dividido por décadas.

Formato: TXT

Tamanho: 8 conjuntos de dados

Frequência de atualização: anualmente (1987-1988; 1995-1996; 2002-2003; 2008-2009; 2017-2018)

***Bucket* Microdados RAIS e CAGED**

<https://www.gov.br/trabalho-e-emprego/pt-br/assuntos/estatisticas-trabalho/microdados-rais-e-caged>

Importância dos dados: As empresas fornecem a Relação Anual de Informações Sociais e, junto com o Cadastro de Geral de Empregados e Desempregados, são fornecidas informações sobre a população economicamente ativa.

Formato: Indisponível

Tamanho: Indisponível

Frequência de atualização: Indisponível

***Bucket* Receita Federal**

<https://dados.gov.br/dados/conjuntos-dados/resultado-da-arrecadacao>

Importância dos dados: Resultado da arrecadação com objetivo de sintetizar informações da economia brasileira.

Formato: CSV; PDF

Tamanho: 134 conjuntos de dados

Frequência de atualização: diariamente

<https://dados.gov.br/dados/conjuntos-dados/repasses-da-arrecadacao-federal>

Importância dos dados: Repasses da união aos estados.

Formato: CSV; PDF; XLSX

Tamanho: 134 conjuntos de dados

Frequência de atualização: diariamente

<https://dados.gov.br/dados/organizacoes/visualizar/ministerio-da-fazenda>

Importância dos dados: Dados relativos a título e tesouro.

Formato: CSV; PDF; XLSX; JSON

Tamanho: 134 conjuntos de dados

Frequência de atualização: diariamente

Bucket INEP

<https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/ana>

Importância dos dados: A Avaliação Nacional de Alfabetização adiciona informações sobre a educação básica da população.

Formato: CSV

Tamanho: 3,300 KB

Frequência de atualização: anualmente

<https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/censo-da-educacao-superior>

Importância dos dados: Evolução anual da educação brasileira (1995-2022)

Formato: CSV

Tamanho: aproximadamente 8 GB

Frequência de atualização: anualmente (1995-2022)

<https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/encceja>

Importância dos dados: Microdados do Exame Nacional para Certificação de Competências de Jovens e Adultos

Formato: CSV

Tamanho: 0.6 GB

Frequência de atualização: anualmente (2014, 2017-2020, 2022)

Arquivo:

<https://drive.google.com/drive/folders/1GDP0qcn8XWYdv0c1w46EK2eV013yXShi?usp=sharing>

Bucket SUS

<https://opendatasus.saude.gov.br/dataset/cnes-cadastro-nacional-de-estabelecimentos-de-saude>

Importância dos dados: Cadastro Nacional de Estabelecimentos de Saúde

Formato: JSON

Tamanho: 832 MB

Frequência de atualização: diariamente

Bucket POF_parceiro

https://drive.google.com/drive/folders/1kB92-Q_pDSIZ4YTxDGy3ygxpylKA19mwJ

Importância dos dados: Pesquisa de Orçamentos Familiares disponibilizado pelo parceiro

Formato: CSV

Tamanho: 900 MB

Frequência de atualização: anualmente

Bucket cnpj

<https://drive.google.com/drive/folders/1hxU0AdkgvT23FRGNp4htNs45WC5mYC1S>

Importância dos dados: CNPJ's das empresas, tanto as que compram quanto as que vendem alimentos

Formato: CSV

Tamanho: 4,582 GB

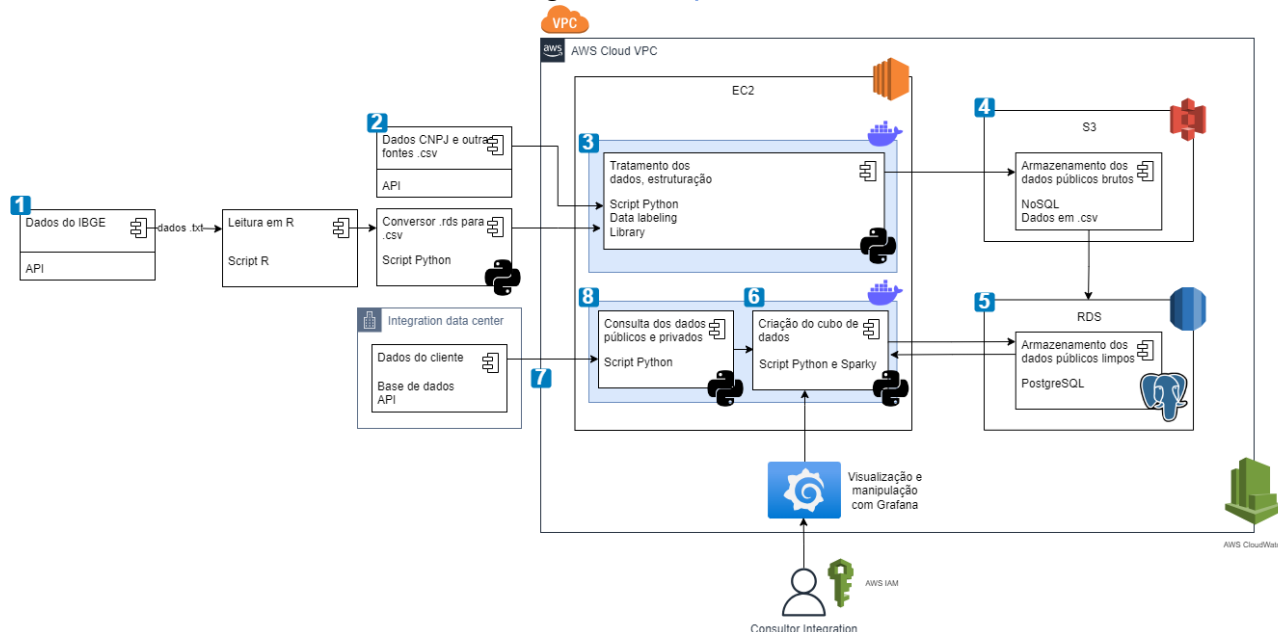
Frequência de atualização: anualmente

5.2 Método

6. Arquitetura Macro

Para a execução do projeto foi necessário a definição de uma arquitetura cloud que será guia para a construção do ambiente na cloud da AWS e definição das tecnologias a serem usadas. O esquema abaixo corresponde a arquitetura realizada na Sprint 1 para apresentação e alinhamento de expectativas com os parceiros de projeto, assim também é esperado uma dinâmica de iteração durante o decorrer das Sprints.

Imagem 1: [Arquitetura](#)



Fonte: Dados dos autores (2023)

Abaixo se encontra uma análise da arquitetura construída:

6.1 Identificação dos dados:

6.1.1 Dados públicos:

São fornecidos pelos canais de pesquisa do governo, tendo como mais frequente o IBGE, através de arquivos .txt em uma pasta no início do projeto. De partida, os dados não apresentavam-se estruturados, sendo fornecido um script em R para leitura desses dados e a criação de tabelas que possibilitam a compreensão. Após a leitura os dados serão convertidos para .rds e posteriormente para .csv e passados para o EC2. Nos dados públicos também se encaixam os dados referentes aos CNPJ no Brasil, que foram fornecidos diretamente em .csv.

Conteúdo: público, dados demográficos coletados pelo governo, que englobam fatores como condição financeira dos moradores de cada região, hábitos de consumos alimentares, despesas, aluguel dentre outros, organizados por fatores como faixa etária,

faixa de renda e região. Os dados de Pesquisa de Orçamentos Familiares (POF) contemplam os seguintes tópicos:

- Primeiros Resultados, que apresenta informações sobre despesas e rendimentos das famílias.
- Avaliação nutricional da disponibilidade domiciliar de alimentos no Brasil.
- Análise do consumo alimentar pessoal no Brasil.
- Análise da Segurança Alimentar no Brasil.
- Perfil das Despesas no Brasil: indicadores selecionados.
- Perfil das Despesas no Brasil: indicadores selecionados de alimentação, transporte, lazer e inclusão financeira.
- Perfil das Despesas no Brasil: indicadores de qualidade de vida.
- Evolução dos indicadores de qualidade de vida no Brasil com base na Pesquisa de Orçamentos Familiares.

Tabelas: Os dados se distribuem em tabelas de acordo com o seu conteúdo, sendo as fornecidas pelo parceiro:

Tabela	Descrição
Aluguel Estimado	Pesquisa sobre os custos estimados de aluguel no Brasil.
Caderneta - Coletiva	Levantamento de informações financeiras coletivas em cadernetas.
Características - Dieta	Estudo das características da dieta alimentar da população.
Condições - Vida	Análise das condições de vida no país.
Consumo - Alimentar	Investigação sobre o consumo de alimentos na sociedade.
Despesa - Coletiva	Coleta de dados sobre despesas coletivas.
Despesa - Individual	Coleta de dados sobre despesas individuais.
Domicílio	Levantamento de informações relacionadas a residências.
Inventário	Pesquisa sobre o inventário de bens e recursos.
Morador - Qualidade - Vida	Avaliação da qualidade de vida dos moradores.
Morador	Coleta de dados sobre os habitantes de uma área específica.
Outros - Rendimentos	Pesquisa de rendimentos não especificados.
Rendimento - Trabalho	Estudo dos rendimentos provenientes do trabalho.
Restrição - Produtos - Serviços - Saúde	Avaliação das restrições no acesso a produtos e serviços de saúde.
Serviço não monetário - POF2	Pesquisa sobre serviços não monetários na Pesquisa de Orçamento Familiar (POF) 2.
Serviço não monetário - POF4	Pesquisa sobre serviços não monetários na Pesquisa de Orçamento Familiar (POF) 4.

6.1.2 Dados do cliente:

Os dados referentes ao cliente serão fornecidos através de uma API hospedada pela Integration, esses dados contém informações privadas sobre os clientes da Integration, portanto, são de natureza sigilosa, não permitindo ficar em posse da Integration após o projeto. Esses dados são obtidos através de queries realizadas pela API da Integration.

Conteúdo: Sigiloso, o conteúdo desses dados engloba informações específicas sobre o parceiro, compreendendo dados como receita gerada por um produto, receita regional, número de vendas, lista de produtos, CNPJ de clientes dentre outros. Esses dados devem variar de cliente para cliente e refletem sua situação no mercado. A solução precisa estar preparada para receber diferentes tipos de dados e estruturas, pois eles podem variar de cliente para cliente.

6.2 Gestão de dados:

6.2.1 Dados públicos:

Ingestão: Os dados são coletados por canais públicos como do Instituto Brasileiro de Geografia e Estatística - IBGE e outros órgãos públicos. Através de uma API, fornecida pelo governo, podem-se ser obtidos em formato .txt, sendo necessário passar por scripts em R para transformá-los em tabelas legíveis e distinguíveis. Esses dados passarão por uma pipeline em python na EC2, juntamente com os dados referentes aos CNPJ, e depois carregados em sua totalidade no S3. Após isso, os dados relevantes para os projetos atuais da empresa devem ser transferidos para o RDS, mantendo os dados mais importantes lá. Os quais posteriormente serão consumidos no EC2 por um código python, o qual será processado com o sparky, e também será parte do cubo de dados.

frequência: A atualização desses dados deve ser feita anualmente, acompanhando a velocidade a qual os censos são atualizados.

quantidade: Espera-se que quando atualizados será em grande quantidade, pois a frequência de atualização não será tão alta, e os dados do governo englobam milhões de pessoas.

6.2.2 Dados privados:

Já para os dados privados os dados serão ingeridos a partir da Api que fornece os dados do cliente em questão para cada projeto. A Api será consultada por um código python rodando no EC2, e posteriormente carregado, em um código python, também na

EC2, utilizando sparky para uma manipulação dos dados. Eles também serão utilizados para a criação de um cubo de dados.

frequência: A atualização desses dados deve ser feita diariamente, gerenciada por um airflow para essa execução, a fim de manter os dados sempre atualizados em relação aos dados fornecidos na api do parceiro.

quantidade: A quantidade de dados, embora grande, será recebida em um quantidade média, pois serão atualizados com uma frequência maior.

6.3 Seleção dos serviços AWS:

Tratamento/processamento de dados:

EC2: realizaremos o tratamento e processamento dos dados pela máquina virtual do Amazon EC2 (Elastic Compute Cloud), através de códigos python. O EC2 é um serviço de nuvem da Amazon que oferece servidores virtuais para hospedar os recursos computacionais para a pipeline de tratamento.

Serão no total quatro códigos alocados na EC2:

Tratamento de dados públicos: Esse código receberá os dados públicos, os processará e os enviará para o AWS S3.

Consulta da api/RDS: Esse código será responsável por consultar a api do parceiro e os códigos guardados no RDS, criar um cubo de dados com essas informações e, passar as informações para o código de tratamento/transformação de dados.

Código de tratamento/ transformação de dados com sparky: Esse código será responsável pelo processamento dos tanto do cliente tanto dos dados públicos, visando alcançar as informações solicitadas pelo parceiro.

Código para o infográfico: Esse código será responsável por criar infográficos com base nos dados fornecidos, com objetivos de fornecer possíveis insights sobre os dados.

Armazenamento:

S3: Por aspectos do pipeline de big data como volume e custos, será utilizado o , serviço que permite o armazenamento de dados de forma flexível e não estruturada. Nele se espera comportar todos os dados públicos com alto volume histórico de datação prezando pela economia em espaço e processamento alcançáveis por uma base de dados não relacional.

RDS: Posteriormente, após a seleção dos dados de interesse para as análises de mercado e negócios, será utilizado o AWS RDS para hospedar um banco de dados relacional Postgre, o qual receberá os dados de mais importância que estavam no S3. Será restrito ao último os dados desejados para a elaboração de tabelas e informações de negócios na próxima etapa de tratamento.

Segurança de dados:

VPC: Possibilita a criação de seções na nuvem da aws, sendo possível controlar o acesso aos serviços da AWS utilizados, criar restrições de acesso e regras de segurança. Sendo possível garantir uma segurança maior à solução.

Monitoramento:

Amazon Cloudwatch: Será utilizado o AWS Cloud Watch para o monitoramento dos serviços mencionados anteriormente, EC2 e RDS. Visualizando informações sobre eles e o status em tempo real.

6.4 Fluxo dos dados:

Os dados possuem diferentes fluxos, mas se conectam futuramente no EC2, segue abaixo os fluxos dos dados até se juntarem:

Dados públicos: Os dados públicos serão retirados dos sites do governo em formato .txt, após isso os dados terão que passar por um Script R para gerar arquivos .rds a partir deles e após isso por um script python para conversão em .csv. Após isso eles serão inseridos no EC2, em um notebook que irá tratar os dados e os enviar para o S3 em sua totalidade. Os dados mais relevantes atualmente para empresa serão selecionados e enviados para o RDS e por fim enviados de volta para o EC2 para criação do cubo de dados.

Dados privados: Os dados privados serão solicitados diretamente na api fornecida pelo parceiro, esse consumo será realizado em código python localizado na EC2 e utilizados para criação do cubo de dados.

Cubo de Dados: Após a formação inicial do cubo de dados os dados passarão por outro código python no EC2, sendo processados e agregados de forma a gerar as informações que o cliente deseja para gerar um relatório. E por fim serão enviados para outro código python no EC2, o qual irá gerar infográficos a partir desses dados.

6.5 Segurança:

Como mencionado anteriormente, uma das ferramentas utilizadas para garantir a segurança do projeto foi o Amazon VPC, que será responsável pelo controle de acesso, criação de restrições e regras de segurança no ambiente da AWS.

Mas além dos cuidados com a segurança da AWS, também foram tomados cuidado com os dados, uma vez que apenas dados públicos serão salvos na base de dados, evitando assim que os dados privados do parceiro obtidos através da api tenham risco de vazarem.

6.6 Monitoramento e gerenciamento:

Como mencionado anteriormente na parte de serviços será utilizado o Amazon CloudWatch para o monitoramento da solução. Dentre os tópicos monitorados vale destacar:

Gasto de recursos: Para garantir que as ferramentas da AWS estejam utilizando os recursos esperados, de forma a evitar gastos indesejados.

Disponibilidade: Verificar a disponibilidade dos serviços, de forma a garantir o funcionamento da solução.

Informações: Monitoramento do que está ocorrendo nas ferramentas da solução, de forma a garantir que o fluxo esteja ocorrendo como planejado.

7. Data lake

7.1 Introdução

Um *Data Lake* é um sistema centralizado que permite armazenar grandes volumes de dados em seu formato natural, seja estruturado, semi-estruturado ou não estruturado. A ideia é que você pode despejar dados de diferentes fontes dentro deste lago e eles estarão disponíveis para análise e processamento com grande flexibilidade.

No nosso caso, o objetivo de criarmos uma pipeline de dados para a Integration, faz com que exista a necessidade da criação de *um data lake*, uma vez que estamos trabalhando com um grande volume de dados e de diferentes fontes. Para isso utilizaremos o S3.

7.2 S3

O S3, significa Simple Storage Solution, e a escolhemos por várias razões. Lançado em 2006, ele oferece um armazenamento de objetos dentro de uma estrutura de pastas que é extremamente eficaz em custo, começando em apenas \$0,023 por gigabyte. Além disso, quanto mais você armazena, mais o custo diminui, sem limitações de capacidade.

Ele proporciona uma maneira barata e confiável de armazenar objetos com acesso de baixa latência e alta capacidade de transferência através do conteúdo do seu *bucket*. Um *Bucket* é um container da AWS S3 para armazenamento de diversas opções de arquivos. Outro ponto forte é a sua integração com serviços como SNS, SQS e Lambda, possibilitando aplicações poderosas orientadas por eventos, podendo utilizar Lambda através de um trigger. Além disso, o S3 oferece mecanismos para transferir dados antigos para armazenamento de longo prazo, reduzindo ainda mais os custos.

7.3 Buckets

O processo para utilizar o S3 é simples. Primeiro, cria-se um *bucket*. Depois, os arquivos são carregados para este *bucket*, com configurações personalizáveis (as quais explicamos a configuração abaixo). É importante notar que, embora o S3 utilize o termo "pastas", ele opera com uma estrutura de "flat file structure". Isso significa que não é como as pastas do Windows; o nome do arquivo é anexado ao nome do arquivo como um prefixo, o que otimiza a organização e o acesso aos dados.


Os *buckets* como dito anteriormente é um instrumento da AWS S3 que se assemelha com um container do docker. Nele você pode fazer o upload de arquivos, como JSON, PNG, HTML, CSV, etc, no caso de nosso projeto o formato escolhido foi o .csv. Se acessado na nuvem através de seu link de acesso, que pode ser definido geograficamente na própria configuração do *Bucket*. Para isso criamos 10 *buckets* que seguem a seguinte configuração.

7.3.1 Configuração dos Buckets

Passo 1: Configuração inicial

Nome do bucket

nome-exemplo

O nome do bucket deve ser exclusivo no namespace global e seguir as regras de nomenclatura do bucket. [Veja as regras para nomenclatura de buckets](#) 

Região da AWS

América do Sul (São Paulo) sa-east-1

Para começar a configuração damos o nome do bucket, (ele tem que ser único nos servidores da AWS) e a localização. Por estarmos fazendo na AWS Workbench não pudemos hospedar nosso bucket em São Paulo, mas esse seria o recomendado.

Passo 2: ACL

Propriedade de objeto [Informações](#)

Controle a propriedade de objetos gravados nesse bucket a partir de outras contas da AWS e o uso de listas de controle de acesso (ACLs). A propriedade do objeto determina quem pode especificar o acesso aos objetos.

☒ ACLs desabilitadas (recomendado)

Todos os objetos nesse bucket são de propriedade dessa conta. O acesso a esse bucket e seus objetos é especificado usando apenas políticas.

☐ ACLs habilitadas

Os objetos nesse bucket podem ser de propriedade de outras contas da AWS. O acesso a esse bucket e seus objetos pode ser especificado usando ACLs.

Desabilitamos ACLs no bucket S3 para simplificar a gestão de permissões.

Passo 3: Bloqueios e privacidade

Configurações de bloqueio do acesso público deste bucket

O acesso público é concedido a buckets e objetos por meio de listas de controle de acesso (ACLs), políticas de bucket, políticas de ponto de acesso ou todas elas. Para garantir que o acesso público a este bucket e todos os seus objetos seja bloqueado, ative a opção de Bloquear todo o acesso público. Essas configurações serão aplicadas apenas a este bucket e aos respectivos pontos de acesso. A AWS recomenda ativar a opção Bloquear todo o acesso público. Porém, antes de aplicar qualquer uma dessas configurações, verifique se as aplicações funcionarão corretamente sem acesso público. Caso precise de algum nível de acesso público a este bucket ou aos objetos que ele contém, é possível personalizar as configurações individuais abaixo para que atendam aos seus casos de uso de armazenamento específicos. [Saiba mais](#)

☒ Bloquear *todo* o acesso público

Ativar essa configuração é o mesmo que ativar todas as quatro configurações abaixo. Cada uma das configurações a seguir são independentes uma da outra.

☒ Bloquear acesso público a buckets e objetos concedidos por meio de *novas* listas de controle de acesso (ACLs)

O S3 bloqueará as permissões de acesso público aplicadas a blocos ou objetos recém-adicionados e impedirá a criação de novas ACLs de acesso público para blocos e objetos existentes. Essa configuração não altera nenhuma permissão existente que permita o acesso público aos recursos do S3 usando ACLs.

☒ Bloquear acesso público a buckets e objetos concedidos por meio de *qualquer* lista de controle de acesso (ACLs)

O S3 ignorará todas as ACLs que concedem acesso público a buckets e objetos.

☒ Bloquear acesso público a buckets e objetos concedidos por meio de *novas* políticas de ponto de acesso e bucket público

O S3 bloqueará novas políticas de bucket e ponto de acesso que concedem acesso público a buckets e objetos. Essa configuração não altera nenhuma política existente que permita o acesso público aos recursos do S3.

☒ Bloquear acesso público e entre contas a buckets e objetos por meio de *qualquer* política de bucket ou ponto de acesso público

O S3 ignorará o acesso público e entre contas para buckets ou pontos de acesso com políticas que concedem acesso público a buckets e objetos.

Para a prova de conceito, optamos por bloquear todo o acesso público ao bucket S3, garantindo segurança e simplicidade na gestão de permissões. A Integration, em um contexto de produção, deve considerar políticas específicas, possivelmente reabilitando ACLs para acesso detalhado, conforme a necessidade do cliente ou a sensibilidade dos dados.

Passo 4: Versionamento de bucket

Versionamento de bucket

O versionamento é um meio de manter múltiplas variantes de um objeto no mesmo bucket. Você pode usar o versionamento para preservar, recuperar e restaurar todas as versões de cada objeto armazenado no bucket do Amazon S3. Com o versionamento, você pode recuperar facilmente ações não intencionais do usuário e falhas da aplicação. [Saiba mais](#)

Versionamento de bucket

- ☒ Desativar
☐ Ativar

Decidimos não ativar o versionamento do *bucket*, considerando a natureza dos dados que estamos manuseando. Como são dados atualizados anualmente, não há necessidade de manter múltiplas versões ao longo do ano, o que poderia resultar em **custos desnecessários**.

Passo 5: Criptografia

Criptografia padrão [Informações](#)

A criptografia no lado do servidor é aplicada automaticamente a novos objetos armazenados nesse bucket.

Tipo de criptografia [Informações](#)

- ☒ Criptografia do lado do servidor com chaves gerenciadas do Amazon S3 (SSE-S3)
- ☐ Criptografia do lado do servidor com chaves do AWS Key Management Service (SSE-KMS)
- ☐ Criptografia de duas camadas no lado do servidor com chaves do AWS Key Management Service (DSSE-KMS)
Proteja seus objetos com duas camadas separadas de criptografia. Para obter detalhes sobre a precificação, consulte os [preços do DSSE-KMS](#) na guia [Armazenamento](#) da [página de preços do Amazon S3](#). [↗](#)

Chave do bucket

O uso de uma chave de bucket do S3 para SSE-KMS reduz os custos de criptografia ao diminuir as chamadas para o AWS KMS. As chaves de bucket do S3 não são compatíveis com o DSSE-KMS. [Saiba mais](#) [↗](#)

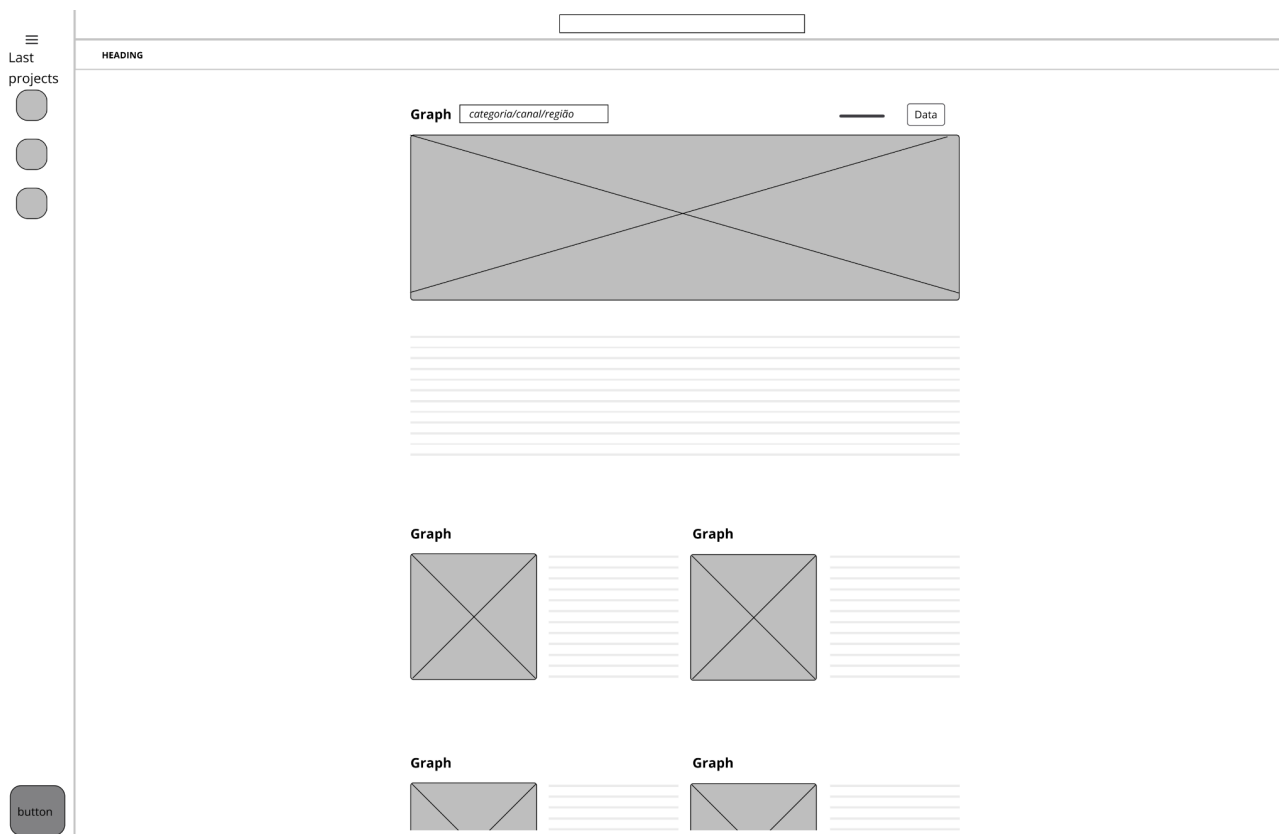
- ☐ Desativar
- ☒ Ativar

Para essa prova de conceito, optamos pela criptografia do lado do servidor fornecida pela própria AWS S3, a SSE-S3. Isso nos dá segurança de forma prática, sem a necessidade de gerenciar as chaves de criptografia manualmente, o que seria necessário se escolhêssemos outras opções como SSE-KMS ou DSSE-KMS.

Para a equipe da Integration, vale apontar que a escolha pelo SSE-S3 é totalmente adequada para essa fase inicial, onde estamos tratando de uma prova de conceito e de dados públicos. No entanto, quando o projeto tomar forma e começar a lidar com um volume maior de dados ou informações mais sensíveis, aí sim faria sentido pensar em algo mais robusto como o SSE-KMS ou DSSE-KMS. Essas opções trazem benefícios adicionais, como um maior controle e auditoria sobre quem está acessando as chaves de criptografia.

8. Wireframe

Um wireframe de baixa fidelidade é uma representação esquemática e simplificada de uma interface ou design de um produto digital. Ele é geralmente criado no estágio inicial do processo de design para esboçar a estrutura básica e a disposição dos elementos, sem detalhes gráficos ou estilísticos. O foco está na funcionalidade e na organização da informação, permitindo uma avaliação rápida e fácil das ideias e conceitos do design. O wireframe da sprint 1 serviu de base para a validação com o parceiro de negócios e primeiro passo para suas posteriores iterações.



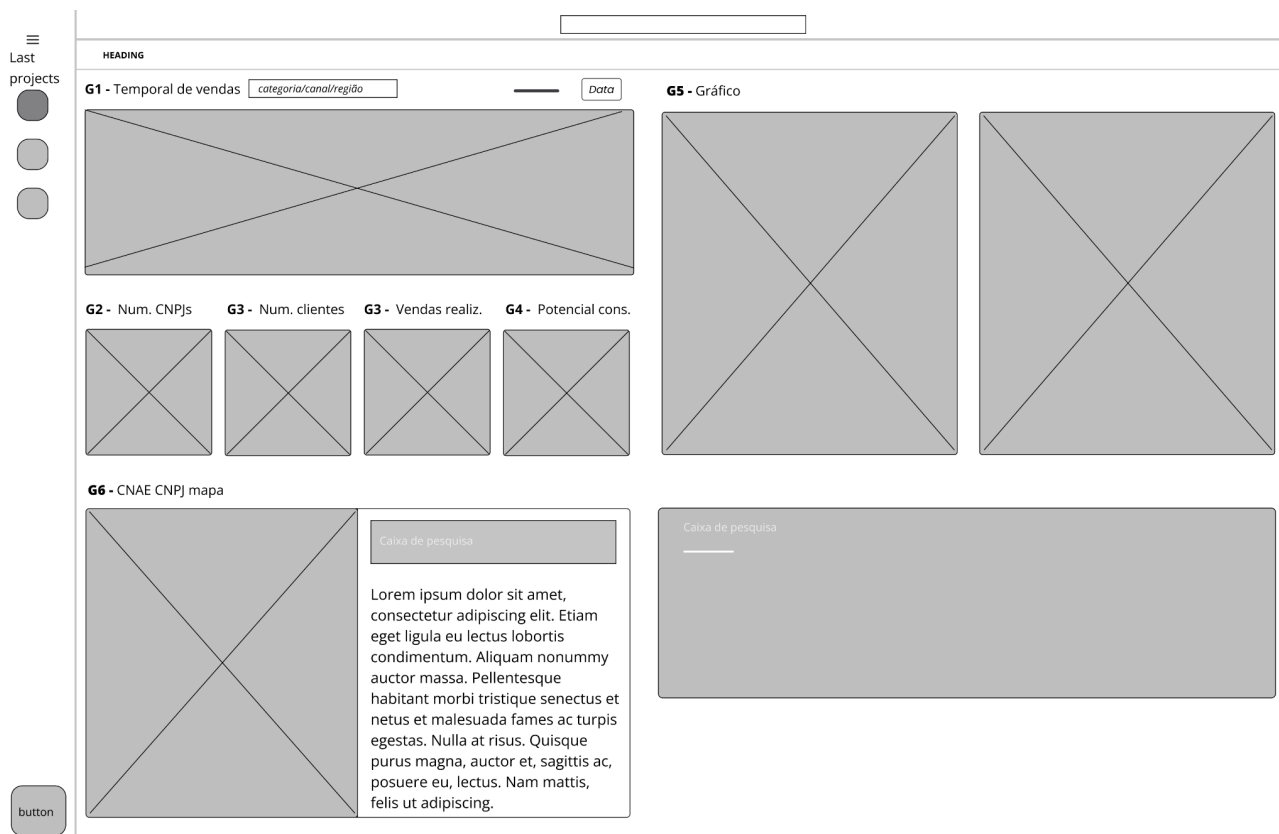
Tela do infográfico do wireframe - desenvolvido pelo autor

8.1 Telas

O objetivo estimado com as interfaces do wireframe é uma visualização fidedigna dos dados provenientes do pipeline de Big Data, contribuindo na tomada de decisão dos consultores, assim como o fácil entendimento das métricas dispostas. Para que esse objetivo seja alcançado dividimos a nossa aplicação em duas telas, o dashboard e o infográfico.

8.2 Dashboard

O Dashboard abriga a visualização inicial, onde temos um sidebar para a exposição de projetos anteriores, assim o consultor tem o fácil acesso como um hub a insights, antes gerados, que poderão ser incorporados ao próximo atendimento de consultoria. Posteriormente a seleção do projeto, abriga-se a tela principal, que apresentará pelo menos 8 gráficos distintos, cada um destacando diferentes informações do mercado ou do cliente. Segmentamos a tela estrategicamente a fim de uma melhor experiência do usuário, contendo áreas destinadas somente a dados do cliente, outra a dados do mercado e região e por último a pesquisa por setor. Finalmente, a Sidebar disponibilizará um botão de acesso ao Infográfico.



Tela do dashboard do wireframe - desenvolvido pelo autor

8.2 Proposta de valor

Para atender aos requisitos do cliente no desenvolvimento do cubo de dados, o principal objetivo é obter e integrar diversas informações essenciais. Primeiramente, é necessário ter acesso ao número de CNPJs por canal e região, assim como os dados de consumo por categoria, canal e região. Além disso, é fundamental obter informações sobre as vendas por CNPJ para cada categoria e região, bem como o mapeamento entre os Canais e os códigos CNAE. Com base nesses dados, o objetivo final é criar um cubo que permita analisar o potencial de mercado de forma mais automatizada, desdobrando para cada combinação de categoria, canal e região: o número total de CNPJs, o número de clientes atendidos, as vendas realizadas e o potencial de consumo.

É importante ressaltar que, ao final do projeto, a entrega não consistirá em uma solução pronta, como um modelo preditivo, mas sim na estruturação e engenharia de dados necessárias para viabilizar uma solução que será implementada pela empresa no futuro. Portanto, o foco está na criação de uma infraestrutura sólida e eficiente para suportar a análise de mercado de forma automatizada e precisa.

8.3 Justificativa das Escolhas de Design

A justificativa das escolhas de design para essa tela é fundamentada na necessidade de proporcionar ao usuário uma experiência intuitiva e eficiente na visualização e interação com os dados apresentados. A presença de três mapas - representando respectivamente a evolução temporal de vendas por canal, a distribuição geográfica dos canais e o relacionamento entre CNAE e CNPJ - foi determinada com base na relevância e na importância estratégica dessas informações para o usuário.

Ao disponibilizar uma caixa de pesquisa para setores, buscamos facilitar a navegação e a localização específica de informações. Essa funcionalidade adiciona um componente de personalização e agilidade à experiência do usuário, permitindo-lhe encontrar dados pertinentes de maneira rápida e direta. A apresentação de opções selecionáveis após a escolha de um setor aprimora a interatividade, proporcionando uma abordagem mais granular e personalizada à exploração dos dados.

A lógica por trás dessas visualizações está ancorada na hierarquia de informações e na facilidade de acesso às diferentes perspectivas dos dados. O mapa temporal de vendas por canal fornece insights sobre tendências ao longo do tempo, enquanto o mapa de canais oferece uma visão espacial das operações. O mapa CNAE-CNPJ, por sua vez, contribui para uma compreensão mais detalhada das relações entre os códigos de atividade e os registros de empresas.

O feedback coletado, seja por meio de testes de usabilidade ou através da análise de dados de interação, foi crucial para refinar e validar as escolhas de design. A partir dessas informações, foi possível ajustar aspectos como a disposição dos elementos na tela, a clareza das indicações visuais e a funcionalidade da caixa de pesquisa. Dessa forma, as decisões de design foram iterativamente refinadas para otimizar a experiência do usuário e garantir que a interface atenda de maneira eficaz às necessidades e expectativas dos usuários finais.

8.4 Técnicas Avançadas

Técnicas avançadas de design de interface de usuário, como a antecipação de ações do usuário e micro-interações, desempenham um papel crucial na criação de experiências digitais envolventes e intuitivas. É válido enfatizar como essas técnicas podem ser aplicadas, incluindo o uso do grid como uma ferramenta importante.

A antecipação de ação do usuário refere-se à capacidade do design de prever e responder às ações prováveis do usuário antes mesmo que elas ocorram. Isso é alcançado através de pistas visuais, animações sutis ou elementos interativos que sugerem o próximo passo lógico. Por exemplo, ao passar o mouse sobre um botão, uma sutil mudança de cor ou sombra pode indicar que o botão é clicável. Essa antecipação ajuda a guiar o usuário de forma intuitiva e a reduzir a fricção na interação.

As micro-interações são pequenas animações ou feedbacks visuais que ocorrem em resposta a uma ação do usuário. Elas fornecem feedback imediato e contextual, tornando a experiência mais dinâmica e envolvente. Por exemplo, ao pressionar um botão de envio em um formulário, uma micro-interação pode exibir uma animação indicando que o formulário está sendo processado. Isso ajuda a manter o usuário informado e engajado durante o processo.

O grid é uma ferramenta fundamental no design de interfaces, mesmo em um nível mais avançado. Ele proporciona uma estrutura organizada e consistente para o layout de elementos na tela. Ao utilizar um grid, é possível alinhar e distribuir elementos de forma equilibrada, garantindo uma aparência profissional e coesa. Além disso, um grid bem elaborado facilita a adaptação do design para diferentes tamanhos de tela e dispositivos, contribuindo para uma experiência responsiva.

Ao combinar essas técnicas, designers podem criar interfaces que não apenas sejam visualmente atraentes, mas também altamente funcionais e intuitivas. A antecipação de ação do usuário e as micro-interações melhoram a usabilidade ao proporcionar feedback imediato e orientação, enquanto o uso eficaz do grid assegura uma organização e alinhamento precisos dos elementos na tela. Essas práticas avançadas resultam em experiências de usuário mais envolventes e eficazes, proporcionando uma vantagem competitiva no desenvolvimento de produtos digitais.

8.5 Feedback e Iterações

O processo de feedback e iterações é essencial no desenvolvimento de qualquer projeto de design, pois permite aprimorar a experiência do usuário com base nas percepções e necessidades reais dos usuários. No contexto da tela que estamos discutindo, o feedback coletado desempenhou um papel crucial na evolução do design e nas decisões tomadas. Inicialmente, o feedback foi obtido por meio de diferentes fontes, como testes de usabilidade, pesquisas de satisfação do usuário e análises de interação. Essas fontes

forneceram informações valiosas sobre a forma como os usuários interagem com a tela, quais elementos eram mais relevantes para eles e quais pontos podiam ser aprimorados.

A partir desse feedback, identificaram-se áreas de melhoria, como a necessidade de uma funcionalidade de pesquisa para facilitar a localização de setores específicos. Esse elemento foi posteriormente incorporado ao design para proporcionar uma experiência mais personalizada e eficiente.

Além disso, observou-se a importância de tornar as visualizações de dados mais claras e intuitivas. Com base nesse insight, foram feitos ajustes na apresentação dos mapas e na forma como as informações eram dispostas na tela, garantindo uma compreensão mais rápida e precisa.

O feedback também influenciou a decisão de adicionar a funcionalidade de seleção de opções após a escolha de um setor na caixa de pesquisa. Isso foi uma resposta direta à necessidade de proporcionar aos usuários uma experiência de navegação mais fluida e direcionada.

Em resumo, o feedback coletado desempenhou um papel fundamental na orientação das decisões de design e nas iterações realizadas. Ele proporcionou insights valiosos sobre as preferências e comportamentos dos usuários, permitindo que o design fosse ajustado de forma a atender de maneira mais eficaz às necessidades e expectativas dos usuários finais. Esse processo iterativo resultou em uma interface mais refinada e intuitiva, proporcionando uma experiência de usuário mais satisfatória e eficiente.

9. Conclusões

10. Referências

11. Anexos