**Descrição do Projeto**

**Contexto**

Você é analista em uma grande loja online. Junto com o departamento de marketing, você compilou uma lista de hipóteses que podem ajudar a aumentar a receita.

Você precisa priorizar essas hipóteses, lançar um teste A/B e analisar os resultados.

**Descrição dos dados**

Dados usados na primeira parte do projeto

/datasets/hypotheses\_us.csv [Baixe o conjunto de dados](https://practicum-content.s3.us-west-1.amazonaws.com/datasets/hypotheses_us.csv)

* Hypotheses — breves descrições das hipóteses
* Reach — alcance do usuário, em uma escala de um a dez
* Impact — impacto nos usuários, em uma escala de um a dez
* Confidence — confiança na hipótese, em uma escala de um a dez
* Effort — os recursos necessários para testar uma hipótese, em uma escala de um a dez. Quanto maior o valor de Effort, mais recursos são necessários para o teste.

Dados usados na segunda parte do projeto

/datasets/orders\_us.csv [Baixe o conjunto de dados](https://practicum-content.s3.us-west-1.amazonaws.com/datasets/orders_us.csv)

* transactionId — identificador do pedido
* visitorId — identificador do usuário que fez o pedido
* date — do pedido
* revenue — do pedido
* group — o grupo de teste A/B ao qual o usuário pertence

/datasets/visits\_us.csv [Baixe o conjunto de dados](https://practicum-content.s3.us-west-1.amazonaws.com/datasets/visits_us.csv)

* date — data
* group — grupo de teste A/B
* visits — o número de visitas na data especificada para o grupo de teste A/B especificado

Certifique-se de pré-processar os dados. Pode haver erros nos conjuntos de dados originais; por exemplo, alguns dos visitantes podem ter entrado no grupo A e no grupo B.

**Parte 1. Priorizando Hipóteses**

O arquivo hypotheses\_us.csv contém nove hipóteses para aumentar a receita de uma loja online com Reach, Impact, Confidence e Effort especificados para cada um.

A tarefa é:

* Aplicar o framework ICE para priorizar hipóteses. Classifique-os em ordem decrescente de prioridade.
* Aplicar o framework RICE para priorizar hipóteses. Classifique-os em ordem decrescente de prioridade.
* Mostre como a priorização de hipóteses muda quando você usa RICE em vez de ICE. Dê uma explicação para as alterações.

**Parte 2. Análise de teste A/B**

Você realizou um teste A/B e obteve os resultados descritos nos arquivos orders\_us.csv e visitors\_us.csv.

Tarefa

Analise o teste A/B:

1. Faça um gráfico da receita acumulada por grupo. Tire conclusões e crie conjecturas.
2. Faça um gráfico do tamanho médio acumulado do pedido por grupo. Tire conclusões e crie conjecturas.
3. Faça um gráfico da diferença relativa no tamanho médio acumulado do pedido para o grupo B em comparação com o grupo A. Faça conclusões e crie conjecturas.
4. Calcule a taxa de conversão de cada grupo como a proporção de pedidos para o número de visitas para cada dia. Trace as taxas de conversão diárias dos dois grupos e descreva a diferença. Tire conclusões e crie conjecturas.
5. Faça um gráfico da diferença relativa na conversão cumulativa para o grupo B em comparação com o grupo A. Tire conclusões e crie conjecturas.
6. Calcule os percentis 95 e 99 para o número de pedidos por usuário. Defina o ponto em que um ponto de dados se torna uma anomalia.
7. Faça um gráfico de dispersão dos preços dos pedidos. Tire conclusões e crie conjecturas.
8. Calcule os percentis 95 e 99 dos preços dos pedidos. Defina o ponto em que um ponto de dados se torna uma anomalia.
9. Encontre a significância estatística da diferença na conversão entre os grupos usando os dados brutos. Tire conclusões e crie conjecturas.
10. Encontre a significância estatística da diferença no tamanho médio do pedido entre os grupos usando os dados brutos. Tire conclusões e crie conjecturas.
11. Encontre a significância estatística da diferença na conversão entre os grupos usando os dados filtrados. Tire conclusões e crie conjecturas.
12. Encontre a significância estatística da diferença no tamanho médio do pedido entre os grupos usando os dados filtrados. Tire conclusões e crie conjecturas.
13. Tome uma decisão com base nos resultados do teste. As decisões possíveis são: 1. Pare o teste, considere um dos grupos o líder. 2. Pare o teste, conclua que não há diferença entre os grupos. 3. Continue o teste.

CONVERSION

|  |  |
| --- | --- |
| RAW | FILTERED |
| 0.00018  0.197 | 0.00046  0.186 |
| O valor-p é consideravelmente menor do que 0.05, então nós rejeitamos a hipótese nula. A análise dos dados brutos mostrou que as taxas de conversão dos grupos possuem diferenças estatísticas significativas. | O valor-p é consideravelmente menor do que 0.05, então nós rejeitamos a hipótese nula. A análise dos dados brutos mostrou que OS VOLUMES MEDIOS DOS PEDIDOS dos grupos possuem diferenças estatísticas significativas. |

ORDER SIZE

|  |  |
| --- | --- |
| RAW | FILTERED |
| 0.255  0.025 | 0.029  0.007 |
| O valor-p é consideravelmente maior do que 0.05, então nós não podemos rejeitar a hipótese nula. A análise dos dados brutos mostrou que as taxas de conversão dos grupos não possuem diferenças estatísticas significativas. | O valor-p é menor do que 0.05, então nós rejeitamos a hipótese nula. A análise dos dados brutos mostrou que OS VOLUMES MEDIOS DOS PEDIDOS dos grupos possuem diferenças estatísticas significativas. |

CONVERSION

|  |  |
| --- | --- |
| RAW | FILTERED |
| 0.01666  0.134 | 0.01090  0.162 |
| O valor-p é consideravelmente menor do que 0.05, então nós rejeitamos a hipótese nula. A análise dos dados brutos mostrou que as taxas de conversão dos grupos possuem diferenças estatísticas significativas. | O valor-p é consideravelmente menor do que 0.05, então nós rejeitamos a hipótese nula. A análise dos dados brutos mostrou que OS VOLUMES MEDIOS DOS PEDIDOS dos grupos possuem diferenças estatísticas significativas. |

ORDER SIZE

|  |  |
| --- | --- |
| RAW | FILTERED |
| 0.673  0.171 | 0.651  -0.057 |
| O valor-p é consideravelmente maior do que 0.05, então nós não podemos rejeitar a hipótese nula. A análise dos dados brutos mostrou que as taxas de conversão dos grupos não possuem diferenças estatísticas significativas. | O valor-p é consideravelmente maior do que 0.05, então nós não podemos rejeitar a hipótese nula. A análise dos dados brutos mostrou que OS VOLUMES MEDIOS DOS PEDIDOS dos grupos não possuem diferenças estatísticas significativas. |