

Práctica 1

Tipos de frecuencia en una distribución bidimensional


Meses Niveles	Enero	Febrero	Marzo	Abril	Mayo
A	4	6	7	8	8
B	3	3	6	5	9
C	9	7	7	13	14

En la tabla conjunta anterior intervienen las dos variables *meses* y *niveles* de forma muy relacionada, quizás demasiado, pero podríamos querer estudiarlas por separado, para poder estudiar sus características individuales. Para eso deberemos usar otro tipo de frecuencias.

Frecuencias marginales

Para estudiar una variable sola en una distribución bidimensional podemos sumar por filas y columnas las frecuencias, obteniendo así las correspondientes a una sola variable.

Copia la tabla del ejemplo en un archivo nuevo de LibreOffice Calc. Para ello selecciona toda la tabla y pide **Edición Copiar**. Después abre un archivo nuevo de Calc, pide **Edición Pegado especial** y elige la modalidad de *formato HTML*. Así obtendrás una copia en la hoja de cálculo para poder sumar filas y columnas.

Para calcular las **frecuencias marginales** deberás sumar por filas y por columnas. Señala la celda inferior a la columna de Enero y súmala con el botón **Autosuma**. 

Deberá darte 16.

Haz lo mismo en los demás meses, arrastrando la fórmula, o con el botón Autosuma. Obtendrás esta fila de sumas: 16, 16, 20, 26 y 31.

Las frecuencias marginales por columnas ya nos dan una información: En Abril y Mayo ha aumentado el número de faltas graves, independientemente de los niveles.

Haz lo mismo con los niveles. Las frecuencias marginales serán 33, 26 y 50 respectivamente.

Las frecuencias marginales por filas nos indican que el nivel C es el más conflictivo.

Las frecuencias marginales pueden aclarar algunos comportamientos que antes podían estar ocultos en la maraña de datos. Completa, si te apetece, el formato de la tabla tal como ha quedado:

Meses Niveles	Enero	Febrero	Marzo	Abril	Mayo	Frec. Marg.
A	4	6	7	8	8	33
B	3	3	6	5	9	26
C	9	7	7	13	14	50
Frec. Marg.	16	16	20	26	31	109

Frecuencias condicionadas

Hemos estudiado cada variable por separado, pero podríamos considerar la influencia de una en la otra viendo cómo cambia la distribución de una de ellas por influencia de la otra. Esto se consigue aislando en la tabla una sola fila o una sola columna. También se pueden condicionar las frecuencias a varios valores, pero es algo más complicado.

Nosotros efectuaremos una operación más potente: Dividiremos, por ejemplo, cada frecuencia entre el total de cada nivel, para ver si la influencia de la primavera ha sido igual de intensa en los tres niveles.

Para lograrlo, haz una copia de la tabla en otra parte (con **Copiar y Pegar**), junto con todos sus totales, y después sustituye cada frecuencia conjunta por su cociente entre el total de su nivel (totales por filas). **Deberás señalar cada celda en la tabla nueva y escribir las operaciones tomando los datos de la tabla primitiva.**

Las celdas de las frecuencias se deberán convertir en porcentajes. Usa el menú **Formato**, elige **Celdas** y la pestaña **Números** y fija el formato en **Porcentaje**. También puedes fijar dos decimales.

Debe quedar así:

Meses Niveles	Enero	Febrero	Marzo	Abril	Mayo	Frec. Marg.
A	12,12%	18,18%	21,21%	24,24%	24,24%	1
B	11,54%	11,54%	23,08%	19,23%	34,62%	1
C	18,00%	14,00%	14,00%	26,00%	28,00%	1

Al final de las filas aparece un 1, pues es suma de porcentajes. Si ves que no lo logras bien con la explicación anterior usa otro procedimiento, pero no sigas hasta haber logrado estos porcentajes dividiendo cada frecuencia entre el total de su fila.

De esta forma vemos que en el nivel A la primavera no ha influido tanto como en los otros niveles (los porcentajes son más semejantes), mientras en B hay dos subidas fuertes en Marzo y Mayo, que es la más fuerte de todas. En C existe una bajada al principio, pero se complica al final, con incrementos suaves. Los tres niveles han reaccionado de forma distinta.

Podemos representar esta distribución en un gráfico tridimensional:

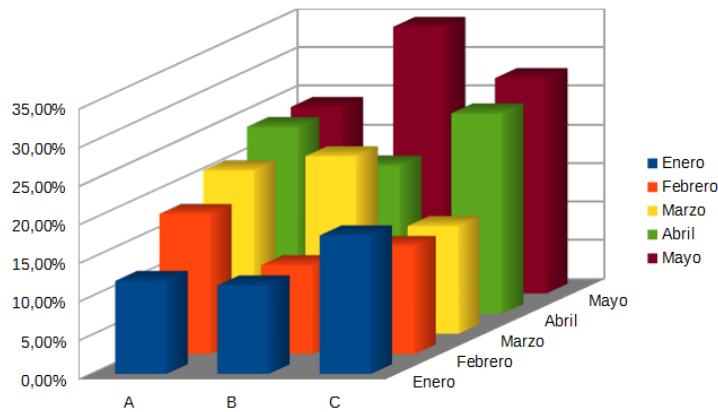
Selecciona toda la nueva tabla, incluidos los rótulos, salvo los 1. Pide **Insertar Gráfico**.

Como tipo de gráfico elige Columnas, en 3D,realista y en profundidad (tercer subtipo)

En el rango activa **que la primera fila y primera columna son etiquetas.**

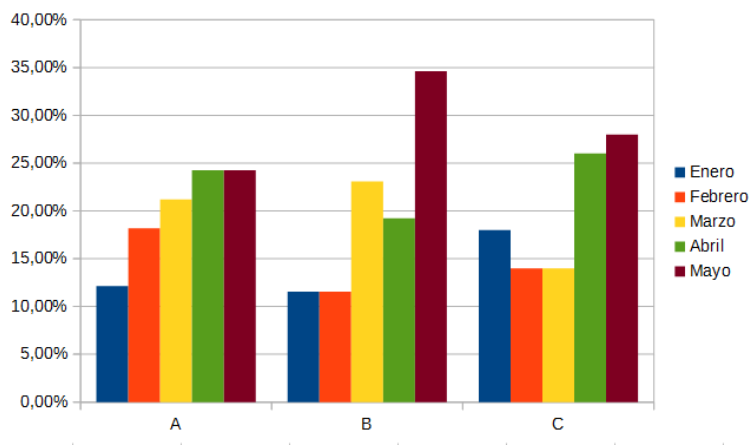
Rellena el título y demás detalles. Termina con **Finalizar.**

Deberás obtener un gráfico parecido a este:



En él se ve que las columnas del nivel B sufren más cambios que las restantes. Aprende a cambiar la orientación de estos gráficos: señala con el ratón todos los puntos de fijación y experimenta los giros hasta que la perspectiva sea la que más te agrade.

También es útil un gráfico de barras normales pero triple. Este lo hemos conseguido **cambiando los datos a filas** en el segundo paso del asistente:



Observamos muy bien en él la evolución de los distintos porcentajes en cada nivel, al venir representados por colores diferentes.

Si te interesa el tema de la verificación de *la independencia* entre el factor **mes** y el factor **nivel**, usa las técnicas de la prueba de independencia incluidas en el apartado **Para ampliar**. Sigue las instrucciones de la herramienta y comprobarás que no se detecta dependencia entre los meses y los niveles: **cada nivel aumenta sus faltas de forma independiente.**