

Relatório – Case Técnico – Ifood Data Analytics

1. Introdução

Este material contém as análises elaboradas para o case técnico para o cargo de Data Analytics no Ifood. O contexto da análise consiste em avaliar os resultados de uma estratégia de cupons como alavanca de crescimento a partir de bases de dados fornecidas com marcação dos clientes que participaram do teste A/B, bem como os pedidos realizados durante o período analisado. O notebook com os resultados obtidos pode ser encontrado em: <https://github.com/kithanos/case-ifood-data-analytics-ab-test>.

2. Indicadores da Campanha

Foram analisados os pedidos dos usuários que participaram do teste A/B no período de dez/18 a jan/19.

2.1. Indicadores Relevantes

Em relação aos indicadores da campanha para analisar se ela possui impacto significativo dentro do período avaliado, podemos relacionar o ticket médio entre os grupos target e controle, e ainda a reincidência de pedidos (se dentro do período compreendido, o usuário realizou mais de um pedido), no primeiro caso, o ticket médio entre os grupos target e controle foram:

```
+-----+-----+
|is_target|    ticket_medio|
+-----+-----+
|  control|47.917294188812164|
|  target |47.809652936992656|
+-----+-----+
```

Aparentemente não há grandes variações entre o público impactado dos 2 grupos no que tange ao ticket médio, ainda sim vale validar esta afirmativa aplicando um método estatístico do teste t de Student, que é usado em testes A/B para verificar se a diferença entre as médias de dois grupos é estatisticamente significativa ou se pode ter ocorrido por acaso. O teste t compara as duas médias levando em conta:

- A média de cada grupo
- A variância dos dados
- O tamanho das amostras

Se o p-valor resultante for menor que 0,05, conclui-se que a diferença entre os grupos é significativa. Aplicando o teste t numa amostra aleatória (10%) da base de dados, obtemos o seguinte valor:

T-statistic: 0.6100

P-value: 0.5419

Não há diferença estatisticamente significativa no ticket médio.

Podemos afirmar diante dos dados que não há impacto significativo do ticket médio no público da campanha. Entretanto isso por si só não reflete os resultados da campanha, afinal a reincidência de pedidos pode indicar que o usuário está fidelizado ao aplicativo, podendo indicar até mesmo principalidade em apps de delivery, vamos repetir as análises, mas olhando para usuários que voltaram a realizar pedidos dentro do aplicativo (em dias diferentes).

is_target	pct_reincidentes
control	0.4762591804402172
target	0.576159356400437

Aqui já temos uma diferença um pouco mais clara, indicando mais reincidência para o público target da campanha, ainda sim vamos analisar estatisticamente, para isto como fizemos uma transformação binária criando a variável de reincidentes, vamos aplicar o teste Z, que segue a mesma lógica do teste t, mas é aplicável para variáveis discretizadas.

Z-statistic: 89.3702

P-value: 0.0000

Diferença significativa na proporção de usuários reativados entre os grupos.

Conforme esperado, há diferença estatística relevante entre o público target da campanha de cupons, indicando que como alavanca de crescimento, a campanha é um sucesso, pois houve mais pedidos pelos usuários impactados, isso pode ser confirmado obtendo a média de pedidos por usuário dos dois públicos.

is_target	pedidos_medio_por_usuario
control	2.803428289212653
target	3.177162625100114

2.2. Viabilidade Financeira

Conforme mencionado no item anterior, em termos de ticket médio para o público da campanha, não houve diferenças significativas, entretanto, houve um percentual de reincidência maior para os usuários do grupo controle, desta forma podemos adotar as seguintes premissas:

- Usuários reativados a mais: $445.473 * (57,61\% - 47,62\%) \rightarrow 44.503$ usuários;
- Número médio de pedidos por usuário reativado: $44.503 * 4,78 \rightarrow 212.723$ pedidos adicionais;
- Ticket Médio: $R\$47,86 * 212.723 \rightarrow R\$ 10.180.922,78$ Valor total dos pedidos
- Comissão IFood: $19,5\% \rightarrow R\$1.985.100$ Comissão IFood
- Custos operacionais e de campanha: R\$ 1.200.000
- Resultado Líquido estimado: R\$ 785.000

Gerando um retorno líquido de R\$ 785 mil. A diferença observada na taxa de reativação indica que a campanha tem alto potencial, com retorno financeiro positivo no

curto prazo e potencial ainda maior no longo prazo, considerando o valor vitalício (LTV) dos usuários reativados, logo podemos concluir que a campanha é viável financeiramente.

2.3. Oportunidades de Melhoria

Apesar do aumento da reincidência nos pedidos, não houve aumento significativo do ticket médio entre o público impactado, há espaço para melhoria neste sentido, segmentando o público-alvo por histórico de ticket médio (clientes com ticket médio mais baixo). Além disso há outros tipos de incentivos que podem trazer custos de campanha semelhantes, mas com objetivos diferentes, como por exemplo frete grátis ou cashback no aplicativo.

3. Segmentação de Usuários

Considerando os dados e as informações disponibilizadas, podemos definir algumas segmentações que fazem sentido para o teste A/B analisado, como por exemplo a de usuários novos vs. antigos que avalia se a campanha funciona melhor para quem está começando ou já é recorrente, a plataforma de origem, a faixa de preço do restaurante, e a região (localização) do cliente, o tempo de entrega que pode impactar a experiência, a repetição de compra (reincidência) e o agendamento de pedidos que avalia se os usuários que agendam pedidos tendem a ter mais uso dos cupons e aumentarem o ticket médio. Vamos avaliar os resultados para algumas destas sugestões.

3.1. Ticket Médio por Plataforma do Usuário

Considerando o sistema operacional do usuário e calculando o ticket médio:

is_target	origin_platform	ticket_medio
control	NULL	50.0
control	ANDROID	43.445326667052576
control	DESKTOP	46.304791060312795
control	IOS	52.66198782006856
control	WINDOWS_PHONE	39.59971987051812
target	ANDROID	43.02639192576987
target	DESKTOP	46.36331193965379
target	IOS	52.76396343255719
target	WINDOWS_PHONE	40.36582655826568

Na maior parte das plataformas, não há diferença significativa no ticket médio, a exceção do Windows Phone (R\$39,60 vs. R\$40,37), podemos analisar pelo teste t se há relevância estatística na hipótese testada para este caso:

T-statistic: 2.1397

P-value: 0.0324

Diferença significativa no ticket médio entre os grupos para WINDOWS_PHONE.

Houve relevância estatística no ticket médio para os usuários do sistema operacional Windows Phone (que no contexto atual não é mais utilizado). Mas que poderia indicar à época do teste um público interessante para o incremento do público-alvo.

Analisando o ticket médio por estado, temos entre os estados com principais diferenças, a seguir:

📌 Estado: SC		📌 Estado: MT	
is_target	ticket_medio	is_target	ticket_medio
control	49.6472029871009	control	39.83599397590361
target	42.412827801639345	target	48.090502512562786

📌 Estado: AL		📌 Estado: AC	
is_target	ticket_medio	is_target	ticket_medio
control	41.203721480843804	target	59.0
target	39.56761582585688	control	10.0

Chama a atenção a diferença no estado do Acre, podemos rodar o teste t para verificar se há relevância estatística para este público:

T-statistic: nan

P-value: nan

Não há diferença estatisticamente significativa no ticket médio para Acre.

Como esperado, apesar da diferença do ticket médio, não há volumetria suficiente que permita afirmar que há relevância estatística para este segmento, além disso, apenas no estado do Mato Grosso houve incremento do ticket médio, podendo estes serem priorizados em campanhas deste tipo.

Analisando agora o tempo de cadastro dos usuários para o ticket médio do público do teste A/B, temos que não há grandes diferenças no ticket médio entre o público-alvo analisado com o tempo de cadastro.

is_target	anos_desde_cadastro	ticket_medio
control	0	46.914980084119684
target	0	46.73190121983345
control	1	49.056597431477066
target	1	48.979091809876394

A principal diferença foi na questão de localização, com o estado do Mato Grosso como destaque positivo, com incremento de 23% no ticket médio para o público-alvo, é possível ainda com algumas outras informações cadastrais do cliente (idade, método de pagamento), segmentar usando algum algoritmo de agrupamento (k-means) e analisar os resultados do teste A/B para os clusters (grupos) separados.

4. Próximos Passos

Em relação aos próximos passos, vale direcionar a campanha para o público mais rentável, escalonando para o público do Mato Grosso por exemplo cujo qual houve aumento do ticket médio, além disso aprimorar a alavanca de reativação com campanhas de cashback ou descontos com pedidos acima do ticket médio (progressivamente acima de 3%, 5%, 7%), e frete grátis para o público reincidente podem ser alternativas interessante para os próximos passos.

O uso de modelos preditivos de propensão para identificar usuários mais propensos a gastarem mais e serem mais propensos a reativação pode ajudar a direcionar a campanha para um público com alto potencial de retorno, está sem dúvida seria a principal recomendação para os próximos passos visando o aumento da receita líquida, para construir estes modelos são necessários variáveis que descrevem o cliente e o seu comportamento, por isso o uso de pesquisas incentivadas dentro do aplicativo (com o recebimento de cupons para os clientes que responderem) pode ser uma forma de obter cada vez mais informações qualitativas e descritivas dos clientes, refinando os modelos de propensão a serem desenvolvidos.