

Guía de Excel: Aleatorización

Este documento le mostrará cómo aleatorizar a los individuos de un conjunto de datos en un grupo de control y un grupo de tratamiento utilizando Microsoft Excel.¹ En este ejercicio, estamos realizando una aleatorización para una evaluación hipotética que busca ver cómo una versión modificada de un correo electrónico afecta al número de destinatarios que hacen clic en un enlace en dicho correo electrónico. En esta evaluación hipotética, los individuos de tratamiento recibirían el correo electrónico modificado y los individuos de control recibirían el correo electrónico de costumbre.

Aleatorización

Para asignar aleatoriamente a los destinatarios del correo electrónico en dos grupos, utilizaremos la función aleatorio() de Excel. Con ella, generaremos un número aleatorio para cada fila, ordenaremos la hoja de cálculo en función de los números generados aleatoriamente y asignaremos la primera mitad de las filas al grupo de control y la otra mitad al grupo de tratamiento.

¿Qué aprenderá a hacer?

1. Asignar valores aleatorios
 - a. Crear la columna "rand1"
 - b. Introducir el commando de aleatorización =aleatorio() y rellenar la columna "rand1"
 - c. Repetir 1a y 1b para crear la columna "rand2"
2. Convertir los resultados del comando de aleatorización en valores
 - a. Copiar las columnas "rand1" y "rand2"
 - b. Pegar como valores
 - c. Eliminar las columnas originales "rand1" y "rand2"
3. Ordenar el conjunto de datos
4. Asignar valores de tratamiento
 - a. Crear la columna "tratamiento"
 - b. Seleccione la primera mitad de las filas y asigne con "0" en la columna "tratamiento"
 - c. Seleccionar la segunda mitad de las filas y asignar con "1" en la columna "tratamiento"

Paso 1) Asignar valores aleatorios

¹ Esta guía fue creada utilizando Microsoft Excel 2021 Versión 16.57 y los comandos de teclado están definidos para los teclados de Windows. Las ubicaciones de las funciones de Excel pueden diferir dependiendo de la versión de Excel que esté utilizando, y los comandos de teclado serán diferentes para otros teclados.

Paso 1a. Cree la columna "rand1". En la hoja de cálculo que incluye las direcciones de correo electrónico de los participantes de la evaluación, añada una nueva columna titulada "rand1".

Paso 1b. Introduzca el comando de aleatorización =aleatorio() y rellene la columna "rand1". En la siguiente fila, escribe el comando: =aleatorio()

Figura 1. Introducir el comando de aleatorización

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

Tras pulsar Enter verá que un número decimal entre 0 y 1 se genera aleatoriamente. Ahora, use este comando en toda la columna para asignar un número aleatorio para cada una de las filas. Puede hacerlo manualmente o mediante comandos de teclado.


Relleno manual: Seleccione la celda donde ha introducido el comando (en este caso, B2) y arrastre el mango de relleno  (el punto que aparece en la esquina inferior derecha) hasta la última fila. (Nota: La función de aleatorización vuelve a calcular cada vez que se abre o se modifica una hoja de cálculo.)

Figura 2. Rellenar manualmente la primera columna de aleatorización

	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		

Rellenar utilizando comandos de teclado. Si el conjunto de datos es tan grande que resulta difícil desplazarse manualmente hasta el final, también puede rellenar toda la columna mediante comandos de teclado. (Vea la Figura 3 más abajo).

- i. Seleccione una celda de una columna rellena (en este caso, "Correo electrónico").
- ii. Pulse Ctrl + flecha abajo. Esto le lleva a la última fila de su conjunto de datos.
- iii. Seleccione la celda a la derecha de la última celda de la columna "Correo electrónico". Esta será la última celda de la columna "rand1".
- iv. Pulse Shift + Ctrl + Flecha arriba. Esto selecciona desde la última celda hasta la celda rellena más cercana. En este caso, B2 que contiene nuestro comando de aleatorización.

v. Rellena el comando pulsando Ctrl + D.

Figura 3. Rellenar la primera columna de aleatorización mediante comandos de teclado

El diagrama muestra cinco etapas de un proceso en Excel:

- I. Seleccione la celda:** Se muestra una hoja de cálculo con la columna 'Email' (A) y 'rand1' (B). La celda A2 ('klm021@gmail.com') está seleccionada.
- II. Pulse Ctrl + Abajo:** El cursor se mueve a la celda A5 ('efg456@yahoo.com').
- III. Seleccione la celda de la derecha:** El cursor se mueve a la celda B5.
- IV. Pulse Mayús + Ctrl + Arriba:** El cursor se mueve a la celda B2.
- V. Pulse Ctrl + D:** El contenido de la celda B2 se copia hacia abajo, reemplazando los valores de la columna 'rand1' en las filas 2 a 5.

Una barra de fórmulas muestra 'A2' con el contenido 'klm021@gmail.com'.

Paso 1c. Repita los pasos 1a y 1b para crear y rellenar la columna "rand2". Cree otra columna titulada "rand2" y repita el proceso realizado para la columna "rand1".

Los datos finales deben contener dos columnas de números generados aleatoriamente.

Figura 4. Creación de una segunda columna aleatoria

	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

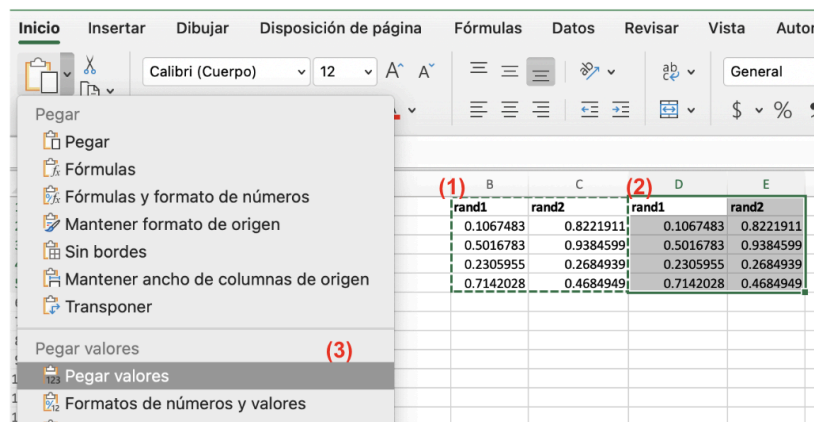
Paso 2) Convertir las salidas del comando de aleatorización en valores

Paso 2a. Copiar las columnas "rand1" y "rand2". Antes de ordenar la hoja de cálculo, primero copiaremos las columnas "rand1" y "rand2" (1).

Paso 2b. Pegar como valores. Seleccione la celda D1 (2), y pegue su selección en las columnas adyacentes, utilizando "Pegar valores" (3). Puede encontrar la opción "Pegar valores" haciendo clic en la flecha hacia abajo junto al portapapeles en la pestaña Inicio. (Véase la figura 5 más abajo).

Una vez que pegamos las dos columnas, vemos que los valores de las columnas originales han cambiado. Esto se debe a que Excel realiza continuamente la función aleatorio() para generar nuevos números aleatorios, pero ordenar la hoja de cálculo sería difícil con números que cambian constantemente. Al tener las nuevas columnas con valores estáticos, podemos ordenar la hoja de cálculo y comprobar si la ordenación ha funcionado correctamente.

Figura 5. Pegar los resultados de la aleatorización como valores

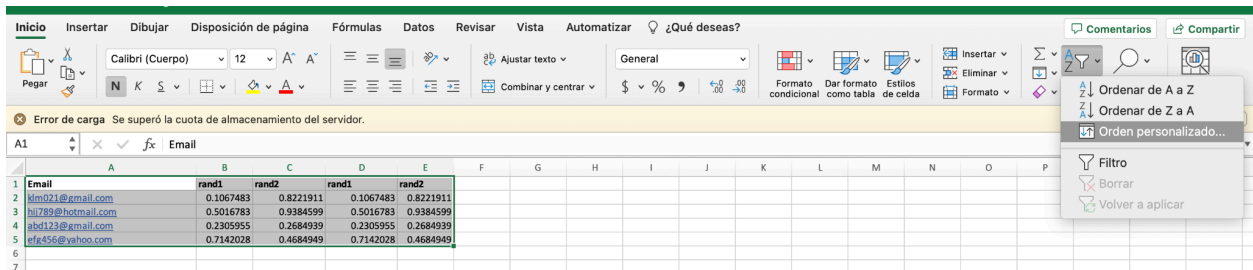


Paso 2c. Eliminar las columnas originales "rand1" y "rand2". Ahora podemos eliminar las columnas originales "rand1" y "rand2" que contienen la función aleatorio() (columnas B y C anteriores). No borre las nuevas columnas con los valores.

Paso 3) Ordenar el conjunto de datos

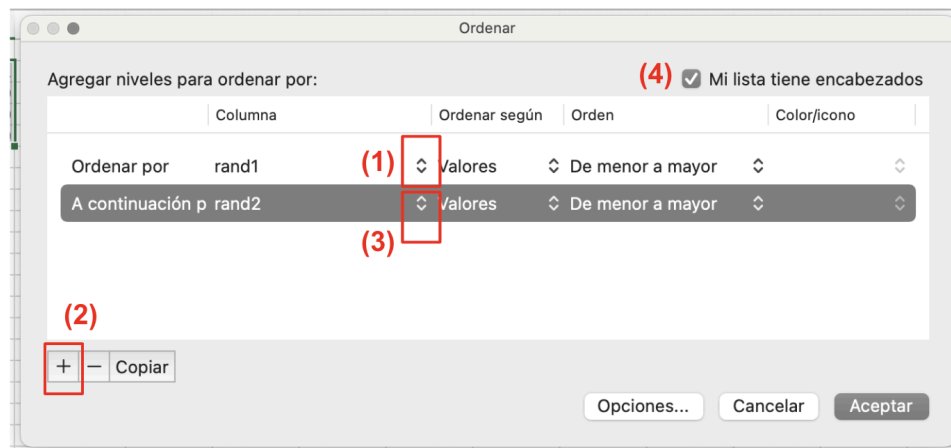
A continuación, ordenaremos de forma personalizada todo el conjunto de datos. La opción "Ordenación personalizada" se encuentra en el menú "Ordenar y filtrar" de la pestaña Inicio.

Figura 6. Acceso a la función "Orden personalizado"



Seleccione ordenar primero por "rand1" (1), haga clic en "+" para añadir otro nivel de orden (2), y luego ordenar por "rand2" (3). Asegúrese de que la opción "Mi lista tiene cabeceras" está marcada (4). (Véase la figura 7 más abajo).

Figura 7. Ordenar por los valores aleatorios



Una vez que pulsamos "OK", vemos que el orden de las direcciones de correo electrónico cambió principalmente en función del orden de "rand1". Aunque los valores de "rand1" casi siempre determinan cómo se ordenan las filas, incluimos "rand2" por si Excel genera dos o más números aleatorios para "rand1" que sean exactamente iguales. En un caso tan raro, el ordenamiento sería imposible, y "rand2" determinaría cómo se ordenan estas filas.

Figura 8. Resultado del orden personalizado

	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

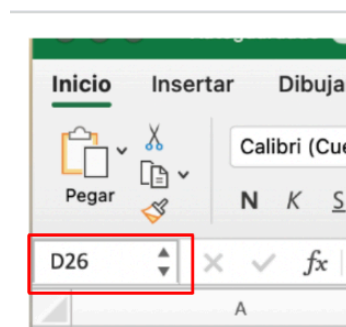
Paso 4) Asignar valores de tratamiento

Paso 4a. Crear la columna "tratamiento". Una vez completada la clasificación, estamos listos para asignar el estado del tratamiento. Añada una nueva columna titulada "tratamiento".

Paso 4b. Seleccione la primera mitad de las filas y asigne con "0" en la columna "tratamiento". Puede seleccionar la primera mitad de las filas manualmente o puede utilizar una estrategia similar utilizando comandos del teclado como se describe en el Paso 1b.

- i. Introduzca "0" en la primera celda debajo del título de la columna "tratamiento". En este caso, D2.
- ii. Seleccione una celda en cualquiera de sus columnas rellenas (en este caso, "Email", "rand1" o rand2.) y seleccione toda la columna haciendo clic en la letra de la columna.
- iii. Observe el recuento en la parte inferior derecha del libro de Excel. De este recuento, reste 1 para tener en cuenta la fila que contiene el título de la columna. Divida este valor por 2.
 - *Ej: El recuento es 51. Después de restar 1 y dividir entre 2, el nuevo valor es 25.*
- iv. Añada 1 a este valor. En el cuadro de selección de celdas, escriba la letra de la columna y este valor. Esto le llevará al punto medio de esta columna. (Véase la figura 9.)
 - *Ej: Después de añadir 1, el nuevo valor es 26. Escriba D26 en el cuadro de selección de celdas. Pulse enter y esto le llevará a D26.*
- v. Pulsa Shift + Ctrl + Flecha arriba para seleccionar hasta la celda donde has introducido "0".
- vi. Pulsa Ctrl + D para rellenar esta selección de celdas con "0"
- vii. Pulsa Shift Ctrl + Flecha abajo para llegar a la última fila rellena con "0".
- viii. Introduce "1" en la celda de debajo de esta fila.

Figura 9. Cuadro de selección de celdas



Paso 4c. Seleccione la segunda mitad de las filas y asigne con "1" en la columna "tratamiento". Repita los pasos anteriores para llenar la segunda mitad de las filas con "1".

Ahora, la aleatorización está completa, y estamos listos para lanzar el ensayo por correo electrónico. Los participantes con un tratamiento de valor "0" recibirán el correo electrónico normal, y los que tengan un valor "1" recibirán la versión modificada.

Figura 10. Valores de tratamiento asignados

	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