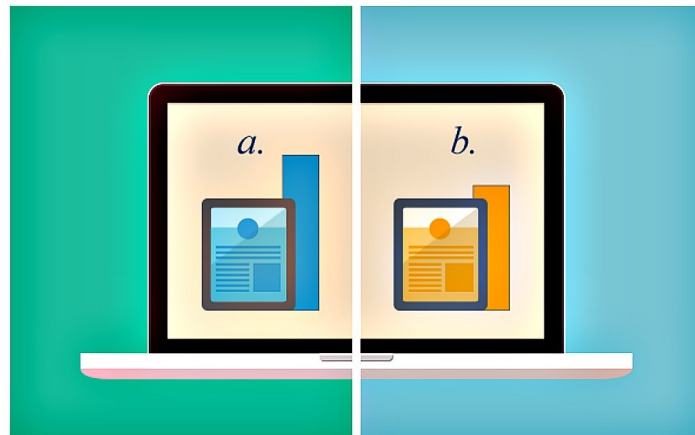


Introdução

Todos os dias trabalhamos para desenvolver novas telas, fluxos, formatos de curso e *features* de produto com um único objetivo: **fazer com que a pessoa usuária utilize o nosso produto para solucionar seus problemas.**

Partindo da lógica de solução de problemas centrada no usuário, os testes são a forma mais eficiente de fazer com que a pessoa usuária participe deste processo, nos ajudando, inclusive, a sermos mais assertivos.



"Testes são a forma mais rápida e mais barata de fazer com que o seu usuário, a pessoa que deve ter o problema resolvido, construa a solução do seu problema junto com você."

Neste material você verá de maneira simples como estruturar o seu teste A/B em uma Ficha de Hipótese, bem como algumas boas práticas que contribuem para que seja possível chegar na validação da sua hipótese ao final dos testes realizados.

A Ficha de Hipótese

A formalização dos testes realizados em uma Ficha de Hipótese faz com que todos os círculos envolvidos na solução do problema do usuário - do desenvolvimento de produto ao conteúdo - tenham conhecimento das modificações que estão sendo realizadas e como estas modificações impactam a nossa pessoa usuária.

Ter visibilidade sobre estas modificações é de extrema importância, tanto para identificar seus impactos na vida do usuário, como para que os demais círculos tenham conhecimento de hipóteses validadas sobre o comportamento dele.

Neste material, utilizaremos a Ficha de Hipótese do teste #60, que buscou validar a hipótese de melhoria do custo de aquisição do usuário com o uso do Visa Checkout, no nosso fluxo de checkout.

1. Qual a sua hipótese?

Todo este processo começa com uma **hipótese**. Uma hipótese normalmente surge de uma análise de comportamento do usuário, seja ela quantitativa ou qualitativa, e se traduz em uma afirmação que relaciona o que você acha que precisa ser modificado e uma suposição de impacto direto que esta modificação pode causar em uma das métricas que refletem o comportamento do nosso usuário ao utilizar o nosso produto.

No nosso exemplo, a seguinte hipótese foi definida:

"A oferta do Visa Checkout com voucher de 30 dias reduz o CAC da empresa X."

Neste teste, devido a uma condição comercial específica, a empresa X ganharia R\$30 por cada cadastro gerado no Visa Checkout, sendo isso denominado "oferta" na hipótese acima.

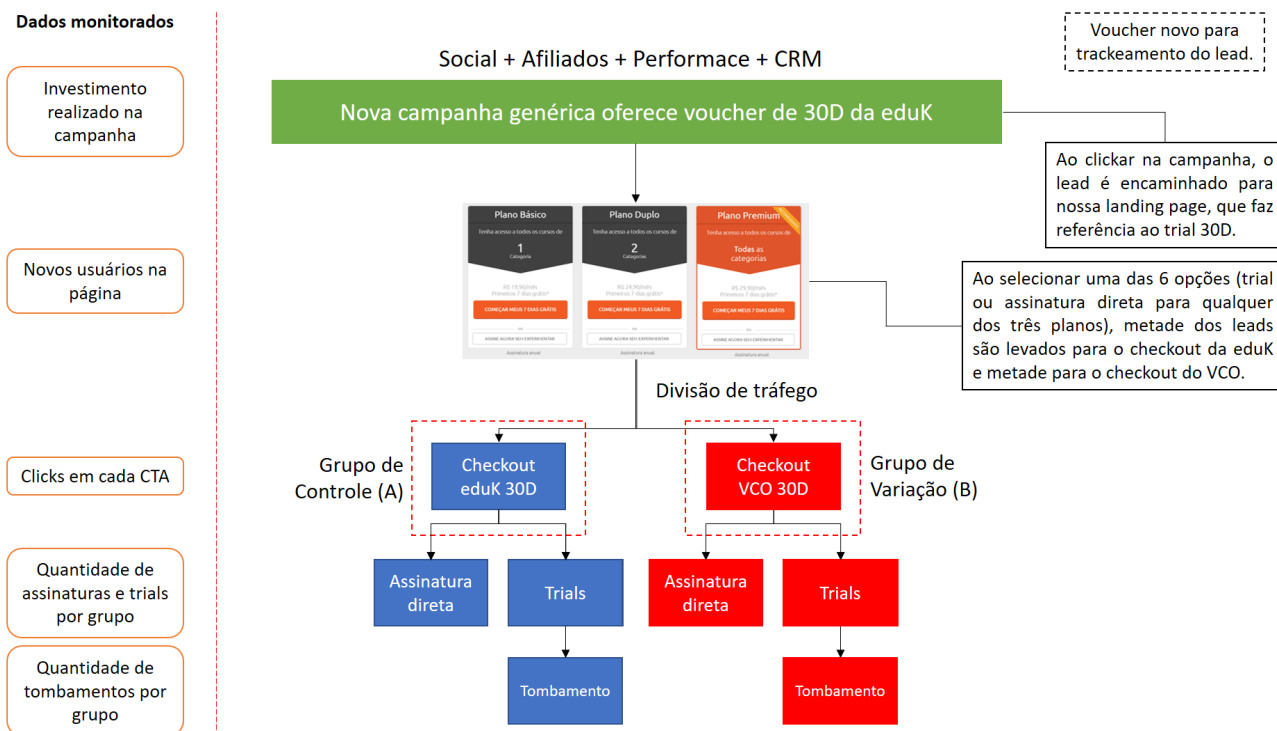
Neste caso, a modificação "Visa Checkout com voucher de 30 dias" reduziria a métrica de aquisição Custo de Aquisição de Cliente (CAC).

2. Descrevendo o seu Experimento

Após definir a hipótese que você quer validar no teste, o próximo passo é descrever o experimento que foi idealizado para validar esta hipótese. Neste ponto, além da contextualização do experimento, as principais perguntas precisam ser respondidas:

2.1. Qual a estrutura do seu experimento?

A estrutura do seu experimento é o que define onde e como ele será realizado. Em geral, a estrutura pode ser resumida em um diagrama que mostra o fluxo de um usuário que estará participando do experimento, como mostrado a seguir:



Como um experimento pressupõe uma modificação em um ponto de contato na experiência do nosso usuário, este passo é importante para que seja possível definir onde esta modificação estará presente e qual o mecanismo aplicado para que uma parte dos usuários continuem vendo o fluxo normal e a outra parte veja o fluxo modificado. Chamamos o primeiro grupo de usuários de "Grupo de Controle" e o segundo grupo de "Grupo de Tratamento ou Variação".

Normalmente, buscamos segmentar o tráfego de usuário no ponto de contato modificado de maneira que a mesma quantidade de usuários esteja no controle e na variação. Caso isso não ocorra, é preciso ter cuidado na normalização dos dados ao final do teste.

2.2. Quem deve participar do seu experimento?

- Usuários novos ou usuários que já acessam a eduK;
- Usuários que entram logados ou deslogados na plataforma;
- Usuários que chegaram de maneira orgânica ou através de uma determinada campanha, às vezes criada para realização do teste.

Atualmente, temos diversos tipos de usuário acessando a nossa plataforma. É importante definir quais usuários devem participar do teste para garantir que determinadas características do usuário não acabem enviesando os resultados do teste.

No nosso exemplo, a seguinte definição foi realizada:

"Não faremos segmentação de usuários, sendo que todo mundo que chegar a página vindo das campanhas e seguir para o fluxo de checkout será sorteado em proporções iguais para um dos grupos."

3. Critérios de validação

Realizamos o teste e temos um resultado. Como saber se a minha hipótese é válida ou não? Neste campo, você definirá quais os critérios você está considerando na análise dos resultados para garantir que, uma vez validada a hipótese, seja possível modificar a nossa experiência e reproduzir na vida real os resultados obtidos com o teste.

3.1. Qual o critério de sucesso do seu teste?

Enquanto a hipótese é uma afirmação aberta que vincula modificação e impacto, o critério de sucesso define quantitativamente qual o mínimo impacto para que valha a pena adotarmos a modificação realizada. Este impacto mínimo normalmente é traduzido em uma variação mínima aceitável em uma determinada métrica avaliada no teste (Métrica de Sucesso). Por exemplo: "Assumiremos como válida a hipótese se variarmos a métrica Taxa de Conversão em, pelo menos, 10%".

Um ponto importante desta etapa é a definição da Métrica de Sucesso do seu teste, momento em que você definirá qual o real impacto que a modificação proposta provocará na experiência do usuário e, por conseguinte, nas métricas de negócio da empresa.

O ideal é que a variação na métrica de sucesso esteja associada aos custos adicionais por manter a variação no ar. No nosso exemplo do teste #60, o racional precisou envolver a receita adicional por cadastro gerado e possíveis impactos na taxa de conversão que a modificação do nosso checkout poderia causar. Você pode conferir a construção deste racional no anexo.

3.2. Quanto tempo vai durar seu teste?

O critério de sucesso está associado a outro ponto muito importante da Ficha de Hipótese que é o tamanho da amostra necessária para garantir que os resultados do teste sejam estatisticamente significantes, o que significa que, uma vez validada a hipótese, podemos garantir que os resultados obtidos irão ocorrer quando a modificação efetuada for disponibilizada para 100% dos usuários.

Existem duas formas de definir se um teste foi concluído:

(a) Quando o tamanho da sua amostra mínima é alcançado. Para calcular o tamanho da amostra mínima, utilizamos uma calculadora de amostra como [esta](#).

Então aplicamos o critério de sucesso na métrica analisada e chegamos à quantidade mínima de usuários participando de cada grupo analisado (Controle e Variação).

Determine a amostra necessária para validar sua hipótese

Taxa de Conversão Atual ?

1.5

%

Expectativa de Melhoria ?

20

%

Significância Estatística ?

95

%

Amostra por variação

26.518

Este é o número mínimo de visitas em cada página para validar a hipótese.

(b) Quando a variação provocada na métrica de sucesso, para mais ou para menos, é suficiente para que o teste tenha significância estatística. Neste caso, utilizamos esta [calculadora](#). Colocamos a quantidade de usuários que participaram até o momento do teste e quantos alcançaram o resultado esperado, e a calculadora mostra se o resultado obtido é significativo ou, se não, quantos usuários ainda são necessários para validar a hipótese. Dispondo dos dados de tráfego no ponto de contato analisado e do tamanho necessário da amostra, chegamos à estimativa de tempo de duração do teste.

Group A is a clear winner (99.9%*)

Group A: **13.5%** conversion

Group B: **10.64%** conversion

	Visitors:	Goals:
Group A:	10000	1350
Group B:	9000	958

Calculate

Qual devemos utilizar? Na verdade, os dois critérios devem ser utilizados e, quando o primeiro for obtido, o teste deve ser concluído. Os testes começam com uma expectativa de variação em uma métrica analisada e, para esta expectativa, devemos calcular a amostra mínima necessária.

No entanto, para mais ou para menos, essa expectativa pode não ser praticada e por isso usamos o critério (b). No nosso exemplo, a modificação produziu uma variação negativa muito superior ao esperado e, por isso, o teste foi concluído antes do esperado.

De maneira geral, variações esperadas muito pequenas resultam em amostras muito grandes e um longo tempo de duração do teste para a significância estatística. Neste ponto, vale o bom senso para definir quais testes devem ser priorizados ou, no limite, para variações muito pequenas, se vale a pena ser realizado.

4. Como o teste será monitorado?

Como uma das nossas premissas é a visibilidade, precisamos saber como o teste será monitorado e onde os seus resultados estarão disponíveis para que a empresa tenha conhecimento. Nesta etapa, é importante definir como será calculada a métrica de sucesso e como ela será monitorada, além das ferramentas utilizadas para este monitoramento. Caso os resultados estejam disponíveis em um dashboard, aqui é o melhor lugar para disponibilizá-lo.

5. Premissas e vieses

Neste campo são relacionadas as premissas que assumimos para realização do teste ou possíveis vieses que podem influenciar no resultado do teste.

Algumas boas referências

TEMPLATE: #60 - Aquisição - Com a oferta VCO a eduK reduz o CAC

Responsável pelo experimento: Fulano de Tal

Objetivo:

Reduzir o CAC da empresa X em, pelo menos, R\$5,00.

Hipótese:

A oferta do Visa Checkout com voucher de 30 dias reduz o CAC da empresa X.

Datas importantes:

data de criação do documento	31/03/2019
data de início do experimento	19/04/2017 às 10hr.
data de conclusão do experimento	20/04/2017 às 19hr.

Resultados do experimento:

	Participantes	Conversões (Trial)	Taxa (Trial)
Tratamento	2587	48	1,8%
Controle	2475	149	6%

Resultados esperados X realizados:

Possíveis Métricas Impactadas	Dados Grupo de Controle	Dados Grupo Tratado
Tx Conversão (Trials)	6%	1,8%
Total trials gerados	149	48
Total assinaturas diretas geradas	2	0

Descrição do experimento:

Serão criadas campanhas específicas via Social, afiliados, performance e CRM que veiculam o voucher de 30 dias grátis oferecido pela eduK. Este voucher é novo e diferente do voucher já utilizado em campanhas anteriores de teste para 30 dias ou campanhas de VCO.

Dados monitorados

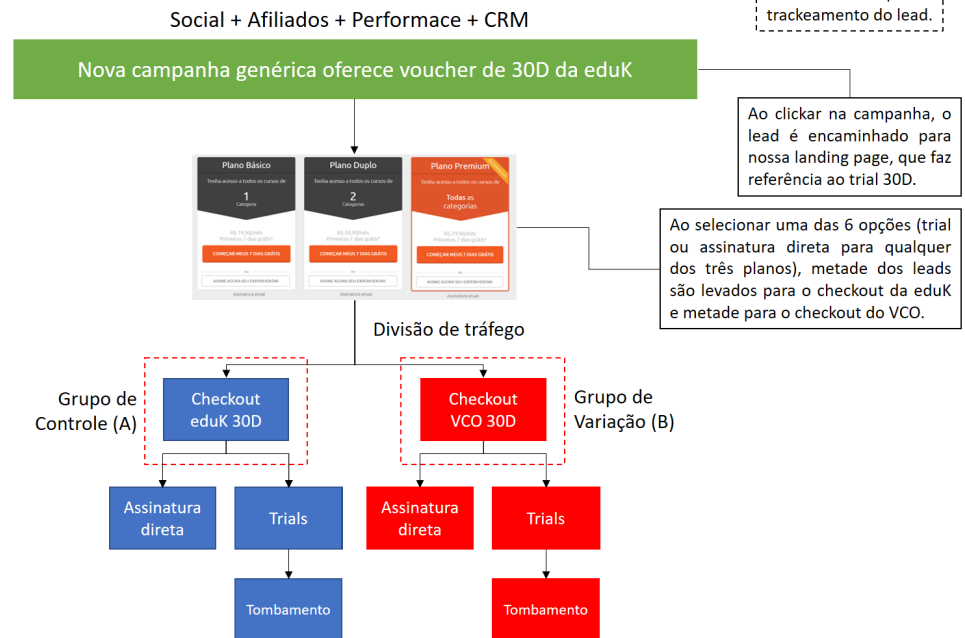
Investimento realizado na campanha

Novos usuários na página

Clicks em cada CTA

Quantidade de assinaturas e trials por grupo

Quantidade de tombamentos por grupo



Estrutura do teste A/B

As campanhas irão direcionar os leads para a landing page. No site, serão identificadas pessoas advindas das campanhas para as quais serão exibidas sempre nos CTAs a mensagem referente ao trial de 30 dias.

Ao avançar no funil, o lead chegará à tela de checkout, onde o tráfego será dividido entre telas de checkout de eduK e tela de checkout utilizando o Visa Checkout.

O grupo de controle é caracterizado como o de pessoas que acessaram a versão do fluxo de checkout da eduK, hipótese que já foi validada como vencedora em testes anteriores que a compararam com a opção de trial de 7 dias.

A variação é caracterizada como o grupo de pessoas que acessaram a versão do fluxo de checkout do Visa Checkout, que oferece também o voucher de 30 dias e oferece o Visa Checkout como gateway de pagamento. Contraturalmente, a Visa paga a eduK R\$30 por cada trial gerado, o que é caracterizado como uma redução de CAC para a empresa.

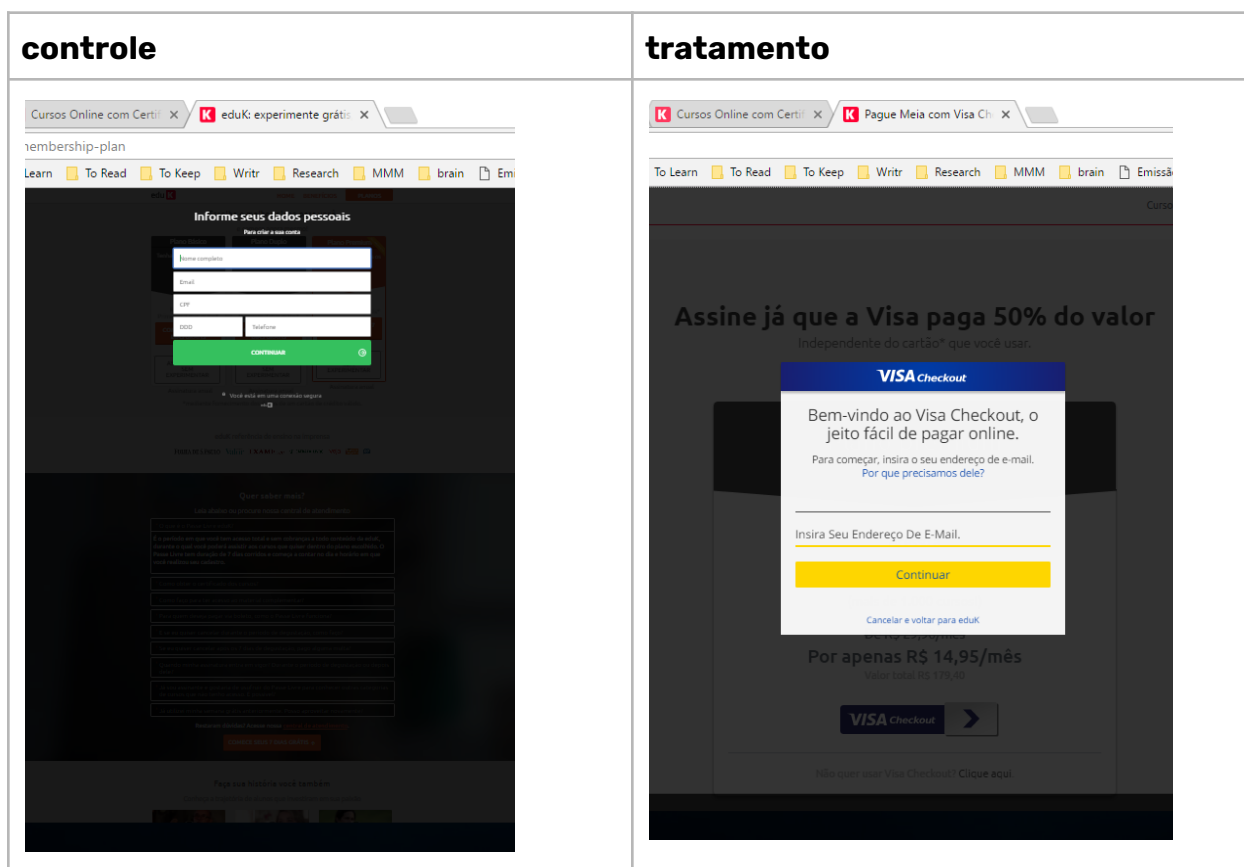
Foi implementado o Visa Checkout para os três tipos de plano e para a assinatura direta. Vale ressaltar que serão computados no dashboard de monitoramento o número de assinaturas diretas geradas pelo Visa Checkout, o que não é realizado

hoje, e a query de aquisição de dados no banco também deve computar estas assinaturas diretas.

Devem deixar de ser veiculados, no período do teste, todos os banners do site que se referem a Visa.

Ponto de atenção! Precisam ser considerados usuários que façam o trial ou assinatura direta e que já possuam conta de Visa Checkout pois, nestes casos, a Visa não paga por lead gerado. Precisaremos contabilizar no final do mês a quantidade de trials/assinaturas diretas geradas que não foram pagas para a eduK, ou assumir um percentual de perda no resultado do teste.

Screenshot das variações:



O teste permanecerá no ar durante o período necessário para ter 95% de [significância estatística](#).

Não faremos segmentação de usuários sendo que todo mundo que chegar à página vindo das campanhas e seguir para o fluxo de checkout será sorteado em proporções iguais para um dos grupos.

No teste realizado anteriormente, a taxa de conversão em trials de 30 dias teve uma melhora de 28,9% em comparação à taxa de conversão em trials para trials de 7 dias.

Crítérios de validação

Validaremos a hipótese caso seja observado que o CAC do trial gerado na variação é, pelo menos, R\$5 mais baixo que o CAC do grupo de controle, considerando o valor pago pela Visa por cada trial gerado.

Para avaliar o CAC, precisaremos identificar o investimento realizado para as campanhas e dividir o valor, de maneira proporcional, entre os grupos de controle e variação. Do valor resultante para o grupo da variação, devemos diminuir do valor investido R\$30 para cada trial ou assinatura direta gerada.

Precisaremos avaliar, após 30 dias do início da campanha, o número final de assinantes gerados por grupo. **Só com este resultado final teremos condição de atestar a validade do experimento.**

Dispondo do investimento real por grupo e do número de assinantes gerados por grupo, é possível calcular e comparar os CACs de cada grupo.

Definimos que o teste foi concluído para uma amostra de, pelo menos, 42.837 visitas em cada tipo de página. Desta forma, estimamos em cerca de 8 dias a duração do teste.

Utilizamos para definir o tamanho da amostra uma [calculadora de amostra](#).

Assumimos como premissa que o Visa Checkout irá diminuir a taxa de conversão da plataforma, mas que o ganho de R\$30 por cadastro gerado compensa a redução. Para definir o critério de resiliência do teste, identificamos qual o máximo de decréscimo na taxa de conversão é aceito. Este máximo considerou os ganhos mínimos que a eduK teria caso a taxa de conversão chegasse neste valor limite.

Para determinar este valor, primeiramente, identificamos a distribuição de assinaturas atuais por plano, quando encontramos os seguintes valores:

Plano básico: 45,39% (Preço: R\$ 19,90)

Plano duplo: 6,01% (Preço: R\$ 24,90)

Plano premium: 48,61% (Preço: 29,90)

Utilizando estes valores, calculamos o ticket médio do nosso usuário atual – sendo ele calculado para 12 parcelas iguais, ou seja, um ticket anual de R\$300,73.

Com a mesma proporção, calculamos o ticket médio anual, assumindo que ganharemos R\$30 por cada ticket gerado, o que resulta em um ticket anual médio de R\$330,73. Assumindo que cerca de R\$5 são perdidos como custo operacional, chegamos no ticket anual médio de R\$325,73.

Feito isso, o próximo passo foi determinar, para uma mesma quantidade de usuários, qual o decréscimo admitido na taxa de conversão, sendo que o limite é, para estes usuários, termos a mesma receita gerada nos casos com e sem Visa Checkout. Nesse caso, identificamos que com o decréscimo de 8%, a receita financeira das duas opções é igual.

Assumindo que outros erros podem influenciar neste processo, adotamos como 15% a máxima variação aceitável na taxa de conversão do teste.

Além disso, assumimos como premissa que a distribuição de planos se mantém a mesma e não é influenciada pelo Visa Checkout.

Monitoramento e métricas

Para acompanhar as variáveis necessárias para a avaliação do experimento, foi configurado o VWO e foram criadas *queries* para requisição de dados de conversão. A divisão de tráfego na página de checkout foi feita utilizando o VWO, que também será utilizado para identificar a conclusão do teste.

Os dados de tombamento serão imputados no VWO, para que possamos avaliar todo o desempenho do teste na ferramenta.

Para análise do teste, iremos realizar as seguintes ações:

1. Trackear o número de pessoas com o voucher novo que chegam na página (VWO);
2. Trackear o número de pessoas que clicam em quaisquer dos CTAs, em cada grupo (VWO - Goal 1 e 2);
3. Trackear o número de pessoas que viraram trial ou assinante direto (parecido com o anterior, só que na nossa base) em cada grupo (Query);
4. Analisar, dentre as pessoas que viraram trial depois de 30 dias, quais usuários viraram assinantes em cada grupo (Query).

Query

```
select t.shopper_reference, t.triggered_by, t.status, t.value_cents,
(select triggered_by from payments.transactions where shopper_reference =
t.shopper_reference and status = 'authorized' limit 1) origin,
t.created_at
from payments.transactions t
where t.shopper_reference in
(select subscription_id from edukflix_api.transaction_metadata where traffic_source
regexp 'test_ab_vco')
and t.status in ('approved', 'authorized')
order by t.id desc;
```

No VWO, os grupos de controle e tratamento serão comparados com relação ao dinheiro gerado por cada um dos grupos. Para viabilizar o cálculo do dinheiro gerado por assinaturas pós-trial (tombamento), o backend da plataforma será integrado com o VWO, de forma a tornar possível a visualização de sucesso ou fracasso do teste em uma única plataforma.

A seguinte lógica irá comparar o ganho de dinheiro por cada grupo de controle:

Checkout eduK

- Assinatura direta: somar valor do plano contratado.
- Trial: Sem valor recebido → Tombamento pós trial: somar valor do plano contratado.

Visa Checkout

- Assinatura direta: somar valor do plano contratado + R\$30.
- Trial: somar R\$ 30 → Tombamento pós trial: somar valor do plano contratado.

Os dados de receita gerados serão imputados no VWO.

Como descrito, o sucesso do teste será analisado através do VWO, assim como os critérios para término e as restrições de decrescimento da taxa de conversão.

Aprendizados/Conclusões

Os resultados do teste podem ser visualizados na tabela descrita no link abaixo:

[Tabela de resultados - Teste VCO](#)

Como principais resultados, vemos que a taxa de conversão para geração de trials é 318% maior no Checkout eduK.

Analisando financeiramente, temos os seguintes resultados:

Investimento na campanha	\$5,582.12		
	eduK	VCO	
Investimento proporcional	\$2,729.31	\$2,852.81	
Assinaturas geradas	76.5	24	Assumindo tombamento de 50%
CAC blended	\$35.68	\$118.87	Considera-se investimento apenas de campanha, mas conversão de todas as fontes de tráfego (Mídia paga, CRM, Social, etc)
Receita Visa	0	1440	Receita de R\$ 30 gerada por cadastro gerado para Visa
CAC blended ajustado	\$35.68	\$58.87	CAC considerando receita gerada pelos cadastros gerados para a Visa

Logo, pode-se verificar que, mesmo ajustando o CAC a partir da receita gerada pelos cadastros da Visa, temos que o CAC do Checkout da Visa é 65% maior que o do Checkout eduK, considerando como premissa que 50% dos trials serão tombados ao final do período de trial.

Possíveis vieses

O teste compara o voucher de 30 dias da eduK com o voucher de 30 dias do Visa Checkout, sendo que uma receita é gerada para cada trial vindo do Visa Checkout. No entanto, ele não valida por que a campanha do Visa Checkout foi mal sucedida entre 27/03 e 31/03.

Outros testes em paralelo poderiam, eventualmente, afetar este teste. Por isso, o teste que compara o Plano único com três planos foi desligado antes do teste ser posto em funcionamento.

Apostas

Saulo = O valor arrecadado pelo grupo de controle (A) é 40% maior que o do grupo da variação (B).

Regis = O valor arrecadado pelo grupo de variação (B) é 15% maior que o do grupo de controle (A).

Lucas = O valor arrecadado pelo grupo de controle (A) é 60% maior que o do grupo da variação (B).

Heider = O valor arrecadado pelo grupo de controle (A) é 20% maior que o do grupo de controle (A).

Recompensa: Sixpack de todos para o ganhador (Não vale Glacial).

Vencedor: Lucas Gomide.