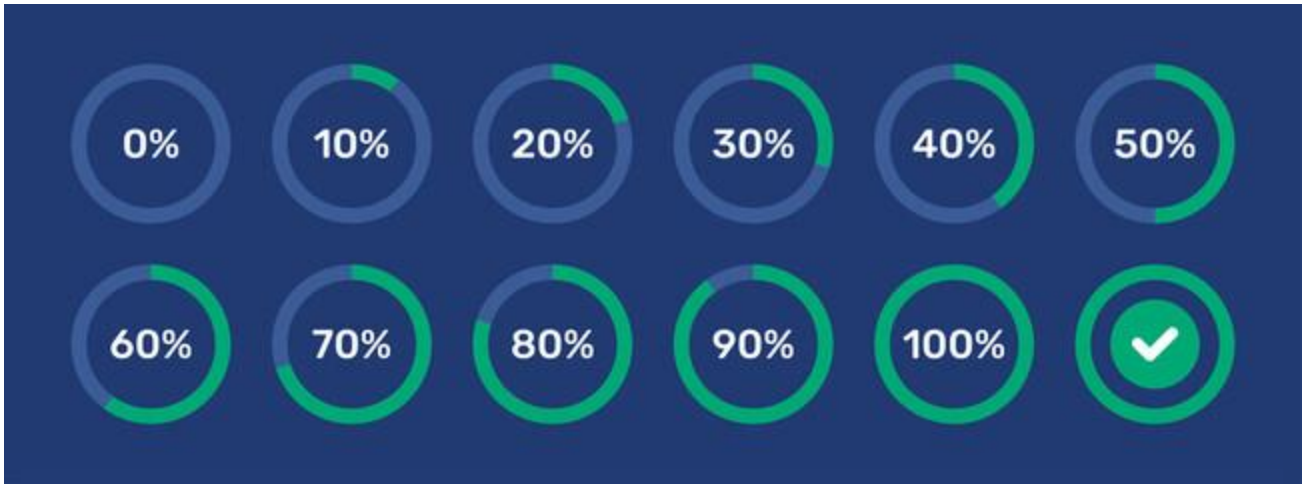


CODIFICACIÓN Y TABULACIÓN DE DATOS E INFORMACIÓN DE MERCADOS

OBJETIVOS DE LA CODIFICACIÓN Y TABULACIÓN DE DATOS

RECETAS DE MARKETING

- La codificación consiste en asignar a todos y cada uno de los tipos de respuesta del cuestionario o del documento de observación un dígito o número que represente a dicha respuesta en un soporte informático, un calculo estadístico...

- El objeto de la codificación es representar todas y cada una de las respuestas de los cuestionarios y de las anotaciones de los documentos de observación por códigos numéricos que faciliten su agrupación.

- Las preguntas cerradas se codifican asignando un número a cada una de las respuestas predeterminadas.

- Para codificar las preguntas abiertas o semi abiertas se procede del siguiente modo:
  - Se leen todas las respuestas dadas a la pregunta
  - Se establecen clases de respuestas
  - Se asigna un código a cada una de las clases de respuestas establecidas.

RECETAS DE MARKETING

- La clasificación de las preguntas abiertas es laboriosa y está sujeta a errores y subjetividades por lo que, sin negar la riqueza de matices que puede aportar, es aconsejable reducirlas al mínimo en los cuestionarios

- LA TABULACIÓN de los datos es el recuento del número de respuestas que ha tenido una pregunta y su presentación en tablas.

- Consiste en resumir los datos en una tabla que sustituya la masa de datos difícilmente manejable, por una representación ordenada de los mismos, que facilite su uso.

- La tabulación no tiene un fin en sí misma, sino que es el medio que permite al investigador tener una visión de conjunto de la información de todos los cuestionarios, al convertir éstos en datos homogéneos que posteriormente serán objeto de análisis.

- DEPURACIÓN DE LOS DATOS: Consiste en detectar aquellos datos que son erróneos, bien por errores en la cumplimentación del cuestionario, o bien por errores en inconsistencia de las respuestas.

- Caben dos opciones ante estos casos: se corrigen los datos “erróneos” , observando de nuevo el elemento o recogiendo otra vez el dato, o, directamente,.se desecha ese dato.

- LAS TABLAS son la disposición conjunta y ordenada de las sumas o totales de las respuestas de una pregunta o de varias relacionadas entre sí, obtenidas de la tabulación de los datos.

TRABAJOS PREVIOS ALA CODIFICACIÓN Y TABULACIÓN DE DATOS

- ***Edición interna***

- La edición interna investiga en forma rigurosa los resultados de la recopilación de datos, el departamento de investigación casi siempre cuenta con personal para llevar a cabo las funciones de edición y codificación.

- ***Edición para evitar inconsistencias***

  - La labor del editor interno es ajustar la respuesta inconsistentes o contradictorias, de modo de que no representen un problema para codificadores y capturistas.

    - Por ej.: suponga que a un entrevistador por teléfono se le indica que entreviste sólo a los electores empadronados que tengan por lo menos 18 años, la revisión por parte del editor puede indicar que el entrevistado tiene 17 años. En este caso la tarea del editor consiste en eliminar esta unidad de muestreo incorrecta, su tarea es aquí asegurar que la unidad de muestreo sea consistente con los objetivos del estudio.

RECETAS DE MARKETING

- El editor interno debe determinar si las respuestas que da un entrevistado a una pregunta son consistentes con las de otras preguntas relacionadas.

- Muchas encuestas emplean preguntas de filtro o se saltan preguntas que dirigen la secuencia de preguntas de acuerdo con las respuestas del entrevistado.

- ***Edición para verificar que las respuestas estén completas***
  - El entrevistado quizás responda sólo a la segunda parte de una pregunta dividida en dos.
  - La ***falta de respuesta a una partida*** es el término técnico que se emplea para asignar una pregunta no respondida en un cuestionario que de otra forma estaría completo.
  - Es preciso indicar las instrucciones del editor, las reglas de decisión específicas para solucionar el problema

- El editor se limita a indicar la falta de respuesta a una partida ofreciendo un mensaje que pida al codificador el registro del “valor faltante” o “espacio en blanco” como la respuesta.

- Cuando la relación entre dos preguntas es importante puede ser necesario que el editor agregue un “valor insertado”. La regla de decisión quizás sea insertar un valor promedio o neutral en cada caso de datos faltantes, o puede ser alternar la elección de las categorías de respuesta utilizadas como valores insertados o seleccionar una respuesta al azar.

- El editor debe decidir si un cuestionario completo es útil, cuando tiene muchas respuestas faltantes quizás no sea adecuado para el análisis de datos planeado, en tal caso el editor lo puede eliminar de la muestra.

- ***Edición de preguntas que se respondieron fuera de orden***
  - Otra tarea del editor es la de reordenar las respuestas de un cuestionario de preguntas abiertas.
  - Por ejemplo: Que el entrevistado ofrezca la respuesta a una pregunta subsecuente en sus comentarios a una pregunta abierta, puesto que el entrevistado identificó la respuesta es probable que el entrevistador haya evitado hacer la pregunta subsecuente para no tener que hacerle repetir la respuesta.

- ***<u>Como facilitar el proceso de codificación</u>***
  - Las actividades de edición ayudan a los codificadores, y se han diseñado varios procesos de edición específicos para facilitar el proceso de codificación. Las decisiones del editor no deben ser arbitrarias, sino basarse en un procedimiento sistemático de reglas fijas para tomar decisiones.

- ***Edición y tabulación de las respuestas "no sé":***
  - A simple vista esta respuesta parece indicar que el entrevistado no está muy familiarizado con la marca, producto o situación, o bien no está seguro y no se ha formado una opinión definida.
  - Esta respuesta legítima de ***"no sé"*** significa ***"sin opinión".***

- Pueden existir otras razones además de la legítima no sé, la respuesta ***"renuente no sé***", se da cuando el ***entrevistado*** simplemente ***no quiere responder a la pregunta*** y desea que el entrevistador deje de preguntarle.

- En algunas situaciones el editor puede separar las respuestas legítimas *"no sé"*, de las otras. Es probable que trate de identificar el significado de la respuesta de *"no sé"* a partir de los otros datos proporcionados en el cuestionario.

RECETAS DE MARKETING

- La tabulación de las respuestas *"no sé"*, requiere que el editor tome una decisión.

- Una alternativa es registrarlas en una categoría por separado, esto ofrece las categorías reales de respuesta, pero puede dar lugar a algunos problemas para calcular el porcentaje.

RECETAS DE MARKETING

- Otra es eliminarlas de la base porcentual.
- Una tercera consiste en distribuir las respuestas de tipo no sé entre las otras categorías, por lo general en forma proporcional

- ***Mecanica de la Edición.***

  - Los datos editados se escriben con un lápiz de color, cuando el espacio en el cuestionario lo permite, se dejan los datos originales para poder identificar los conceptos originales en una edición subsecuente.

- ***Prueba previa de la Edición:***
  - El análisis de las respuestas para la prueba previa puede identificar instrucciones inadecuadas o redacción incorrecta de las preguntas.

- ***Desventajas de la Edición: Subjetividad***:
  - Para hacer un trabajo de edición apropiado, el editor debe ser inteligente, tener experiencia y ser objetivo.
  - El analista de la investigación debe crear un procedimiento sistemático para evaluar los cuestionarios de manera que el editor cuente con reglas de decisión definidas

ELABORACIÓN DE UN CÓDIGO MAESTRO

- **Codificación**

  - ***Codificación:*** es el proceso de identificar y clasificar cada respuesta con una calificación numérica u otro símbolo.

  - La asignación de símbolos numéricos permite la transferencia de los datos de los cuestionarios a la computadora. Los códigos se definen como reglas para interpretar, clasificar y registrar los datos.

RECETAS DE MARKETING

- Los investigadores organizan los datos codificados en:
    - Campos: es un conjunto de caracteres que representa un solo tipo de dato.
    - Registros: es un conjunto de campos relacionados.
    - Archivos: es un conjunto de registros relacionados

- **Matriz de datos**
  - Es una disposición rectangular de datos en hileras y columnas.
    - Las ***hileras*** representan registros para casos individuales que reflejan las unidades fundamentales del análisis.
    - Cada ***columna*** representa un campo en particular.
  - La interacción de una hilera y una columna indica donde capturar un número u otro código asignado con respecto a una variable en particular.
  - La entrada directa de datos es el dispositivo de entrada más común para el almacenamiento de datos.

- ***Elaboración de códigos***

  - Cuando la pregunta es cerrada, la cantidad de categorías que requieren de un código, se determina durante la etapa de diseño del cuestionario.

  - Muchos programas reconocen que un campo en blanco o cierto símbolo indica un valor faltante.

  - Debe tomarse en cuenta la elección del programa de computación al seleccionar los códigos para los casos “sin respuesta”.

RECETAS DE MARKETING

- Existen dos reglas básicas para la creación de códigos:

  - Las categorías de codificación deben ser ***"exhaustivas",*** es decir, es preciso proporcionar categorías de codificación para todos los temas, objetos o respuestas. Quizás requiera la inclusión de una categoría de código "otros" para asegurarse que se incluyan todas las categorías.

  - Las categorías de codificación deben ser ***"mutuamente excluyentes"*** e "independientes" es decir, que no debe haber superposición entre las categorías para asegurar que un tema o respuesta pueda incluirse en una sola categoría.

- Cuando un cuestionario es muy estructurado, las categorías pueden codificarse antes de recopilar los datos.

- Cuando las preguntas suelen ser abiertas, no es posible establecer un marco de referencia para clasificar las respuestas antes de la recopilación de los datos.

- Esta situación requiere de decisiones cuidadosas sobre la determinación de las categorías después de realizar el proceso de edición y se conoce como “postcodificación” o simplemente “codificación”.

- ***Precodificacion de preguntas cerradas***
  - El entrevistador encierra en un círculo la calificación numérica codificada como respuesta a la pregunta.
  - La precodificación, puede realizarse si el investigador sabe cuales serán las categorías de respuestas antes de que tenga lugar la recopilación de datos.

RECETAS DE MARKETING

- Después de diseñar el cuestionario e identificar las respuestas estructuradas, la codificación se convierte en un proceso de rutina, en algunos casos las respuestas previamente determinadas, se basan en esquemas de clasificación estandarizados.

- Las entrevistas telefónicas asistidas por computadora, y las encuestas interactivas por computadora, requieren de la precodificación.

- Cambiar el marco de referencia de la codificación después de iniciar el proceso de entrevista es muy difícil porque requiere de cambios en los programas de computación.

- ***Codificación de preguntas abiertas***

  - Las preguntas abiertas se usan cuando el investigador no cuenta con una hipótesis clara acerca de las respuestas, que serán numerosas y muy variadas. El propósito de codificarlas es reducir la gran cantidad de respuestas individuales a unas cuantas categorías individuales a las que es posible asignar códigos numéricos.

  - Las respuestas similares deberán colocarse en una categoría general y tener el mismo código.

- La creación de códigos en estas situaciones debe reflejar el juicio del investigador. La creación de los códigos se basa en pensamientos y no sólo en palabras. El resultado final es una lista en forma abreviada y ordenada de todos los comentarios y pensamientos dados como respuesta a las preguntas.

RECETAS DE MARKETING

- Los investigadores llevan a cabo una tabulación de prueba para identificar las respuestas al pie de la letra de aproximadamente 20% de los cuestionarios terminados, posteriormente establecen categorías de códigos que reflejan el juicio de la persona que lo crea.

- La tabulación de prueba es una muestra pequeña de la cantidad total de respuestas a una pregunta en particular.

- Su propósito es la identificación preliminar de la estabilidad y distribución de las respuestas que determinaran como establecer un esquema de códigos.

RECETAS DE MARKETING

- Cada pensamiento separado se codifica una sola vez. Cuando un pensamiento esta compuesto por más de una palabra o frase sólo se codifica la palabra o frase más específica.

- Después de tabular las respuestas básicas, el investigador debe determinar cuantas categorías de respuesta serán aceptables.

- ***Elaboración del esquema de códigos***

  - Los esquemas de códigos no deben ser muy elaborados. La tarea del codificador consiste sólo en resumir los datos, después sigue la elaboración de un esquema de códigos. El investigador debe decidir si lo revisará y si los códigos son apropiados para dar respuesta a las preguntas de la gerencia. Un esquema preliminar con demasiadas categorías siempre puede compactarse o reducirse en un momento posterior durante el análisis.

  - Si los códigos iniciales son muy abstractos y sólo se establecen unas cuantas categorías, la revisión de los códigos para obtener afirmaciones más concretas será difícil, a menos que se registren los datos en bruto.

RECETAS DE MARKETING

- Los codificadores con experiencia agrupan las respuestas bajo encabezados generalizados que son pertinentes para la pregunta de investigación.
- Los códigos deben ser consistentes.
- Las categorías deben ser suficientemente ambiguas para que los codificadores no clasifiquen las partidas en forma diferente.

- ***Libro de códigos***
  - Este, identifica cada variable en el estudio y su ubicación en el disco suave u otro medio de entrada.
  - Con el libro de códigos, el investigador puede identificar la descripción, el nombre de código y el campo de cualquier variable.

- Se identifica a cada entrevistado con un número de identificación de cuestionario. Así los errores que se descubren durante el proceso de tabulación pueden buscarse en el cuestionario para verificar la respuesta.

- Cuando hay varias respuestas para una sola pregunta, los códigos deben distribuirse en varios campos debido a la posibilidad de respuestas múltiples.

- ***Codificación de producción.***

  - El proceso de transferir los datos o el formato de recopilación de datos después de recabarlos se conoce como codificación de producción.

  - Dependiendo de la forma de recopilación de datos los códigos pueden escribirse directamente en el instrumento o en una hoja de códigos especial.

  - La codificación debe llevarse a cabo en una oficina central, de modo que un supervisor pueda ayudar a resolver los problemas de interpretación.

- ***Edición y codificación*** combinadas:
  - Con frecuencia, la persona que codifica el cuestionario, lleva a cabo ciertas funciones de edición.
- ***Procesamiento de datos por computadora***:
  - La mayor parte de los estudios que tienen tamaños de muestra grandes utilizan una computadora para el procesamiento de datos.

- ***Captura de datos:*** es el proceso de transferir los datos a las computadoras.
  - Las entrevistas telefónicas asistidas por computadora, almacena y tabula automáticamente las respuestas conforme se recopilan, lo que reduce en gran medida los errores que ocurren durante el proceso de edición y codificación.
  - Para los cuestionarios muy estructurados puede utilizarse el sistema de rastreo óptico a fin de leer el material directamente de cuestionarios sensibles marcados.

RECETAS DE MARKETING

- Cuando no se utilizan los métodos anteriores, el procesamiento de datos empieza con la ***captura***. Se transfieren los datos codificados de los cuestionarios u hojas de códigos a un disco duro o suave.

- Para asegurar una exactitud del 100% en la transferencia de los códigos a la cinta un segundo capturista verifica el trabajo y revisa que los datos capturados sean precisos.

- Si se cometió un error, el verificador corrige la entrada de los datos. Este proceso de verificación de los datos nunca se lleva acabo por la misma persona que capturo los datos originales.

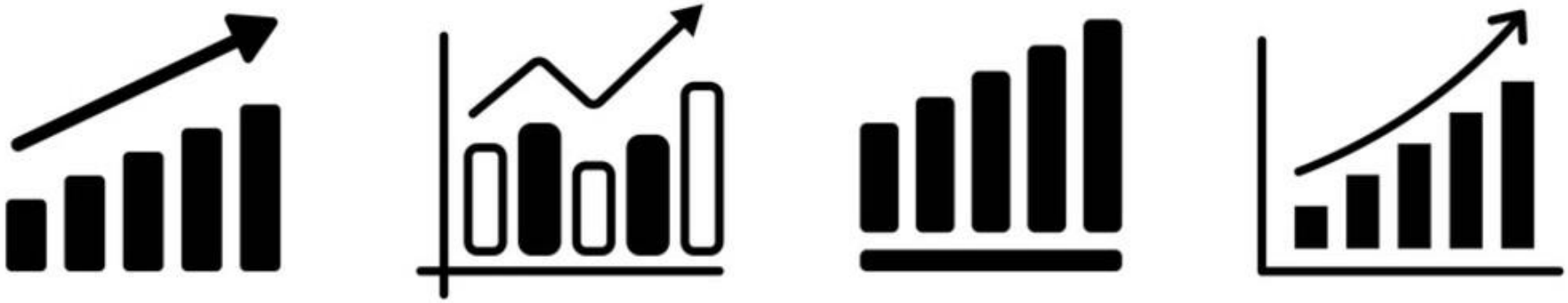
- Los capturistas prefieren tener los datos en hojas de codificación, esto casi siempre aumenta el tiempo y esfuerzo requeridos para la codificación. Realizarla con el cuestionario elimina la necesidad de transferir las respuestas a la hoja de códigos. Los recursos particulares del proyecto dictaran la fuente que debe utilizarse para la entrada de datos.

- **Recodificacion**

  - En muchas situaciones, es más fácil dar entrada a los datos en bruto en la computadora con el uso de la recodificación del cuestionario y después programar la computadora para recodificar algunos de ellos.

- ***Revisión de errores***

  - La última etapa en el proceso de codificación es la revisión y verificación de errores o depuración de datos, para asegurar que todos los códigos sean legítimos.



TABULACIÓN DE DATOS

- **Tabulación de datos**
  - El proceso de tabulación consiste en el recuento de los datos que están contenidos en los cuestionarios.
  - En este proceso incluimos todas aquellas operaciones encaminadas a la obtención de resultados numéricos relativos a los temas de estudio que se tratan en los cuestionarios.
  - Se requiere una previa codificación de las respuestas obtenidas en los cuestionarios.

RECETAS DE MARKETING

- Realizamos tabulación, codificación y diseño de gráficos con datos biográficos, de consumo o de opinión.

- Los resultados serán presentados en tablas y/o mapas gráficos que expliquen las relaciones existentes entre las diversas variables analizadas.

- Esta presentación se adecuará a la petición de nuestros clientes mediante análisis estadísticos de datos, grabados por nosotros o por terceros, análisis bivariantes, análisis multivariantes, tests de contraste de hipótesis

- Tras la recogida de información en una investigación, nos encontraremos con una serie de datos sobre diferentes variables de los individuos de una muestra.

- El primer paso para comenzar a analizar los datos es el de organizarlos de manera que podamos ver las características de los diferentes valores que han tomado las observaciones.

- El tipo de organización de los datos depende del tipo de variable. A continuación se describen las técnicas de organización más habituales para cada tipo de variable.

- **Variables cualitativas**

  - Las variables cualitativas (ordinales o no) toman valores no numéricos. Para tener una idea de los datos recogidos, pueden contabilizarse las diferentes ocurrencias de los distintos valores, dando lugar a los tres conceptos siguientes.

- **Frecuencia absoluta**

  - La frecuencia absoluta es el número de apariciones en una muestra de un determinado valor de una variable. Si la variable puede tomar valores en el conjunto , la frecuencia absoluta del valor o categoría i se denota por

  - La suma de las frecuencias de todos los valores de la variable equivale a una cuenta de los elementos en la muestra, por lo que esa suma debe ser igual al tamaño muestral.

$$\sum_{i=1}^{n} f_i = n$$

- **Frecuencia relativa**

  - La frecuencia relativa nos indica las ocurrencias de un valor determinado, pero no nos indica la proporción o tamaño relativo respecto a los otros valores. Para esto se utilizan las frecuencias relativas, que no son otra cosa que la relación entre la frecuencia absoluta de cada valor y el tamaño muestral n. Estas frecuencias se denotan como .

  - La suma de las frecuencias relativas de los valores de una muestra son *tantos por uno*, y en consecuencia su suma tiene que ser la unidad.

- **Porcentajes**
  - Otra forma de representar la misma información que proporciona la frecuencia relativa es utilizar porcentajes, es decir, *tantos por cien* en lugar de *tantos por uno*.
  - En este caso, la suma de los porcentajes deberá ser el 100%.

- **Variables cualitativas ordinales**
  - En las variables cualitativas ordinales se pueden utilizar los mismos conceptos de frecuencia absoluta, relativa o porcentajes, pero además se pueden utilizar tres conceptos relacionados que incluyen una acumulación.
  - La **frecuencia absoluta acumulada** para una variable con valores se define para cada valor i de la variable de la siguiente forma:

- Es decir, para el valor de posición i, se consideran (se suman) las frecuencias de todos los valores menores, además de la del propio valor.

- La frecuencia absoluta acumulada del mayor valor coincide con el tamaño muestral n, ya que se consideran las frecuencias de todos los valores.

- De manera similar, la **frecuencia relativa acumulada** para una variable con valores se define para cada valor i de la variable de la siguiente forma:

- Las frecuencias relativas acumuladas se pueden calcular alternativamente a partir de las frecuencias absolutas acumuladas:

- Este tipo de medidas no tiene una interpretación clara si no hay un orden en los valores de la variable.

- **Variables cuantitativas discretas**
  - Para las variables cuantitativas discretas se pueden utilizar las mismas técnicas de tabulación de las variables cualitativas ordinales.
  - No obstante, cuando el número de valores numéricos de la variable es muy grande, pueden agruparse las medidas en intervalos, tal y como se hace con las variables cuantitativas continuas, que se ven a continuación.

- **Variables cuantitativas continuas**

  - Aunque los datos de las variables cuantitativas continuas se pueden tabular al igual que en los casos anteriores, el gran número de valores posibles que suelen tener aconseja agruparlas en intervalos contiguos denominados **intervalos de clase**.

  - Cada intervalo de clase tiene un valor representativo, que es el valor intermedio, al que se denomina **marca de clase**.

  - Los intervalos deben cumplir las siguientes normas:

- Cada dato debe pertenecer solo a un intervalo.
- Todos los datos deben estar incluidos en algún intervalo.
  - NOTA: La tabulación en intervalos es útil para la representación y comprensión de los datos. No obstante, a la hora de hacer cálculos de medidas como medias o varianzas y en general para análisis de tipo matemático, deben utilizarse los datos originales, sin la tabulación.

GRACIAS.

RECETAS DE MARKETING.

Bulevar Vicente Blasco Ibáñez 19 5º E.

12.003 Castellón de la Plana.

Tel. 617997667.

direccion@recetasdemarketing.com

www.recetasdemarketing.com