

## Nodo Partitioning

Considere como ejemplo el siguiente conjunto de datos.

| JuegaGolf | Panorama | Temp     | Humedad | Viento    |
|-----------|----------|----------|---------|-----------|
| No        | Lluvioso | Caliente | Alta    | FALSO     |
| No        | Lluvioso | Caliente | Alta    | VERDADERO |
| Si        | Nublado  | Caliente | Alta    | FALSO     |
| Si        | Soleado  | Templado | Alta    | FALSO     |
| Si        | Soleado  | Frio     | Normal  | FALSO     |
| No        | Soleado  | Frio     | Normal  | VERDADERO |
| Si        | Nublado  | Frio     | Normal  | VERDADERO |
| No        | Lluvioso | Templado | Alta    | FALSO     |
| Si        | Lluvioso | Frio     | Normal  | FALSO     |
| Si        | Soleado  | Templado | Normal  | FALSO     |
| Si        | Lluvioso | Templado | Normal  | VERDADERO |
| Si        | Nublado  | Templado | Alta    | VERDADERO |
| Si        | Nublado  | Caliente | Normal  | FALSO     |
| No        | Soleado  | Templado | Alta    | VERDADERO |

| Row ID | S JuegaGolf | S Panorama | S Temp   | S Humedad | B Viento |
|--------|-------------|------------|----------|-----------|----------|
| Row0   | No          | Lluvioso   | Caliente | Alta      | false    |
| Row1   | No          | Lluvioso   | Caliente | Alta      | true     |
| Row2   | Si          | Nublado    | Caliente | Alta      | false    |
| Row3   | Si          | Soleado    | Templado | Alta      | false    |
| Row4   | Si          | Soleado    | Frio     | Normal    | false    |
| Row5   | No          | Soleado    | Frio     | Normal    | true     |
| Row6   | Si          | Nublado    | Frio     | Normal    | true     |
| Row7   | No          | Lluvioso   | Templado | Alta      | false    |
| Row8   | Si          | Lluvioso   | Frio     | Normal    | false    |
| Row9   | Si          | Soleado    | Templado | Normal    | false    |
| Row10  | Si          | Lluvioso   | Templado | Normal    | true     |
| Row11  | Si          | Nublado    | Templado | Alta      | true     |
| Row12  | Si          | Nublado    | Caliente | Normal    | false    |
| Row13  | No          | Soleado    | Templado | Alta      | true     |

## Configuración

Dialog - 3:2 - Partitioning

File

First partition Flow Variables Memory Policy

Choose size of first partition

☐ Absolute
 ☒ Relative[%]
 ☐ Take from top
 ☐ Linear sampling
 ☐ Draw randomly
 ☐ Stratified sampling
 ☐ Use random seed

100

70

B Viento

1,604,372,002,3

OK Apply Cancel ?

## Particiones

| Row ID | S JuegaGolf | S Panorama | S Temp   | S Humedad | B Viento |
|--------|-------------|------------|----------|-----------|----------|
| Row0   | No          | Lluvioso   | Caliente | Alta      | false    |
| Row1   | No          | Lluvioso   | Caliente | Alta      | true     |
| Row2   | Si          | Nublado    | Caliente | Alta      | false    |
| Row3   | Si          | Soleado    | Templado | Alta      | false    |
| Row4   | Si          | Soleado    | Frio     | Normal    | false    |
| Row5   | No          | Soleado    | Frio     | Normal    | true     |
| Row6   | Si          | Nublado    | Frio     | Normal    | true     |
| Row7   | No          | Lluvioso   | Templado | Alta      | false    |
| Row8   | Si          | Lluvioso   | Frio     | Normal    | false    |

Este modo coloca las filas superiores en la primera tabla de salida y el resto en la segunda tabla.

| Row ID | S JuegaGolf | S Panorama | S Temp   | S Humedad | B Viento |
|--------|-------------|------------|----------|-----------|----------|
| Row9   | Si          | Soleado    | Templado | Normal    | false    |
| Row10  | Si          | Lluvioso   | Templado | Normal    | true     |
| Row11  | Si          | Nublado    | Templado | Alta      | true     |
| Row12  | Si          | Nublado    | Caliente | Normal    | false    |
| Row13  | No          | Soleado    | Templado | Alta      | true     |

## Configuración

Dialog - 3:2 - Partitioning

File

First partition | Flow Variables | Memory Policy

Choose size of first partition

☐ Absolute

☒ Relative[%]

☐ Take from top

☒ Linear sampling

☐ Draw randomly

☐ Stratified sampling

☐ Use random seed

## Particiones

| Row ID | S JuegaGolf | S Panorama | S Temp   | S Humedad | B Viento |
|--------|-------------|------------|----------|-----------|----------|
| Row0   | No          | Lluvioso   | Caliente | Alta      | false    |
| Row1   | No          | Lluvioso   | Caliente | Alta      | true     |
| Row3   | Si          | Soleado    | Templado | Alta      | false    |
| Row4   | Si          | Soleado    | Frio     | Normal    | false    |
| Row6   | Si          | Nublado    | Frio     | Normal    | true     |
| Row7   | No          | Lluvioso   | Templado | Alta      | false    |
| Row9   | Si          | Soleado    | Templado | Normal    | false    |
| Row10  | Si          | Lluvioso   | Templado | Normal    | true     |
| Row13  | No          | Soleado    | Templado | Alta      | true     |

Este modo siempre incluye la primera y la última fila y selecciona las filas restantes linealmente en toda la tabla (por ejemplo, cada tercera fila). Esto es útil para reducir la muestra de una columna ordenada mientras se mantienen los valores mínimo y máximo.

| Row ID | S JuegaGolf | S Panorama | S Temp   | S Humedad | B Viento |
|--------|-------------|------------|----------|-----------|----------|
| Row2   | Si          | Nublado    | Caliente | Alta      | false    |
| Row5   | No          | Soleado    | Frio     | Normal    | true     |
| Row8   | Si          | Lluvioso   | Frio     | Normal    | false    |
| Row11  | Si          | Nublado    | Templado | Alta      | true     |
| Row12  | Si          | Nublado    | Caliente | Normal    | false    |

## Configuración

Dialog - 3:2 - Partitioning

File

First partition | Flow Variables | Memory Policy

Choose size of first partition

☐ Absolute

☒ Relative[%]

☐ Take from top

☐ Linear sampling

☒ Draw randomly

☐ Stratified sampling

☐ Use random seed

OK Apply Cancel ?

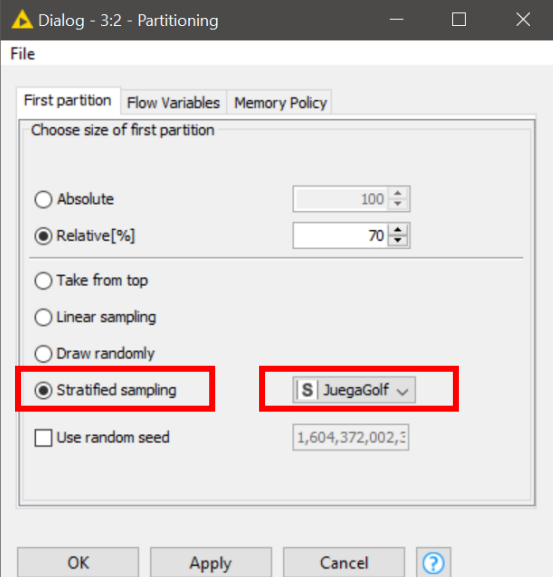
## Particiones

| Row ID | <b>S</b> JuegaGolf | <b>S</b> Panorama | <b>S</b> Temp | <b>S</b> Humedad | <b>B</b> Viento |
|--------|--------------------|-------------------|---------------|------------------|-----------------|
| Row0   | No                 | Lluvioso          | Caliente      | Alta             | false           |
| Row1   | No                 | Lluvioso          | Caliente      | Alta             | true            |
| Row3   | Si                 | Soleado           | Templado      | Alta             | false           |
| Row4   | Si                 | Soleado           | Frio          | Normal           | false           |
| Row9   | Si                 | Soleado           | Templado      | Normal           | false           |
| Row10  | Si                 | Lluvioso          | Templado      | Normal           | true            |
| Row11  | Si                 | Nublado           | Templado      | Alta             | true            |
| Row12  | Si                 | Nublado           | Caliente      | Normal           | false           |
| Row13  | No                 | Soleado           | Templado      | Alta             | true            |

Muestreo aleatorio de todas las filas, puede especificar opcionalmente una semilla fija.

| Row ID | <b>S</b> JuegaGolf | <b>S</b> Panorama | <b>S</b> Temp | <b>S</b> Humedad | <b>B</b> Viento |
|--------|--------------------|-------------------|---------------|------------------|-----------------|
| Row2   | Si                 | Nublado           | Caliente      | Alta             | false           |
| Row5   | No                 | Soleado           | Frio          | Normal           | true            |
| Row6   | Si                 | Nublado           | Frio          | Normal           | true            |
| Row7   | No                 | Lluvioso          | Templado      | Alta             | false           |
| Row8   | Si                 | Lluvioso          | Frio          | Normal           | false           |

## Configuración



Dialog - 3:2 - Partitioning

File

First partition Flow Variables Memory Policy

Choose size of first partition

☐ Absolute 100

☒ Relative[%] 70

☐ Take from top

☐ Linear sampling

☐ Draw randomly

☒ Stratified sampling

☐ Use random seed 1,604,372,002,3

OK Apply Cancel ?

## Particiones

| Row ID | <b>S</b> JuegoGolf | <b>S</b> Panorama | <b>S</b> Temp | <b>S</b> Humedad | <b>B</b> Viento |
|--------|--------------------|-------------------|---------------|------------------|-----------------|
| Row1   | No                 | Lluvioso          | Caliente      | Alta             | true            |
| Row3   | Si                 | Soleado           | Templado      | Alta             | false           |
| Row5   | No                 | Soleado           | Frio          | Normal           | true            |
| Row7   | No                 | Lluvioso          | Templado      | Alta             | false           |
| Row8   | Si                 | Lluvioso          | Frio          | Normal           | false           |
| Row9   | Si                 | Soleado           | Templado      | Normal           | false           |
| Row10  | Si                 | Lluvioso          | Templado      | Normal           | true            |
| Row11  | Si                 | Nublado           | Templado      | Alta             | true            |
| Row12  | Si                 | Nublado           | Caliente      | Normal           | false           |

Esta opción genera un muestreo estratificado, es decir, la distribución de valores en la columna seleccionada se retiene (aproximadamente) en las tablas de salida. Opcionalmente, puede especificar una semilla fija.

| Row ID | <b>S</b> JuegoGolf | <b>S</b> Panorama | <b>S</b> Temp | <b>S</b> Humedad | <b>B</b> Viento |
|--------|--------------------|-------------------|---------------|------------------|-----------------|
| Row0   | No                 | Lluvioso          | Caliente      | Alta             | false           |
| Row2   | Si                 | Nublado           | Caliente      | Alta             | false           |
| Row4   | Si                 | Soleado           | Frio          | Normal           | false           |
| Row6   | Si                 | Nublado           | Frio          | Normal           | true            |
| Row13  | No                 | Soleado           | Templado      | Alta             | true            |

