



ClearSale

intelligence to move\_

# sumário

1. Backtest: O que é
2. Backtest: Problemas Atuais
3. Discovery: Validação de hipóteses
4. Backtest: Proposta de MVP

[Material Completo](#)

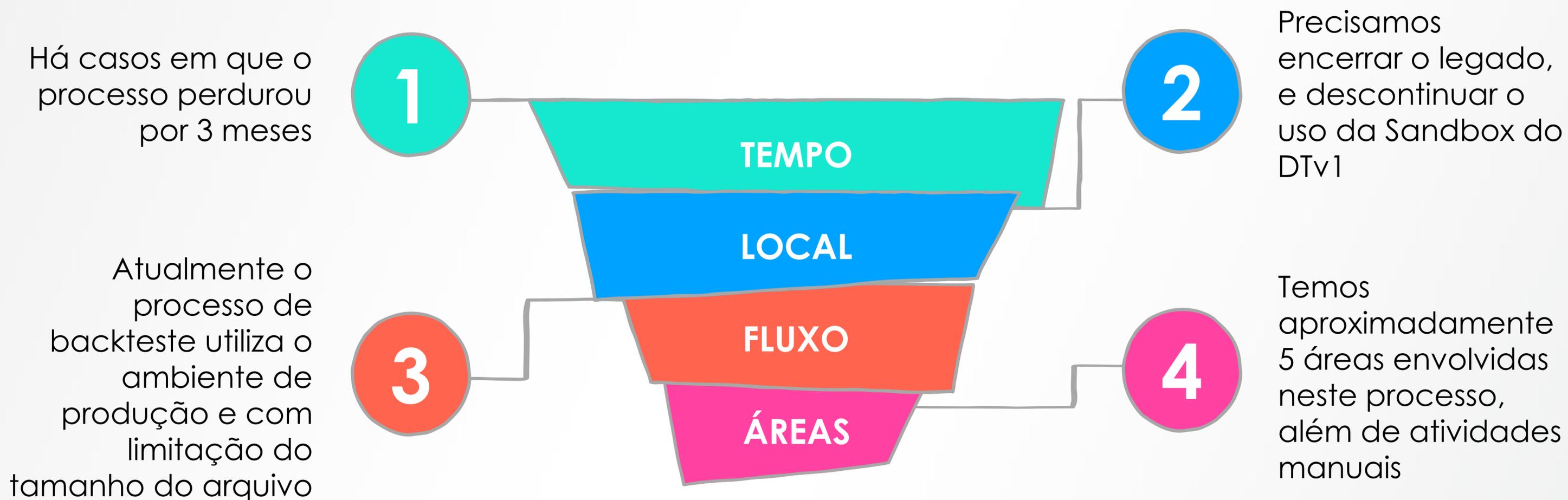
# Backtest

## O que é



# Backtest

## Problemas Atuais



# Backtest: Discovery

## Validando Hipóteses

### HISTÓRICO

Durante o mês de julho realizamos entrevistas com pessoas/áreas envolvidas no processo



### ANÁLISES

Realizamos algumas análises de cenário atual e possibilidades



### DADOS

Levantamos informações sobre processos atuais

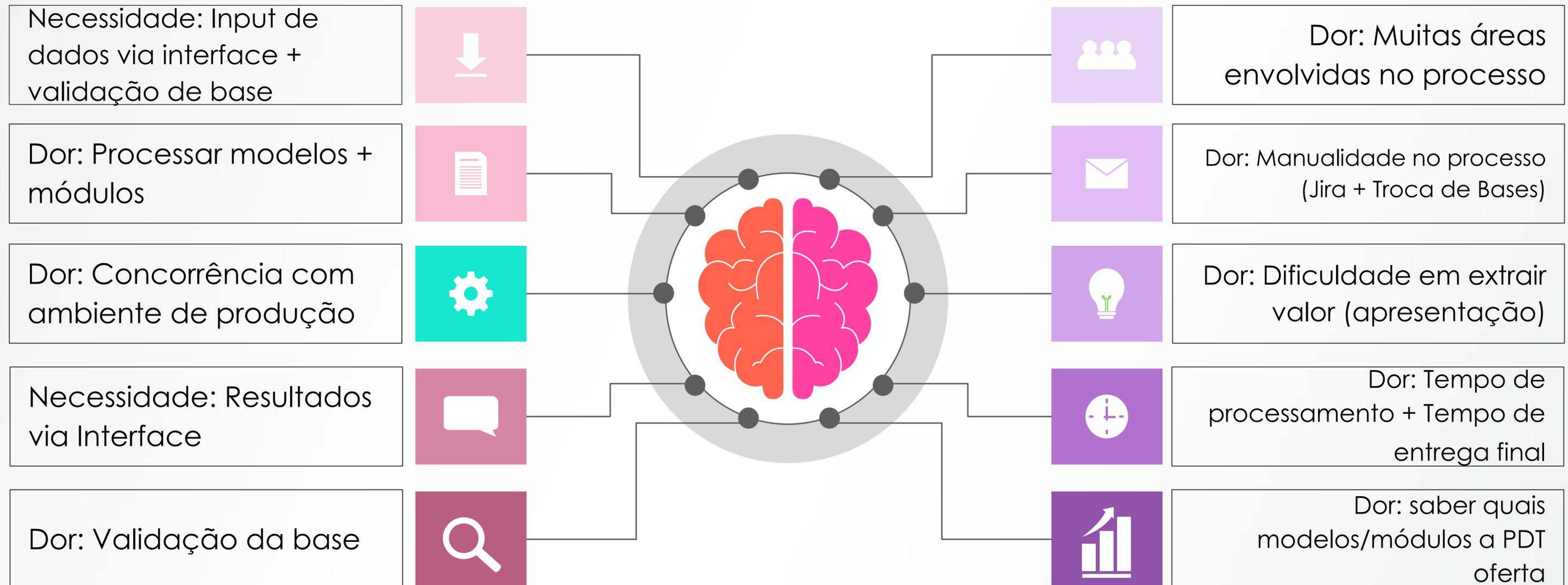


### PROPOSTA

A partir do levantamento realizado, realizamos uma proposta de MVP e possibilidades



# Backtest: Discovery Insights





01

## Problema

1. Tempo entre solicitação e entrega. 2. Concorrência com ambiente de produção e utilização do legado. 3. Limitação do processamento (20k).



02

## Objetivo

Tratando como uma feature da plataforma DT. Ganhar tempo automatizando. Escalabilidade. Processo padronizado e self service, pois não depende apenas do time de BT - qualquer analista consegue realizar o processo e estudos



03

## Oportunidades

Time de Pré-Sales conseguir tomar decisão de maneira mais rápida e assertiva. Auxiliar time Comercial em UpSell. Apresentar resultados de maneira visual, auxiliando times na construção de materiais para pitch de vendas



04

## Como medimos o sucesso

Redução do tempo. Escalabilidade. Automatização

# Backtest

## Proposta de MVP



### BASE

- Implementar infraestrutura básica para o backtest, incluindo a integração com fontes de dados e configurações iniciais.
- Desenvolver uma interface de usuário simples para configurar e iniciar um backtest básico.
- Validação da qualidade dos dados (quantidade de nulls e dados inválidos para cada campo recebido).



### FLEXIBILIDADE

- Permitir a personalização de parâmetros e estratégias de backtest para atender às necessidades específicas dos usuários (módulos + modelos ASP).
- Permitir upload de arquivo acima de 20k linhas.



### USABILIDADE

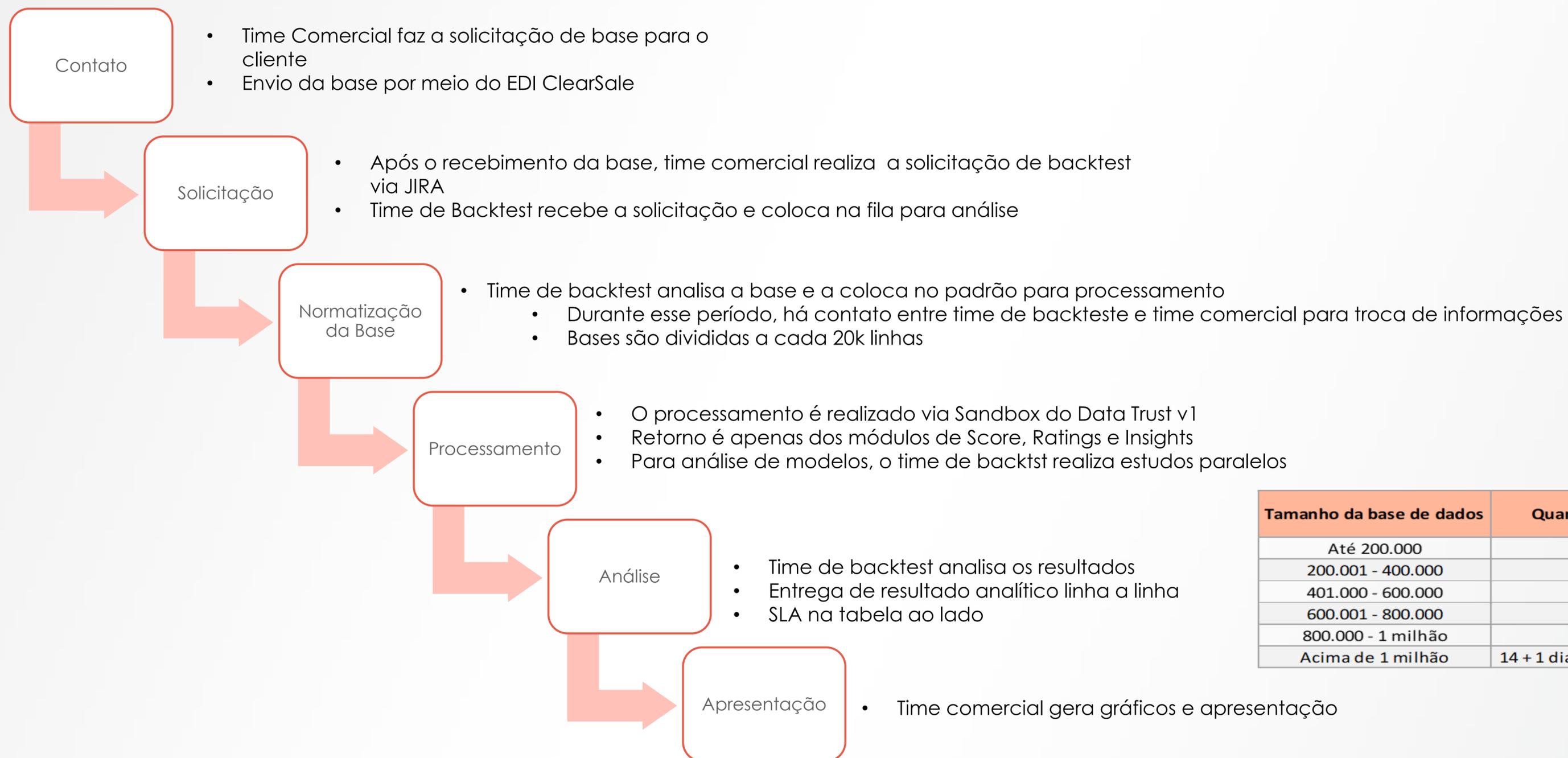
- Otimizar o desempenho do backtest para lidar com grandes volumes de dados de forma mais eficiente.
- Introduzir recursos de paralelização e distribuição para acelerar a execução do backtest.
- Aprimorar a interface de usuário para tornar a configuração e a análise dos resultados mais intuitivas e fáceis de usar.



### RESULTADOS

- Desenvolver recursos de relatórios básicos para analisar e visualizar os resultados do
- Possibilidade de acesso a base analítica para compartilhamento com cliente ou manipulação interna.

# Backtest As Is



Tamanho da base de dados	Quantidade de dias úteis
Até 200.000	10
200.001 - 400.000	11
401.000 - 600.000	12
600.001 - 800.000	13
800.000 - 1 milhão	14
Acima de 1 milhão	14 + 1 dia a cada 200.000 registros

# Backtest To Be (MVP)

Contato

- Time Comercial faz a solicitação de base para o cliente
- Envio da base por meio do EDI ClearSale

Acesso e  
processamento

- Time comercial acessa a Interface web da PDT e realiza upload do arquivo
- A capacidade de processamento será de 100k linhas (MVP)
- O tempo para o processamento e retorno dos resultados será apenas o de processamento do arquivo (algumas horas)

Resultados e  
Apresentação

- O retorno para o MVP será dos módulos de Score, Ratings e Insights + 3 modelos de Score e 3 modelos de Insights
- Os resultados serão disponibilizados via dashboard na interface da PDT, com possibilidade de exportação para que o time comercial realize a apresentação, bem como planilha analítica

Análise

- Caso o time comercial entenda a necessidade de maior aprofundamento no estudo, o time de backstst poderá ser acionado para complementar as análises

# sumário (material completo)

1. Backtest: O que é
2. Backtest: Problemas Atuais
3. Backtest: Discovery
4. Backtest: Proposta de MVP
5. Backtest: Eventos e Métricas
6. Anexos e Materiais

# sumário

- 1. Backtest: O que é**
2. Backtest: Problemas Atuais
3. Backtest: Discovery
4. Backtest: Proposta de MVP
5. Backtest: Eventos e Métricas
6. Anexos e Materiais

# Backtest

## Contexto

O backtest da Plataforma Data Trust é **indicado para provar o valor do produto para o cliente e ajudá-lo a definir o ponto de corte do modelo de score e até mesmo as regras de corte para aplicar em um modelo de Decisão** (quando o cliente comprar o componente de decisão customizado, é necessária a realização do backtest).

A intenção do backtest é **tentar reproduzir o que teria acontecido na operação, se o cliente tivesse utilizado o modelo no passado**. Portanto, para fazer um backtest nós precisamos que o cliente já tenha uma operação em andamento, ou seja, não seja um cliente que está iniciando seu negócio recentemente, que possua base de dados com as informações dos pedidos (CPF, E-mail, Celular, CEP, Data - sendo que os únicos dados obrigatórios são CPF e Data) e que essas informações tenham pelo menos uma distância de 3 meses da data atual, dado que precisamos da **marcação da fraude** para entender a aderência do produto (Variável resposta).

Além de ser estrategicamente muito importante para a empresa como um todo, ela acaba sendo muito importante para a tomada de decisão e estratégia dos gestores.

Esse processo consiste em obter através de uma simulação, considerando uma base de dados do passado, os resultados que seriam apresentados de nossos modelos preditivos. Em muitos casos, ainda é possível que o cliente nos compartilhe uma variável resposta, que nos permite medir o KS, que pode ser utilizado para embasar o valor que o mesmo trará a sua operação ao contratar a plataforma. O Teste não paramétrico de Kolmogorov-Smirnov (Ks) é utilizado para comparar o quão próximo os dados amostrais estão de uma distribuição de probabilidade de referência, logo a variável resposta.

### **1. Análise dos ratings e insights**

*Ratings - força do vínculo entre duplas de dados (documento/email; documento/telefone; documento/cep; e-mail/telefone; e-mail/cep; telefone/cep), sendo classificado como:*

- 0 - não informado
- 1 - vínculo fraco
- 2 - vínculo médio
- 3 - vínculo forte

*Insights - são mais de 400 informações sobre os dados recebidos do cliente, por exemplo: quantidade de CPFs que utilizam esse CEP para compras; CPF com alta incidência de compras positivas (não fraude); os dados informados foram vistos pela primeira vez entre 1 e 2 anos; entre outras informações.*

### **2. Análise de aderência - para telefones/emails/cep**

Indica do total de pedidos analisados, quantos a ClearSale conhece o telefone, e-mail e cep, dos conhecidos quantos são hotphone, hotmail e hotcep e indica também o alcance adicional, ou seja, dos dados informados quantos o cliente não conhecia o hot, mas a ClearSale conhecia.

### **3. Análise do Modelo de Score - cálculo do KS e CSI**

A base é dividida e ordenada em 20 faixas de tamanho igual, cada uma contendo 5% dos pedidos, sendo que a ordenação será feita pelo score, então as primeiras faixas terão a concentração dos pedidos bons, e as últimas faixas dos pedidos maus. O ideal é que os pedidos maus estejam acumulados nas últimas faixas, pois quanto maior o score, maior a probabilidade de ser fraude.

KS (Kolmogorov-Smirnov) - É o quanto o modelo consegue separar os pedidos bons dos maus. Pensando em um gráfico que tenha duas curvas, uma com os pedidos bons acumulados e outra com os pedidos maus acumulados, o KS é a maior distância entre essas duas curvas. Essa é uma das métricas estatísticas mais conhecidas e utilizadas para análise de aderência do modelo de score.

CSI20 - o CSI é uma métrica criada na ClearSale, e ele representa o percentual de pedidos maus acumulados nas faixas com score mais alto. Então, se tivermos um resultado de CSI20 = 60%, significa que dentro de 20% dos pedidos com score mais alto estão 60% dos pedidos maus, então se tivesse um corte de reprovação nessa faixa, 60% da fraude estaria sendo reprovada.

# Backtest

## Fluxo Atual

Para que seja iniciado o processo, a área comercial da empresa negocia com o cliente o envio e formato da base.

O cliente solicitante deve seguir as orientações presentes [neste tutorial](#) para envio da base através da [ferramenta EDI](#), seguindo um template padrão.

A amostra deve ser aleatória e representar o cenário real do negócio do cliente, como taxa de fraude/cbk e taxa de aprovação mensal. O ideal é ter pelo menos 6 safras na base de dados e que elas tenham distância mínima de 3 meses da data atual e máxima de um ano, por exemplo: em maio/2022 as melhores safras para backtestar o DataTrust seria de setembro/2021 até fevereiro/2022. A base de dados não deve possuir menos de 1000 registros e mais que 100.000 registros. A variável resposta precisa ser significativa, ou seja, ter pelo 100 casos.

Após o recebimento da base, o time comercial salva o arquivo em diretório específico e solicita a análise via [formulário no Jira](#).

 Backtest

### Solicitação de Backtest

Atenção!

**Pausa de Backtest período BlackFriday: do dia 16/11/22 ao dia 30/11/22.**

Os backtests solicitados até o dia 31/10 serão entregues antes do dia 16/11.

Solicitações realizadas após o dia 31/10 serão iniciadas a partir do dia 01/12/22, portanto fiquem atentos com as datas e solicitem seus backtest até o final de outubro.

Data de Solicitação

Selecionar data 

E-mail do Solicitante \*

Área do Solicitante \*

A Entidade já utiliza algum produto da ClearSale? \*

ID oportunidade (Sales Force)

Entidade - Frente \*

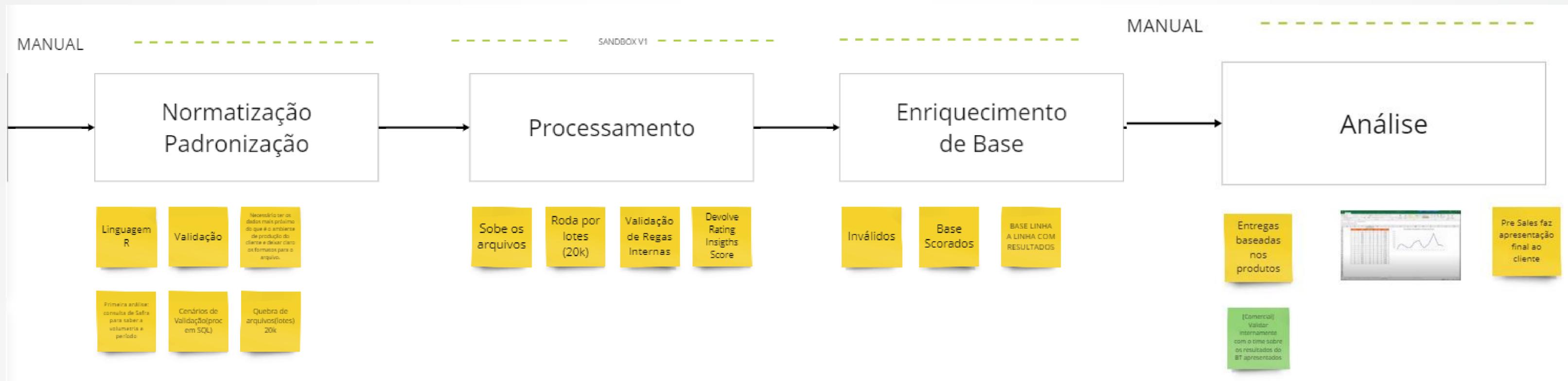
Essa Entidade já é atendida por alguma Squad? Se sim, qual o nome da Squad? \*

Produto(s) que deseja backtestar (campo obrigatório) \*

# Backtest

## Fluxo Atual

Após a solicitação chegar ao time de backtest, inicia-se o processo de normalização da base, processamento e finalização. Entende-se por processamento online, o processamento baseado na API que por sua vez, utiliza dados e informações que são trafegadas pela aplicação em produção real-time.



# Backtest

## Entregáveis

Os entregáveis da **equipe de backtest** para o produto DataTrust são:

- Sem variável resposta: estudo de DataTrust sem análise de aderência do modelo (KS e CSI) e resultado analítico (linha a linha);
- Com variável resposta: estudo de DataTrust com análise de aderência do modelo (KS e CSI) e resultado analítico (linha a linha);

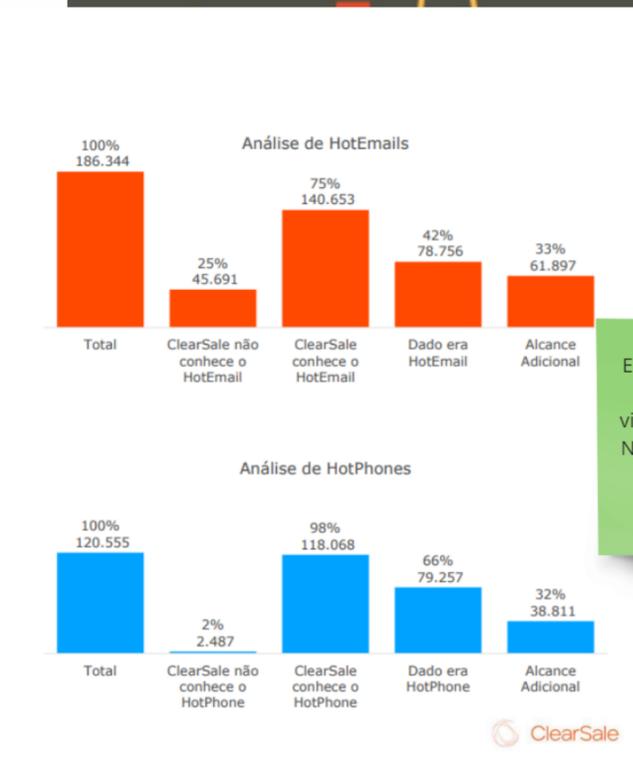
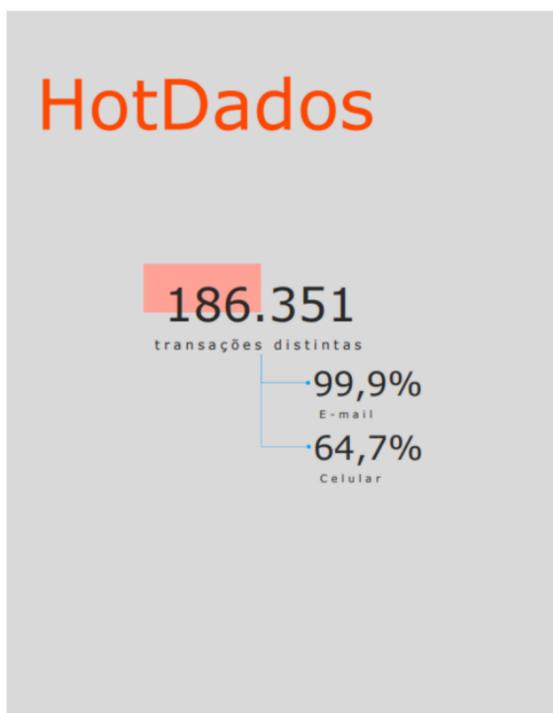
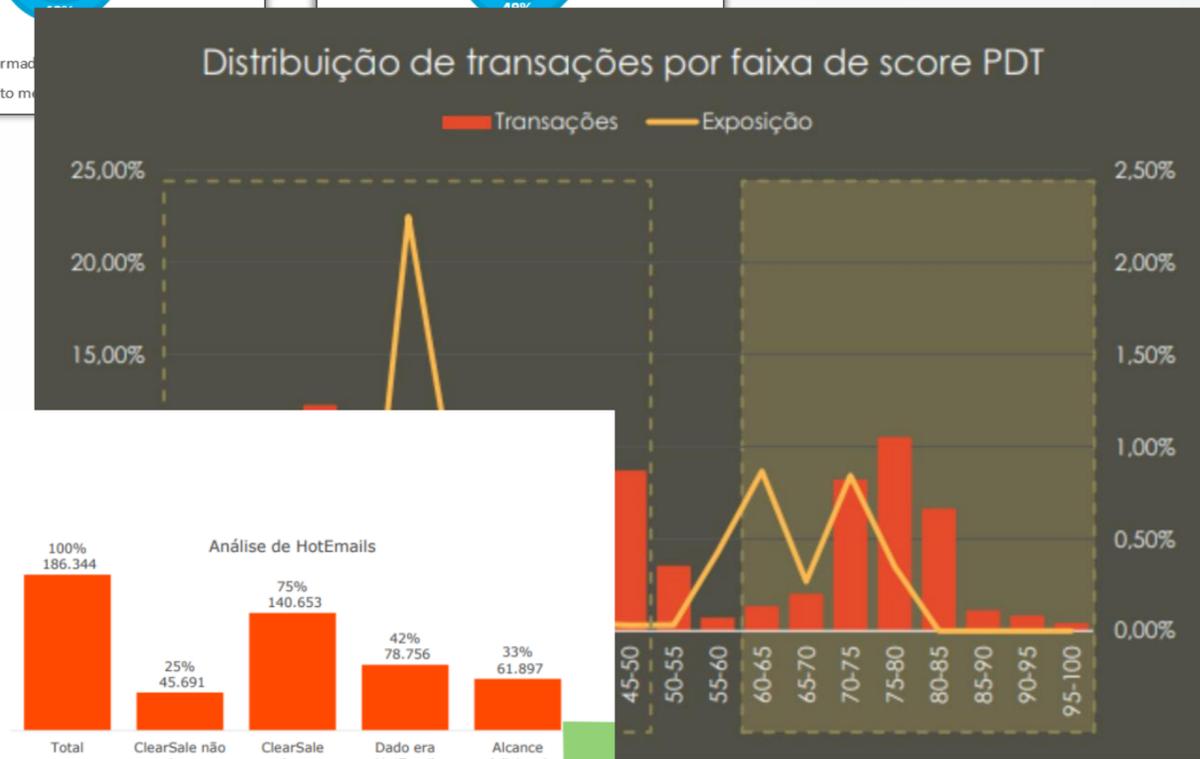
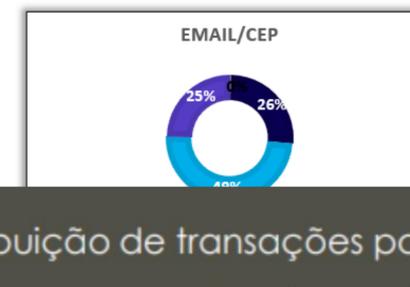
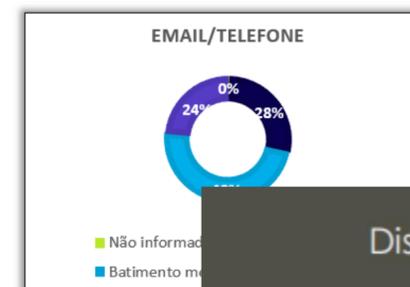
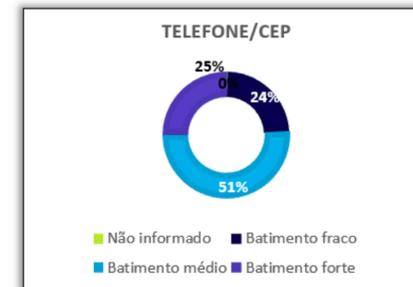
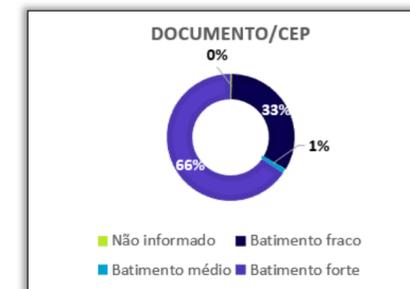
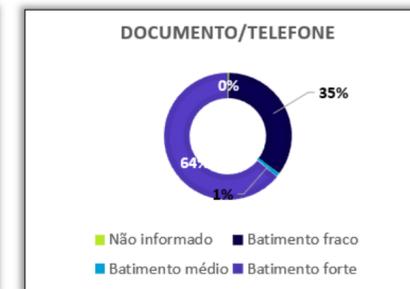
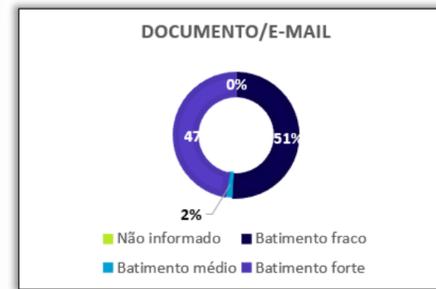
Após retorno da equipe de backtests a área comercial deve:

- Montar apresentação com os resultados do backtest (quando necessário);
- Alinhar com a equipe de backtests anteriormente a agenda, para reuniões de apresentação com o cliente.
- Especificar como deve ser a análise do retorno dos estudos construídos pela Modelagem;
- A equipe comercial precisa estar apta para analisar os resultados e entender o ganho que os mesmos levarão para a operação do cliente, para que assim consiga apresentar da melhor maneira o valor do nosso produto.
- Listar os graficos/visões que devemos ter para a análise e prova de valor do produto;
- Criar a apresentação/storytelling do backtest.

# Backtest

## Entregáveis

Após a entrega do time de Backtest, o time de Pre-Sales, fornece ao cliente um relatório das análises geradas e assim, os mesmos podem comparar com dados previamente conhecidos, validando se os resultados foram satisfatórios ou não. Esta prática, traz para este potencial cliente, a primeira prova de valor ao utilizar nosso produto e também a confiança de prosseguir com o processo de contratação.



## Quando não fazer

O backtest do DataTrust é indicado para provar o valor do produto para o cliente e ajudá-lo a definir o ponto de corte do modelo de score.

A ideia é **tentar reproduzir o que teria acontecido na operação, se ele tivesse utilizado o modelo no passado**. Portanto, para fazer um backtest nós precisamos que o cliente já tenha uma operação em andamento, ou seja, não seja um cliente que está iniciando agora, que possua base de dados com as informações dos pedidos (CPF, E-mail, Celular, CEP, Data - sendo que os únicos dados obrigatórios são CPF e Data) e que essas informações tenham pelo menos uma distância de 3 meses da data atual, dado que precisamos da marcação da fraude para entender a aderência do produto.

**Não é recomendado fazer backtest da Plataforma DataTrust se o cliente não tiver base de dados com no mínimo 1.000 pedidos/transações ou se o cliente não tiver marcação de fraude (variável resposta)**, apesar de ser possível realizar (nesse caso será retornado o estudo sem a análise de aderência do modelo KS e CSI).

# sumário

1. Backtest: O que é
- 2. Backtest: Problemas Atuais**
3. Backtest: Discovery
4. Backtest: Proposta de MVP
5. Backtest: Eventos e Métricas
6. Anexos e Materiais

## Quais problemas queremos resolver?

Visando, neste primeiro momento, **auxiliar o time de Pré-Sales a atuar proativamente na construção de valor para o cliente**, precisamos reestruturar a forma que realizamos os backtests e também os motores que realizam essa operação, pois temos um grande gargalo em relação à **manualidade/passagem de bastão** e ao **tempo que esse processo se desenvolve**, além da limitação técnica presente no atual modelo.

Jaqueline de Moura Oliveira [commented 6 de mar.](#)

Alinhamento realizado com a Mariana em 06/03/2023.

Macro etapas, responsáveis e prazos:

Definição Amostra | Trust Score | 1 dia

Backtest | Backtest | 10 dias

Análise | Trust Score | 3 dias

Início previsto em 10/03/2023.

Jaqueline De Moura Oliveira 26/05 12:38

JO

Boa tarde, Mari. Tudo bem?

Compartilho aqui a apresentação que utilizarei na reunião de Segunda com o Elias, ok?!

Caso tenha alguma consideração fique a vontade para me pontuar.



BancoABC\_MigracaoFFR\_DT.pdf ...



Mariana Coralina do Carmo Jun 16, 9:40 AM

Elias, bom dia! Tudo bem?

Dúvida: vcs chegaram a apresentar o backtest para o Banco ABC?



Elias Rodrigues Dos Santos Jun 16, 11:42 AM

Mariiiii!!!

Bem sim graças a Deus e com você? Tudo bem?

Mari, vc nem vai acreditar.

O cara que pediu o backtest entrou de férias, e disse que no retorno falaríamos. nessa semana, ele saiu da empresa

Tamanho da base de dados	Quantidade de dias úteis
Até 200.000	10
200.001 - 400.000	11
401.000 - 600.000	12
600.001 - 800.000	13
800.000 - 1 milhão	14
Acima de 1 milhão	14 + 1 dia a cada 200.000 registros

## Quais problemas queremos resolver?

Também precisamos nos atentar para as atuais limitações que temos internamente. Atualmente, o fluxo é **realizado via Sandbox do Data Trust v1**, e passa pela fila de processamento, muitas vezes causando uma fila de espera, dando morosidade na execução. Além disso, há o problema de capacidade de processamento.

Dentre alguns menos impactantes, temos um específico que nos traz grande preocupação: **concorrem com a aplicação que é utilizada pelo cliente**. Isto põe grande risco para a saúde de nossa aplicação e com um potencial de gerar downtime parciais ou totais.

**Dependendo do tamanho do arquivo (quantidade de registros), para execução de um BT, ele será quebrado em arquivos menores, pois o processamento só comporta rodar arquivos de até 20 mil registros** e leva cerca de 2h30 para finalizar. Cada arquivo é startado e processado individualmente e em fila, **não permitindo o processamento paralelamente ou pausar seu processamento**.

# sumário

1. Backtest: O que é
2. Backtest: Problemas Atuais
- 3. Backtest: Discovery**
4. Backtest: Proposta de MVP
5. Backtest: Eventos e Métricas
6. Anexos e Materiais

### Como faremos

As variáveis são muitas, mas as principais:

*Dados*

*Conhecer necessidades do cliente*

*Conhecimento das necessidades Clear*

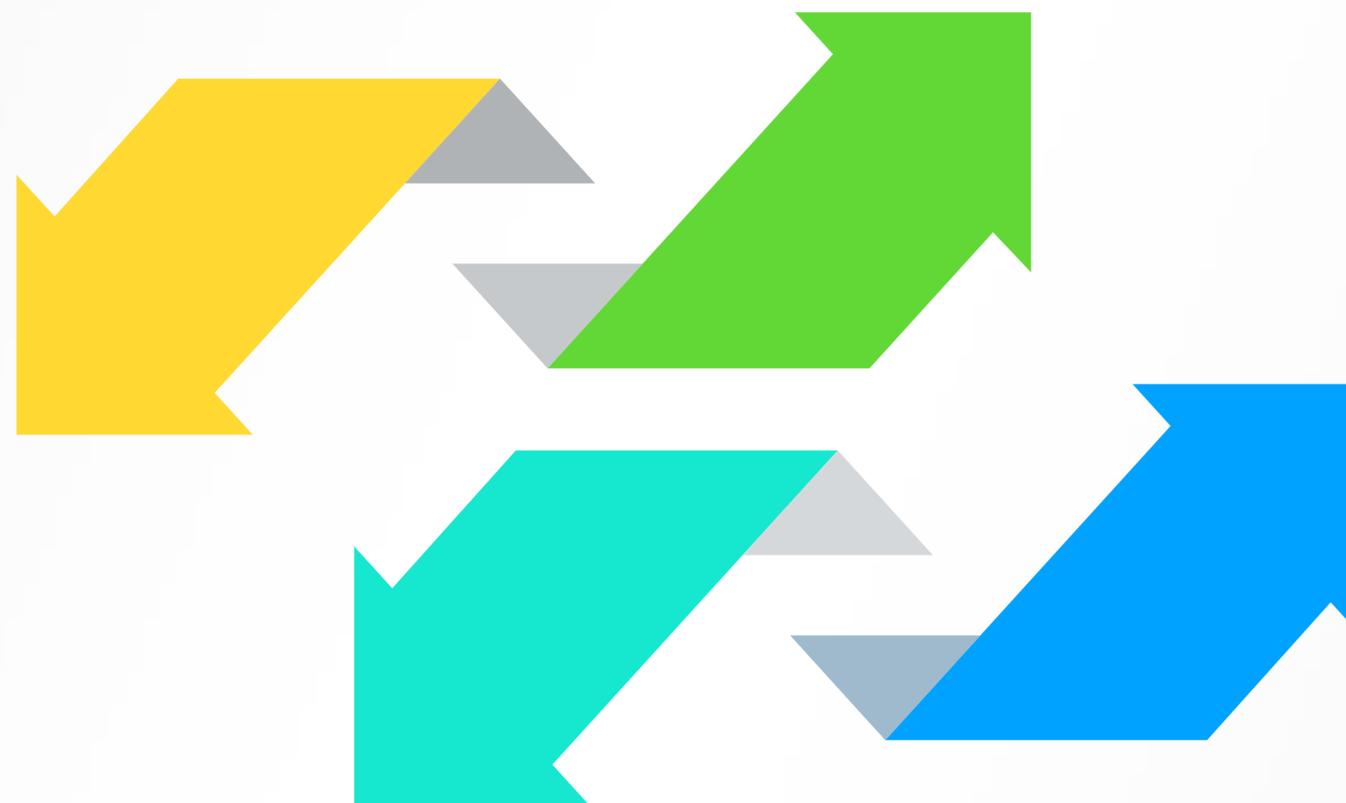
*(Direcionamento estratégico)*

*Benchmarking*

*Dados*

*Feeling*

...



### Entrevistas Qualitativas

Tem por objetivo buscar uma compreensão profunda de necessidades e dores de áreas internas, para nos ajudar a direcionar esforços nos próximos ciclos.

# Discovery Backtest

## Entrevistas

Entrevistado	Área	Principais Dores	O que é importante	Demais Pontos
Tamyris Marconi	Pre Sales	Tamanho do arquivo (ideal: +100K) processo de compartilhamento de base é ruim/demorado SLA	Entender o que cada modelo faz, para poder escolher sozinha (ter descritivo em tela) Poder usar qq modelo da clear Modelo Card4	Difícil perceber riscos para usuário final (riscos do cliente usando) Quem faz a apresentação é o pre sales e não modelagem Atenção: BT não tem para Inside
Rafaela Souza	Data Field	Tamanho do arquivo (ideal: +100k) Pensar numa maneira melhor da extração dos resultados Linhas inválidas: a devolutiva deveria ser imediata (às vezes, mesmo após a tratativa, ainda sim o arquivo não é processado por completo) Terceira dor: só PDT é diferente (que precisa ser na sandbox. Os demais são feitos diretamente em banco)	Ter todos os componentes da PDT disponíveis para backtest Ter maior visibilidade de componentes da PDT e tipos de validação que ela faz Poder usar qq modelo da clear não faz sentido usar bureau (mas os modelos da PDT precisam deles) ter todas as infos num único lugar	Ponto de atenção: ter todos os modelos em um único lugar. Hoje, pega score, ratings e insights do DT e os demais, no ASP. Bagunça todo o processo Quer pensar em backtest para qq componente da PDT, mas hj só testam score, ratings e insights por limitação técnica (da sandbox). E quando tem um modelo diferente, faz o estudo separado
Jaqueline Oliveira	Analytics	Tamanho do arquivo	Ter todos os componentes da PDT disponíveis para backtest	Score customizado la no asp = isso já está pronto com a estrutura que fizemos para modelo 9198 e card4

# Discovery Backtest

## Entrevistas

Entrevistado

Área

O que pode ser feito

 Bruno  
 Joel/Gustavo  
 Rocha

Analytics

Usar o C.A.O.S (direto no DW), que é muito mais performático, tornando o backtest mais agil

- CAOS ainda não tem todos os modelos configurados na pipeline
- Gargalo: fazer em batch

Testar vários modelos ao mesmo tempo: GRANDE NECESSIDADE

Hoje faz no transacional e loop

Pensar em dados retroativos (drift)

Tabela de entrada: DW (e da para mudar)

O que podemos fazer: interface web ser o caminho para o usuário final: tamyres inputa dos dados e da start na pipe que for configurada.

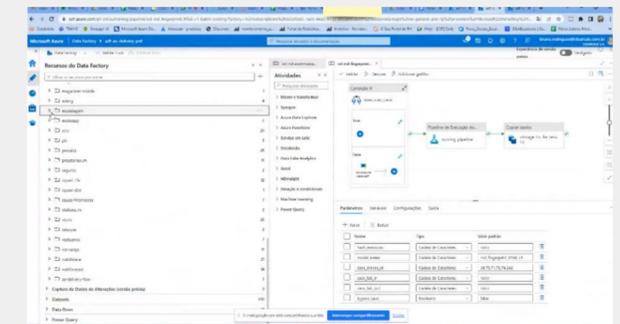
Precisamos pensar como serão gravados os dados (melhor: base fria). Para podermos gerar o visual e o reporte dos dados

Separar lote de bt de prd

Para socre já tem processo alternativo que vai mais rapido (para os demais ainda vai direto na api)

Ideal: ter tudo num lugar só

Não vai para bureau: o que pode ser feito é salvar os dados em uma tabela específica para BT (HUB salva em uma tabela e BT pega no banco)



# Discovery Backtest

## Entrevistas

Entrevistado	Área	Problemas atuais
Rodolfo Luna	TI	<p>O "autoscale" do backtest é manual.</p> <p>Bureaus: se não precisasse de bureau rodaria direto no banco (visão de futuro ainda vai focar em bureau)</p> <p>Tempo: entendem ser um processo lento (em decorrência dos modelos + bureau = demora para processar). Quase que não tem como escapar disso</p> <p>Processar grandes arquivos</p> <p>Pensar em liimitação de front</p> <p>Subir do servidor do cliente para o nosso</p> <p>Testar um bom numero para não derrubar a API (10 threads)</p> <p>Vai até 30, apenas em casos urgentes</p> <p>Etapa de validação = backtest que faz</p> <p>Fila de lote do sandbox, mas a API é a mesma</p> <p>Pensar na exclusão dos dados, caso seja solicitado pelo cliente</p> <p>Apagar direto no banco, se a consulta for direto no banco</p> <p>Para nós: ter usuário/entidade específica</p> <p>Time de bt: nem olham o excel</p> <p>Select no banco</p> <p>Para a entidade de BT, ter cache de 90 dias</p>

# Discovery Backtest

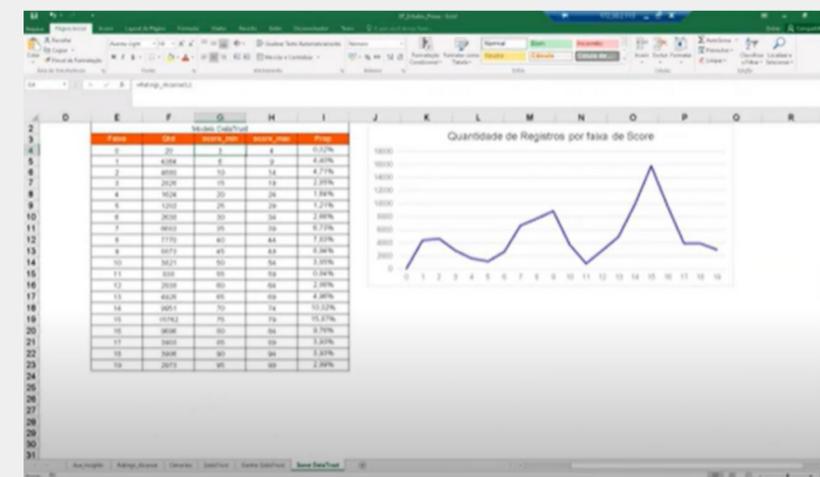
## Entrevistas

Entrevistado    Área    O que é importante

Joyce Costa

Modelagem  
(Analista)

Excel, TXT e CSV (base original)  
Recebimento: EDI (via soluções/comercial)  
Devolutiva tbm (nos diretórios - pasta na rede)  
Tratamento em R para validação dos campos/tratamento de base  
Base "ruim" depende da situação: se for muito ruim, volta para o cliente  
A depender de como for, o time mesmo faz o tratamento  
Ideal trazer exatamente para o cliente como os campos devem ser preenchidos.  
Procs são as mesmas dentro de sempre  
Não tem diferença  
Consulta da safra  
Ideal: safra de 3 meses anteriores à consulta  
Quebra da base de 20 em 20k  
Sobe na sandbox (via tela)  
Extração dos resultados  
primeira entrega: inválidos (não escorados)  
Retornos e entregas salvas em diretório



# Discovery Backtest Bench (SEON)

Add New Batch Test

Mariana Coralina M

### HOW TO RUN A BATCH TEST

- 1 Upload the CSV or XLSX file. You have 1190 records left to upload.
- 2 The file will be processed after we receive it. This may take longer, depending on the amount of data in your batch.
- 3 The results will be delivered to your selected email address and loaded onto the Transactions page

### FORMATTING GUIDELINES i

```
email;transaction_id;user_name;user_country;phone_number
gikobe1772@gmail.com;ufce82172;"frozege
kobe34141";US;14432907167
dudaomed@yahoo.com;413f03013150;distrakteduck89;UK;447828140031
matixd@gmail.com;e15410509e3f;puzzling_barracuda;DE;491745387291
```

Download sample files: [CSV](#) | [XLSX](#) [Learn more on SEON Docs](#)



email



Hi there fraud fighter,  
Great news! We've successfully completed your batch test.  
Head to the [Batch test page](#) on the SEON Admin Panel to view and download the results.  
Do not hesitate to [contact us](#) if you have any questions.  
Best regards,  
SEON team

# Discovery Backtest

## Bench (Strip)

### Backtesting rules

As a general philosophy to rule analysis, there's a trade-off between preventing fraud and blocking good transactions or false positives. Backtesting helps identify rules that fall within your risk appetite or strike the right balance between disputes prevented and any increase in false positives. To estimate the impact of a rule, you can backtest combinations using transaction data from the last six months via the Radar Dashboard and conduct more targeted analyses to understand how the rule would have performed if recently in place.

### Backtesting in the Dashboard

#### Add a rule for **placing payments in review**

Payments matching this rule will be placed into review, unless they already matched an allow or block rule.

```
Review if :card_country: = 'BR' AND :auths_per_card_number_daily: > 3
```

# Discovery Backtest

## Matriz CSD

### Certezas

Precisamos aumentar nossa capacidade de processamento (pelo menos 100k)

Informação tem que ficar no DW para time de backteste poder pegar

Limitação de processamentos em simultâneo

Precisamos de forte alinhamento com interface

Se for necessário enviar algum campo específico para o modelo do asp que não exista na nossa transaction, CORE agora aceita envio de campos genéricos,

### Suposições

entidade de teste com score padrão

entidade de teste com outros scores  
isso para poder usar mais de um score para o teste

Maiores dores hoje dizem respeito à utilização dos modelos \*teste com tipos de modelos diferentes

Não sabemos qual o modelo mais usado

Bruno Joel: analytics criando modelos no DW - visão de BT para o futuro

Podemos ter cache de 90 dias para a entidade teste

### Dúvidas

Interfaces: PDs para construir visualização (fazer wire)

entidades configuradas certinhas para ter o melhor uso (para não gastar muito)

precisamos de ambiente isolado? Sim, mas não para agora

concorrência com wile, biac, lote: o quanto podemos paralelizar?

É viável colocar limitação de horário?

# Discovery Backtest

## Governança

	<b>Responsáveis</b>	<b>Pessoas chaves (Stakeholders)</b>
	Desempenha as tarefas e é responsável pela execução delas	Tem a responsabilidade final pela conclusão e aprovação das tarefas
Program Management		William Awaji
Product Management	Mariana Carmo	
Design / PDs		
Desenvolvedores		Rodolfo Luna
Usuários		Pre Sales; Business Analytics (Denis Leite Andrade); TrustScore

# Discovery Backtest

## Insights e Aprendizados

### Necessidades

Processar um backteste de forma ágil sem interação no banco de dados

Disponibilizar cenários pré-configurados na plataforma de teste para simular como ela ajudaria a resolver os problemas do cliente

Utilização do CAOS futuramente

Próprio cliente realizar um teste em lote futuramente

Todos os modelos/módulos da PDT

Possibilidade de usar várias bases/modelos diferentes para ter o melhor resultado

Permitir que o usuário escolha o módulo que deseja utilizar para processar o arquivo

Sistema deve permitir cancelar o processamento

Validação dos campos obrigatórios de acordo com os módulos escolhidos

Visualizar os resultados do arquivo em dashboards para ter um panorama geral das transações

Permitir subida de arquivo grande via csv

### Dores

Backteste rodado na mão

Não tem como gerenciar backteste

Dificuldade de analisar o detalhe de cada transação pela qtd de linhas do arquivo

Limitação do excel em qtd de linhas ao subir um arquivo mto grande

Concorrência com ambiente de produção

Oportunidade muitas vezes ainda não está no Sales Force

Preenchimento MANUAL do formulário

Fluxo de trabalho em paralelo na plataforma JIRA

Complexidade de extrair métricas

Complexidade de extrair valor "consolidado"

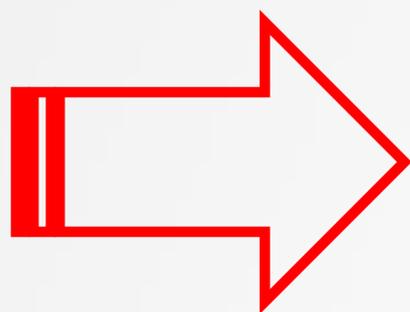
Lentidão ao executar

# sumário

1. Backtest: O que é
2. Backtest: Problemas Atuais
3. Backtest: Discovery
- 4. Backtest: Proposta de MVP**
5. Backtest: Eventos e Métricas
6. Anexos e Materiais

## Quais problemas queremos resolver?

Em resumo, podemos classificar em dois focos principais:



### 1. Aumentar Conversão/Upsell

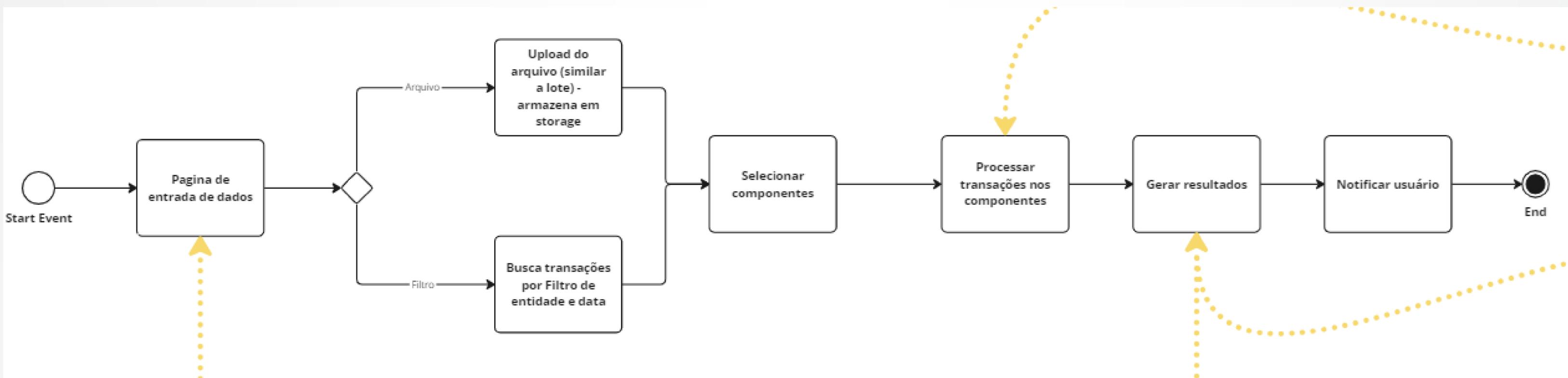
1. Potencializar os resultados do BT
2. Reduzir tempo tempo de entrega de valor ( TTV)
  1. *Prover ao time comercial/vendas a área de backtest da PDT*

### 1. Diminuir CAC

1. Consumo de infraestrutura para execução BT
2. Reduzir passagem de bastão entre áreas
3. Reduzir tempo de processamento/utilização de recursos
  1. *Via Consulta em Lote e/ou CAOS*

## Como vamos resolver?

Neste momento, fica recomendado a construção de novos sistemas apartados que viabilizarão atender todas as condições necessárias para o futuro do negócio (processo padronizado e self-service). Estes novos componentes, por sua vez, mais performáticos, irão possibilitar uma maior entrega de valor para o cliente, além de reduzir o consumo de nossas máquinas nos servidores Azure, logo gerando um menor custo para a companhia.



# Backtest

## Desafios

### Riscos

#### Modelos utilizados

- Atualmente não são todos os modelos que possuem a possibilidade de processamento online (plataforma CAOS)
- Precisamos entender como capturar o modelo do ASP.
- Na PDT, ainda necessitamos da consulta a bureaus
- Paralelismos das consultas X performance
- Horários permitidos para execução

### Dependências

#### Orquestração dos Bureaus

- Através das consultas geradas internamente ao Data Factory, não será possível manter a consulta dos Bureaus, portanto, deverá ser exposto ao time de backtest estas informações (caso seja feito diretamente com a ferramenta CAOS).

Configuração e distinção de acessos ao lote será feita de acordo com a integração com SSO.

#### Ambiente de produção

Tendo em vista que a solução completa que pode levar um maior tempo para a conclusão, optamos por realizar a seleção estratégica de cenários e casos de uso que irão compor nosso MVP. Entende-se por MVP, aquelas que dentre todas as necessidades, oferecerão minimamente algum valor tangível a nossos clientes.

Neste ponto podemos ter como objetivo, que na menor unidade de tempo, possamos entregar a possibilidade de se basear na estrutura da consulta em lote, para os seguintes casos de uso:

### **A.Retirar dependência do time de Backtest para:**

- a.Time de Pré-Sales conseguir tomar decisão de maneira mais rápida e assertiva.
- b.Auxiliar time Comercial em UpSell.
- c.Sugerir construção de funil e decisão.
- d.Apresentar resultados de maneira visual e com a possibilidade de download (com foco em auxiliar times na construção de materiais para pitch de vendas).
- e.Estruturação do fluxo de demonstração (que vem ocorrendo diretamente na interface da PDT), concedendo uma jornada aderente às necessidades do time comercial e também, considerando os impactos que podem gerar para as bases da ClearSale com a ingestão de dados "sujos".

# Backtest

## Possibilidades de Entrega

- **Estabelecendo a Base**

- Implementar infraestrutura básica para o backtest, incluindo a integração com fontes de dados e configurações iniciais.

- Desenvolver uma interface de usuário simples para configurar e iniciar um backtest básico.

- Realizar testes de integração e validação dos dados para garantir a qualidade e a confiabilidade dos resultados.

- **Aprimorando a Precisão e Flexibilidade**

- Aumentar a precisão do backtest, considerando diferentes cenários e fatores, atendendo as necessidades do time de backtest..

- Permitir a personalização de parâmetros e estratégias de backtest para atender às necessidades específicas dos usuários.

- **Melhorando a Eficiência e Usabilidade**

- Otimizar o desempenho do backtest para lidar com grandes volumes de dados de forma mais eficiente.

- Introduzir recursos de paralelização e distribuição para acelerar a execução do backtest.

- Aprimorar a interface de usuário para tornar a configuração e a análise dos resultados mais intuitivas e fáceis de usar.

- **Validando e Integrando Resultados**

- Implementar ferramentas e recursos para validar os resultados do backtest em comparação com dados reais.

- Desenvolver recursos de relatórios detalhados para analisar e visualizar os resultados do backtest.

- Integrar os resultados do backtest com outras partes do produto, como venda de novos módulos ou estratégias automatizadas de negociação.

- **Aprendendo e Expandindo**

- Coletar feedback dos usuários e realizar iterações com base nas necessidades e nas melhorias identificadas.

- Pesquisar e implementar técnicas avançadas de backtest, como machine learning e análise preditiva

# Backtest

## Possibilidades de Entrega

- **Análise da base recebida pelo cliente:**
  - Distribuição da população por safra
  - Risco por safra (quando tivermos variável resposta. Se mais de 1, uma visão geral (OU OU OU) + uma visão por tipo de var).
  - Validação da qualidade dos dados (quantidade de nulls e dados inválidos para cada campo recebido).
  - Perfil de cliente – idade, região, etc.
- **Escolha dos modelos/vars a serem backtestados:**
  - Possibilidade de escolher os modelos/insights/variáveis mais aderentes (segmentos, tipos de fraude, dados, etc.) – podemos ter um dicionário na ferramenta, tipo:
    - § Modelo 1: segmento de onboarding de emissão de cartão, dados necessários, dados opcionais, dados mais impactantes na performance na performance do modelo.
    - § Modelo 2: segmento de onboarding de telecomunicações, dados necessários, dados opcionais, dados mais impactantes na performance na performance do modelo.
  - Cálculo do custo do backtest antes de iniciar a escoragem.
- **Backtest**
  - Possibilidade de rodar os modelos e backtests em paralelo.
  - Sem quebras de bases (ideal termos uma possibilidade de subirmos bases bem grandes para evitar esse tipo de processo).
- **Retornos**
  - Distribuição da população para cada score backtestado.
  - Gráficos e valores de KS para cada marcação de variável resposta recebida (Se mais de 1, uma visão geral (OU OU OU) + uma visão por tipo de var)
  - Visão de tabelão, com os acumulados de maus por quebra (tipo tabela e gráficos gerados pela Modelagem atualmente). No caso de recebimento de var resposta, se mais de 1, uma visão geral (OU OU OU) + uma visão por tipo de var.
  - Exemplos de bons e maus (scores baixos e altos).
  - Possibilidade de acesso a base analítica para compartilhamento com cliente ou manipulação interna.
  - Decisão

# Backtest

## [MVP] Foco da primeira entrega

- **Estabelecendo a Base**

- Implementar infraestrutura básica para o backtest, incluindo a integração com fontes de dados e configurações iniciais.
- Desenvolver uma interface de usuário simples para configurar e iniciar um backtest básico.
- Realizar testes de integração e validação dos dados para garantir a qualidade e a confiabilidade dos resultados.
- Validação da qualidade dos dados (quantidade de nulls e dados inválidos para cada campo recebido).

### Via entidade TESTE

- **Aprimorando a Precisão e Flexibilidade**

- Permitir a personalização de parâmetros e estratégias de backtest para atender às necessidades específicas dos usuários.

- **Eficiência e Usabilidade**

- Otimizar o desempenho do backtest para lidar com grandes volumes de dados de forma mais eficiente.
- Introduzir recursos de paralelização e distribuição para acelerar a execução do backtest.
- Aprimorar a interface de usuário para tornar a configuração e a análise dos resultados mais intuitivas e fáceis de usar.

- **Validando e Integrando Resultados**

- Desenvolver recursos de relatórios detalhados para analisar e visualizar os resultados do backtest (via Interface já desenhada para visão de Gestor + trabalho do time de BI).
- Possibilidade de acesso a base analítica para compartilhamento com cliente ou manipulação interna.
- Visão de tabelão, com os acumulados de maus por quebra (tipo tabela e gráficos gerados pela Modelagem atualmente).



## Solicitação e Recebimento da Base

V1: Via time Pre Sales (EDI) - Pré-Sales faz upload via Interface (dentro da interface terá um pequeno tutorial de objetivos e resultados esperados por um BT, com informações sobre o modelo)  
V2: Time Backtest (Diretório) Pre sales com input no Salesforce  
V3: Próprio Cliente



## Processamento

Interface processa consulta e devolve CSV ao cliente  
V1: API (threads)  
V2: Ambiente Isolado  
V3: CAOS



## Resultados

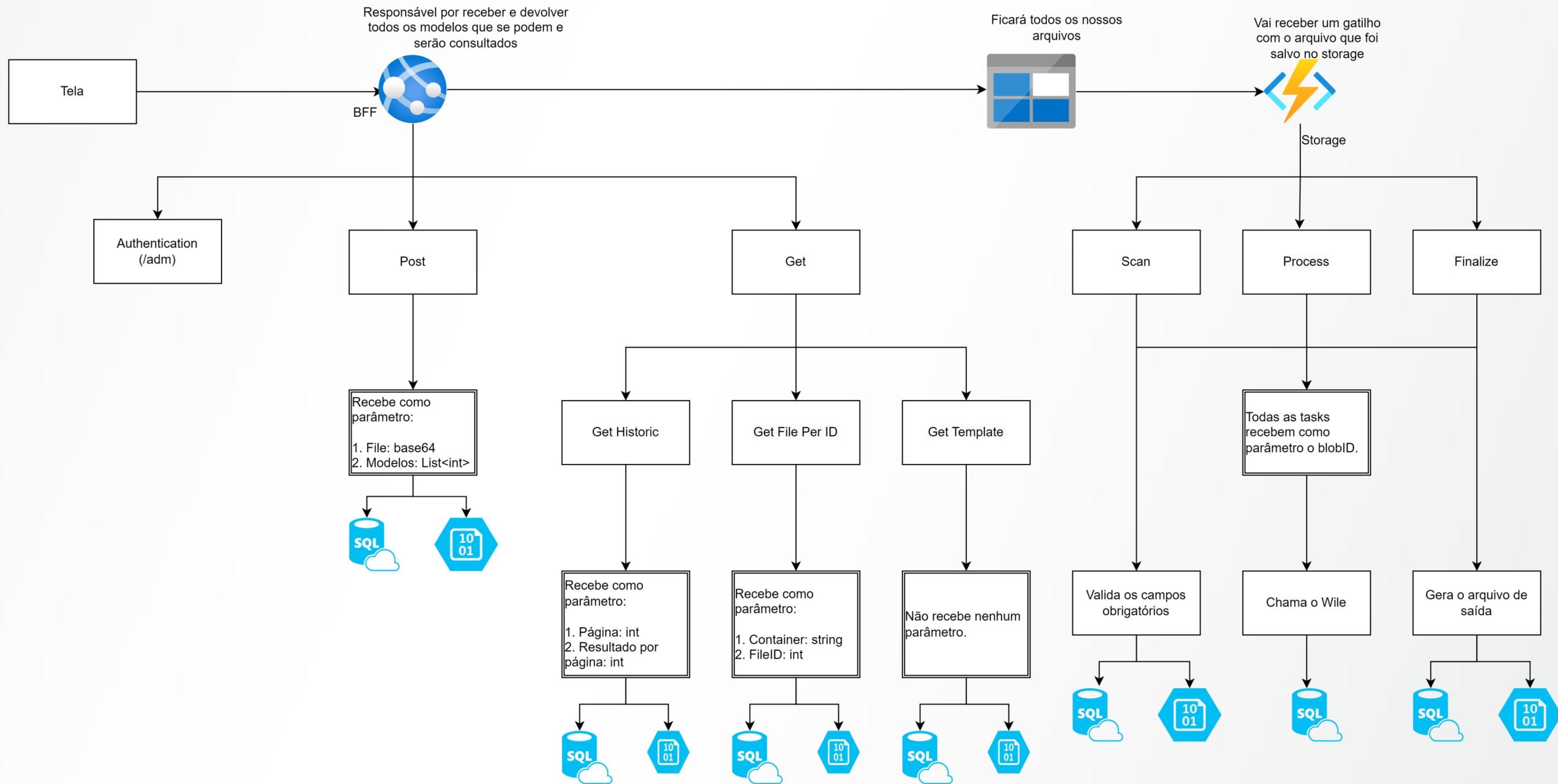
V1: Dash e excel  
V2: Envio por e-mail + Interface  
Para time de BT: O retorno dos resultados também será feita pelo sistema EDI e via planilha

# Backtest Fluxos

Perguntas que precisam ser respondidas de forma rápida:  
 1. Vamos usar o mesmo recurso que se encontra do Batch hoje?  
 Sim

O que vamos criar o zero:  
 1. Projetos do Scan (podendo usar muito do que já temos hoje),  
 Process e finalize

Testes:  
 1. Gostaria de fazer um teste usando o gRPC ao invés do Rest



# Backtest

## [MVP] Regras de Negócio

- **Estabelecendo a Base**

- Seguir regras de negócio da Consulta em Lote (formatações do arquivo, regras de fluxos de tela, regras de upload, regras de processamento, regras de retentativas, etc.)

- Seguir base da PDT: a estrutura proposta deve ficar “dentro” da PDT (ou “embaixo”). Isso significa que não devemos ter uma estrutura separada, mas sim, onde todas as funcionalidades que vamos pensar seja colocadas na interface, para termos um local único.

- **Premissas básicas**

1. Entidade específica para Backtest + Usuário específico interno para bloqueio
2. Autenticação: ROLE (a área de backtest na interface deve aparecer apenas para a entidade de backtest e seus usuários).
3. Precisa ter Data de referência e Data de Consulta
4. Capacidade de processamento maior que o lote (precisa ser maior que 20k – pelo menos 100k)
5. Consulta a bureaus: cache de 90 dias
6. Consulta de dados via base fria (transaction data – processo novo do Core)
  1. Lógica de chamar o CAD (pq precisa de bureau)
7. Consulta de módulos e modelos
  1. Pensar na arquitetura pelo modelo (para deixar pronto)
8. Apresentação visual dos resultados
  1. Ter exportação dos gráficos em PDF
9. Não passar pela PDT e chamar os modelos direto (Wile e ASP)
10. Não precisa de transaction
11. Ter dados no DW ou outra base fria para equipe de backtest ter acesso
12. Não precisamos pensar em billing (cobrança) para Backtest, mas precisamos ter em mente que: ter identificação de cliente, para poder apagar dados e para poder cobrar em caso de estudos pagos
13. Ter restrição de número de requisições (máx 5k, por exemplo)
14. Modelos utilizados em backtest: 9203 (card4), 9100, 9115, 9219, 9144

# Processo de Backtest

## Protótipo - Home

Olá, William!

Bem vindo de volta à Plataforma Data Trust.

### Consulta Individual

Valide a força de vínculo dos dados utilizando nome, CPF, CEP e E-mail.

Consultar

### Histórico de Transações

Acesse a lista de todas as últimas transações.

Ver histórico

### Busca rápida

Busque por CPF ou Transaction ID para encontrar uma transação única ou transações de um único usuário nos últimos 30 dias.

CPF

Transaction ID

Digite aqui

Buscar

### Consulta em Lote

Economize tempo enviando um único arquivo e tenha o processamento rápido ao seu alcance.

Enviar arquivo

### Background Check

Checke os antecedentes de pessoas ou empresas.

Checkar antecedentes

### FAQ

Veja as perguntas recorrentes ou envie sua dúvida para nós.

Acessar ajuda

### Área de Estudos

Realize estudos enviando um único arquivo e obtenha insights a partir dos resultados.

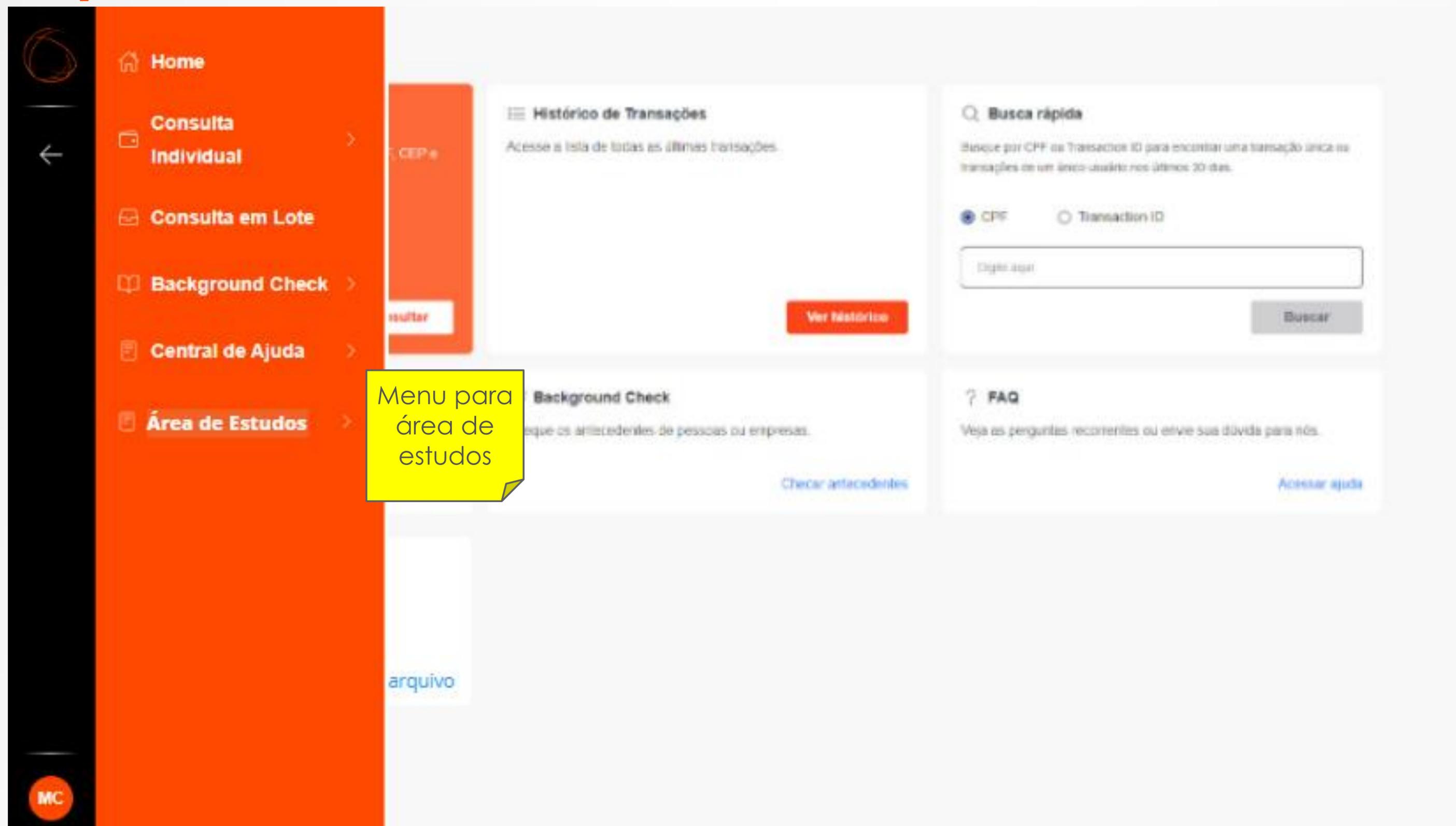
Enviar arquivo

Card para  
área de  
estudos

Separar usuário por role (só entidade x vai ser backtest e só usuário x pode acessar a role de backtest)

# Processo de Backtest

## Protótipo - Menu



# Processo de Backtest

## Protótipo – Tela de Upload



**Upload de Arquivo**

[Dashboard](#) [Histórico de Consultas](#)

**Visão de Dashboard**

**Baixar template**

**Aumento do tamanho do arquivo**

Clique nessa área para escolher um arquivo ou o arraste para cá  
 O arquivo deve estar no formato .CSV, conter no máximo XXXX linhas, não ultrapassando o tamanho

**Módulos + Modelos com tooltip de informação**

Selecione os módulos e/ou modelos que serão executados para o estudo

- Selecionar Todos
- Score
  - Score padrão (XTrust)
  - Card 4
  - Score Desconhecido
- Insights
  - Insights padrão
  - Insights SIM Swap
  - Insights Trust Card
  - Ratings

**Enviar para Análise**

**MC**

# Processo de Backtest

## Protótipo – Tela de Upload



← Voltar Baixar template

**Upload de lote** Dashboard Histórico de Consultas

  
 CelularInvalidoComCaracterEspecialEEspacoEmBranco.csv  
 0 MB

Seleccione os módulos e/ou modelos que serão executados para o estudo

- Selecionar Todos<sup>o</sup>
- Score
- Score padrão (XTrust)<sup>o</sup>
- Card 4<sup>o</sup>
- Score Desconhecido<sup>o</sup>
- Insights
- Insights padrão<sup>o</sup>
- Insights SIM Swap<sup>o</sup>
- Insights Trust Card<sup>o</sup>
- Ratings<sup>o</sup>

Enviar para análise: não precisa do modal de confirmação Enviar para Análise

Possibilidade de selecionar como quiser

MC

# Processo de Backtest

## Protótipo – Tela de Histórico

← Voltar
Baixar template
ⓘ

### Histórico

Exibindo os últimos 10 arquivos

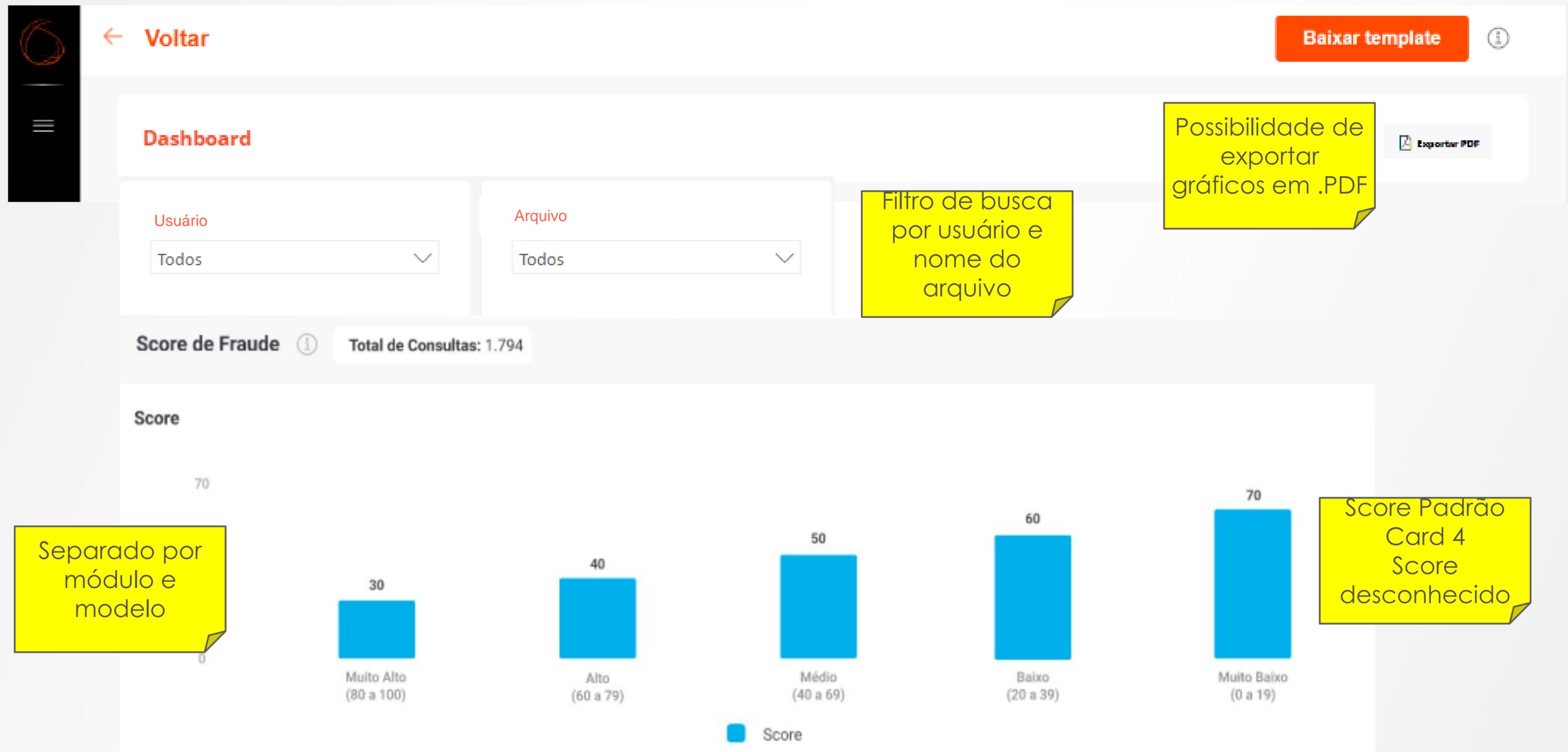
Arquivo	Data	Transações		
Teste - Consulta em Lote.csv 0.001174 Mb	Enviado em 26 jul. 2023 às 04:11	10 válidas • 9 inválidas	Concluído	↓
Clear Sale.csv 0.005989 Mb	Enviado em 24 jul. 2023 às 06:19	51 válidas • 51 inválidas	Concluído	↓
Teste - Consulta em Lote.csv 0.001174 Mb	Enviado em 21 jul. 2023 às 13:45	10 válidas • 9 inválidas	Concluído	↓
Teste - Consulta em Lote.csv 0.001174 Mb	Enviado em 19 jul. 2023 às 13:36	10 válidas • 9 inválidas	Concluído	↓
Teste - Consulta em Lote.csv	Data	Transações		

Tela de Histórico igual da Consulta em Lote



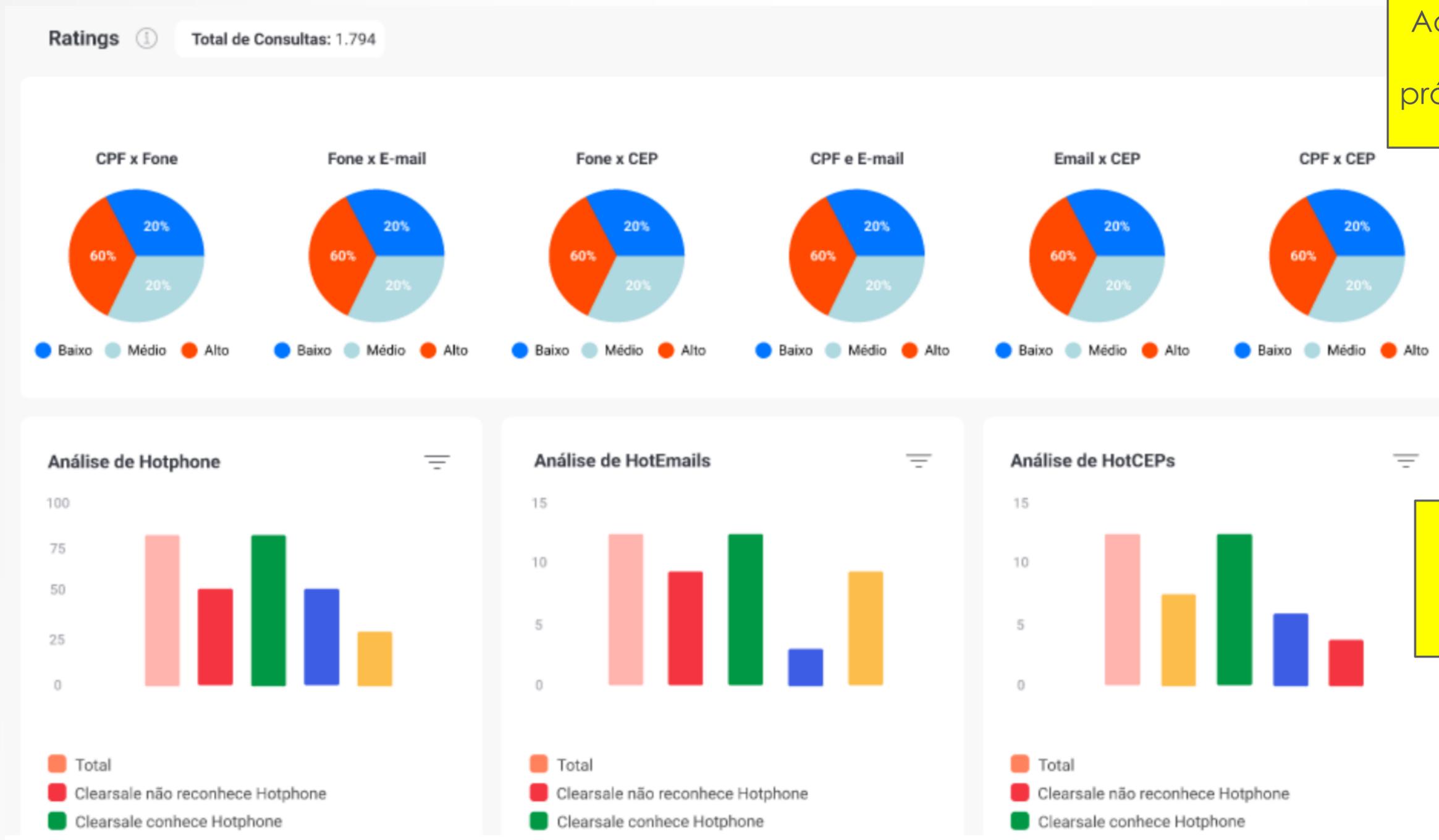
# Processo de Backtest

## Protótipo – Tela de Dashboard



# Processo de Backtest

## Protótipo – Tela de Dashboard



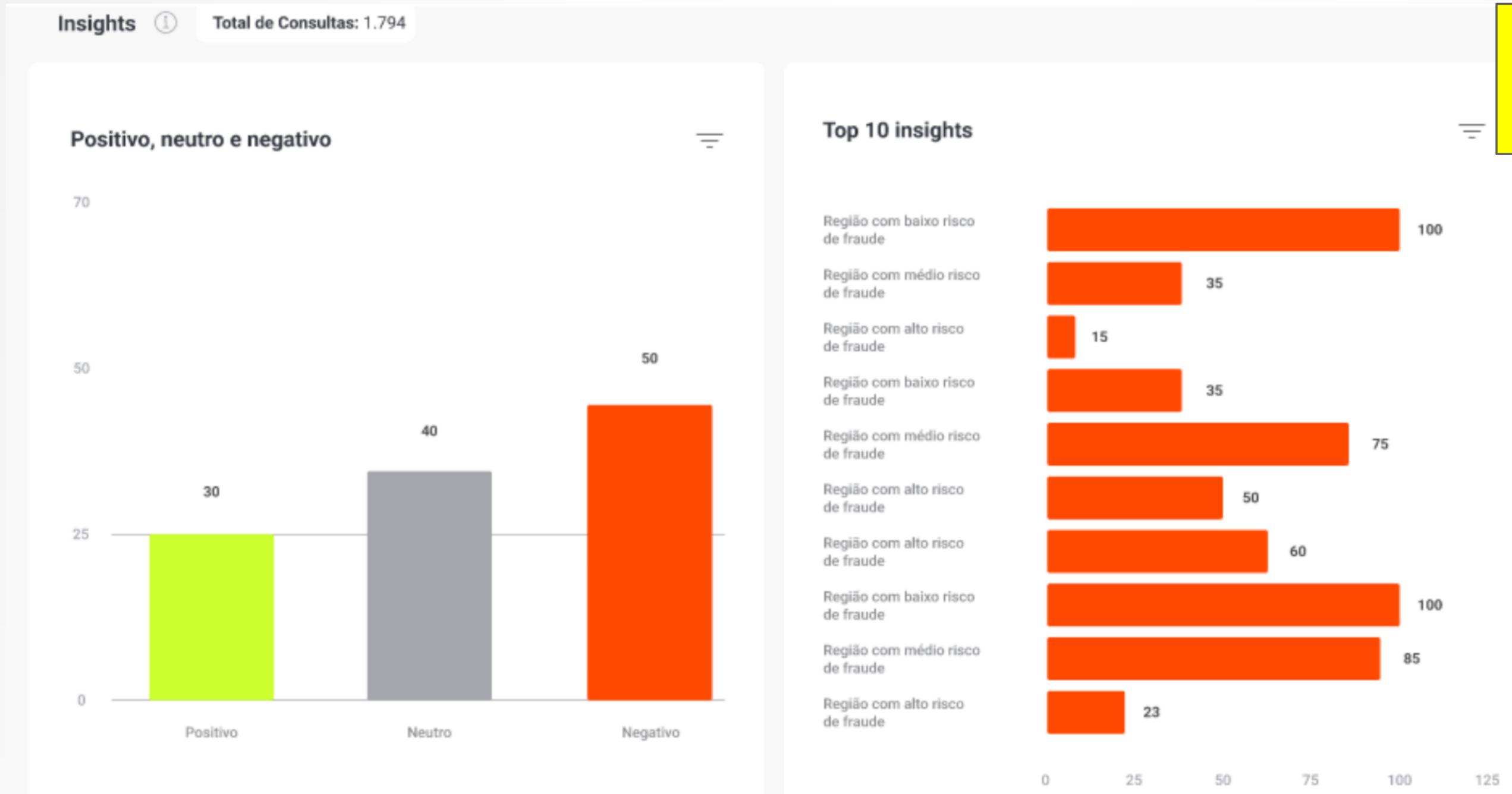
Ao rolar a tela, aparece o próximo módulo

Ter visão de análise de dados (como esta)

# Processo de Backtest

## Protótipo – Tela de Dashboard

Ter visão de análise de dados (como esta)



# Processo de Backtest

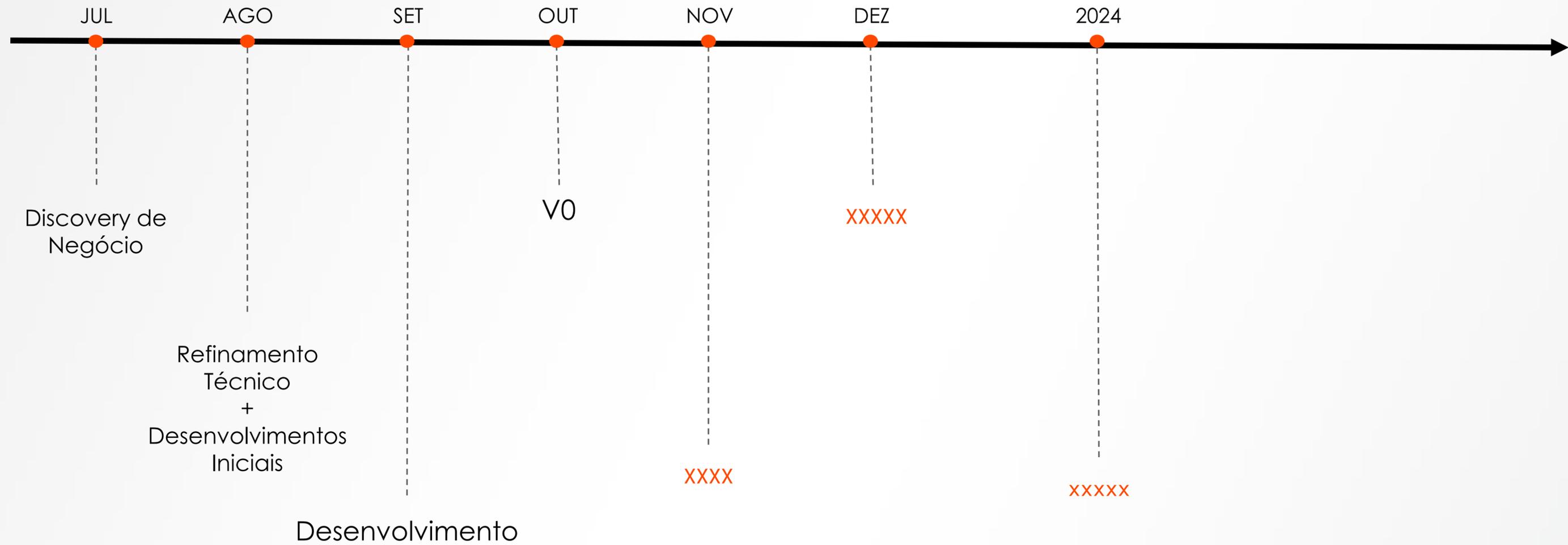
## Protótipo – Tela de Dashboard



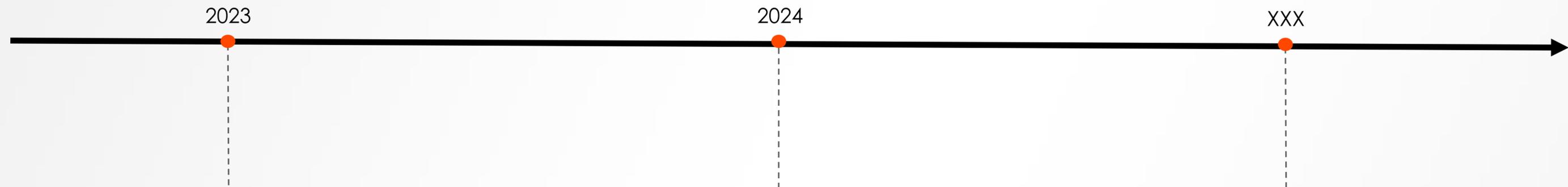
# sumário

1. Backtest: O que é
2. Backtest: Problemas Atuais
3. Backtest: Discovery
4. Backtest: Proposta de MVP
- 5. Backtest: Eventos e Métricas**
6. Anexos e Materiais

# Backtest Eventos



# Backtest Evolução



BT Upsell/Lote:  
fornecer insumos  
para que o time  
comercial apresente  
para o cliente  
resultados dos  
módulos ainda não  
contratados

A equipe comercial precisa estar apta para analisar os resultados e entender o ganho que os mesmos levarão para a operação do cliente, para que assim consiga apresentar da melhor maneira o valor do nosso produto

Ajustes no processo atual, visando automatizar/facilitar o processo de ponta a ponta (demandante - equipes internas - cliente final)

Entendimento e planejamento de integração das duas plataformas (SF e Jira), tendo em vista que o workflow do time de backtest é todo realizado pelo Jira, uma integração entre as mesmas removeria bastante processos manuais e assim, podemos reduzir nosso tempo total de onboarding de novos clientes

Cliente final realizando backtest

# Consulta em Lote

## Métricas de Sucesso

### TBT

- Definir o que será métrica de sucesso para o squad
- Definir o que será métrica de sucesso para o negócio
- Definir métricas de usabilidade
- Entrada em PRD = metrifique e colete feedback

Aumento de receita da PDT



Diminuir tempo de processamento



Facilitar venda de novos componentes



Aumentar TAM da PDT



# Consulta em Lote

## Evolução

Com o foco em reduzir as dores dos clientes atuais, bem como trazer novas experiências para os novos clientes que passarão a utilizar o produto, temos que XXXX.

### Feedback

Feedback com clientes: oportunizar espaços para coleta de relatos e melhorias de experiência. Além de coleta de dados via Lucky Orange (navegação, autonomia da jornada, usabilidade). Focar nas dores coletadas da utilização da solução legada.

### Funcionalidades

Análise de funcionalidades que podem agregar valor ao processo de backtest..  
Pausa do arquivo  
Processamento concomitante de arquivos  
Notificações

### Métricas

Experiência do usuário:  
Métricas: extrações dos dados, relatórios, gráficos, dashboards  
Métricas de produto (interno): analytics

# Processo de Backtest

## Sobre

Problema	Como vamos resolver?	Como saberemos que resolvemos esse problema?	Como vamos medir
<p>Precisão: Certificar-se de que o processo de backtest seja válido e representativo da situação real de produção</p>	<p>Utilização de modelos que vão ao encontro das necessidades do cliente</p>	<p>METRIFICAÇÃO DO FUNIL DE VENDAS</p>	
<p>Recursos: Gerenciar e otimizar os recursos para garantir que o backtest seja executado de forma eficiente</p>	<p>Criação de ambiente específico Automatização fluxo BackTest com todos componentes e modelos disponíveis</p>	<p>Diminuir CAC</p>	<p>Cumprir 100% dos SLAs acordados no fluxo de atendimento</p>
<p>Resultados: Garantir que as partes interessadas internas e externas entendam os resultados do backtest, suas limitações e suas implicações</p>	<p>Apresentação dos resultados de maneira visual</p>	<p>Ajudar time de pré sales a atuar proativamente na construção de valor para o cliente</p>	
<p>Áreas envolvidas: Integrar o backtest com outros sistemas e processos pode ser complexo e requer coordenação com outras equipes técnicas, o que demanda muito tempo até a entrega final</p>	<p>Automatização de processos</p>	<p>DIMINUIR TIME to VALUE</p>	<p>Reduzir em X% esforço de gestão manual do backtest</p>

# Processo de Backtest

## Pessoas e Áreas

**Objetivo: Automatização fluxo BT com todos componentes e modelos disponíveis self service**

Pessoas e áreas envolvidas

Squad Evolução

Field Data Science

Pre Sales

Analytics

Entregas e responsabilidades de cada área

Squad Evolução: Backtest Automatizado

Field Data Science:

Pre Sales: Cliente final

Analytics: Ambiente para processamento (CAOS)

# sumário

1. Backtest: O que é
2. Backtest: Problemas Atuais
3. Backtest: Discovery
4. Backtest: Proposta de MVP
5. Backtest: Eventos e Métricas
- 6. Anexos e Materiais**

# Regras de Negócio

## Visão Geral

A intenção do backtest é tentar reproduzir o que teria acontecido na operação, se o cliente tivesse utilizado o modelo no passado. Portanto, para fazer um backtest nós precisamos que o cliente já tenha uma operação em andamento, ou seja, não seja um cliente que está iniciando seu negócio recentemente, que possua base de dados com as informações dos pedidos (CPF, E-mail, Celular, CEP, Data - sendo que os únicos dados obrigatórios são CPF e Data) e que essas informações tenham pelo menos uma distância de 3 meses da data atual, dado que precisamos da marcação da fraude para entender a aderência do produto (Variável resposta).

A funcionalidade de backtest deve incluir os seguintes elementos chave:

- 1. Importação de dados históricos:** Os usuários devem ser capazes de importar dados históricos relevantes, como histórico de transações, de forma fácil e eficiente.
- 2. Configuração de parâmetros:** A ferramenta deve permitir que os usuários configurem os parâmetros de suas estratégias de negociação, como modelos e módulos..
- 3. Simulação e execução:** O backtest deve executar a estratégia em dados históricos, aplicando os parâmetros definidos pelos usuários para identificar oportunidades de negociação e calcular o desempenho hipotético.
- 4. Análise de desempenho:** A funcionalidade deve fornecer métricas e estatísticas relevantes para avaliar o desempenho da estratégia, como retorno total, taxa de acerto, entre outros..
- 5. Visualização de resultados:** Os resultados do backtest devem ser apresentados de forma clara e visualmente atraente, permitindo que os usuários compreendam rapidamente o desempenho da estratégia e possam identificar áreas para otimização.
- 6. Testes robustos e confiáveis:** A ferramenta de backtest deve ser testada e validada para garantir sua confiabilidade e precisão. Ela também deve ser capaz de lidar com diferentes frequências de dados e períodos de tempo para a realização dos testes.
- 7. Otimização:** O produto pode incluir recursos avançados de otimização, como busca por parâmetros ideais, para auxiliar os usuários a melhorarem suas estratégias com base em diferentes cenários de mercado.
- 8. Integração com a plataforma de negociação:** Idealmente, a funcionalidade de backtest pode ser integrada à plataforma de negociação, permitindo que os usuários executem suas estratégias em tempo real após a validação dos resultados do backtest.
- 9. Educação e suporte:** O produto pode oferecer recursos educacionais para ajudar os usuários a entenderem como realizar backtests de maneira eficiente e interpretar os resultados.

- 1. Dados históricos adequados:** Certificar-se de que haja um conjunto representativo de dados históricos sobre transações e eventos relacionados à fraude para que o backtest seja significativo e relevante.
- 2. Capacidade de recriar cenários passados:** O sistema deve ser capaz de reconstruir situações passadas usando os dados históricos e aplicar as regras e modelos antifraude existentes na época. Isso permite avaliar como as decisões teriam sido tomadas na prática.
- 3. Análise de resultados:** O backtest deve fornecer métricas e relatórios que ajudem a entender a eficácia das estratégias antifraude em cenários passados. Isso pode incluir taxas de detecção de fraudes, falsos positivos, taxas de aprovação de transações legítimas e outras métricas relevantes.
- 4. Capacidade de ajuste:** O backtest deve permitir a realização de ajustes nas regras e modelos antifraude com base nos resultados obtidos. Isso pode envolver a introdução de novas regras, modificação de pesos em modelos de machine learning, entre outras ações.
- 5. Validação de decisões passadas:** O backtest permite que a equipe verifique se as decisões tomadas pelo sistema em situações passadas foram corretas, ou seja, se os casos identificados como fraudes eram realmente fraudulentos e se as transações legítimas foram corretamente aprovadas.
- 6. Garantia de privacidade e conformidade:** Certificar-se de que o backtest é conduzido de acordo com as políticas de privacidade e regulamentos aplicáveis, garantindo que dados sensíveis e informações confidenciais sejam devidamente protegidos.
- 7. Capacidade de automação:** O backtest pode ser uma tarefa complexa e trabalhosa, portanto, é desejável que seja possível automatizar o processo tanto quanto possível, tornando-o uma rotina recorrente no ciclo de desenvolvimento do produto.
- 8. Avaliação do impacto de atualizações:** Se novas regras ou modelos forem introduzidos, o backtest pode ser usado para avaliar o impacto dessas mudanças no desempenho geral do sistema.

## Requisitos Funcionais e Não Funcionais

### Requisitos Funcionais

- 1. Importação de Dados:** Permitir que os usuários importem dados históricos, como arquivos CSV.
- 2. Configuração de Parâmetros:** Permitir que os usuários ajustem os parâmetros do backtest, como modelos e módulos a serem testados.
- 3. Execução do Backtest:** Oferecer a capacidade de executar o backtest com base nas configurações definidas pelos usuários e fornecer os resultados correspondentes.
- 4. Métricas de Desempenho:** Calcular automaticamente as métricas de desempenho, como retorno do modelo selecionado, utilização de diferentes módulos, entre outros.
- 5. Análise de Resultados:** Apresentar os resultados do backtest de forma clara e visualmente atraente, com gráficos e tabelas que mostram o desempenho histórico da estratégia.
- 6. Exportação de Resultados:** Permitir que os usuários exportem os resultados do backtest para compartilhamento ou análise externa, em formatos como CSV ou PDF.
- 7. Segurança dos Dados:** Garantir que todos os dados e resultados de backtest sejam tratados com segurança e privacidade, cumprindo as regulamentações de proteção de dados aplicáveis.
- 8. Suporte:** Forneça referências e documentação de suporte, como manuais do usuário, FAQs e guias de boas práticas para ajudar os usuários a utilizar a funcionalidade de backtest de forma eficaz.

## Requisitos Funcionais e Não Funcionais

### Requisitos Não Funcionais

- 1. Desempenho:** Garantir que a funcionalidade de backtest seja rápida e responsiva, mesmo quando lidando com grandes volumes de dados históricos.
- 2. Precisão:** Assegurar que os cálculos e resultados do backtest sejam precisos e confiáveis, minimizando erros e discrepâncias.
- 3. Escalabilidade:** Certificar-se de que a funcionalidade possa lidar com um grande número de usuários simultâneos e volumes crescentes de dados.
- 4. Confiabilidade:** Assegurar que a funcionalidade de backtest seja robusta e estável, evitando falhas e interrupções inesperadas.
- 5. Usabilidade:** Projetar a interface do usuário de forma intuitiva e amigável, facilitando a configuração dos parâmetros do backtest e a interpretação dos resultados.
- 6. Compatibilidade:** Garantir que a funcionalidade seja compatível com diferentes sistemas operacionais, navegadores e dispositivos, para alcançar um amplo público de usuários.
- 7. Documentação e Suporte:** Fornecer documentação abrangente e suporte ao usuário para ajudar os usuários a entender e utilizar efetivamente a funcionalidade de backtest

# Processo de Backtest

## Casos de Uso

- 1. Realização de um Backtest Básico:** O usuário insere os parâmetros básicos (base de dados e seleção do modelo e módulos). O backtest é executado, e o usuário visualiza os resultados para avaliar o desempenho da estratégia dentro do contexto selecionado.
- 2. Teste de Estratégias Avançadas:** Neste caso de uso, o usuário configura uma estratégia mais complexa, utilizando variável resposta, bem como diferentes modelos. O backtest é executado para avaliar o desempenho da estratégia em cenários mais abrangentes.
- 3. Análise de Sensibilidade:** O usuário realiza um backtest, variando certos parâmetros-chave (modelos). O objetivo é entender como o desempenho da estratégia é afetado por essas mudanças e identificar os melhores ajustes.
- 4. Comparação de Estratégias:** O usuário realiza vários backtests com diferentes estratégias. Em seguida, compara os resultados para determinar qual abordagem teve melhor desempenho com a utilização de um determinado módulo ou modelo.
- 5. Exportação de Resultados:** Após a conclusão de um backtest, o usuário exporta os resultados em formato CSV ou PDF.
- 6. Acesso a Histórico de Backtests Anteriores:** O usuário pode acessar e revisitar os resultados de backtests anteriores para análise comparativa ou referência ao desenvolver novas estratégias.

# Processo de Backtest

## Personas

### Persona 1 - Interno ClearSale (PreSales)

#### NOME E FOTO

**NATALIA LIMA**

#### PERFIL

Comunicativo

35  
anos

Pro-  
ativa

Perfil de  
Liderança

#### Nível de senioridade



#### Conhecimento sobre a Plataforma Data Trust



#### Conhecimento de programação



#### COMPORTAMENTOS

Visão  
holística

Acompanha e  
está por  
dentro das  
novidades de  
mercado

#### NECESSIDADES

Velocidade  
em realizar  
estudos

Processamento  
de grandes  
bases

Apresentações  
para cliente

Visualização  
gráfica dos  
resultados

Arquivo para  
detalhamento  
dos resultados

Todos os  
modelos e  
módulos da  
PDT

Interface  
amigável

# Processo de Backtest

## Personas

### Persona 2 - Interno ClearSale (Analytics)

#### NOME E FOTO

***PEDRO SILVA***

#### PERFIL

Cientista de Dados

28 anos

#### COMPORTAMENTOS

Analítico

Conhecimento de negócio

#### NECESSIDADES

Velocidade em realizar estudos

Processamento de grandes bases

Apresentações para cliente

Visualização gráfica dos resultados

Arquivo para detalhamento dos resultados

Todos os modelos e módulos da PDT

Interface amigável

#### Nível de senioridade



#### Conhecimento sobre a Plataforma Data Trust



#### Conhecimento de programação





# Processo de Backtest

## FAQ

1) No fluxo AS IS, o processamento acontece através da Sandbox do DT V1 e contempla somente os módulos de Score, Ratings e Insights. Já na proposta de fluxo TO BE, o processamento será via Interface Web da PDT. Este processamento irá acontecer diretamente no ambiente de produção? Se sim, fizemos um mapeamento de impactos e riscos previstos (por exemplo: risco de sujar a base de produção, custos, impactos em performance caso tenhamos altas demandas de backtest e também volumetrias, etc.)?

R: A arquitetura que vem sendo desenhada não irá passar pelo ambiente de produção da Plataforma. Como o backtest não é uma transação, não precisamos passar pela Plataforma, e acessar diretamente nosso repositório de módulos/modelos (Wile e Biac).

No que diz respeito a impactos na base, está mapeada a tratativa para isso.

Para a questão de impactos de performance, temos mapeado a restrição de requisições que o banco suporta e vamos colocar essa limitação em nosso processamento para que não tenhamos impactos relacionados à alta volumetria.

2) No fluxo AS IS, o time de Backtest precisa adequar a base, que recebe do Comercial, para colocar no padrão de processamento. Como fica este processo no fluxo TO BE? A solução automática já prevê esta tratativa ou teremos alguma alteração prevista no template do arquivo atualmente utilizado?

R: Em conversa com o time de backtest e time de pre sales, entendemos que se conseguirmos disponibilizar um arquivo de template e instruções de preenchimento atendemos à necessidade de não precisar de ajuste de base. Além disso, a ferramenta prevê o retorno de um arquivo com os dados inválidos, informando o que está errado (por exemplo: CPF incorreto).

3) Para lotes acima de 100K, como faremos? Os times responsáveis precisarão separar em diversos arquivos e rodar um por vez?

R: Neste momento estamos analisando o impacto que uma arquivo "gigante" tem no navegador web (inserir um arquivo de 1 milhão de linhas e analisar o tempo que demora para fazer o upload). Se não tivermos impactos de lentidão, vamos deixar sem limitador de tamanho. Se tivermos impactos negativos, vamos deixar com 100k que foi o número acordado com o time de backtest e pre sales. E sim, caso tenhamos uma base maior, ela precisará ser dividida.

4) Existirá a possibilidade de processamento simultâneo e orquestração de filas? Ou a gestão disso será feita pelos times responsáveis, de modo a garantir que teremos o processamento e execução de um lote / entidade que está sendo 'backtestada' por vez?

R: Existirá a possibilidade de processamento simultâneo, mas colocaremos também um limitador de arquivos processados simultaneamente. Caso esse limite seja ultrapassado, teremos a orquestração de filas.

5) Fizemos um mapeamento / levantamento de quantos módulos e quantos modelos temos atualmente? O MVP irá atender 3 módulos e 2 modelos, a princípio. Mas quantos módulos e modelos temos ao todo e que precisaremos evoluir / implementar depois?

R: Atualmente temos 7 modelos que são testados para a PDT, sendo 2 modelos os mais usados (em 80% dos casos). São esses dois que iremos incluir no MVP. No que diz respeito aos módulos, temos 10 módulos, sendo 3 os mais usados (e que entrarão no MVP).

*Os módulos e modelos para o MVP foram acordados entre backtst e pre sales*

6) Como fica a tabela de SLAs frente ao novo fluxo automatizado? Tais SLAs não incluem a geração dos gráficos e apresentações atualmente, certo?

R: No caso do fluxo automatizado, não teremos mais esse SLA pois devolveremos em "tempo real" os resultados, via arquivo em excel analítico, linha a linha e via dashboard.

O SLA na proposta que estamos montando seria apenas o tempo do processamento do arquivo, ou seja, algumas horas.

Depois, ficará a cargo do time de pre sales montar a apresentação.

Esse SLA apresentado pelo time de backtst envolve o recebimento da base, tratativas, processamento e retorno dos resultados (excel analítico linha a linha).

# Materials Data Field

- SLA: <https://docs.google.com/presentation/d/1VJq73jCCWjQyn6vEKotyj5Qr6Ql4D8f2/edit?usp=sharing&oid=101456555703908036471&rtpof=true&sd=true>
- Treinamento: [https://drive.google.com/file/d/1fgu-w1JKL6ZeGKz6xehWt5VT1Zxz\\_5Ym/view?usp=drive\\_link](https://drive.google.com/file/d/1fgu-w1JKL6ZeGKz6xehWt5VT1Zxz_5Ym/view?usp=drive_link)
- Repositorio: <https://drive.google.com/drive/folders/1VRzQxJ7uarVbAlsnfQzv39Y9sSb6N228>
- Formulário: <https://clearsale.atlassian.net/jira/core/projects/FDS/form/1>
- Exemplo apresentações PreSales

# Materials

- Zendesk: <https://clearsale.zendesk.com/hc/pt-br/articles/12900596496027-Processo-de-Backtest-Plataforma-Data-Trust>
- Discovery v0:  
[https://docs.google.com/document/d/1rwT8Kms8RsBzgCF3z7QgAK\\_UQwAECiBDhhBOnuLyVv0/edit#heading=h.2m3bgc52ub5e](https://docs.google.com/document/d/1rwT8Kms8RsBzgCF3z7QgAK_UQwAECiBDhhBOnuLyVv0/edit#heading=h.2m3bgc52ub5e)
- Processo V0:  
[https://docs.google.com/presentation/d/1mAHXp6QPYPqWRCw\\_Py86tXXNzcNB0RWeZJQiBKmQgo/edit#slide=id.g12a9e93b8d7\\_0\\_70](https://docs.google.com/presentation/d/1mAHXp6QPYPqWRCw_Py86tXXNzcNB0RWeZJQiBKmQgo/edit#slide=id.g12a9e93b8d7_0_70)
- Mapeamento Dores de Clientes:  
<https://docs.google.com/spreadsheets/d/1YrHHf9ooyCF2D9izR60qYZfz33AmJOWD/edit#gid=1436370632>
- Uso de bureaus:  
[https://docs.google.com/presentation/d/1QqNMWsHXJRfvppAJbFq2AqbHBTVF0I6jvi\\_GGeTVjwA/edit#slide=id.g1dc7324887c\\_0\\_91](https://docs.google.com/presentation/d/1QqNMWsHXJRfvppAJbFq2AqbHBTVF0I6jvi_GGeTVjwA/edit#slide=id.g1dc7324887c_0_91)
- <https://miro.com/app/board/uXjVPRd6tk0=/>

# Processo de Backtest

## Documentos

**Objetivo: Automatização fluxo BT com todos componentes e modelos disponíveis self service**

Modelos utilizados em backtest: 9203 (card4), 9100, 9115, 9219, 9144

Tabela modelos:

<https://docs.google.com/spreadsheets/d/1sRcWrnFTLRry5iQFy0hcbaDEVCTPzGza/edit#gid=784169238>

# Processo de Backtest

## Objetivos



•**Personas:** Soluções/Pré-venda/Comercial/Growth

•**Objetivos:**

- Facilitar a vida de Soluções na tomada de decisão
- Facilitar a vida do Comercial em UpSell
- Sugerir Construção de FUNIL e decisão
- Ter opção de download dos resultados para montar material para uso em pitch de vendas.

**1. Facilitação de UpSelling:** O MVP fornece ao time de vendas dados tangíveis e resultados visuais que demonstram a eficácia do produto. Esses resultados são cruciais para argumentos de venda, especialmente em discussões de UpSelling, onde a ênfase está em demonstrar valor adicional aos clientes existentes.

**2. Aprimoramento do Material de Vendas:** Com a visualização clara dos resultados do backtest, a equipe de vendas pode criar materiais de pitch mais convincentes. Isso inclui apresentações, folhas de dados e estudos de caso que destacam os benefícios específicos do produto, baseados em dados reais.

**3. Automação e Eficiência:** A automação das etapas de execução de estudos de backtest reduz significativamente o trabalho manual, agilizando o processo de pré-venda. Isso permite que a equipe de vendas se concentre mais na interação com o cliente e menos em tarefas operacionais.

**4. Decisões Baseadas em Dados:** O MVP fornece uma base de dados sólida para tomada de decisões, tanto internamente quanto para os clientes. Isso inclui a identificação de oportunidades de UpSelling com base no desempenho do modelo em diferentes cenários.

**5. Interface Amigável e Acessível:** Ao oferecer uma interface intuitiva, o MVP garante que os membros da equipe de vendas possam facilmente entender e utilizar os resultados do backtest, mesmo sem profundo conhecimento técnico.

**6. Redução do Ciclo de Vendas:** Com informações mais precisas e relevantes disponíveis rapidamente, o MVP pode ajudar a encurtar o ciclo de vendas, permitindo uma resposta mais rápida às necessidades e perguntas dos clientes.

**7. Demonstração de Valor Personalizado:** O MVP permite simulações baseadas nos dados específicos de cada cliente, tornando a proposta de valor mais relevante e personalizada, um fator crucial no processo de UpSelling.

**8. Melhoria Contínua do Produto:** O feedback e os insights obtidos através do MVP podem ser usados para aprimorar continuamente tanto o produto quanto a estratégia de vendas, assegurando que a plataforma permaneça alinhada às necessidades do mercado.

**Conclusão:**

O MVP de backtest não é apenas uma ferramenta operacional, mas um componente estratégico que empodera as equipes de Pré-Vendas e Vendas com dados e insights valiosos. Ele permite uma abordagem mais consultiva e orientada por dados para vendas, especialmente em situações de UpSelling, reforçando a confiança dos clientes na plataforma e no valor que ela agrega aos seus negócios. A longo prazo, essa abordagem não apenas aumenta as vendas, mas também fortalece as relações com os clientes, promovendo uma parceria mais profunda e baseada em confiança.



ClearSale

intelligence to move\_