

Guia Excel: Aleatorização

Este documento mostra como aleatorizar indivíduos de um conjunto de dados em um grupo de controle e um grupo de tratamento usando o Microsoft Excel.¹ Neste exercício, vamos realizar uma aleatorização para uma avaliação hipotética que busca verificar como uma versão modificada de um e-mail afeta o número de destinatários que clicam em um link presente nesse e-mail. Nesta avaliação hipotética, os indivíduos do grupo de tratamento receberiam o e-mail modificado e os indivíduos do grupo de controle receberiam o e-mail normal.

Aleatorização

Para dividir aleatoriamente os destinatários do e-mail em dois grupos, usaremos a função aleatório() do Excel. Com ela, vamos gerar um número aleatório para cada linha, ordenar a planilha com base nos números gerados aleatoriamente e atribuir a primeira metade das linhas ao grupo de controle e a outra metade ao grupo de tratamento.

O que você vai aprender?

1. Atribuir valores aleatórios
 - a. Criar a coluna "rand1"
 - b. Inserir o comando de aleatorização =aleatório() e preencher a coluna "rand1"
 - c. Repetir 1a e 1b para criar a coluna "rand2"
2. Converter os resultados do comando de aleatorização em valores
 - a. Copiar as colunas "rand1" e "rand2"
 - b. Colar como valores
 - c. Excluir as colunas originais "rand1" e "rand2"
3. Organizar o conjunto de dados
4. Atribuir valores de tratamento
 - a. Criar a coluna "tratamento"
 - b. Selecionar a primeira metade das linhas e atribuir "0" na coluna "tratamento"
 - c. Selecionar a segunda metade das linhas e atribuir "1" na coluna "tratamento"

Passo 1) Atribuir valores aleatórios

Passo 1a. Crie a coluna "rand1". Na planilha que inclui os endereços de e-mail dos participantes da avaliação, adicione uma nova coluna intitulada "rand1".

Passo 1b. Insira o comando de aleatorização =aleatório() e preencha a coluna "rand1". Na próxima linha, digite o comando: =aleatório()

¹ Este guia foi criado usando o Microsoft Excel 2021 Versão 16.57 e os comandos de teclado são definidos para o Windows. Os locais das funções do Excel podem variar dependendo da versão do Excel que você está usando, e os comandos de teclado serão diferentes para outros sistemas operacionais.

Figura 1. Inserir o comando de aleatorização

	A	B
1	Email	rand1
2	klm012@gmail.com	=RAND()
3	hij789@hotmail.com	
4	abd123@gmail.com	
5	efg456@yahoo.com	

Ao pressionar Enter, um número decimal entre 0 e 1 será gerado aleatoriamente. Agora, use este comando em toda a coluna para atribuir um número aleatório a cada uma das linhas. Você pode fazer isso manualmente ou com atalhos de teclado.


Preenchimento manual: Selecione a célula onde você inseriu o comando (neste caso, B2) e arraste a alça de preenchimento  (o ponto que aparece no canto inferior direito) até a última linha. (Observação: A função de aleatorização recalcula cada vez que uma planilha é aberta ou modificada.)

Figura 2. Preenchimento manual da primeira coluna de aleatorização

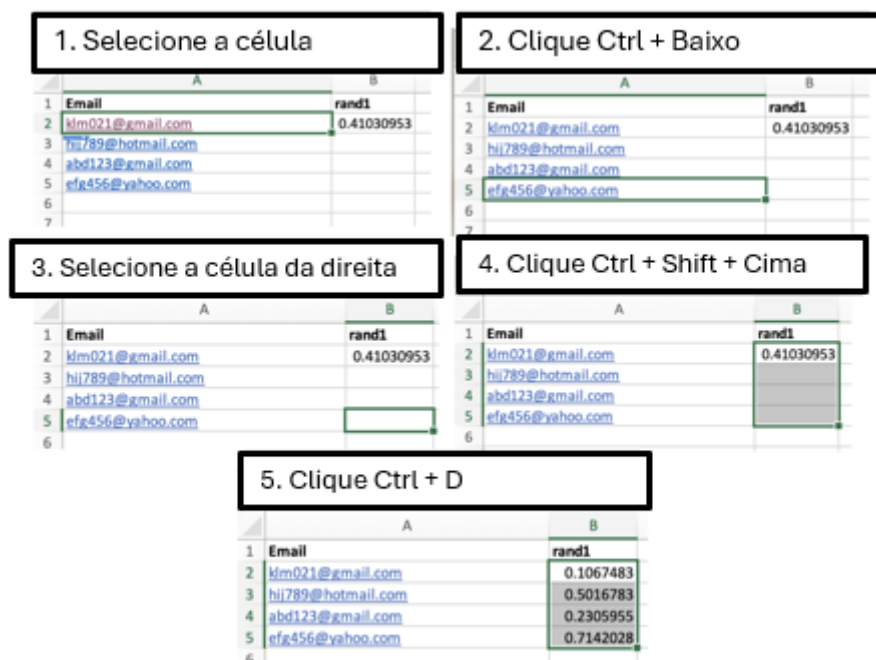
	A	B
1	Email	rand1
2	klm012@gmail.com	0.6064885
3	hij789@hotmail.com	
4	abd123@gmail.com	
5	efg456@yahoo.com	

	A	B
1	Email	rand1
2	klm012@gmail.com	0.0496934
3	hij789@hotmail.com	0.7574394
4	abd123@gmail.com	0.1248301
5	efg456@yahoo.com	0.2576304
6		

Preencher usando atalhos do teclado. Se o conjunto de dados for tão grande que seja difícil rolar manualmente até o final, você também pode preencher a coluna inteira usando comandos do teclado. (Veja a Figura 3 abaixo).

1. Selecione uma célula em uma coluna preenchida (neste caso, "E-mail").
2. Pressione Ctrl + seta para baixo. Isso o leva à última linha do seu conjunto de dados.
3. Selecione a célula à direita da última célula na coluna "E-mail". Esta será a última célula da coluna "rand1".
4. Pressione Shift + Ctrl + Seta para cima. Isso seleciona da última célula até a célula preenchida mais próxima. Neste caso, B2 que contém nosso comando de randomização.
5. Preencha o comando pressionando Ctrl + D.

Figura 3. Preenchimento da primeira coluna de aleatorização usando atalhos do teclado



Passo 1c. Repita os passos 1a e 1b para criar e preencher a coluna "rand2". Crie outra coluna intitulada "rand2" e repita o processo realizado para a coluna "rand1".

Os dados finais devem conter duas colunas de números gerados aleatoriamente.

Figura 4. Criar uma segunda coluna de aleatorização

	A	B	C
1	Email	rand1	rand2
2	klm012@gmail.com	0.5810108	0.227079
3	hij789@hotmail.com	0.72175	0.3546413
4	abd123@gmail.com	0.7540408	0.0100365
5	efg456@yahoo.com	0.2549857	0.8285137

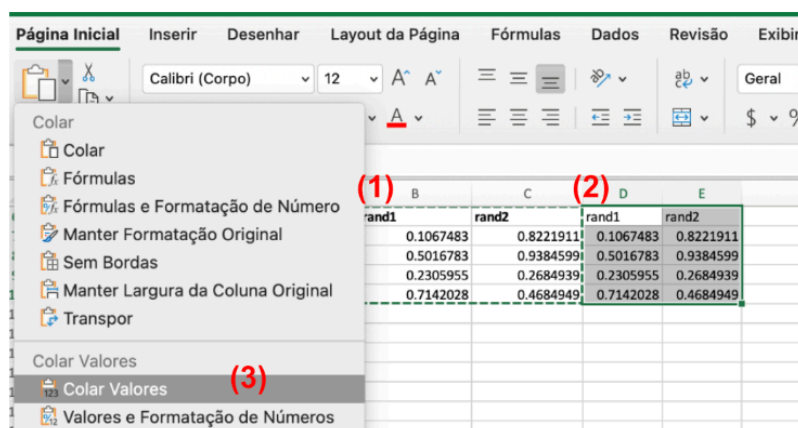
Passo 2) Converter as saídas do comando de aleatorização em valores

Passo 2a. Copiar as colunas "rand1" e "rand2". Antes de ordenar a planilha, copiaremos primeiro as colunas "rand1" e "rand2" (1).

Passo 2b. Colar como valores. Selecione a célula D1 (2) e cole sua seleção nas colunas adjacentes, usando "Colar valores" (3). Você pode encontrar a opção "Colar valores" clicando na seta para baixo ao lado da área de transferência na guia Página Inicial. (Veja a figura 5 abaixo).

Depois de colarmos as duas colunas, vemos que os valores das colunas originais mudaram. Isso ocorre porque o Excel executa continuamente a função aleatório() para gerar novos números aleatórios, mas ordenar a planilha seria difícil com números que mudam constantemente. Ao ter as novas colunas com valores estáticos, podemos ordenar a planilha e verificar se a ordenação funcionou corretamente.

Figura 5. Colar os resultados da aleatorização como valores

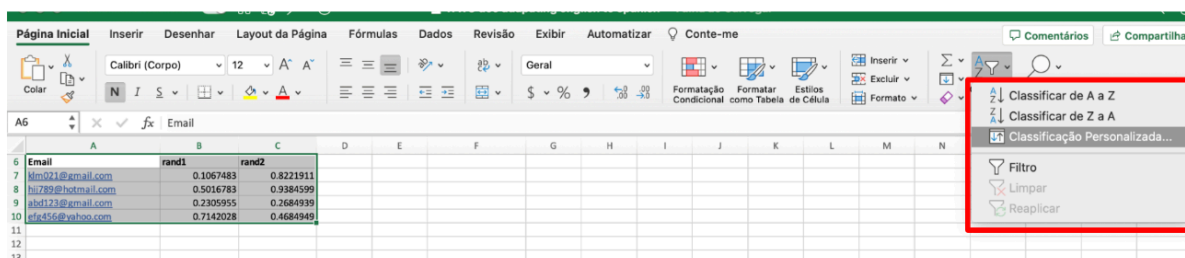


Passo 2c. Excluir as colunas originais "rand1" e "rand2". Agora podemos excluir as colunas originais "rand1" e "rand2", que contêm a função aleatório() (colunas B e C acima). Não exclua as novas colunas com os valores.

Passo 3) Classificar o conjunto de dados

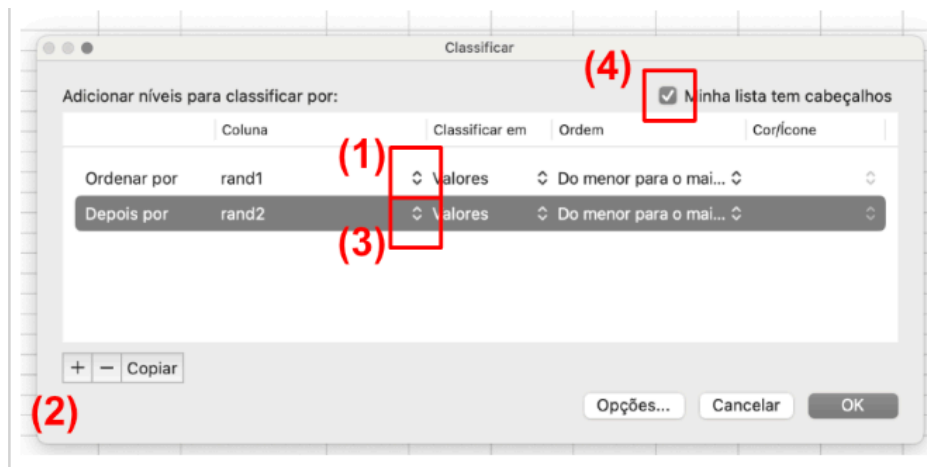
Em seguida, classificaremos de forma personalizada todo o conjunto de dados. A opção "Classificação personalizada" está no menu "Classificar e Filtrar" na guia Início.

Figura 6. Acessar a função "Classificação personalizada"



Selecione para classificar primeiro por "rand1" (1), clique em "+" para adicionar outro nível de classificação (2) e, em seguida, classifique por "rand2" (3). Certifique-se de que a opção "Minha lista tem cabeçalhos" esteja marcada (4). (Consulte a Figura 7 abaixo).

Figura 7. Classificar pelos valores aleatórios



Após clicar em "OK", vemos que a ordem dos endereços de e-mail mudou principalmente em função da ordem de "rand1". Embora os valores de "rand1" quase sempre determinem como as linhas são ordenadas, incluímos "rand2" caso o Excel gere dois ou mais números aleatórios para "rand1" que sejam exatamente iguais. Em um caso tão raro, o ordenamento seria impossível, e "rand2" determinaria como essas linhas seriam ordenadas.

Figura 8. Resultado da classificação personalizada

	A	B	C
1	Email	rand1	rand2
2	abd123@gmail.com	0.128533	0.900317651
3	klm012@gmail.com	0.2338778	0.690948625
4	efg456@yahoo.com	0.5686872	0.020733377
5	hij789@hotmail.com	0.8373926	0.255109258

Passo 4) Atribuir valores de tratamento

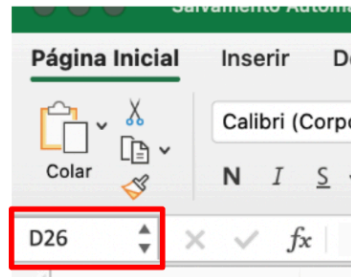
Passo 4a. Criar a coluna "tratamento". Uma vez concluída a classificação, estamos prontos para atribuir o status do tratamento. Adicione uma nova coluna intitulada "tratamento".

Passo 4b. Selecione a primeira metade das linhas e atribua "0" na coluna "tratamento". Você pode selecionar a primeira metade das linhas manualmente ou usar uma estratégia similar usando comandos do teclado, como descrito no Passo 1b.

1. Insira "0" na primeira célula abaixo do título da coluna "tratamento". Neste caso, D2.
2. Selecione uma célula em qualquer uma das suas colunas preenchidas (neste caso, "Email", "rand1" ou rand2.) e selecione a coluna inteira clicando na letra da coluna.
3. Observe a contagem na parte inferior direita da pasta de trabalho do Excel. Desta contagem, subtraia 1 para levar em conta a linha que contém o título da coluna. Divida este valor por 2.
 - Ex: A contagem é 51. Depois de subtrair 1 e dividir por 2, o novo valor é 25.
4. Adicione 1 a este valor. Na caixa de seleção de células, digite a letra da coluna e este valor. Isso o levará ao ponto médio desta coluna. (Veja a figura 9.)
 - Ex: Depois de adicionar 1, o novo valor é 26. Digite D26 na caixa de seleção de células. Pressione Enter e isso o levará a D26.

5. Pressione Shift + Ctrl + Seta para cima para selecionar até a célula onde você inseriu "0".
6. Pressione Ctrl + D para preencher esta seleção de células com "0".
7. Pressione Shift + Ctrl + Seta para baixo para chegar à última linha preenchida com "0".
8. Insira "1" na célula abaixo desta linha.

Figura 9. Caixa de seleção de células



Passo 4c. Selecione a segunda metade das linhas e atribua "1" na coluna "tratamento". Repita os passos anteriores para preencher a segunda metade das linhas com "1".

Agora, a aleatorização está completa e estamos prontos para lançar o teste por e-mail. Os participantes com um valor de tratamento "0" receberão o e-mail normal, e os que tiverem um valor "1" receberão a versão modificada.

Figura 10. Valores de tratamento atribuídos

	A	B	C	D
1	Email	rand1	rand2	treatment
2	abd123@gmail.com	0.128533	0.900317651	0
3	klm012@gmail.com	0.2338778	0.690948625	0
4	efg456@yahoo.com	0.5686872	0.020733377	1
5	hij789@hotmail.com	0.8373926	0.255109258	1