

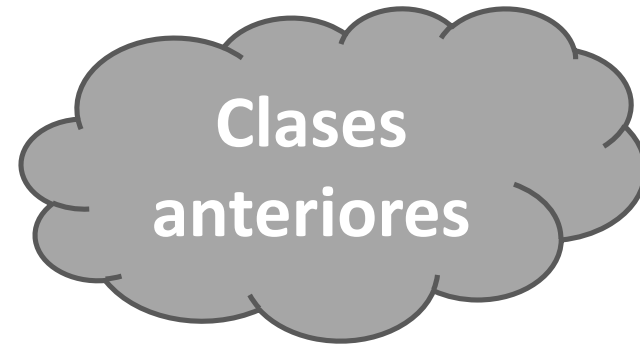


Universidad Nacional de Buenos Aires
Facultad de Ciencias Sociales
Licenciatura en Ciencias de la Comunicación
Metodología y Técnicas de la Investigación Social

Encuesta y Matriz de datos

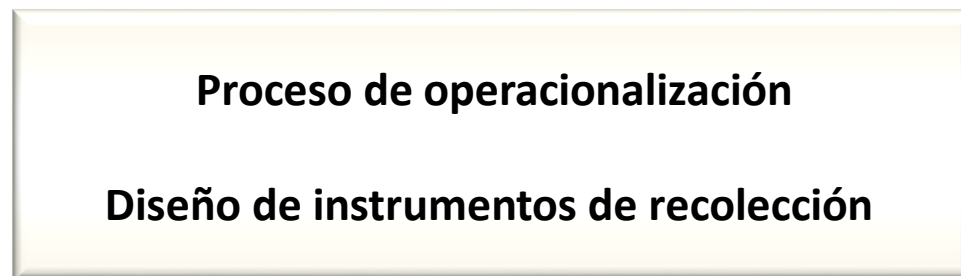
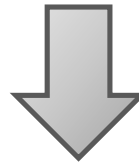
Dra. Jérica Pla

Hasta ahora....



Proceso de investigación «científico»: ir y venir entre el momento empírico o descriptivo y el momento teórico.

Cuestión central: ¿cómo se establece este tránsito entre el plano de los hechos y el de las ideas? (Samaja, 2004; Salvia, 1997).



El dato...

Resultado de un proceso de elaboración, en el que entran en juego teorías y conocimientos previos del objeto.



Nos alejamos de la definición «del latín datum», lo que se da.
Antecedente necesario para llegar al conocimiento exacto de una cosa o para deducir las consecuencias legítimas de un hecho.

Dato:

Evidencias empíricas que se encuentran en estrecho vínculo con las hipótesis teóricas que han orientado las prácticas de descubrimiento y validación.

No son la expresión de las cosas sino lo que hacemos visible de las cosas a partir de una mirada predeterminada (Salvia, s/f)



La encuesta...

Historia... a partir de mediados del siglo XX se Rice, Thorndike, Thustone, Likert comienzan a medir las actitudes y no ya a usar datos secundarios para el estudio demográfico....

Es un método de investigación basado en una serie de preguntas dirigidas a sujetos que constituyen una muestra representativa de una población con la finalidad de describir y/o relacionar características personales, permitiendo así generalizar las conclusiones

ENTONCES.... A diferencia de un censo, se aplica solo a una porción de la población.....

Diferentes tipo de muestreo.... Diferentes posibilidades de inferir resultados



¿Censo o encuesta?

Pensar...

- Limite de la población
- Recursos económicos
- Cantidad de variables a explorar
- Sesgos / errores



¿porque encuesta...?

Busca obtener mediciones cuantitativas de cualidades tanto objetivas como subjetivas de la población

NO MIDE SOLO CUESTIONES CUANTITATIVAS....

::::::::::> Recordar: Variables cualitativas y variables cuantitativas

*Por otro lado, Reflexionar sobre el proceso de aplicación....
Lenguaje, situación de administración, la encuesta como
momento de comunicación (Bourdieu)*



Establecer un instrumento estándar que garantice la aplicación de las mismas preguntas, planteadas de la misma forma, a todos lxs encuestadxs

Esto garantiza la agrupación y comparación de resultados entre los elementos de la muestra, y permite la inferencia estadística sobre los resultados



PREGUNTAS CERRADAS Dan la opción de elegir entre varias alternativas al entrevistado

Por ejemplo:

- ¿Miras habitualmente televisión de aire? SI /NO

En este caso la pregunta es por habitualmente.... Y es una pregunta dicotómica.

- ¿Cuántas veces a la semana miras televisión por aire?

- Todos los días⁰

- 5 o 6 días

- 3 o 4 días

- 2 días



- Menos de un día

- Todos los días

- Habitualmente

- Siempre

- Sólo cuando tengo ganas



Recordar que el sistema de categorías debe ser EXHAUSTIVO y MUTUAMENTE excluyente



PREGUNTAS ABIERTAS Permiten que le entrevistadx pueda expresarse de una manera mucho más amplia, completa, profunda.

Por ejemplo: ¿Por qué considera que el programa de televisión XXX no es adecuado para su horario?

La respuesta ha esto va ser muchas palabras que el cliente crea conveniente escribirlas para expresar su opinión.

PROBLEMA ¿Cómo medir cuantitativamente?

*Codificar
Establecer futuras preguntas cerradas.*

Principal fin: Exploratorio.

Análisis de
contenido
cuantitativo

Si tengo muchas preguntas abiertas quizás debería haber aplicado una estrategia cualitativa...





¿De donde salen las preguntas?

Recordar: Proceso de
operacionalización (proceso de
investigacion=)



La definición de los objetivos de investigación es el paso más importante de un diseño de investigación por encuesta y puede hacer que tu investigación sea un éxito o un fracaso.

Cada elemento (PREGUNTA) de la encuesta debe hacer referencia a este diseño.

La ***determinación de la muestra*** es el siguiente paso decisivo en la determinación de un diseño por encuesta y depende de muchos factores.

- **Precisión.** Deseas tratar de entrevistar a la más amplia base posible de personas.

Pero

La cantidad no siempre es la respuesta.

Si investigamos sobre un detergente quizás solo queremos encuestar a quienes lo utilizan pero para una encuesta política, sobre la que todo el mundo puede tener una opinión válida, nos conviene una muestra representativa



¿CÓMO HACER UN CUESTIONARIO?

1. Tener en cuenta los objetivos:

- ¿Qué se quiere averiguar? Explicitar los objetivos
- ¿Cuáles son los puntos clave que me debe aclarar la encuesta?
- Elaborar un temario de la información que se pretende obtener (contenidos obligatorios y opcionales)

2. Investigar sobre el tema:

- Acudir a las fuentes más cercanas (amigos, compañeros, familiares, etc.) o a especialistas e indagar: quién, cuánto, dónde, cómo, dónde, qué, cuál, con qué frecuencia, etc.
- Esto es necesario para elaborar alternativas útiles y válidas de respuesta en las preguntas

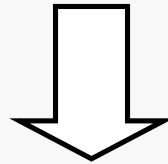
3. Redacción de las preguntas (fraseo):

- Considerar la forma en que entenderán las preguntas los encuestados ⇒ lenguaje adaptado al grupo objetivo del estudio



¿CÓMO HACER UN CUESTIONARIO?

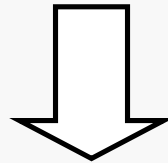
- Revisar ambigüedad y complejidad de las preguntas \Rightarrow ¿se entiende y de cuántas maneras distintas se entiende?
- ¿Puede ofender o molestar esta pregunta?
- ¿Cómo se plantea esta pregunta en una conversación normal?
- ***En el caso de elevados tamaños de muestra los encuestadores no pueden dar aclaraciones para que se entiendan las preguntas, por lo que cada una de ellas debe ser perfectamente clara***



ii Hay que probar el instrumento diseñado !!



El paso al análisis de datos: otra etapa del proceso de investigación



Dato



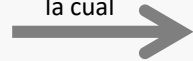
Un **dato** es una construcción compleja que, por consecuencia, posee una **estructura interna (tripartita)**, la cual es su **contenido formal invariable**, es decir, está presente en todo dato.



se organiza en

Matriz de datos

la cual



hace visible esa estructura y el marco analítico del investigador (Ragin, 2007)

Sintetizando...

En el proceso de creación de la matriz de datos se hace visible :

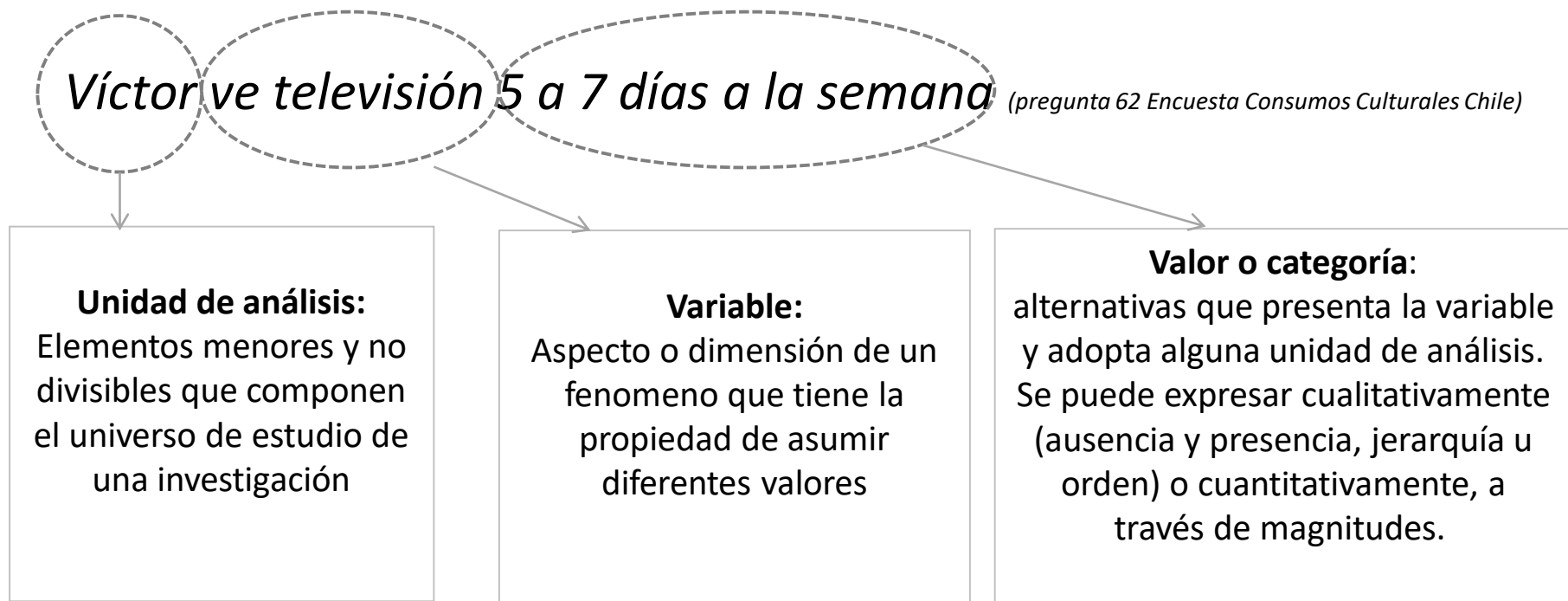
- ✓ *el proceso de la ciencia (Samaja (2004) y ...*
- ✓ *las operaciones teórico / metodológicas involucradas en todo proceso de investigación (Di Virgilio, 2006 – 2007)*



La estructura tripartita del dato

3 componentes del dato

Unidad de análisis – Variable - Valor (o categoría)



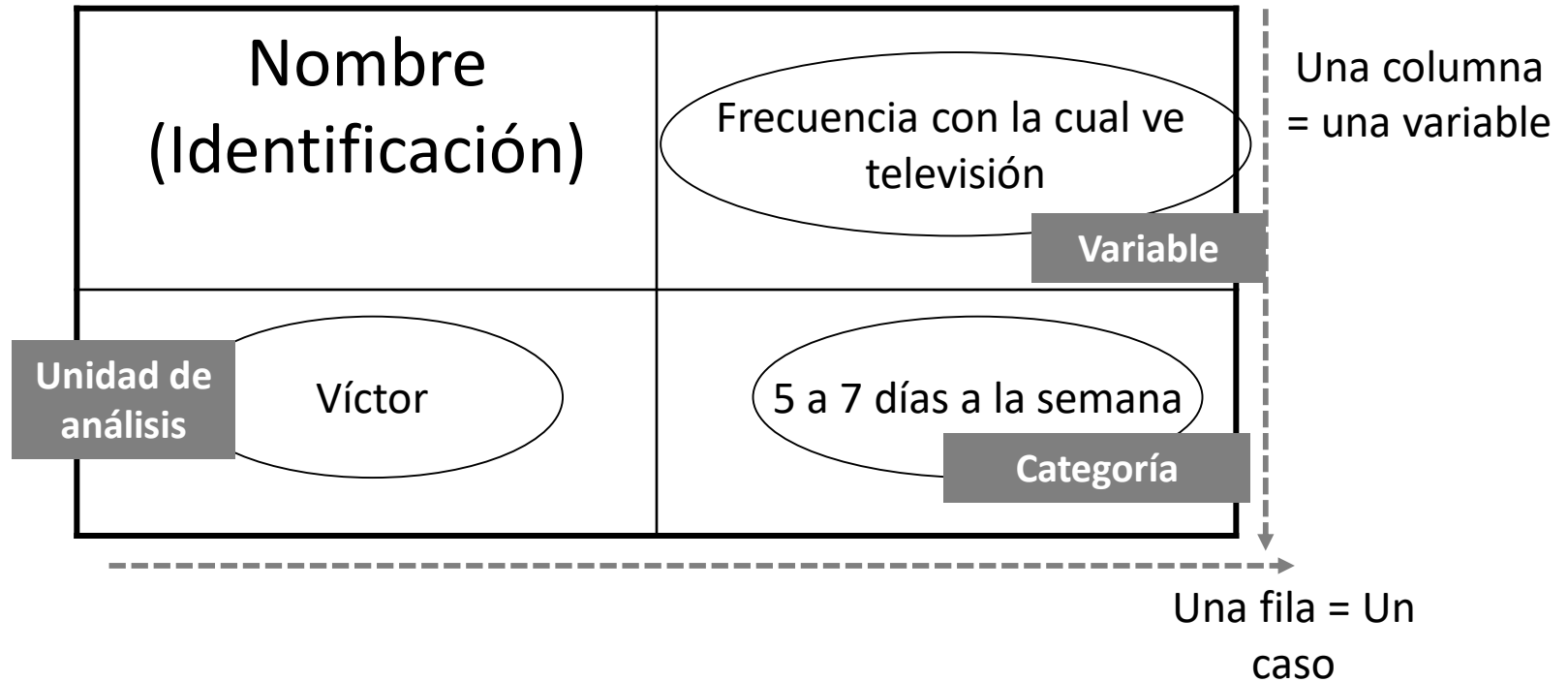
Tres elementos que constituyen un dato a partir de las relaciones que se establecen entre sí.

A su vez

La matriz de datos



La matriz de datos hace visible la estructura del dato



La matriz tendrá tantos casos como casos tenga la muestra y tantas variables como preguntas tenga la encuesta (o ítems del instrumento de registro). En cada celda aparece el valor que dicha unidad de análisis adquiere en dicha variable

La matriz de datos: codificación I

“Frecuencia con la cual ve televisión” ::::> Variable de naturaleza cualitativa, ordinal

Codificada “A priori”:

1. 5 a 7 días a la semana
2. 4 días a la semana
3. 3 días a la semana
4. 2 días a la semana
5. 1 días a la semana
6. Nunca

En las celdas de la matriz de datos, para el caso de variables cualitativas se «cargan» los códigos de las categorías (hoy ya no tan asi....)

Nombre (Identificación)	Frecuencia con la cual ve televisión
Victor	1

Analisis de
contenido
cuantitativo

Las categorías pueden ser codificadas, antes, durante o después de efectuar el trabajo de campo

La matriz de datos: codificación II

La codificación **no** es necesaria en el caso de las variables de naturaleza cuantitativa, pues se usa el numero que toma la forma de valor en determinada variable, por ejemplo...

“Edad en años cumplidos” :::> Variable de naturaleza cuantitativa

Víctor tiene 35 años (Módulo RPH Encuesta Consumos Culturales Chile)

Nombre (Identificación)	Edad en años cumplidos
Victor	35

Luego se las puede reagrupar en categorías... objetivos de investigación

La matriz de datos: armado

Una vez realizada la codificación de los datos, se puede proceder a la confección de la matriz de datos

Edad y Módulo “Revista” Encuesta de consumos culturales (preguntas 53 a 55 inclusive)

Identificación	Edad en años cumplidos	P53 ¿Con qué frecuencia diría usted que lee revistas?	P54 ¿De qué manera accede a las revistas que lee habitualmente?	P55_1 ¿Qué tipo de revistas acostumbra usted a leer? Respuesta 1	P55_2 ¿Qué tipo de revistas acostumbra usted a leer? Respuesta 2
Emanuel	28	4	2	3	1
Sabrina	39	1	3	4	10
Karina	35	7			
Esteban	56	6	1	11	

- **Esteban tiene 56 años, lee revistas una vez al mes (código 6), compra (código 1) y acostumbra a leer revistas de cine (código 11)**
- **Karina tiene 35 años y no lee revistas (código 7 nunca)**

La evaluación de la matriz de datos I

- ✓ Principio de comparabilidad
- ✓ Principio de clasificación
- ✓ Principio de integridad

Tres principios fundamentales que se derivan lógicamente de las nociones de “estructura tripartita del dato” y “matriz de datos”

Galtung (1966) – Barenger (2009)

Principio de comparabilidad

Todas las unidades de análisis han de ser medidas en las mismas variables y cada una de las combinaciones de una variable determinada con una unidad de análisis debe *tener sentido*. Es decir debe ser verdadero o falso que una combinación de unidad de análisis – variable presenta un valor determinado.

Si la unidad de análisis es una nación o país, la variable «P64 Disponibilidad de tiempo libre», y sus categorías «mucho tiempo libre» «poco tiempo libre» «suficiente tiempo libre», cualquier combinación no sería ni verdadera ni falsa, simplemente no tendría sentido.

La evaluación de la matriz de datos II

Principio de clasificación:

Para cada par unidad de análisis – variable, deberá haber un valor posible (exhaustividad) y sólo uno (exclusión mutua) en el cual deba ser ubicado.

Si lógicamente se supone que habrán distintos valores para una misma unidad, en la matriz deberán existir tantas “columnas” como clasificaciones se pretendan hacer de las unidades.

P53 ¿Con qué frecuencia diría usted que lee revistas?

1. Todos los días
2. Casi todos los días
3. 2 – 3 veces por semana
4. Una vez a la semana
5. Cada dos semanas
6. Una vez al mes
7. Ocasionalmente
8. No / Nunca

1. P55 ¿Qué tipo de revistas acostumbra usted a leer?
2. Actualidad
3. Hogar / Jardinería / diseño / decoración
4. Deportes
5. ...
6. Geografía
7. Religiosas

Leer cada dos semanas (categoría 5) puede ser considerado “ocasionalmente” (categoría 7). Debe revisarse y también revisarse que no se marquen más de una opción.

En el cuestionario se señala que pueden elegir dos opciones, por lo tanto se deben incorporar dos columnas

La evaluación de la matriz de datos III

Principio de integridad:

Para cada par unidad de análisis – variable debe hallarse empíricamente un valor. En términos prácticos no deben quedar celdas vacías, o al menos debe intentarse que estas sean las menos.

Con este principio se evalúa el trabajo empírico de producción del dato y del “llenado de la matriz”, y no el trabajo lógico de definición de su estructura.

P62: ¿Usted ve televisión?

5 a 7 días a la semana

1. 4 días a la semana
2. 3 días a la semana
3. 2 días a la semana
4. 1 días a la semana
5. Nunca

En caso de no haber ninguna respuesta marcada, debe tomarse una decisión, por ejemplo codificar ese valor “perdido” con un código (por ejemplo 999: No responde) (Hérrnandez Sampieri; Férnandez Collado; Baptista Lucio, 2006: 392)

Formas de examinar la matriz de datos I

Por fila

Se examina para cada unidad de análisis que porcentaje de celdas están vacías (se puede llegar a decidir la eliminación total del caso)

Por columna

Se examina para cada variable el porcentaje de valores perdidos. Puede concluirse la eliminación de la variable de los análisis

Un alto numero de casos perdidos (más del 10%) indica que el instrumento tiene problemas. Un grado de «admisibilidad máxima» (Galtung, 1965) sería de menos del 5%.

Los valores perdidos pueden reducirse con instrumentos que motiven al participante y no sean muy largos, con instrucciones claras y capacitación a los encuestadores.

Formas de examinar la matriz de datos I

Completada la matriz de datos, y con anterioridad al procesamiento y análisis de los datos, el investigador debe tomar decisiones explícitas sobre los casos “sin datos”.

Principio de ignorabilidad fuerte

Tratar la ausencia de información como una situación aleatoria en la producción del dato (no es problema sistemático, ni de operacionalización, y los errores se compensan entre si)

Principio de ignorabilidad débil

Se supone que la variable no cumplía con ser exhaustiva y se arma un nuevo código para la categoría residual (se supone que ésta no altera la operacionalización de conceptos).

Formas de examinar la matriz de datos II

Razón sistemática teóricamente sustantiva

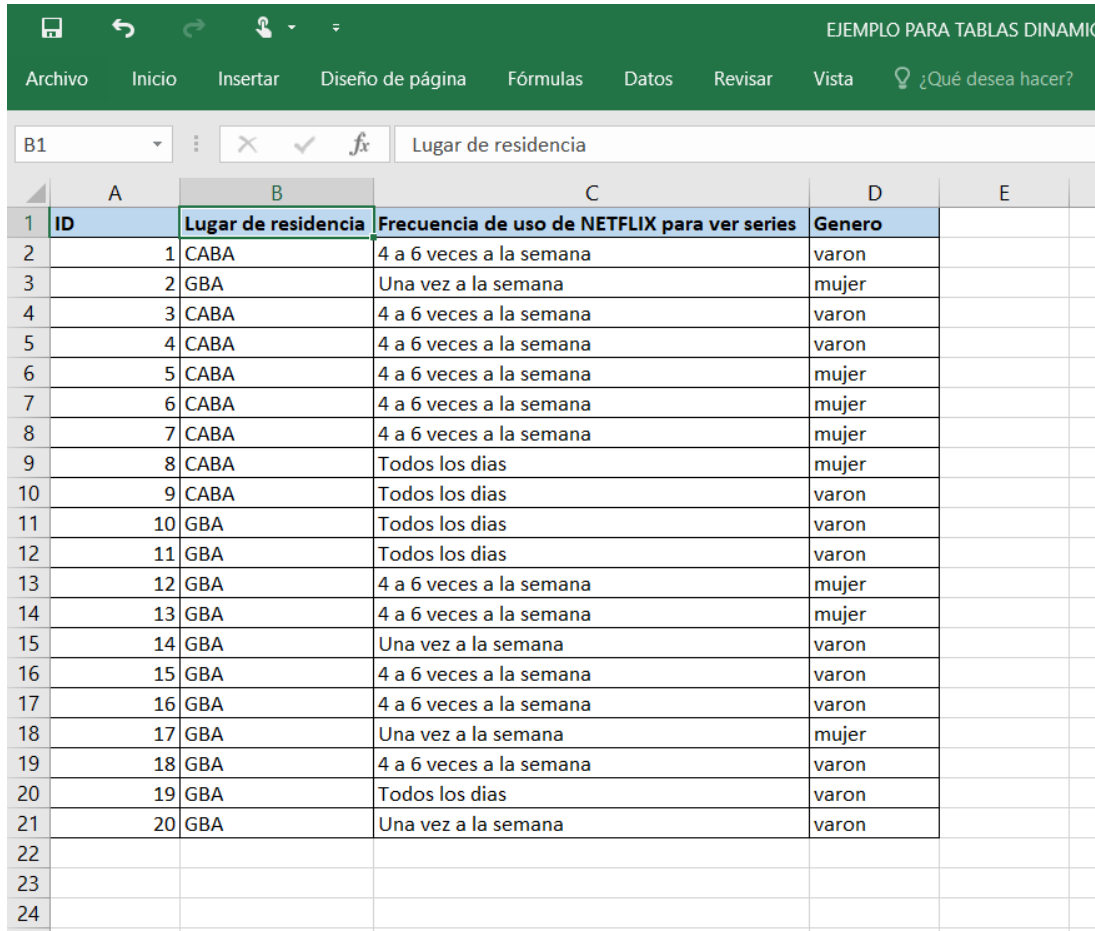
Si la ausencia de información se concentra en muchas variables para un mismo conjunto de unidades, se supone que existió dicha razón que la produjo.

Principio de intrapolación de datos - Imputación de datos faltantes

Si la ausencia de información para una variable tiene una frecuencia importante (mayor al 10%) y se presenta sin ningún patrón de regularidad para un conjunto específico de unidades. Se identifica el patrón mediante un modelo estadístico, que debe ser satisfactorio.

El análisis de los datos: otra etapa del proceso de investigación

La base de datos: la matriz



EJEMPLO PARA TABLAS DINAMICAS

Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista ¿Qué desea hacer?

B1 X ✓ fx Lugar de residencia

	A	B	C	D	E
1	ID	Lugar de residencia	Frecuencia de uso de NETFLIX para ver series	Genero	
2	1	CABA	4 a 6 veces a la semana	varon	
3	2	GBA	Una vez a la semana	mujer	
4	3	CABA	4 a 6 veces a la semana	varon	
5	4	CABA	4 a 6 veces a la semana	varon	
6	5	CABA	4 a 6 veces a la semana	mujer	
7	6	CABA	4 a 6 veces a la semana	mujer	
8	7	CABA	4 a 6 veces a la semana	mujer	
9	8	CABA	Todos los dias	mujer	
10	9	CABA	Todos los dias	varon	
11	10	GBA	Todos los dias	varon	
12	11	GBA	Todos los dias	varon	
13	12	GBA	4 a 6 veces a la semana	mujer	
14	13	GBA	4 a 6 veces a la semana	mujer	
15	14	GBA	Una vez a la semana	varon	
16	15	GBA	4 a 6 veces a la semana	varon	
17	16	GBA	4 a 6 veces a la semana	varon	
18	17	GBA	Una vez a la semana	mujer	
19	18	GBA	4 a 6 veces a la semana	varon	
20	19	GBA	Todos los dias	varon	
21	20	GBA	Una vez a la semana	varon	
22					
23					
24					

El Excel asume la forma de matriz de datos

En este caso no usó códigos (les puede pasar a ustedes

Matriz de datos y distribución de frecuencias

EJEMPLO PARA TABLAS DINAMICAS - Excel							
Compartir							
F20							
	A	B	C	D	E	F	G
1	ID	Lugar de residencia	Frecuencia de uso de NETFLIX para ver series	Genero			
2	1	CABA	4 a 6 veces a la semana	varon			
3	2	GBA	Una vez a la semana	mujer			
4	3	CABA	4 a 6 veces a la semana	varon			
5	4	CABA	4 a 6 veces a la semana	varon			
6	5	CABA	4 a 6 veces a la semana	mujer			
7	6	CABA	4 a 6 veces a la semana	mujer			
8	7	CABA	4 a 6 veces a la semana	mujer			
9	8	CABA	Todos los días	mujer			
10	9	CABA	Todos los días	varon			
11	10	GBA	Todos los días	varon			
12	11	GBA	Todos los días	varon			
13	12	GBA	4 a 6 veces a la semana	mujer			
14	13	GBA	4 a 6 veces a la semana	mujer			
15	14	GBA	Una vez a la semana	varon			
16	15	GBA	4 a 6 veces a la semana	varon			
17	16	GBA	4 a 6 veces a la semana	varon			
18	17	GBA	Una vez a la semana	mujer			
19	18	GBA	4 a 6 veces a la semana	varon			
20	19	GBA	Todos los días	varon			
21	20	GBA	Una vez a la semana	varon			
22							
23							
24							

Etiquetas de fila		Cuenta de Frecuencia de uso de NETFLIX para ver series	
4 a 6 veces a la semana		11	
Todos los días		5	
Una vez a la semana		4	
Total general		20	

Etiquetas de fila		Cuenta de Frecuencia de uso de NETFLIX para ver series	
4 a 6 veces a la semana		55,0%	
Todos los días		25,0%	
Una vez a la semana		20,0%	
Total general		100,0%	

Diferencias entre matriz y tabla de contingencia: frecuencias absolutas

EJEMPLO PARA TABLAS DINAMICAS - Excel

Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista ¿Qué desea hacer? Compartir

F18

	B	C	D	E	F	G	H	I	J	K
1	Lugar de residencia	Frecuencia de uso de NETFLIX para ver series	Genero							
2	CABA	4 a 6 veces a la semana	varon							
3	GBA	Una vez a la semana	mujer							
4	CABA	4 a 6 veces a la semana	varon							
5	CABA	4 a 6 veces a la semana	varon							
6	CABA	4 a 6 veces a la semana	mujer							
7	CABA	4 a 6 veces a la semana	mujer							
8	CABA	4 a 6 veces a la semana	mujer							
9	CABA	Todos los días	mujer							
10	CABA	Todos los días	varon							
11	GBA	Todos los días	varon							
12	GBA	Todos los días	varon							
13	GBA	4 a 6 veces a la semana	mujer							
14	GBA	4 a 6 veces a la semana	mujer							
15	GBA	Una vez a la semana	varon							
16	GBA	4 a 6 veces a la semana	varon							
17	GBA	4 a 6 veces a la semana	varon							
18	GBA	Una vez a la semana	mujer							
19	GBA	4 a 6 veces a la semana	varon							
20	GBA	Todos los días	varon							
21	GBA	Una vez a la semana	varon							
22										

Cuenta de Genero Etiqueta

Etiquetas de fila mujer varon Total general

4 a 6 veces a la semana	5	6	11
Todos los días	1	4	5
Una vez a la semana	2	2	4
Total general	8	12	20

Diferencias entre matriz y tabla de contingencia: frecuencias relativas

EJEMPLO PARA TABLAS DINAMICAS - Excel					Herramientas de tabla dinámica				
Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista Analizar Diseño					¿Qué desea hacer?				
G18									
	A	B	C	D	E	F	G	H	I
1	ID	Lugar de residencia	Frecuencia de uso de NETFLIX para ver series	Genero					
2	1	CABA	4 a 6 veces a la semana	varon					
3	2	GBA	Una vez a la semana	mujer					
4	3	CABA	4 a 6 veces a la semana	varon					
5	4	CABA	4 a 6 veces a la semana	varon					
6	5	CABA	4 a 6 veces a la semana	mujer					
7	6	CABA	4 a 6 veces a la semana	mujer					
8	7	CABA	4 a 6 veces a la semana	mujer					
9	8	CABA	Todos los días	mujer					
10	9	CABA	Todos los días	varon					
11	10	GBA	Todos los días	varon					
12	11	GBA	Todos los días	varon					
13	12	GBA	4 a 6 veces a la semana	mujer					
14	13	GBA	4 a 6 veces a la semana	mujer					
15	14	GBA	Una vez a la semana	varon					
16	15	GBA	4 a 6 veces a la semana	varon					
17	16	GBA	4 a 6 veces a la semana	varon					
18	17	GBA	Una vez a la semana	mujer					
19	18	GBA	4 a 6 veces a la semana	varon					
20	19	GBA	Todos los días	varon					
21	20	GBA	Una vez a la semana	varon					
22									
23									
24									

Cuenta de Genero		Etiquetas de columna		
Etiquetas de fila		mujer	varon	Total general
4 a 6 veces a la semana		5	6	11
Todos los días		1	4	5
Una vez a la semana		2	2	4
Total general		8	12	20

Cuenta de Genero		Etiquetas de columna		
Etiquetas de fila		mujer	varon	Total general
4 a 6 veces a la semana		62,5%	50,0%	55,0%
Todos los días		12,5%	33,3%	25,0%
Una vez a la semana		25,0%	16,7%	20,0%
Total general		100,0%	100,0%	100,0%

De un total de 8 mujeres, 5 usan NETFLIX para ver series en una frecuencia de 4 a 6 veces por semana, lo que equivale al 62,5% del total de las mujeres

$$8 = 100\%$$

$$5 = X (5 * 100 / 8)$$

$$X = 62,5\%$$

¡Gracias!

Dra. Jérica Pla

