

BIBLIOTECOMETRÍA (Librametrics).

EVALUACIÓN DE COLECCIONES, UN ESTUDIO DE CASO

Prof. Cáser O. Archuby (Coord)
UNLP / FAHCE
coa@isis.unlp.edu.ar

Temas BCI:

Relación entre la edad del material bibliográfico y la circulación, relación entre el tamaño del fondo bibliográfico y la circulación (Ley del 80/20, Trueswell), relación entre la edad del material bibliográfico y su presencia en el fondo bibliográfico. Análisis comparado de las distribuciones estadísticas del fondo bibliográfico y de la circulación, según la edad del material.

Temas matemáticos e informáticos:

Sucesos, variables aleatorias, variables agrupadas, marcas de clases, frecuencias absolutas y relativas, frecuencias acumuladas, medidas de centralización, distribución estadística de las edades (lognormal), logaritmos, operaciones matriciales (producto escalar de vectores, producto de matrices), estructura de datos, funciones de Excel, automatización del cálculo.

Temas tratados en Parte 1:

Estructura de datos, reconocimiento de la matriz de datos, interpretación de los datos, estadística descriptiva y cálculos preliminares.

Parte 2:

- **Automatización del cálculo**
- **Reducción de la probabilidad de error y del esfuerzo**
- **Estadística inferencial mono y bivariada**
- **Análisis de resultados**
- **Ley del 80/20**

Requisitos previos: Haber alcanzado los objetivos de la Parte 1.

Consigna:

Estudiar el texto y las Hojas de Cálculo de los Libros Datos originales y Resolución Parte 1, repitiendo todos los cálculos sin ayuda del texto ni de las planillas, usando sólo los datos copiados en otra Hoja y resolviendo sobre la misma la Parte 1, como parte de los requisitos previos de esta Parte 2. Luego, a partir del Libro Sólo Datos, resolver la Parte2 de la misma manera, con la guía del texto y las resoluciones del libro Cálculos.

Auto-evaluación

En todas las instancias de la resolución tomar nota de las dificultades y clasificarlas de acuerdo a su origen temático y grado de dificultad:

Temas:

Matemáticos
Informáticos
BCI

Grados De Dificultad

1(dificultad baja), 2(dificultad media) ó 3(dificultad alta)

Informar el resultado a los docentes.

INTRODUCCIÓN

En la Parte 1 adquirimos familiaridad con la estructura y significado de los datos y realizamos algunas operaciones básicas de cálculo y conteo lo que nos permitió adquirir familiaridad con las capacidades para la operatoria matemática básica que ofrece Excel.

También adquirimos familiaridad con precisiones terminológicas que permiten eliminar la posibilidad de definiciones ambiguas en los procedimientos, tema de imprescindible presencia en la ingeniería de procesos automáticos.

En este tramo del estudio trabajaremos sobre la matriz de datos modificada que resultó de la resolución de la Parte 1. En esta matriz todos los datos son numéricos por lo que el *tipo de datos* permitirá hacer todos los cálculos necesarios y, por otra parte, la *estructura matricial de los datos* facilitará la automatización de los procesos de cálculo y el uso de funciones de Excel.

Por último, se presentará y destacará el valor de la abstracción como base de la modelización matemática que facilitará el uso de funciones de Excel y, también, como base de la generalización de los procedimientos, lo que ampliará su radio de aplicación.

En lo referido a la abstracción resaltaremos dos niveles de avance en tal sentido, el primero nos permitirá expresar nuestro problema bibliotecario en términos estadísticos y el segundo, el más avanzado, en términos del álgebra matricial; esto último permitirá el uso de poderosas funciones matemáticas de Excel, ya que sus diseñadores las incluyeron en la biblioteca de funciones con el objeto de lograr el más amplio rango de aplicaciones, para lo que deberemos expresar nuestro problema con la mayor generalidad posible y esto requiere alcanzar el máximo nivel de abstracción.

En concreto, si en lugar de describir nuestro problema BCI en términos BCI como documentos o préstamos, lo hacemos en términos estadísticos y algebraicos como variables aleatorias, frecuencias, vectores y matrices, estaremos usando el lenguaje común de la ciencia y la técnica aplicada a cualquier tema y así veremos

facilitado el uso de Excel y la automatización de procedimientos con la consecuencia de menos trabajo y menos errores

Como toda autoevaluación tiene por objetivo respondernos preguntas, comenzaremos nuestro trabajo formulando preguntas de interés para la evaluación de la colección, en general, y para establecer relaciones entre la edad y la circulación de los documentos, en particular.

Comentario sobre el método

Adherimos al postulado “Lo que se usa no se olvida y lo que no se usa no debe enseñarse”, debido al Dr. Luis Santaló, que es un resumen de nuestra postura y se reflejará en nuestro material ya que sólo incluiremos *repetidamente* aquellos temas que aparecen *repetidamente* durante la resolución de problemas de la BCI.

PREGUNTAS PRELIMINARES y CÁLCULOS PREVIOS NECESARIOS

Pregunta 1

**Sobre el final de la primera parte respondimos a la pregunta
"Es muy grande la proporción de la colección que no salió del estante?"**

➤ **Cálculos previos necesarios:**

La proporción de la colección que no salió del estante está dada por el cociente entre la cantidad de documentos que nunca fueron prestados y la cantidad total de documentos, de la colección en estudio. El porcentaje de documentos que nunca fueron prestados es la expresión de la proporción en forma porcentual.

Por lo tanto, necesitamos ambas cantidades subrayadas y para calcularlas seguiremos el siguiente paso:

Procedimiento

Paso 1

Cálculo de la cantidad de documentos que nunca fueron prestados.

Todas las *celdas* de la *columna* C desde C5 hasta C34 contienen cantidades de documentos, de distintas edades, que nunca fueron prestados, por lo tanto sumaremos la columna C desde la *fila* 5 hasta la 34.

En C35 ingresamos: =Suma (C5:C34). El resultado será **606**.

Comentario:

Hemos visto que el valor 10 de C5 corresponde a la cantidad de documentos cuya edad está entre 1 y 5 años que nunca fueron prestados. (Esta expresión está en términos bibliotecológicos)

En el ámbito abstracto de la estadística, diríamos que un [suceso](#) (*Documento con edad entre 1 y 5 años que nunca fue prestado*) se produjo con una [frecuencia absoluta](#) de 10. Cada valor de las celdas de la columna sumada contiene la frecuencia absoluta del evento correspondiente.

606 es la frecuencia absoluta del evento "Documento nunca prestado"

*Habitualmente las frecuencias absolutas estarán referidas a sucesos relacionados con valores de alguna variable aleatoria, en adelante **VA**.*

Paso 2

Cálculo de la cantidad de documentos de la colección.

Todas las celdas de la matriz, desde C5 hasta AA34, contienen cantidades de documentos de diversas edades (*valores de la VA*) que fueron prestados diversa cantidad de veces (*frecuencias*). Ningún documento está repetido y todos están presentes por lo que si sumamos todas las celdas tendremos la cantidad total de documentos.

En AB37 ingresamos: = Suma (C5:AA34). El resultado será **1294**.

Más adelante veremos que, si bien este procedimiento de cálculo es el más simple y directo, hay otros que ofrecen resultados parciales útiles para responder otras preguntas.

Paso 3

Cálculo de la proporción de documentos que nunca fueron prestados, con respecto al total de documentos de la colección en estudio.

El valor de la *proporción* de una parte con respecto al todo es el cociente de la parte dividida por el todo, en este caso: **606 / 1294**. El resultado es **0,468**. Ver celda C36.

0,468 es la frecuencia relativa del evento "Documento nunca prestado"

Si no disponemos de otra información podremos decir que, como la *frecuencia relativa* tiende a la *probabilidad*, **0,468** será la probabilidad de que el año próximo no se preste un documento de la colección, para el caso de conservarse el contexto.

La expresión de la proporción como porcentaje, habitualmente llamada "el porcentaje" se obtiene multiplicando el resultado anterior por 100. El resultado es **46,8%**. Ver celda C36

Pregunta 2

Cuál es la edad media de la colección?

➤ **Cálculos previos necesarios:**

La edad media de la colección es la Media estadística de la Variable Aleatoria Edad, (en adelante VA Edad) y se calcula, como cualquier otra media, **sumando todos los valores de la VA Edad**, que son las edades de cada documento de la colección y luego dividiendo esa suma por la cantidad de documentos de la colección, **1294**, que ya conocemos por la respuesta a la Pregunta 1.

Para el cálculo de las sumas de las edades no podemos usar los valores de la VA Edad de la columna A porque no son numéricos y están expresados en forma de datos agrupados en intervalos de clase.

En su lugar tomaremos como valor de la VA Edad a la Marca de Clase de la VA Edad, que vemos en la columna B, desde B5 hasta B34, como vimos en Parte 1.

La **marca de clase** es el punto medio del intervalo de clase y se calcula haciendo el promedio de los extremos, para el primer intervalo sería 3 ya que:

$$= (1+5)/2 = 3$$

Definiciones recursivas y Automatización (Dificultad 1=matemático)

La definición aritmética de la marca de clase como semisuma de los extremos del intervalo es correcta, además de simple, fácilmente entendible y práctica para el *cálculo manual*, pero eso significaría 30 sumas y divisiones, con incremento de la carga de trabajo manual y, consecuentemente, de la probabilidad de cometer errores.

La causa está en que la definición es dependiente de los datos y por ello cambia la definición cuando cambian los datos (que son los extremos del intervalo).

La base de la automatización de procesos de cálculo está en hallar **definiciones comunes para todos los datos**, a partir de propiedades cumplidas por todos los datos, con el requerimiento obvio de que los datos no formen parte de la definición.

Observemos los datos:

Los extremos inferiores de cada intervalo de clase son iguales al anterior más 5

➤ (1 – 6 – 11 – 16 -- 136 – 141 – 146).

Lo mismo ocurre para los extremos superiores

➤ (5 – 10 – 15 -- 140 – 145 – 150)

y también para cada punto interior al intervalo, por ejemplo, *el punto medio, que es la marca de clase*: 3 – 8 – 13 -

Entonces, definiremos cada marca de clase a partir de la anterior como:

- *Cada marca de clase es igual a la anterior más 5, ó más general,*
- *Cada marca de clase es igual a la anterior más el tamaño de la clase*

Ambas definiciones de tipo recursivo son válidas para todas las clases excepto la primera, ya que no tiene anterior, y por ello deberemos calcularla manualmente

Procedimiento para el cálculo del punto medio (marca de clase) con las funciones Excel:

Paso 1 Ingresar $= (1+5)/2$, en la celda B5 (Cálculo de la primera marca)

Paso 2 Ingresar $= B5 + 5$ en la celda B6 (Definición general de las marcas)

Paso 3 "Arrastrar" la celda B6 hasta la celda B34 (Cálculo automático)

Resultado: Marcas de clase desde 3 hasta 148 en las celdas B5 hasta B34.

Observar las fórmulas de cada marca de clase.

Como es habitual en estadística las VA con valores repetidos se presentan en forma de valores asociados a frecuencias, en este caso, las VA son las edades y tales frecuencias son las cantidades de documentos cuyas edades están entre 1 y 5 (Marca de clase 3) años y están dispersas en la fila 5, desde C5 hasta AA5, según sea la cantidad de préstamos correspondiente.

Todos los valores de la fila son cantidades de documentos, por lo cual, si sumamos la fila tendremos la cantidad total de documentos cuya edad está entre 1 y 5 años.

**Esta cantidad es la frecuencia absoluta del valor de la variable aleatoria.
En términos de la Bibliotecología es la cantidad de documentos que
tienen una determinada antigüedad**

Procedimiento:

Paso 1

· Ingresar en la celda AB5: = Suma (C5:AA5)

El resultado es 40, lo que significa que hay 40 documentos cuyas edades están entre 1 y 5 años. Como para los cálculos tomaremos como edad del intervalo a la marca de clase, diremos que hay 40 documentos cuya edad está *alrededor* de 3 años.

➤ Como no olvidamos que estamos buscando la suma de las edades de todos los documentos, observamos que la suma de las edades de estos 40 documentos será, entonces, $40 \times 3 = 120$ años

Paso 2

Si repetimos el proceso tendremos las cantidades de documentos correspondientes a cada edad, entonces, **"Arrastramos"** la celda AB5, hasta AB34. Veremos valores que van desde 40, en AB5, hasta 6, en AB34.

Como en el Paso 1, observamos que este valor 6 de AB4 significa que hay 6 documentos cuya edad está entre 146 y 150 años, *alrededor o aproximadamente* 148 años.



Como para los cálculos tomaremos el valor 148 la marca de clase que se registra en la celda B34, diremos que hay 6 documentos cuya edad es 148, por lo que la suma de sus edades será igual a $6 \times 148 = 888$ años.

Paso 3 (Suma manual)

Vemos que el producto de cada valor de la columna B por su correspondiente de la columna AB ($B5 \times AB5, B6 \times AB6, \dots, B34 \times AB34$), nos da la suma de la cantidad de años de todos los documentos de la misma edad, por lo que si sumáramos todos los años así calculados tendríamos la búsqueda **suma de la edad de todos los documentos** de la colección. Para hacerlo, agregamos una columna auxiliar, la columna AD, de la siguiente manera:

Ingresar en $AD5 = AB5 \times B5$, el resultado será 120
"Arrastrar" AD5 hasta AD34. En AD34 el resultado será 888.

Ingresar en $AB35 = \text{SUMA}(AB5:AB34)$. Resultado = 1294

La suma de la cantidad de documentos de cada edad nos da la cantidad total de documentos.

Ingresar en $AD35 = \text{SUMA}(AD5:AD34)$. Resultado = 49852

La suma de la cantidad de años de los documentos de cada edad nos da la cantidad total de años de todos los documentos de la colección

Ingresar en $AD36 = AD35/AB35$. Resultado = 38,5

La división de la cantidad total de años de los documentos dividida por la cantidad de documentos, nos da 38,5 años que es el promedio de años de los documentos al que llamaremos Edad Media de la Colección.

Veamos otra forma de hacerlo, más complicada pero con más futuro y poder(+automatización).

Paso 3 (Suma automática)

Si observamos la suma las celdas desde AD5 hasta AD34 y la fórmula de cada una de sus celdas, veremos que corresponde a la siguiente fórmula:

$$\mathbf{B5 \times AB5 + B6 \times AB6 + B7 \times AB7.....+ B34 \times AB34}$$

Vemos que hemos multiplicado el primer elemento de una columna por el primer elemento de la otra, luego el segundo por el segundo, hasta finalizar multiplicando el último por el último y, finalmente, hemos sumado todos los productos.

Esto que hemos hecho es una operación matemática clásica muy importante, llamada producto escalar de vectores, que es un caso particular de la multiplicación de matrices, ya que el vector es un caso particular de matriz que tiene una sola fila o una sola columna.

Excel dispone de una función llamada MMULT (Matrix Multiplication) para multiplicar matrices que cumplan las propiedades que requiere tal multiplicación.

Nuestros vectores son parte de las columnas B (B5 a B34, marca de clase de las edades) y AB (AB5 a AB34, suma del total de documentos de cada edad) que son matrices de 30 filas por 1 columna. (Tenemos 1 matriz columna de marcas de clase de 30 filas y 1 matriz columna de total de documentos por edad, de 30 filas)

La multiplicación de dos matrices requiere el cumplimiento de la siguiente propiedad:

➤ *la cantidad de columnas de la primera debe ser igual que la cantidad de filas de la segunda;*

Nuestros vectores no cumplen esta propiedad ya que ambos tienen la misma cantidad de filas y columnas y eso impide usar la función MMULT.

La solución al problema es convertir una de las dos columnas en fila ya que entonces tendrá 30 columnas y 1 fila y la otra 30 filas y 1 columna, por lo que la cantidad de filas de una será igual a la cantidad de columnas de la otra, como

requiere la multiplicación de matrices y, consecuentemente, la función MMULT de Excel.

Vamos a convertir el *vector columna* de las edades de la columna B en un *vector fila*, en una operación llamada *transposición de matrices*, que puede realizarse mediante el siguiente

Procedimiento:

Paso 3.1 Copiar las celdas B5 a B34

Paso 3.2 Seleccionar la celda C40. (Podría ser cualquier otra, siempre que fuera parte de una fila con 30 celdas vacías)

Paso 3.3 Presionar el botón derecho y seleccionar Pegado Especial, Valores y Transponer.

Paso 3.4 Aceptar.

Vemos que los valores de las edades de las celdas B5 - B34 aparecen en las celdas C40 - AF40. La columna B (matriz de 30 filas y una columna) ahora se transformó en una matriz de 1 fila y 30 columnas

En la columna AB, bajo el encabezamiento TD (Total Documentos), en las celdas AB5 – AB34 aparecen los valores de la cantidad de documentos correspondientes a cada edad con la forma de una matriz de 30 filas y 1 columna

Ahora se cumple con el requerimiento:

La cantidad de columnas de la matriz de las edades (C40-AB40) es igual que la cantidad de filas de la matriz de los totales de documentos por edad (AB5-AB34), por lo que podemos multiplicarlas.

Paso 4

Ingresar en AD38: =MMULT (C40:AF40;AB5:AB34), ó =MMULT(C40:AF40|AB5:AB34), la diferencia entre las dos formas de la ecuación es el separador (; ó |).

También puede invocarse a la función MMULT y en Matriz 1 seleccionar la fila C40 a AF40 que corresponde a las edades de los documentos, y en la Matriz2 la columna AB que corresponde al total de documentos por cada edad.

El resultado será: **49852** años, que es la suma de las edades de todos los documentos. Es el mismo valor que la celda AD35 que corresponde al mismo cálculo pero de forma manual

Paso 5 Cálculo de la edad media de la colección

Ya disponemos de la *suma de las edades de todos los documentos* (49852 años) y de la *cantidad de documentos* (1294), por lo que la edad media será:

$$\text{Edad media} = 49852 / 1294 = \mathbf{38,5 \text{ años.}} \text{ Ver celda AD36.}$$

NOTA:

Llamamos "suma manual" o "manual" a la que requiere mayor participación nuestra y "automática" o "semiautomática", a la que requiere menor participación.

Comentario:

Las expresiones anteriores (*suma de las edades de todos los documentos, cantidad de documentos*) **pertenecen al mundo concreto de la BCI.** En el primer avance de abstracción, el de la estadística, diremos que calcularemos la media de la VA Edad cuyos valores están asociados a frecuencia, aplicando la siguiente fórmula:

$$\mathbf{X_{media} = (\sum x f_x) / \sum f_x.}$$

donde en el producto $\sum x f_x$, las x corresponden a las edades, que son los valores de la VA, (B5 a B34) y f_x a las cantidades de documentos, que son las frecuencias de los valores de la VA (AB5 a AB34).

Pregunta 3

Cuál es la edad media de los documentos que nunca fueron prestados?

➤ **Cálculos previos necesarios:**

La edad media de los documentos que nunca fueron prestados es igual a la *suma de las edades de tales documentos*, dividida por la *cantidad de los mismos*.

La cantidad de documentos que nunca fueron prestados, según edad, está registrada en la columna C35, desde C5 hasta C34, por lo que su total es la suma de la columna. Ver C35(cálculo ya realizado en la respuesta a la PREGUNTA 1).

Procedimiento

Paso 1 En C35, ingresar =suma (C5:C34). El resultado es 606, lo que significa que 606 documentos no fueron prestados durante el año del estudio.

La suma de las edades de estos 606 documentos se calcula con el mismo procedimiento que usamos para calcular la suma de las edades de toda la colección. La diferencia sólo está en el vector de cantidad de documentos que multiplicaremos por el vector de las marcas de clase de las edades.

En respuesta a la Pregunta 2, multiplicamos las marcas de clase de las edades de la columna B por las cantidades totales de documentos de cada edad registradas en la columna AB.

En la respuesta a la Pregunta 3 las cantidades de documentos sólo incluirán a los que nunca fueron prestados, cuyas cantidades están en la columna C, por lo

que multiplicaremos B por C en lugar de B por AB , como hicimos en la **Pregunta 2**.

El procedimiento es el mismo, sólo cambia una de las dos variables del cálculo, tanto para la suma manual como para la automática.

Ver resultados, fórmulas y comentarios en la columna AE.

El significado del resultado es distinto ya que en lugar de obtener la media de la colección obtendremos la media de la porción de la colección que no circuló.

La comparación de ambas medias puede aportar información sobre una posible relación entre la edad del material bibliográfico y su circulación.

En la ejecución del procedimiento debe elegirse las celdas auxiliares y las depositarias de los resultados parciales y finales de modo que la planilla conserve la mayor claridad posible.
Si se considera necesario puede usarse una Hoja auxiliar como borrador.
En todos los casos es conveniente colocar comentarios.

Pregunta 4

Cuál es la edad media de los documentos que fueron prestados alguna vez?

➤ **Cálculos previos necesarios:**

La edad media de los documentos que fueron prestados alguna vez es igual al resultado de la división entre la suma de las edades por la cantidad de tales documentos, por lo que necesitamos calcular ambas cantidades.

Procedimiento:

La cantidad de documentos que fueron prestados alguna vez es igual a la cantidad total de documentos (AB35) menos la cantidad de documentos que nunca fueron prestados (C35 de Pregunta 1).

La suma de las edades de los documentos que fueron prestados alguna vez es igual a la suma de las edades de todos los documentos (AD35, AD38, Pregunta 2), menos la suma de las edades de los documentos que nunca fueron prestados (AE35, AE38 Pregunta 3)

Paso 1

Ingresar $=AB35 - C35$, en la celda B43, o en otra celda que parezca conveniente. El resultado es 688, que es la cantidad de documentos que fueron prestados alguna vez.

Paso 2

Ingresar en B44 $= AD35 - AE35$. El resultado es 20654, que es la suma de las edades de los referidos 688 documentos.

Paso 3

Ingresar en B45 $= B44 / B43$. El resultado es 30, que es la media de la edad de los documentos que fueron prestados alguna vez.

Observación:

Se repite aquí la situación planteada en el último párrafo del Paso 2 de la respuesta a la [Pregunta 1](#), ya que el procedimiento usado es el más simple y directo para responder la Pregunta 4, pero no proporciona información que facilite la respuesta a otras preguntas como las preguntas 8, 9 y 10.

Pregunta 5

Cuál es el porcentaje de documentos de cada edad?

➤ **Cálculos previos necesarios:**

El porcentaje de documentos de cada edad es igual a la división de la cantidad de documentos de cada edad por la cantidad total de documentos, multiplicada por 100. Necesitamos, entonces, la cantidad de documentos de cada edad y la cantidad total de documentos.

En la Pregunta 2, Paso 1 y Paso 2 del Procedimiento, calculamos las cantidades de documentos para cada edad que registramos en la columna AB, desde la fila 5 hasta la 34. La cantidad total de documentos, 1294 (AB35), es conocida desde la Pregunta 1.

Partiendo del Libro Pregunta 2 sólo queda, entonces, dividir cada valor de las celdas AB5 hasta AB34, por el valor de la celda AB35.

Procedimiento

Paso 1 Ingresar en $AF5 = AB5 / AB\$35 * 100$ (disminuir decimales hasta 1)

Resultado: 3,1%

Significado: El 3,1 por ciento de los documentos tiene edad entre 1 y 5 años.

Paso 2 "Arrastrar" AF5 hasta AF34.

Paso 3 **Verificación:** Sumar los porcentajes desde AF5 hasta AF34.

Resultado: 100 (AF35)

Comentario:

**Prestar atención al signo \$ en AB\$35.
Ya que el mismo fija la fila 35 para el "arrastre"**

Pregunta 6

Cuál es el porcentaje de documentos que tiene 30 años o menos?

➤ **Cálculos previos necesarios:**

El porcentaje de documentos que tiene 30 años o menos es igual a la cantidad de documentos que tiene 30 años o menos dividida por la cantidad total de documentos y multiplicada por 100.

Necesitamos, entonces, la cantidad de documentos que tiene 30 años o menos y la cantidad total de documentos.

- La cantidad total de documentos es conocida (1294, AB35).
- La cantidad de documentos que tiene 30 años o menos es igual a la cantidad de documentos cuya edad va de 1 a 30 años, inclusive, por lo que debemos sumar la cantidad de documentos cuya edad está entre 1 y 5 años, más los de 6 a 10 años, etc. Hasta llegar a la edad de 26 a 30 años

Procedimiento 1:

Partimos del Libro Pregunta 5

Ingresar en B48: =suma (AB5:AB10) / AB35 * 100.

Resultado: 60%

Significado: El 60% de los documentos tiene una edad igual o inferior a 30 años
Consecuentemente el 40% tiene más de 30 años.

Procedimiento 2: (Usando resultados de la respuesta a la [Pregunta 5](#))

Ingresar en C48: = suma (AF5:AF10).

Resultado: 60%

Significado: El 60% de los documentos tiene una edad igual o inferior a 30 años
Consecuentemente el 40% tiene más de 30 años.

Observación 1:

Hasta la pregunta 6 hemos contado cantidades de documentos, totales y parciales, para lo que hemos dispuesto de una matriz en la que cada elemento es una cantidad de documentos y corresponde a una *frecuencia* de la conjunción de los *valores* de las dos VA objeto de nuestro estudio: **Edad** y **Cantidad de préstamos**.

No disponemos, entonces, de la cantidad de préstamos total ni de cantidades de préstamos parciales.

PRÉSTAMOS

Pregunta 7

Cuál es el porcentaje de préstamos correspondientes a documentos de menos de 21 años de edad?

➤ **Cálculos previos necesarios:**

Para calcular el porcentaje que responde a la pregunta se requiere la **cantidad total de préstamos** y la **cantidad de préstamos correspondiente a documentos cuya edad no supere los 20 años**

Observación 2:

La pregunta es equivalente a:

“Cuál es el porcentaje de préstamos correspondiente a documentos cuya edad está entre 1 y 20 años?”

Observación 3:

Revisar en Parte 1, los valores registrados en L4 y L5 que se interpretan como:

Tenemos que 6 documentos (L5), fueron prestados 9 veces (L4), cada uno por lo que se realizaron $6 \times 9 = 54$ préstamos.

La cantidad de préstamos correspondiente al material bibliográfico de una de las edades (una fila) se obtiene sumando los productos de cada elemento de la fila elegida por su correspondiente de la fila 4.

Para calcular el porcentaje de préstamos correspondiente a documentos de menos de 21 años de edad, se requiere calcular **la cantidad de préstamos de documentos de esas edades** y la **cantidad total de préstamos**. El porcentaje será el resultado de dividir la cantidad parcial por la total y, luego, multiplicar por 100.

Observación 4:

La pregunta no hace referencia al porcentaje de documentos que tienen menos de 21 años de edad sino a su **circulación**, pero sería interesante conocer cuál es el porcentaje de la colección (costo) al que corresponde el porcentaje de circulación (beneficio) de la pregunta con el fin de establecer la relación costo beneficio de la porción de la colección en estudio.

Procedimiento 1:

- **Cálculo semiautomático** (30 operaciones)

Comentario

Este procedimiento hace uso del *producto escalar de vectores* cuyo nombre proviene de su característica principal, ya que se trata de una operación entre vectores que no tiene como resultado un vector, sino un *escalar*, un número. En nuestro caso el escalar resultante será igual a la cantidad de préstamos correspondiente a un grupo de edades, por ejemplo 1 a 5 años.

El producto escalar es una operación compuesta por una cantidad de productos y sumas que, en nuestro caso, serán 25 productos, uno para cada cantidad de préstamos por documento, y las correspondientes 24 sumas, 49 operaciones en total; como veremos más adelante, si quisiéramos calcular la cantidad total de préstamos, ver Observación 7, seguramente deberemos repetir la operación 30 veces, una para cada grupo de edades, lo que resultaría en **1470 operaciones simples** de suma y multiplicación si trabajáramos con una calculadora, que se reducirán a **30 operaciones complejas**, una para cada edad, usando el producto escalar de vectores. Las operaciones para el nivel del usuario de Excel serán **dos**: la definición del primer producto escalar con fijación del vector de fila 4 y el correspondiente arrastre.

Más adelante veremos que, usando MMULT, en forma de *operación matricial*, las 30 operaciones de producto escalar, pueden ser reducidas a UNA, como veremos en:

Procedimiento 2: Cálculo automático.

Paso 0 : Ingresar en B49 =suma (AF5:AF8).

Resultado: **36%**.

Significado: El 36% de la colección tiene menos de 21 años de edad.

Paso 1 Repasar la **Pregunta 2**, prestando atención a la variante “Suma Automática”, en particular al papel de la abstracción, que convierte una operación bibliotecológica en una operación matemática, el *producto escalar* de dos vectores, y así permite el uso de la función MMULT de Excel, previa preparación de los datos.

En este caso se trata de multiplicar dos *vectores fila*, pero varias veces, ya que aunque uno de los vectores permanece fijo, el de la fila 4 que registra cantidad de préstamos por documento, el otro varía para cada edad entre 1 y 20 años, filas 5 a 8.

Por ejemplo:

Debemos *multiplicar en forma escalar* el vector cuyos valores están registrados en las celdas que van de **C4 a AA4**, por el vector cuyos valores están registrados en las celdas que van de **C5 a AA5**, luego **C6 a AA6**, **C7 a AA7** y, finalmente, **C8 a AA8**.

Observación 5:

Recordemos que la función MMULT de Excel multiplica dos matrices tales que la primera tenga la misma cantidad de columnas que la cantidad de filas de la segunda, como exige la multiplicación de matrices. Los dos *vectores fila* que debemos multiplicar, son matrices de 1 fila por 25 columnas por lo que deberemos convertir uno de ellos en una matriz de 25 filas por 1 columna. Si bien no es un requerimiento matemático, elegiremos para esta conversión al vector registrado en las celdas C4 hasta AA4 porque se repetirá en cada multiplicación.

Observación 6:

Cada multiplicación dará como resultado un valor numérico (un escalar, no un vector, de allí el nombre de *producto escalar*), como el valor 197 de la celda AC5, que es la cantidad de préstamos de los documentos cuya edad está entre 1 y 5 años, que fue calculada ingresando =MMULT(C5:AA5;AG\$5:AG\$29) en AC5

Observación 7:

Aunque no lo necesitamos para responder esta pregunta, parece conveniente repetir el procedimiento para todas las edades, no sólo para las requeridas, ya que el "arrastré" requerido para la operación exige muy poco esfuerzo y el resultado parece potencialmente valioso.

Paso 2

2.1 Copiar C4:AA4 y pegar en AG5, con Pegado especial-Valores-Transponer
Resultado: La matriz fila de cantidades de préstamos registrada desde C4 hasta AA4 ha sido convertida en la matriz columna desde AG5 hasta AG34, con los mismos valores.

2.2 Ingresar =MMULT (C5:AA5;AG\$5:AG\$29) en AC5

Resultado: **197**

Significado: Hubo 197 préstamos de documentos cuya edad estaba entre 1 y 5 años.

Paso 3 "Arrastrar" la celda AC5 hasta AC34

Resultado: 30 valores en las celdas AC5 a AC34.

Significado: Cada valor corresponde a la cantidad de préstamos realizado con los documentos de cada grupo de edades, por ejemplo, el valor 2063 de la celda AC8 significa que se realizaron 2063 préstamos con documentos cuyas edades estaban entre 16 y 20 años.

Observar que se ha "fijado" (utilización del signo \$) las celdas AG5 hasta AG29.

Paso 4 Ingresar =SUMA (AC5:AC34), en AC35.

Resultado: 7625

Significado: Cantidad total de préstamos.

Paso 5

5.1 Ingresar en B50, =SUMA (AC5:AC8)

Resultado: 4535

Significado: Cantidad de préstamos de documentos con edad entre 1 y 20 años

5.2 Ingresar en B51 =B50/AC35*100

Resultado: 59%

Significado: El 59% de los préstamos se realizó con documentos cuya edad no superaba los 20 años o menores a 21 años.

Observación 8:

(Rever Observación 4) El **36%** (B49) de los documentos concentra el **59%** (B51) del préstamo, hecho que sustenta la hipótesis que asocia menor edad con mayor circulación y puede interpretarse como que la parte más nueva de la colección tiene mejor rendimiento.

Sería interesante establecer esta relación para cada grupo de edades que resulte de interés y sus correspondientes acumulados.

Procedimiento 2: Cálculo automático (1 operación matricial)

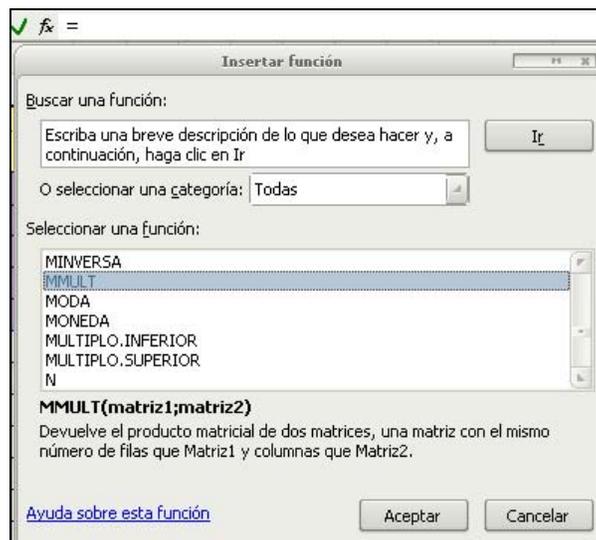
Comentario:

En la columna AC, bajo el encabezamiento TP (Total Préstamos) registramos las 30 cantidades de préstamos correspondientes a cada grupo de edades, calculadas mediante 30 productos escalares, cada uno de los cuales dio como resultado un número. Los 30 valores numéricos registrados en las celdas AC5:AC34 conforman un *vector*, lo que significa que el resultado final de las 30 operaciones es un *vector columna*, o una matriz de 30 filas y 1 columna.

La [multiplicación de matrices](#) del álgebra matricial permite resolver nuestro problema y el uso de la función MMULT de Excel como *fórmula matricial*, permite calcularlo con una sola operación cuyo resultado será una matriz de 30 filas por 1 columna (vector columna) que almacenaremos en la columna AH, cuyo encabezamiento será TP (Total Préstamos) igual que en la columna AC y cuyos 30 componentes serán, obviamente, los mismos.

Paso 1 En la columna AH, encabezada con la abreviatura TP, seleccionaremos las celdas AH5 hasta AH34, una para cada grupo de edades.

Paso 2 Seleccionar Insertar Función y seleccionar MMULT



Paso 3 Seleccionar como Matriz 1 a toda la matriz, desde C5 hasta AA34.

Paso 4 Seleccionar como Matriz 2 a las celdas AG5 hasta AG29

Observación 9:

La cantidad de columnas de Matriz 1 debe ser igual que la cantidad de filas de Matriz 2. Matriz 1 es una matriz de 30 x 25 (30 filas por 25 columnas). Matriz 2 es una matriz de 25 x 1 (25 filas por 1 columna). La matriz producto de la multiplicación tendrá "**la dimensión de los extremos**". La multiplicación de dos matrices, una de **30 x 25** y la otra de **25 x 1** dará una matriz de **30 x 1**

Paso 5 Presionar Ctrl. + Shift + Enter

En versiones de Excel puede exigirse presionar F2 antes de Presionar Ctrl. + Shift + Enter

Resultado parcial: Vector columna de 30 componentes (matriz de 30 x 1), desde AH5 hasta AH34, con valores idénticos a los correspondientes de la columna AC.

Significado: Cada valor corresponde a la cantidad de préstamos realizada con los documentos de cada grupo de edades, por ejemplo, el valor 2063 de la celda AC8 significa que se realizaron 2063 préstamos con documentos cuyas edades estaban entre 16 y 20 años.

Paso 6 En la celda C53, ingresar =SUMA (AH5:AH8)/AH35*100

Resultado: **59%**

Significado: El 59% de los préstamos se realizó con documentos cuya edad no superaba los 20 años (o era menor a 21 años).

Pregunta 8

Cuál es el porcentaje de préstamos correspondiente a documentos prestados más de 10 veces?

Observación 1:

En este caso se pide el cálculo de porcentaje de préstamos, como en la pregunta anterior, pero se identifica al conjunto de documentos a partir de la frecuencia de su circulación, mientras que en el caso anterior se los identificaba a partir de una propiedad intrínseca del documento, como es la edad.

➤ **Cálculos previos necesarios:**

El porcentaje de préstamos correspondiente a documentos prestados más de 10 veces es igual a la **cantidad de préstamos correspondiente a documentos prestados más de 10 veces**, dividida por la **cantidad total de préstamos** y, luego, multiplicado el resultado por 100.

Como ya conocemos la cantidad total de préstamos (7625, celda AC35), nos abocaremos a calcular la cantidad de préstamos correspondiente a documentos prestados más de 10 veces.

Para calcular la cantidad de préstamos correspondiente a los documentos prestados una determinada cantidad de veces, **necesitamos conocer la cantidad de documentos prestados esa cantidad de veces**.

Esa cantidad se obtiene sumando la columna correspondiente por lo que haremos la suma de todas las columnas aunque sólo necesitemos, para esta pregunta, desde la columna N hasta la AA.

Insertar en C35 =SUMA (C5:C34). Como vimos en Pregunta 1, el resultado es 606

“Arrastrar” C35 hasta AA35. Vemos que N35 es 11, lo que significa que 11 documentos fueron prestados 21(N4) veces, generando un total de $11*21=231$ préstamos. U35 es 6, lo que significa que 6 documentos fueron prestados 40(U4) veces, generando un total de $6*40=240$ préstamos, y así para toda la fila 35.

Ahora conocemos la cantidad de documentos prestados cada cantidad de veces.

Observación 2:

La cantidad requerida es igual a la suma de las cantidades de préstamos correspondientes a los documentos que han sido prestados 11 veces ó más, en nuestro caso 21 veces ó más, por lo que debemos calcular las cantidades de préstamos para 21, 22, 25, 27....hasta 83, y sumarlas.

Paso 0

Insertar en C35 =SUMA (C5:C34). Como vimos en Pregunta 1, el resultado es 606

“Arrastrar” C35 hasta AA35. Vemos que N35 es 11, lo que significa que 11 documentos fueron prestados 21(N4) veces, generando un total de $11*21=231$ préstamos. U35 es 6, lo que significa que 6 documentos fueron prestados 40(U4) veces, generando un total de $6*40=240$ préstamos, y así para toda la fila 35. Ahora conocemos la cantidad de documentos prestados cada cantidad de veces, por lo que podemos calcular la correspondiente cantidad de préstamos

Paso 1 Comenzaremos con la última columna (83 préstamos por cada documento de la columna AA.)

Ingresar en la celda AA36, =AA35*AA4

Resultado: 166, ya que hay 2 documentos (AA35), que fueron prestados 83 veces (AA4)

Paso 2 Arrastrar AA36 hacia la izquierda, hasta C36.

Resultado: Un vector cuyas componentes ocupan las celdas entre D36 y AA6, con valores que comienzan con 137 en D36 (=D35*D4) y terminan con 166 en AA36 (=AA35*AA4). Estos valores corresponden a las cantidades de préstamos correspondientes a cada cantidad de préstamos por unidad documental registradas entre D4 y AA4.

Observación 3

De acuerdo con la observación anterior, para responder a la pregunta sólo necesitamos los valores de la fila 36 desde las columnas N hasta AA (en la fila 4, valores 21 a 83). Los demás valores fueron calculados porque no demandaron mayor trabajo ni acrecentaron la probabilidad de error y podrían ser útiles para responder otras preguntas. Para calcular la **cantidad de préstamos correspondiente a documentos prestados más de 10 veces**, sólo necesitamos los valores de las celdas N36 a AA36.

Paso 3

Para calcular el porcentaje:

Ingresar en B36, =SUMA (N36:AA36)/AC35*100

Resultado: 68%

Significado: El 68% de los 7625 préstamos se concentra en documentos que fueron prestados más de 10 veces cada uno que, en este caso, es equivalente a decir que el 68% de los préstamos se concentra en documentos que fueron prestados 21 veces o más.

Observación 4

El alto porcentaje del préstamo que muestra el resultado anterior, sugiere que sería interesante agregar una nueva pregunta que permita medir el conjunto de documentos que concentra casi el 70% de la circulación, por ejemplo:
Cuál es el porcentaje de los documentos que cubren el citado 68% del préstamo.

También sería interesante conocer el promedio de edad de tales documentos

Pregunta 9

Cuál es el porcentaje de documentos prestados más de 5 veces?

Observación 1

En este caso se pide el cálculo del porcentaje de documentos, a diferencia de la pregunta anterior en la que se pedía el porcentaje de préstamos pero, como en la pregunta anterior, ***se identifica el conjunto de documentos a partir de la frecuencia de su circulación***

➤ **Cálculos previos necesarios:**

Cantidad de documentos prestados más de 5 veces
(desde 6 préstamos hasta 83, celdas I35 hasta AA35)
Cantidad total de documentos (ya conocida: 1294)

Procedimiento:

Sumar desde I35 hasta AA35, dividir por AB35 y multiplicar por 100
Ver celda B52

Significado: Algo más de la cuarta parte de la colección tuvo más de 5 préstamos en el año.

Pregunta 10

Cuál es el porcentaje de documentos prestados más de "n" veces?

Comentario:

En este caso "n" es la cantidad de préstamos que, en nuestro caso, toma valores desde 0 hasta 83, por lo que más de "n" significa 1, 2, 3 hasta 84 que es más de 83, pero 84 no tiene interés ya que la cantidad de documentos con 84 préstamos y más, es CERO.

Nótese que, como "n" comienza en cero, "más de cero" es cualquier valor mayor que cero, por lo que la cantidad de documentos prestados "más de cero veces" será la *suma* de las cantidades de documentos prestados 1 vez, 2 veces, hasta 83 veces.

Esa cantidad puede calcularse como la suma que acabamos de describir o restando a la cantidad total de documentos (AB35, valor 1294), la cantidad de documentos que nunca fueron prestados (C35, valor 606).

Lo mismo para "más de 1 vez" y para todas las demás cantidades de préstamos, por lo que, previamente, necesitamos:

➤ **Cálculos previos necesarios:**

Cantidad, porcentaje y porcentaje acumulado de documentos con 1 préstamo, 2 préstamos, hasta 83.

Cantidad total de documentos.

Procedimiento:

(Partimos del Libro Pregunta 9)

Para las cantidades de documentos con 1,2,....83 préstamos, ver celdas D35 hasta AA35

Paso 1

En la celda D39 ingresar $=D35/AB35*100$.

Resultado: 10,6.

Significado: El 10,6% de los documentos fueron prestados 1 vez

Paso 2 Arrastrar desde D39 hasta AB39.

Resultado: 10,6% en D39, para los documentos prestados una vez, hasta 0,2% en AA39, para los documentos prestados 83 veces

Paso 3 Ingresar en AB39 = SUMA (D39:AA39).

Resultado: 53,2%.

Significado: el 53,2% de los documentos fue prestado alguna vez.

Comentario:

Sumando AB39 y C36, debe dar al 100% ya que el 100% de los documentos fue prestado alguna vez o ninguna vez.

Paso 4 Ingresar en D41, =100-C36 (el 100% menos el 46,8 que nunca fue prestado)

Resultado: 53,2 (igual que AB39), porcentaje de documentos que fueron prestados alguna vez (más de cero veces).

Paso 5 Ingresar en E41 =D41-D39.

Resultado: 42,6%, porcentaje de documentos prestados más de una vez

Paso 6 Arrastrar E41 hasta AB41.

Resultado: desde 53,2 en D41 hasta 0,0 en AB41

Significado: Porcentajes de documentos prestados más de "n" veces,
con "n" = 0 hasta 83

Comentario: Vemos en L41 que el 20% de los documentos fueron prestados más de ocho veces

Pregunta 11

Se verifica la ley de Trueswell en este caso?

➤ **Cálculos previos necesarios:**

Identificación de los documentos más prestados y acumulación hasta llegar al 20%

Cálculo del porcentaje de préstamos correspondiente a ese 20% de los documentos

Procedimiento:

Paso 1 En la fila 37, calcular el porcentaje de préstamos correspondiente a cada columna

Ingresar en AA37, = AA36/\$AC35*100.

Resultado: 2,2

"Arrastrar" AA37 hasta D37 (=1,80)

En la fila 38 calcular el acumulado de la fila 37, desde la columna AA hasta la columna D.

Ingresar en AA38, =AB38+AA37

Resultado: 2,2

"Arrastrar AA38 hasta D38 (=100)

Paso 2 Ver en fila 38 el acumulado del porcentaje de préstamos, desde 2,2%, correspondiente a 83 préstamos, hasta 100%, correspondiente a 1 préstamo.

En la celda L38 vemos 82% que es el porcentaje de préstamos correspondientes a los documentos prestados 9 veces o más (o más de 8 veces).

INTERPRETACIÓN

Existe un conjunto de documentos que comprende el 20% de la colección (los documentos con 9 préstamos o más) que concentra el 82% del préstamo.

Esto muestra que la clase 3 de esta colección tiene una fuerte aproximación a la **“Ley del 80-20” o “Ley de Trueswell”**, que dice que con el 20% de la colección se atiende el 80% del préstamo.

Pregunta 12

La distribución estadística de la edad de la colección corresponde al modelo lognormal?

Partimos del Libro Pregunta 7

➤ **Cálculos previos necesarios:**

Graficar con tipo de gráfico Dispersión X-Y, tomando como X a la marca de clase de la edad de la columna B y como Y a la cantidad de documentos de la columna AB

INTERPRETACIÓN

Ver Gráfico.

La distribución Lognormal es una distribución asimétrica en la que la cola izquierda, la correspondiente a edades bajas, es siempre ascendente y corta, mientras que la cola derecha es siempre descendente y muy larga.

En este caso la distribución del material bibliográfico sigue a la distribución lognormal hasta la edad de 43 años. Luego, a partir de 48 años crece, cuando debería continuar decreciendo, por lo que se aparta de la distribución lognormal.

Significado bibliográfico de la cola descendente: Cuanto mayor es la edad del material menor es la cantidad del mismo en la colección, a partir del punto de mayor concentración de material, en este caso 18 años.

Es necesario investigar las causas de la permanencia de material antiguo en la colección. Como una de las causas podría ser que se trata de material de uso habitual, debemos estudiar el comportamiento de la circulación por edad. Justamente este estudio no permitirá responder la siguiente pregunta.

Pregunta 13

La distribución estadística de la edad del material circulante corresponde al modelo lognormal?

Graficar con tipo de gráfico Dispersión X-Y, tomando como X a la marca de clase de la edad de la columna B y como Y a la cantidad de préstamos de la columna AC

INTERPRETACIÓN

El gráfico nos muestra buen ajuste a la distribución lognormal, aún para las edades en las que el fondo bibliográfico se aparta.

Pregunta 14

Se observa correspondencia entre las dos distribuciones anteriores?

Las diferencias a partir de los 48 años son claros indicadores que nos muestran que el comportamiento de la colección se aparta del correspondiente a la circulación, por lo que se impone investigar las causas.

Pregunta 15

Cuál es la moda de la edad de la colección?

➤ **Cálculos previos necesarios:**

Identificar el valor de la variable aleatoria con mayor frecuencia

El intervalo modal es 16 – 20 ó la marca de clase 18 (frecuencia =192, celda AB8)

Pregunta 16

Cuál es la moda de la edad de la colección con respecto a la circulación?

➤ **Cálculos previos necesarios:**

Identificar el valor de la variable aleatoria con mayor frecuencia

El intervalo modal es 16 – 20 ó la marca de clase 18 (frecuencia =2063, celda AC8)

Pregunta 17

Cuáles otras preguntas le parece necesario responder para completar la evaluación cuantitativa de la colección?

Formule las preguntas, desarrolle las respuestas, interprete los resultados y justifique las preguntas con la información obtenida con las respuestas.