

Curso de Estadística Aplicada a las Ciencias Sociales

Tema 4. Tablas y gráficos

Capítulos 2 y 3 del manual

Tema 4. Tablas y gráficos

Introducción

1. Distribución de frecuencias
2. Representaciones gráficas

Resumen

Introducción

- Los datos “brutos” no son útiles

- Ejemplo: sexo y edad

- ◆ Edad:

29, 22, 21, 20, 20, 24, 21, 20, 23, 22, 26, 20, 21, 23,
22, 22, 23, 20, 23, 24

- ◆ Sexo:

V, M, V, V, M, V, M, V, M, M, M, V, V, M, V, V, M, M,
V, M.

Introducción (2)

Ejemplo del fichero HOGARES.XLS (en diskette del manual)

| Nº | TMUN | CAPITAS | L14 | SOC | SITPROF | EDC | NPER | M2V | IT |
|----|------|---------|-----|-----|---------|-----|------|-----|--------|
| 1 | 4 | 4 | 2 | 3 | 1 | 3 | 1 | 0 | 90000 |
| 2 | 4 | 2 | 0 | 7 | 2 | 2 | 1 | 54 | 62091 |
| 3 | 4 | 6 | 4 | 3 | 2 | 2 | 1 | 94 | 113379 |
| 4 | 4 | 4 | 2 | 5 | 2 | 4 | 2 | 90 | 137168 |
| 5 | 4 | 3 | 1 | 3 | 2 | 3 | 1 | 82 | 156860 |
| 6 | 4 | 5 | 0 | 5 | 2 | 1 | 1 | 100 | 213841 |
| 7 | 2 | 4 | 1 | 1 | 4 | 2 | 2 | 75 | 146374 |
| 8 | 4 | 6 | 3 | 5 | 1 | 7 | 2 | 0 | 600000 |
| 9 | 4 | 1 | 0 | 7 | 1 | 1 | 1 | 0 | 48000 |
| 10 | 4 | 5 | 3 | 5 | 1 | 3 | 1 | 0 | 212800 |

Introducción (3)

- Esto “no nos sirve para nada” en “bruto”
- Necesitamos “resumir” los datos para que sean comprensibles
- Esto es lo que hace la estadística descriptiva
- En este tema: resumir con tablas (distribuciones de frecuencias) y gráficos

1. Distribución de frecuencias

- Llamamos **distribución de frecuencias** a una tabla donde se resumen los datos de una variable
- Es una manera “sencilla” de agrupar mucha información en unos pocos datos comprensibles
- Contenido básico: ¿qué valores toma la variable, y qué proporción de los sujetos tiene cada valor?
- Lo hacemos de manera diferente para variables discretas y continuas

Tema 4. Tablas y gráficos

Introducción

1. Distribución de frecuencias

1.1. Variables discretas

1.2. Variables discretas con EXCEL

1.3. Variables continuas

1.4. Variables continuas con EXCEL

2. Representaciones gráficas

Resumen

1.1. Distribución de frecuencias para variables discretas

■ Recordatorio variables discretas son:

- ◆ Todas las cualitativas (región, ciudad, partido al que vota, sentimiento religioso, ve o no ve la tele...)
- ◆ Todas las de escala ordinal: opinión sobre algo (muy buena, buena, regular, mala, muy mala)
- ◆ Algunas variables cuantitativas: las que tienen pocos valores (número de personas en el hogar, número de partidos políticos en el ayuntamiento...)

1.1. Distribución de frecuencias para variables discretas

- Tenemos una variable con un número N de observaciones (casos)
- La variable tiene una serie de valores diferentes
- **Frecuencia absoluta** de un valor: el número de sujetos que tienen ese valor (n_i)
- **Frecuencia relativa** de un valor: la proporción que representan los sujetos que tienen ese valor sobre el total de los sujetos de la muestra o población ($f_i = n_i / N$)
- Distribución de frecuencias: tabla con clases, frecuencias absolutas y relativas

1.1. Distribución de frecuencias para variables discretas

- Ejemplo con una variable cualitativa (escala nominal): tipo de accidente en el que murieron los fallecidos en accidente de tráfico en carretera (no ciudades), en España en 2004

| Tipos de accidente | Muertos (n_i) | Frecuencia relativa (f_i) |
|----------------------------------|-------------------|-------------------------------|
| Salida de la vía | 1466 | 0,382 |
| Colisión frontal | 693 | 0,180 |
| Colisión lateral y frontolateral | 826 | 0,215 |
| Colisión trasera y múltiple | 303 | 0,079 |
| Atropello a peatón | 319 | 0,083 |
| Vuelco | 70 | 0,018 |
| Otro tipo de accidente | 164 | 0,043 |
| Suma | 3841 | 1,000 |

- Frecuencias relativas suman 1

1.1. Distribución de frecuencias para variables discretas

- Ejemplo con una variable de escala ordinal: opinión sobre los impuestos sobre el tabaco que tienen los entrevistados en una encuesta

| Los impuestos actuales sobre el tabaco diría Ud. que son... | | |
|---|-------------------------|-------------------------------|
| | Entrevistados (n_i) | Frecuencia relativa (f_i) |
| Excesivos | 966 | 0,578 |
| Adecuados | 332 | 0,199 |
| Bajos | 372 | 0,223 |
| Total | 1.670 | 1 |

1.1. Distribución de frecuencias para variables discretas

- Ejemplo con una variable cuantitativa discreta (CAPITAS en el fichero HOGARES)
- Puede incluir las **frecuencias acumuladas**

| Variable CAPITAS (número miembros hogar) | | | | | |
|--|---------------------|---------------------|------------------------|---------------------|--|
| Valor | Frecuencias | | Frecuencias acumuladas | | |
| | absolutas (n_i) | relativas (f_i) | Absolutas (N_i) | Relativas (F_i) | |
| 1 | 6 | 0,08 | 6 | 0,08 | |
| 2 | 11 | 0,15 | 17 | 0,23 | |
| 3 | 11 | 0,15 | 28 | 0,37 | |
| 4 | 20 | 0,27 | 48 | 0,64 | |
| 5 | 15 | 0,20 | 63 | 0,84 | |
| 6 | 8 | 0,11 | 71 | 0,95 | |
| 7 | 3 | 0,04 | 74 | 0,99 | |
| 8 | 0 | 0,00 | 74 | 0,99 | |
| 9 | 1 | 0,01 | 75 | 1,00 | |
| N | 75 | 1 | | | |

1.1. Distribución de frecuencias para variables discretas

- Frecuencias relativas acumuladas permiten calcular:

- ◆ Frecuencia relativa valores menores que x
- ◆ Frecuencia relativa valores mayores que x ($1-F_i$)
- ◆ Frecuencia relativa entre dos valores (F_i-F_j)

| Variable CAPITAS (número miembros hogar) | | | | |
|--|---------------------|---------------------|------------------------|---------------------|
| Clase | Frecuencias | | Frecuencias acumuladas | |
| | absolutas (n_i) | relativas (f_i) | Absolutas (N_i) | Relativas (F_i) |
| 1 | 6 | 0,08 | 6 | 0,08 |
| 2 | 11 | 0,15 | 17 | 0,23 |
| 3 | 11 | 0,15 | 28 | 0,37 |
| 4 | 20 | 0,27 | 48 | 0,64 |
| 5 | 15 | 0,20 | 63 | 0,84 |
| 6 | 8 | 0,11 | 71 | 0,95 |
| 7 | 3 | 0,04 | 74 | 0,99 |
| 8 | 0 | 0,00 | 74 | 0,99 |
| 9 | 1 | 0,01 | 75 | 1,00 |
| | 75 | 1 | | |

1.2. Distribución de frecuencias para variables discretas: con EXCEL

- Abrimos fichero municipios.xls
- Tomamos nota del rango de valores que nos interesa. Por ejemplo: celdas donde están las provincias (c2:c285) (O le damos nombre de rango)

| | A | B | C | |
|----|--------------------|------------|-------------|-----|
| 1 | Nombre Municipio | Código INE | Provincia | Tot |
| 2 | Abenójar | 13001 | Ciudad Real | |
| 3 | Agudo | 13002 | Ciudad Real | |
| 4 | Ajofrín | 45001 | Toledo | |
| 5 | Alameda de la Sa | 45002 | Toledo | |
| 6 | Albacete | 2003 | Albacete | |
| 7 | Albaladejo | 13004 | Ciudad Real | |
| 8 | Alberca de Záncar | 16007 | Cuenca | |
| 9 | Alcalá del Júcar | 2007 | Albacete | |
| 10 | Alcaraz | 2008 | Albacete | |
| 11 | Alcaudete de la Ja | 45006 | Toledo | |
| 12 | Alcázar de San Ju | 13005 | Ciudad Real | |
| 13 | Alcolea de Calatra | 13007 | Ciudad Real | |
| 14 | Aldea del Rey | 13009 | Ciudad Real | |
| 15 | Alhambra | 13010 | Ciudad Real | |
| 16 | Almadén | 13011 | Ciudad Real | |
| 17 | Almarcha | 13013 | Ciudad Real | |

1.2. Distribución de frecuencias para variables discretas: con EXCEL

- Vamos a una zona "libre" de la hoja de cálculo
- Escribimos en columna los nombres de las provincias
- A la derecha del primer valor escribimos la función: =CONTAR.SI(RANGO;VALORES)
- En este caso =CONTAR.SI(c\$2:c\$285;c288)

| | A | B | C | D | E | F |
|-----|-------------------|-------|-------------|------------------------------|------|---|
| 282 | Yeste | 2086 | Albacete | 3781 | 1922 | |
| 283 | Yuncler | 45203 | Toledo | 2287 | 1160 | |
| 284 | Yuncos | 45205 | Toledo | 4776 | 2517 | |
| 285 | Yunquera de Henao | 19331 | Guadalajara | 2147 | 1079 | |
| 286 | | | | | | |
| 287 | | | | | | |
| 288 | | | Albacete | =CONTAR.SI(C\$2:C\$285;C288) | | |
| 289 | | | Ciudad Real | | | |
| 290 | | | Cuenca | | | |
| 291 | | | Guadalajara | | | |
| 292 | | | Toledo | | | |
| 293 | | | | | | |

15

1.2. Distribución de frecuencias para variables discretas: con EXCEL

- Copiamos la fórmula en las celdas inferiores
- El resultado: una tabla con los valores y sus frecuencias absolutas

| | A | B | C | D |
|-----|-------------------|-------|-------------|------|
| 285 | Yunquera de Henao | 19331 | Guadalajara | 2147 |
| 286 | | | | |
| 287 | | | | |
| 288 | | | Albacete | 47 |
| 289 | | | Ciudad Real | 64 |
| 290 | | | Cuenca | 36 |
| 291 | | | Guadalajara | 24 |
| 292 | | | Toledo | 113 |
| 293 | | | | |

1.2. Distribución de frecuencias para variables discretas: con EXCEL

- Frecuencias relativas:
 - Calcular N bajo columna de frec. Absolutas con =SUMA(rango)
 - Calcular cada frecuencia relativa: n_i / N
 - Truco para hacerlo más rápido: copiar fórmula con \$

| | D | E | F | G |
|-----|---|-------------|------------------|--------------|
| 289 | | | | |
| 290 | | Cuenca | 36 | =f290/f\$295 |
| 291 | | Ciudad Real | 63 | =f291/f\$295 |
| 292 | | Toledo | 113 | =f292/f\$295 |
| 293 | | Albacete | 47 | =f293/f\$295 |
| 294 | | Guadalajara | 24 | =f294/f\$295 |
| 295 | | Total | =suma(f290:f294) | =f295/f\$295 |

17

1.2. Distribución de frecuencias para variables discretas: con EXCEL

Ya tenemos nuestra distribución de frecuencias!!!

| | | |
|-------------|-----|-------|
| Cuenca | 36 | 0,127 |
| Ciudad Real | 63 | 0,223 |
| Toledo | 113 | 0,399 |
| Albacete | 47 | 0,166 |
| Guadalajara | 24 | 0,085 |
| Total | 283 | 1,000 |

- Podemos cambiar aspectos "estéticos"
- Podemos añadir rótulos explicativos...

| Municipios de Castilla-La Mancha de más de 1000 habitantes | | |
|--|--------|---------------------|
| | Número | Frecuencia relativa |
| Cuenca | 36 | 0,127 |
| Ciudad Real | 63 | 0,223 |
| Toledo | 113 | 0,399 |
| Albacete | 47 | 0,166 |
| Guadalajara | 24 | 0,085 |
| Total | 283 | 1,000 |

1.2. Distribución de frecuencias para variables discretas: con EXCEL

Con variables de escala ordinal o de intervalo, podemos añadir las frecuencias acumuladas

En columnas adyacentes copiamos fórmulas como las siguientes

| | E | F | G | H | I |
|----|---------|------------------|-------------|------------|-----------|
| 1 | | Variable CAPITAS | | | |
| 2 | | | | Acumuladas | |
| 3 | Valores | Absolutas | Relativas | Absolutas | Relativas |
| 4 | 1 | 6 | =+F4/F\$13 | =+F4 | =+G4 |
| 5 | 2 | 11 | =+F5/F\$13 | =+H4+F5 | =+I4+G5 |
| 6 | 3 | 11 | =+F6/F\$13 | =+H5+F6 | =+I5+G6 |
| 7 | 4 | 20 | =+F7/F\$13 | =+H6+F7 | =+I6+G7 |
| 8 | 5 | 15 | =+F8/F\$13 | =+H7+F8 | =+I7+G8 |
| 9 | 6 | 8 | =+F9/F\$13 | =+H8+F9 | =+I8+G9 |
| 10 | 7 | 3 | =+F10/F\$13 | =+H9+F10 | =+I9+G10 |
| 11 | 8 | 0 | =+F11/F\$13 | =+H10+F11 | =+I10+G11 |
| 12 | 9 | 1 | =+F12/F\$13 | =+H11+F12 | =+I11+G12 |
| 13 | Total | =SUMA(F4:F | =+F13/F\$13 | | |
| 14 | | | | | |

1.2. Distribución de frecuencias para variables discretas: con EXCEL

Este es el resultado final

Aspectos "estéticos" se pueden arreglar

○ Aquí no nos vamos a preocupar de eso

| Valores | Variable CAPITAS | | Acumuladas | |
|---------|------------------|-----------|------------|-----------|
| | Absolutas | Relativas | Absolutas | Relativas |
| 1 | 6 | 0,08 | 6 | 0,08 |
| 2 | 11 | 0,15 | 17 | 0,23 |
| 3 | 11 | 0,15 | 28 | 0,37 |
| 4 | 20 | 0,27 | 48 | 0,64 |
| 5 | 15 | 0,20 | 63 | 0,84 |
| 6 | 8 | 0,11 | 71 | 0,95 |
| 7 | 3 | 0,04 | 74 | 0,99 |
| 8 | 0 | 0,00 | 74 | 0,99 |
| 9 | 1 | 0,01 | 75 | 1,00 |
| Total | 75 | 1,00 | | |

1.3. Distribución de frecuencias de variables continuas

- Variables continuas: variables cuantitativas (escala de intervalo), con muchos valores
- No podemos hacer distribución de frecuencias como antes: demasiados valores distintos
- Ejemplo: variable GTINE en fichero HOGARES. Cada hogar tiene un valor diferente. Tabla: 1, 1, 1, 1,....
- ¿Cómo hacer?: Agrupar los datos por intervalos
- Cada clase no es ya un valor sino un intervalo de valores

1.3. Distribución de frecuencias de variables continuas

Ejemplo con variable GTINE en fichero HOGARES

| | Absolutas | Relativas |
|-------------------|-----------|-----------|
| | n_i | f_i |
| Hasta 50.000 | 1 | 0,01 |
| 50.001 a 100.000 | 10 | 0,13 |
| 100.001 a 150.000 | 9 | 0,12 |
| 150.001 a 200.000 | 12 | 0,16 |
| 200.001 a 250.000 | 11 | 0,15 |
| 250.001 a 300.000 | 11 | 0,15 |
| 300.001 a 350.000 | 3 | 0,04 |
| 350.001 a 400.000 | 1 | 0,01 |
| 400.001 a 450.000 | 6 | 0,08 |
| 450.001 a 500.000 | 5 | 0,07 |
| 500.001 a 550.000 | 1 | 0,01 |
| 550.001 a 600.000 | 0 | 0,00 |
| 600.001 a 650.000 | 2 | 0,03 |
| 650.001 a 700.000 | 1 | 0,01 |
| 700.001 a 750.000 | 1 | 0,01 |
| 750.001 a 800.000 | 0 | 0,00 |
| 800.001 a 850.000 | 0 | 0,00 |
| 850.001 a 900.000 | 1 | 0,01 |
| | 75 | 1 |

1.4. Distribución de frecuencias de variables continuas: con EXCEL

Escribimos en columna los valores que definen los límites de las clases en las que queremos dividir la escala (OJO: eso es distinto de la marca de clase)

- Marcamos las celdas que están a la derecha de la columna con los valores, donde van a ir las frecuencias absolutas
- Presionamos F2 y escribimos en la celda superior del rango marcado la función:
 - =FRECUENCIA(DATOS;GRUPOS)

1.4. Distribución de frecuencias de variables continuas: con EXCEL

Datos: donde están los valores cuya frecuencia queremos contar

- Grupos: dónde están los valores que sirven de punto de corte entre clases

- Ejemplo:

=frecuencia(gtine;a2:a20)

| | A | B | C |
|----|--------|---------------------------|---|
| 1 | | | |
| 2 | 50000 | =frecuencia(gtine;a2:a20) | |
| 3 | 100000 | FRECUENCIA(datos; grupos) | |
| 4 | 150000 | | |
| 5 | 200000 | | |
| 6 | 250000 | | |
| 7 | 300000 | | |
| 8 | 350000 | | |
| 9 | 400000 | | |
| 10 | 450000 | | |
| 11 | 500000 | | |
| 12 | 550000 | | |
| 13 | 600000 | | |
| 14 | 650000 | | |
| 15 | 700000 | | |
| 16 | 750000 | | |
| 17 | 800000 | | |
| 18 | 850000 | | |
| 19 | 900000 | | |
| 20 | 950000 | | |
| 21 | | | |

1.4. Distribución de frecuencias de variables continuas: con EXCEL

Para terminar:
May+Ctrl+Intro
(FRECUENCIA es una "fórmula matricial")

- Resultado es este

| | A | B |
|----|--------|----|
| 1 | | |
| 2 | | |
| 3 | 50000 | 1 |
| 4 | 100000 | 10 |
| 5 | 150000 | 9 |
| 6 | 200000 | 12 |
| 7 | 250000 | 11 |
| 8 | 300000 | 11 |
| 9 | 350000 | 3 |
| 10 | 400000 | 1 |
| 11 | 450000 | 6 |
| 12 | 500000 | 5 |
| 13 | 550000 | 1 |
| 14 | 600000 | 0 |
| 15 | 650000 | 2 |
| 16 | 700000 | 1 |
| 17 | 750000 | 1 |
| 18 | 800000 | 0 |
| 19 | 850000 | 0 |
| 20 | 900000 | 1 |
| 21 | 950000 | 0 |

1.4. Distribución de frecuencias de variables continuas: con EXCEL

Frecuencias relativas y frecuencias acumuladas (absolutas y relativas): igual que con variables discretas

| | A | B | C | D | E |
|---|--------|-----------|------------|-------------|-----------|
| 1 | | Absolutas | Relativas | Absol. Acum | Rel. Acum |
| 2 | 50000 | 1 | =+B2/B\$21 | =+B2 | =+C2 |
| 3 | 100000 | 10 | =+B3/B\$21 | =+D2+B3 | =+E2+C3 |
| 4 | 150000 | 9 | =+B4/B\$21 | =+D3+B4 | =+E3+C4 |
| 5 | 200000 | 12 | =+B5/B\$21 | =+D4+B5 | =+E4+C5 |
| 6 | 250000 | 11 | =+B6/B\$21 | =+D5+B6 | =+E5+C6 |
| 7 | 300000 | 11 | =+B7/B\$21 | =+D6+B7 | =+E6+C7 |

1.4. Distribución de frecuencias de variables continuas: con EXCEL

- Este es el resultado final
- Como siempre, aspectos "estéticos" aparte

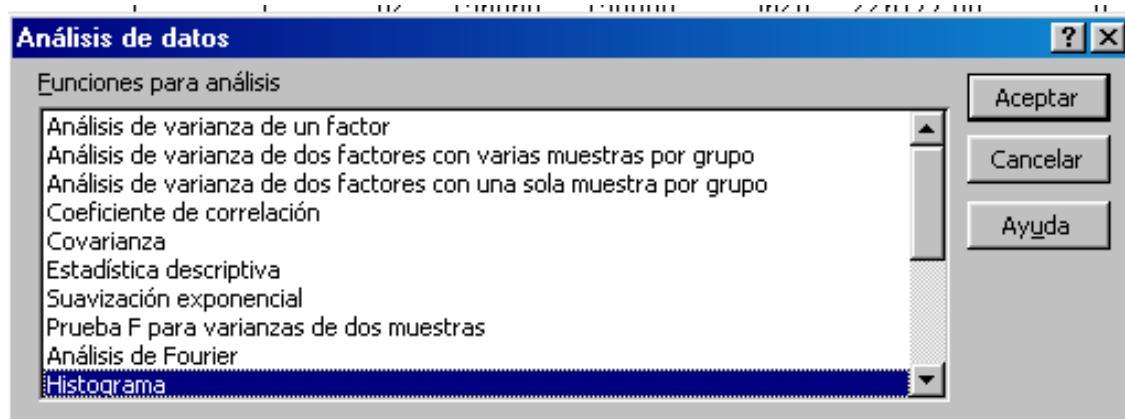
| DISTRIBUCIÓN DE FRECUENCIAS VARIABLE GTINE | | | | |
|--|-----------|-----------|-------------|-----------|
| | Absolutas | Relativas | Absol. Acum | Rel. Acum |
| 50000 | 1 | 0,01 | 1 | 0,01 |
| 100000 | 10 | 0,13 | 11 | 0,15 |
| 150000 | 9 | 0,12 | 20 | 0,27 |
| 200000 | 12 | 0,16 | 32 | 0,43 |
| 250000 | 11 | 0,15 | 43 | 0,57 |
| 300000 | 11 | 0,15 | 54 | 0,72 |
| 350000 | 3 | 0,04 | 57 | 0,76 |
| 400000 | 1 | 0,01 | 58 | 0,77 |
| 450000 | 6 | 0,08 | 64 | 0,85 |
| 500000 | 5 | 0,07 | 69 | 0,92 |
| 550000 | 1 | 0,01 | 70 | 0,93 |
| 600000 | 0 | 0,00 | 70 | 0,93 |
| 650000 | 2 | 0,03 | 72 | 0,96 |
| 700000 | 1 | 0,01 | 73 | 0,97 |
| 750000 | 1 | 0,01 | 74 | 0,99 |
| 800000 | 0 | 0,00 | 74 | 0,99 |
| 850000 | 0 | 0,00 | 74 | 0,99 |
| 900000 | 1 | 0,01 | 75 | 1,00 |
| 950000 | 0 | 0,00 | 75 | 1,00 |
| Totales | 75 | 1,00 | | |

Tema 4- Tablas y gráficos

27

1.4. Distribución de frecuencias de variables continuas: con EXCEL ALTERNATIVA: Herramientas-Análisis de Datos, "Histograma"

(Si "Análisis de datos" no aparece en menú Herramientas: Herramientas-Complementos-Herramientas para análisis-Aceptar)



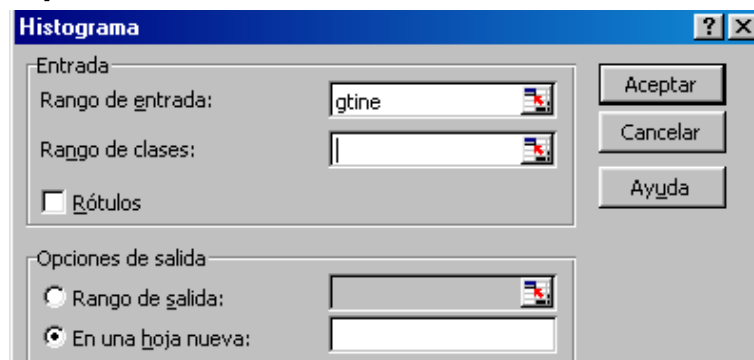
28

1.4. Distribución de frecuencias de variables continuas: con EXCEL

Importante: empezar en la hoja en la que están los datos

Podemos especificar Rango de clases y Rango de salida, pero OJO: siempre en la misma hoja

- Si no especificamos Rango de clases: Excel elige los puntos de corte entre las clases



29

1.4. Distribución de frecuencias de variables continuas: con EXCEL

Ejemplo con variable GTINE, dejando a EXCEL que escoja las clases

- Podemos añadir, como en ejemplos anteriores: frecuencias relativas, y frecuencias acumuladas (absolutas y relativas)

| <i>Clase</i> | <i>Frecuencia</i> |
|--------------|-------------------|
| 48586 | 1 |
| 152032,875 | 19 |
| 255479,75 | 26 |
| 358926,625 | 12 |
| 462373,5 | 8 |
| 565820,375 | 4 |
| 669267,25 | 3 |
| 772714,125 | 1 |
| y mayor... | 1 |

Ejercicios recomendados

- Ejercicios de exámenes
 - ◆ Feb02: 2-a
 - ◆ Jun02: 2-a
 - ◆ Feb03: 2-a
 - ◆ Sep03: 4-1
 - ◆ Feb04: 3-a
 - ◆ Jul04: 3-a
 - ◆ Feb05, Jun05: 2

Tema 4. Tablas y gráficos

Introducción

1. Distribución de frecuencias
2. **Representaciones gráficas**

Resumen

2. Representaciones gráficas

- Otra forma de resumir información de una variable
- Un gráfico
- Diferentes tipos
- Algunas diferencias también para variables discretas o continuas
- También alguna diferencia entre variables cualitativas o cuantitativas (orden de los valores)

Tema 4. Tablas y gráficos

Introducción

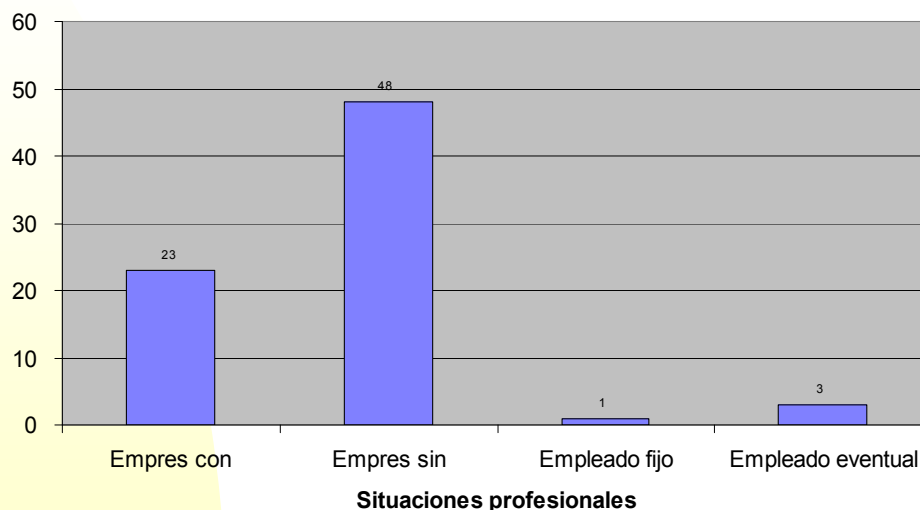
1. Distribución de frecuencias
2. Representaciones gráficas
 - 2.1. **Diagrama de barras**
 - 2.2. Diagrama de Pareto
 - 2.3. Pictograma
 - 2.4. Histograma
 - 2.5. Polígono de frecuencias
 - 2.6. Interpretar los gráficos

Resumen

2. Representaciones gráficas

2.1. Diagrama de barras

Diagrama de barras variable SITPROF



2. Representaciones gráficas

2.1. Diagrama de barras

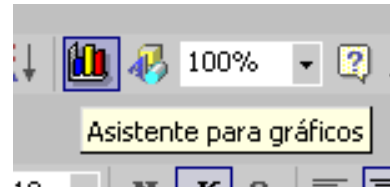
- Sólo variables discretas
- Cada valor de la variable: una barra
- Altura de la barra: frecuencia del valor
- Eje de ordenadas (vertical): pueden ser frecuencias absolutas o relativas
- Orden de los valores de izquierda a derecha:
 - ◆ Variables cualitativas: puede ser cualquiera
 - ◆ Variables de escala ordinal, o de intervalo: normalmente hay un “orden natural” que se sigue en el gráfico (como en distribuciones de frecuencias)

Representaciones gráficas

2.1. Diagrama de barras

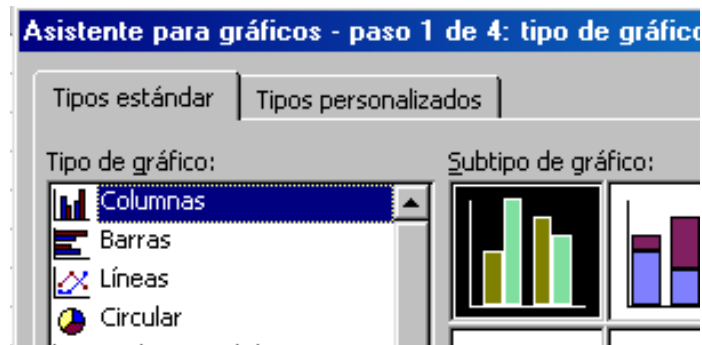
○ Ejemplo en EXCEL

- Marcar datos en distribución de frecuencias (absolutos o relativos)



- Menú de herramientas:

- Asistente para gráficos: Tipo de gráfico: Columnas



37

2. Representaciones gráficas

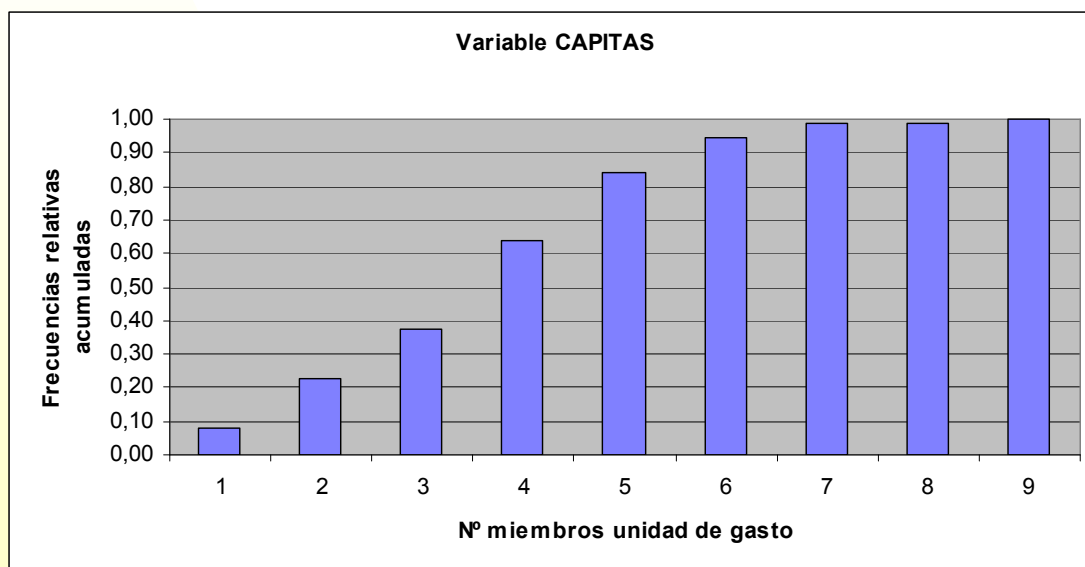
2.1. Diagrama de barras

- Opciones sobre títulos...
- Opción sobre dónde colocar gráfico
- Complicaciones: tipo, rótulos, colores...
- OJO: el objeto es TRANSMITIR INFORMACIÓN

2. Representaciones gráficas

2.1. Diagrama de barras

- Variables con escala ordinal o de intervalo: admiten diagrama de barras de frecuencias acumuladas

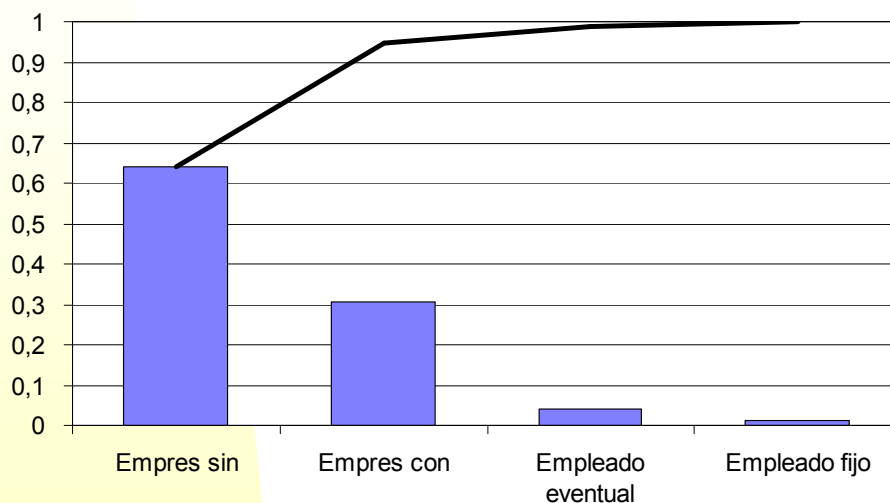


39

2. Representaciones gráficas

2.2. Diagrama de Pareto

Diagrama de Pareto variable SITPROF



2. Representaciones gráficas

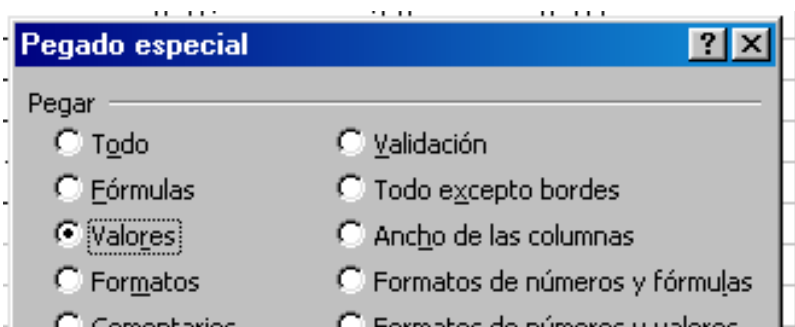
2.2. Diagrama de Pareto

- También para variables discretas
- Diagrama de barras con algunas peculiaridades
 - ◆ Orden de los valores: de más frecuente a menos frecuente
 - ◆ Además de barras: línea que representa las frecuencias acumuladas

2. Representaciones gráficas

2.2. Diagrama de Pareto

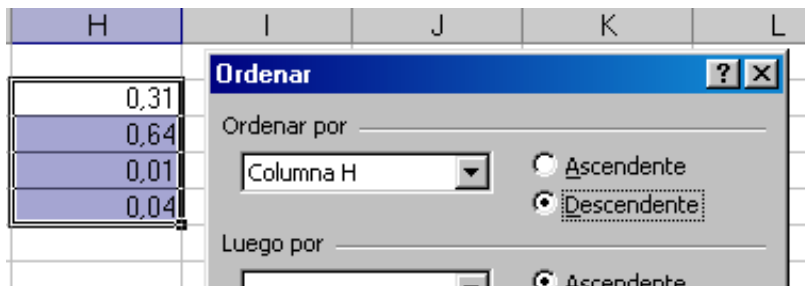
- Ejemplo en EXCEL
- Copiar datos distribución frecuencias
- Pegar datos usando Editar-Pegado Especial
- Escoger "valores" (pega los resultados en lugar de las fórmulas)



2. Representaciones gráficas

2.2. Diagrama de Pareto

- Reordenar datos (marcando bloque, Datos-Ordenar, Descendente)

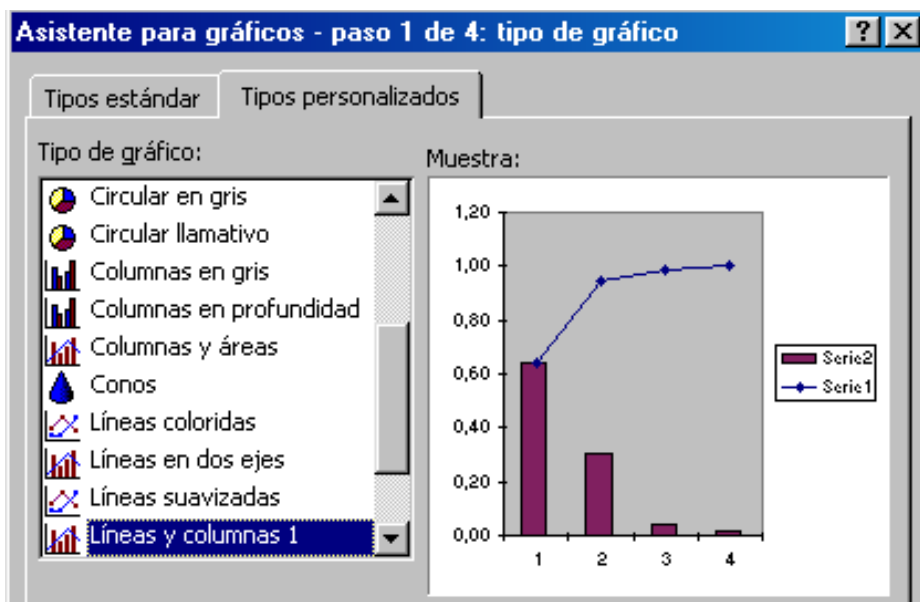


- Añadir columna frecuencia acumulada (como ejemplos anteriores)
- Marcar bloque con las dos columnas

2. Representaciones gráficas

2.2. Diagrama de Pareto

- Gráficos- Tipos personalizados- Líneas y columnas



2. Representaciones gráficas

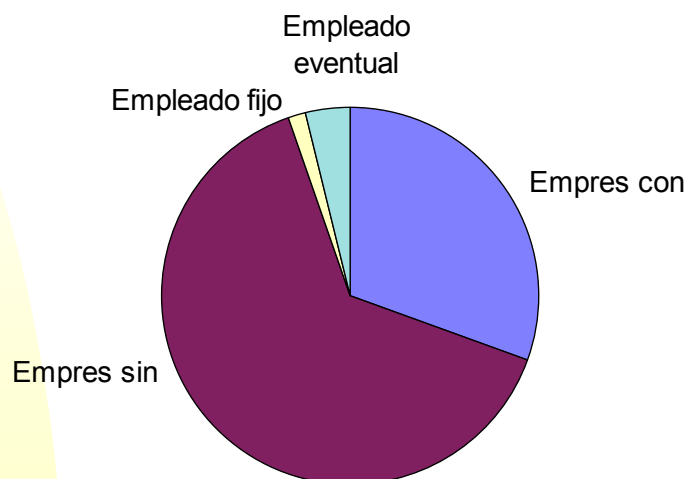
2.2. Diagrama de Pareto

- Resto del proceso: igual que otros gráficos
 - Títulos
 - Ejes
 - Nombres de valores
 - Colores...

2. Representaciones gráficas

2.3. Pictograma

Pictograma variable SITPROF



2. Representaciones gráficas

2.3. Pictograma

- Variables discretas
- Ejemplo en EXCEL: como diagrama de barras
 - Marcar datos de variable (absolutos o relativos)
 - Icono gráfico, Tipo de gráfico: **CÍRCULO**
 - Elegir si queremos ver datos
 - Elegir página
- OJO OTRA VEZ A LAS TENTACIONES

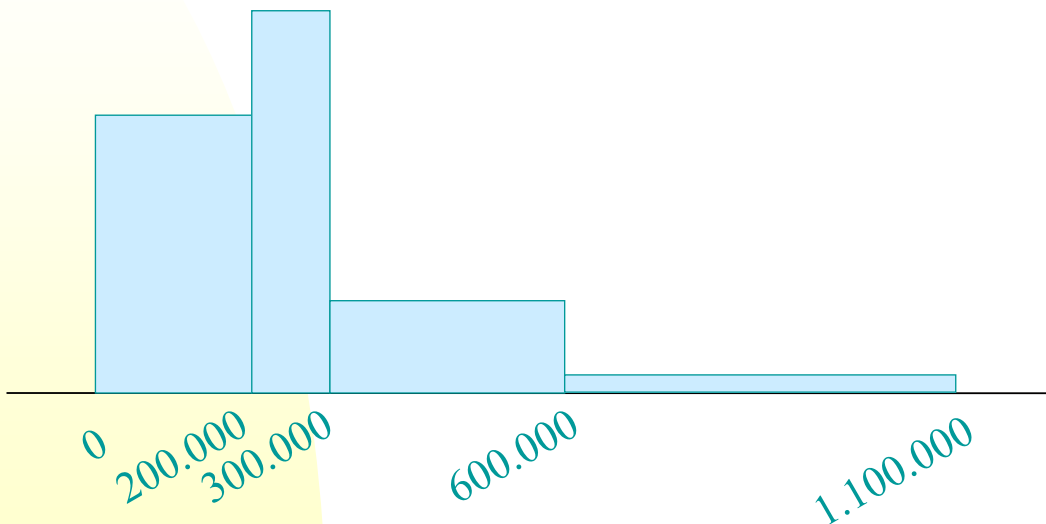
2. Representaciones gráficas

2.4. El histograma

- Sólo para variables continuas
- Frecuencias representadas por áreas
- Modelo original:
 - clases de diferente tamaño.
 - Eje X; rectángulos con base de diferente longitud
 - Eje vertical no tiene sentido

2. Representaciones gráficas

2.4. El histograma



Tema 4- Tablas y gráficos

49

2. Representaciones gráficas

2.4. El histograma

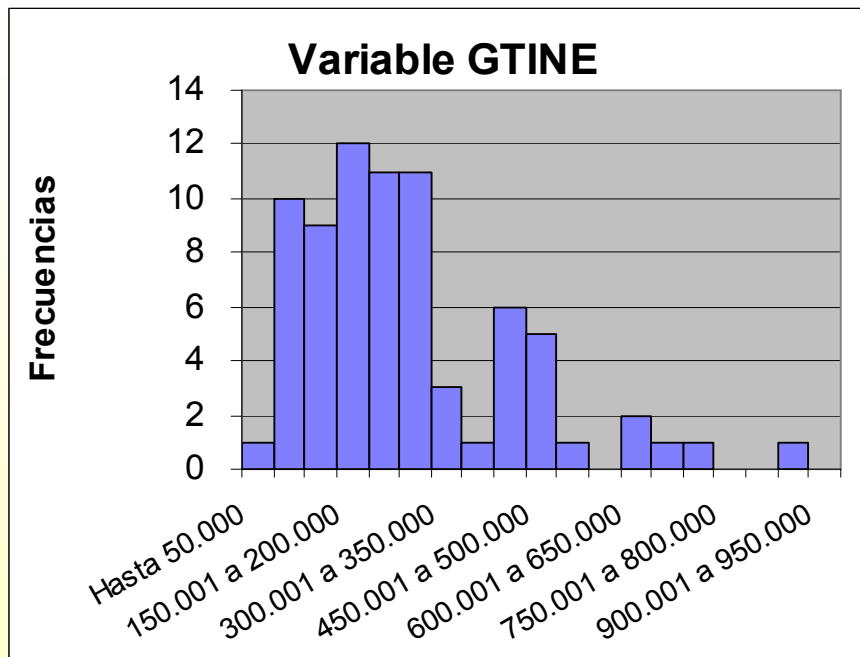
- Lo más usual: clases del mismo tamaño → la altura de los rectángulos proporcional a frecuencia (como en diagrama de barras)
- Similar a diagrama de barras, excepto:
 - Barras son contiguas
 - Rótulos no en valores sino en líneas de división (no en EXCEL)
- Como en diagrama de barras, forma no cambia por usar frecuencias absolutas o relativas

Tema 4- Tablas y gráficos

50

2. Representaciones gráficas

2.4. El histograma

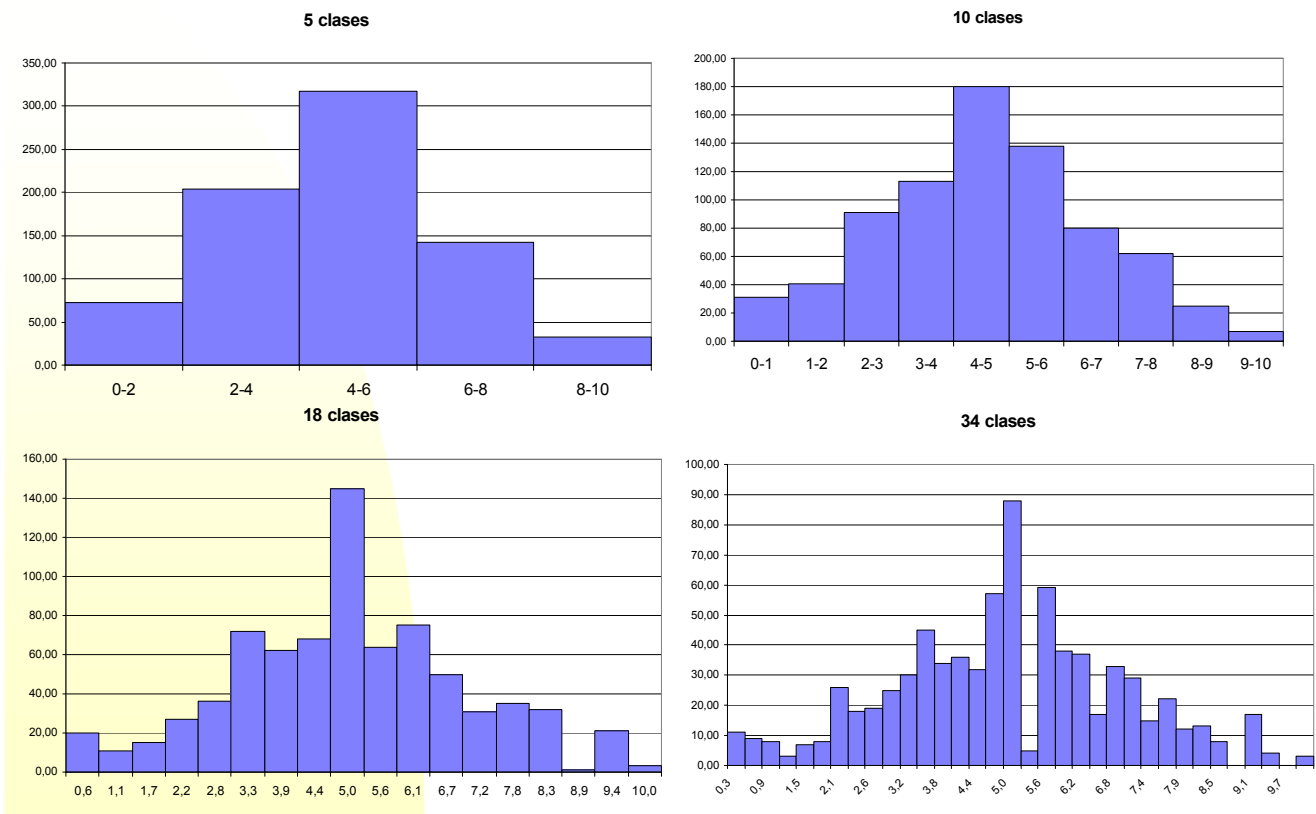


2. Representaciones gráficas

2.4. El histograma

- La forma del histograma Sí cambia según el número de las clases
- Ejemplo: variable NOTAS en fichero NOTAS (OJO: primero “arreglar” fichero NOTAS)
 - 4 histogramas con 5, 10, 18 y 34 clases.

2.4. El histograma



2. Representaciones gráficas

2.4. El histograma

- Por tanto: a la hora de hacer un histograma es **MUY IMPORTANTE** la elección de las clases
- Una regla: empezar con pocas y a partir de resultados ir aumentando
- Para una población grande: el número de clases podrían ser muchas, y muy estrechas
- Tendencia: el histograma como una curva

Representaciones gráficas

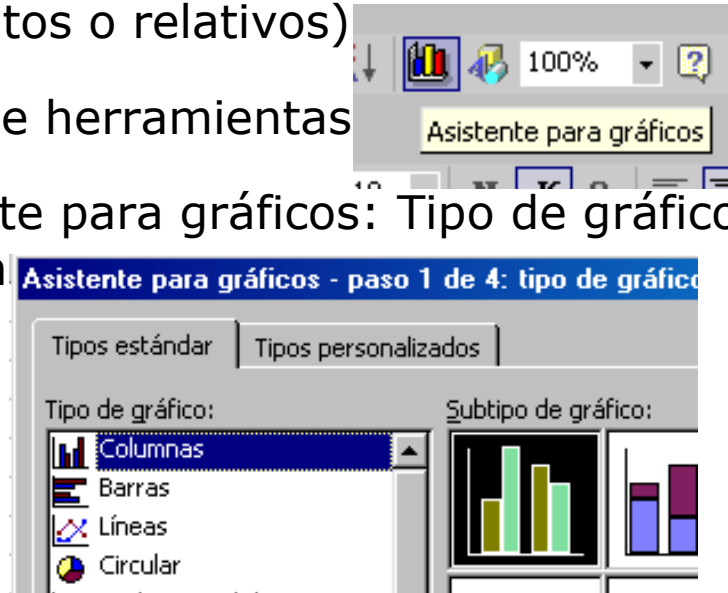
2.4. Histograma con EXCEL

○ Comienzo: igual que diagrama de barras

- Marcar datos en distribución de frecuencias (absolutos o relativos)

- Menú de herramientas

- Asistente para gráficos: Tipo de gráfico: Columnas

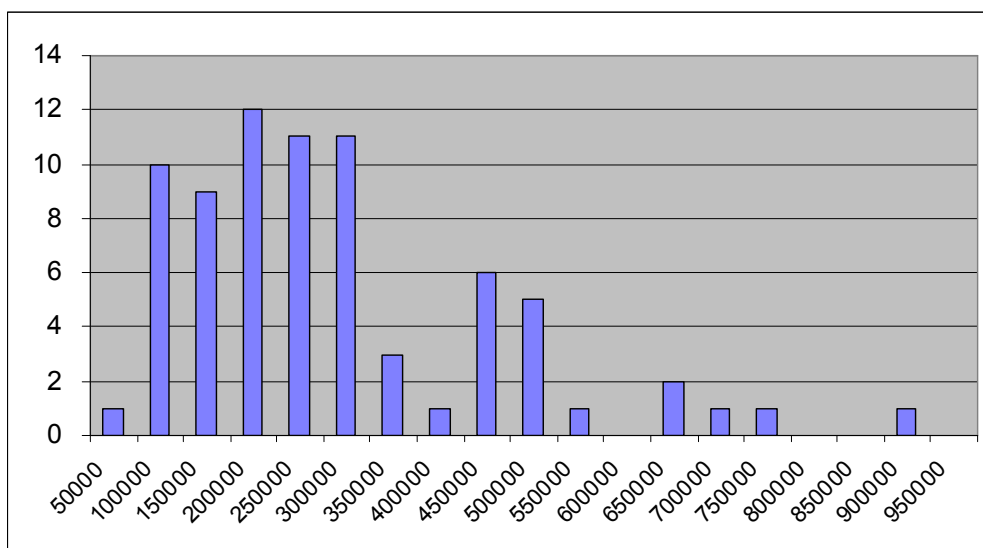


55

Representaciones gráficas

2.4. Histograma con EXCEL

○ Primer resultado:



Representaciones gráficas

2.4. Histograma con EXCEL

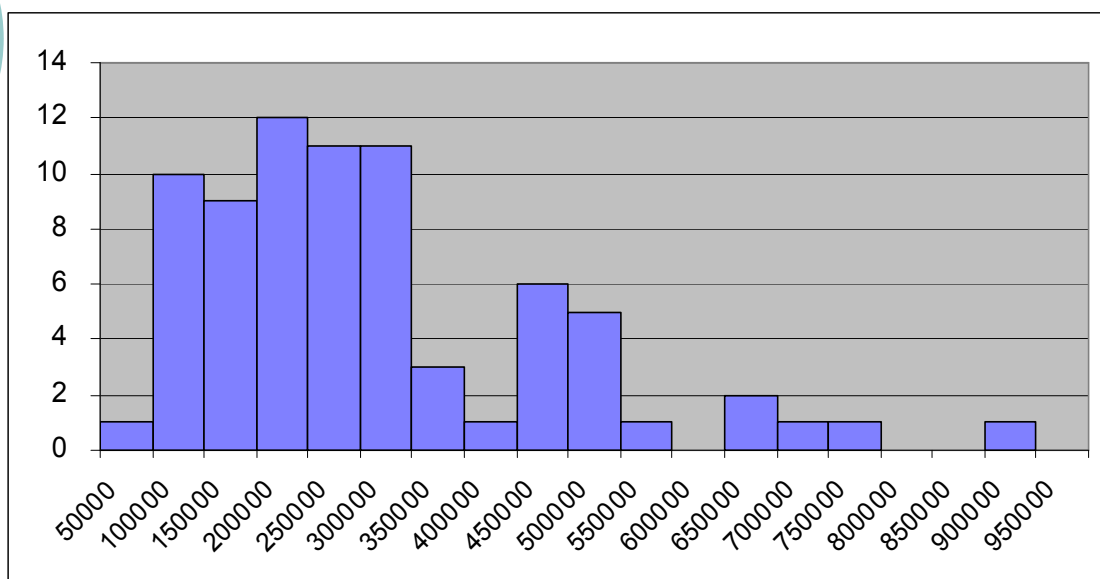
○ "Arreglarlo":

- Seleccionar el gráfico
- Click sobre barras (queda "seleccionada")
- Click derecho- Menú- "Formato de serie de datos"
- Pestaña "Opciones"- Ancho de Rango: Valor cero

Representaciones gráficas

2.4. Histograma con EXCEL

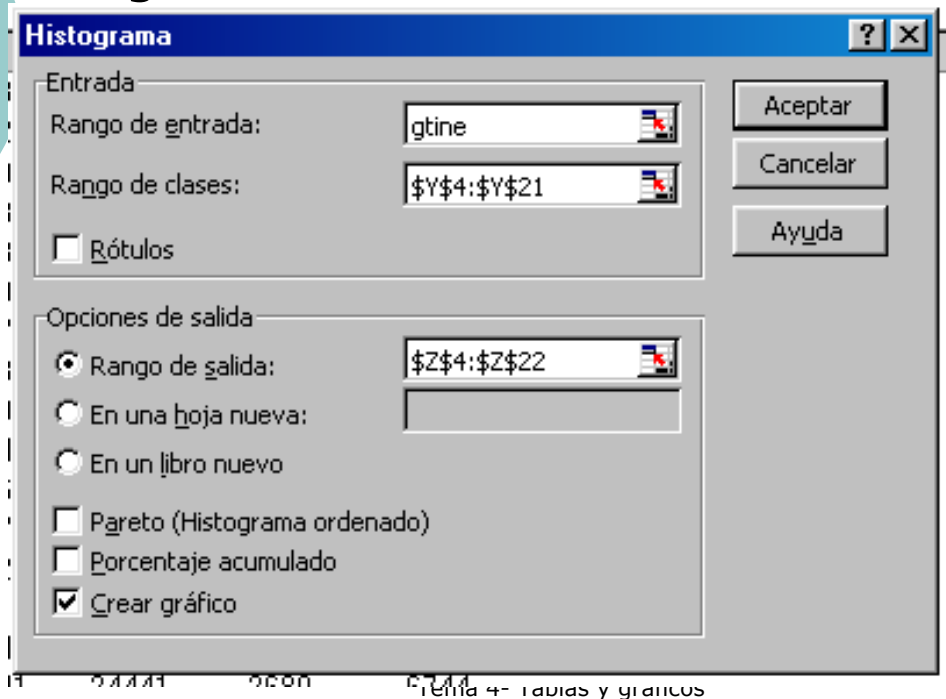
○ Arreglado:



2. Representaciones gráficas

2.4. Histograma con EXCEL

- Alternativa: Herramientas-Análisis de datos-Histograma

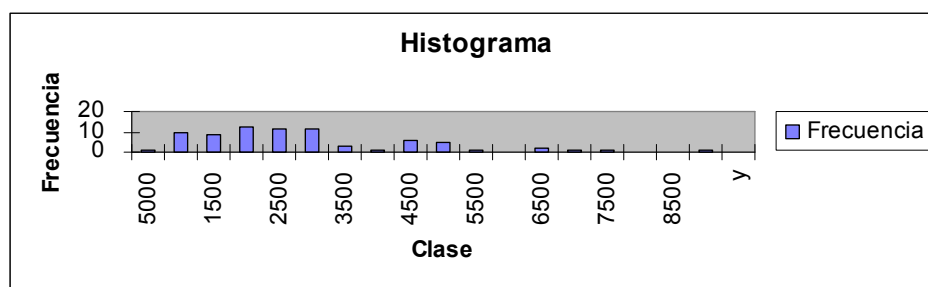


59

2. Representaciones gráficas

2.4. Histograma con EXCEL

- Como antes: podemos elegir nosotros el rango de clases, o dejarle a EXCEL que "elija" él
- Importante: Marcar "Crear gráfico"
- Resultado inicial puede ser "raro" ("estirar")
- Hacer barras contiguas: como transparencia 61

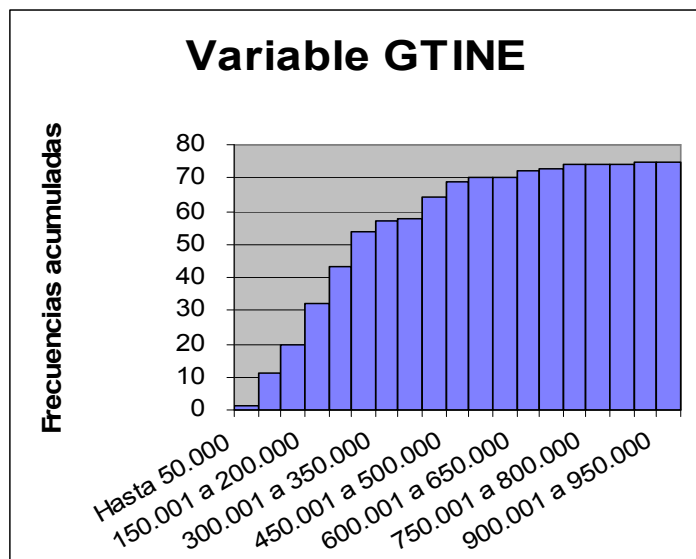


60

2. Representaciones gráficas

2.4. El histograma

- Histograma de frecuencias acumuladas
- Igual que con diagrama de barras: crear datos de frecuencias acumuladas, y usar esos datos para generar gráfico
- Luego eliminar huecos entre barras



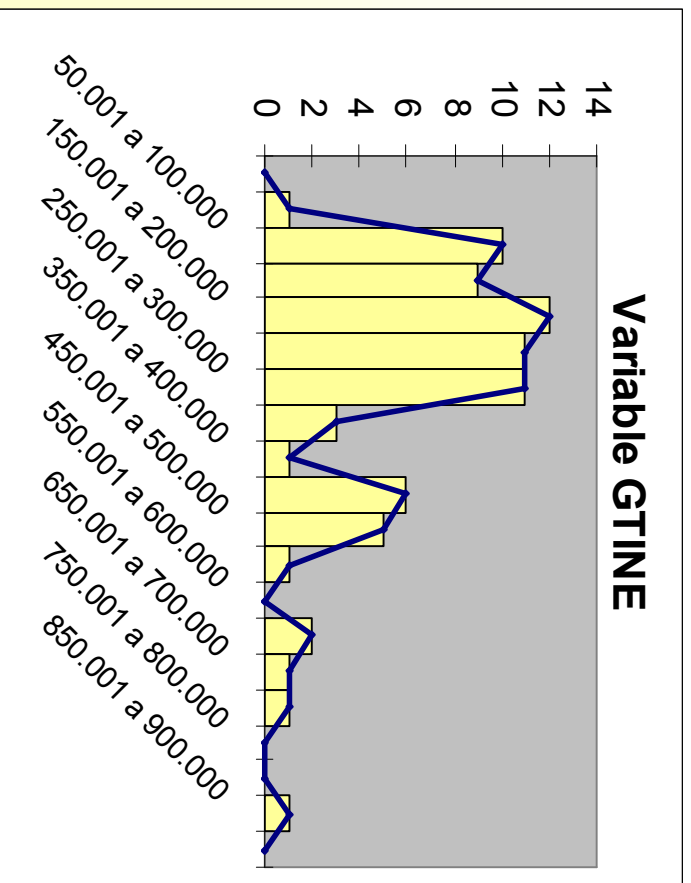
2. Representaciones gráficas

2.5. El polígono de frecuencias

- Sólo para variables continuas
- Básicamente equivalente a histograma
- Líneas que unen puntos medios de bases superiores de rectángulos
- En EXCEL: igual que histograma pero usando Tipo de gráfico: líneas

2. Representaciones gráficas

2.5. El polígono de frecuencias

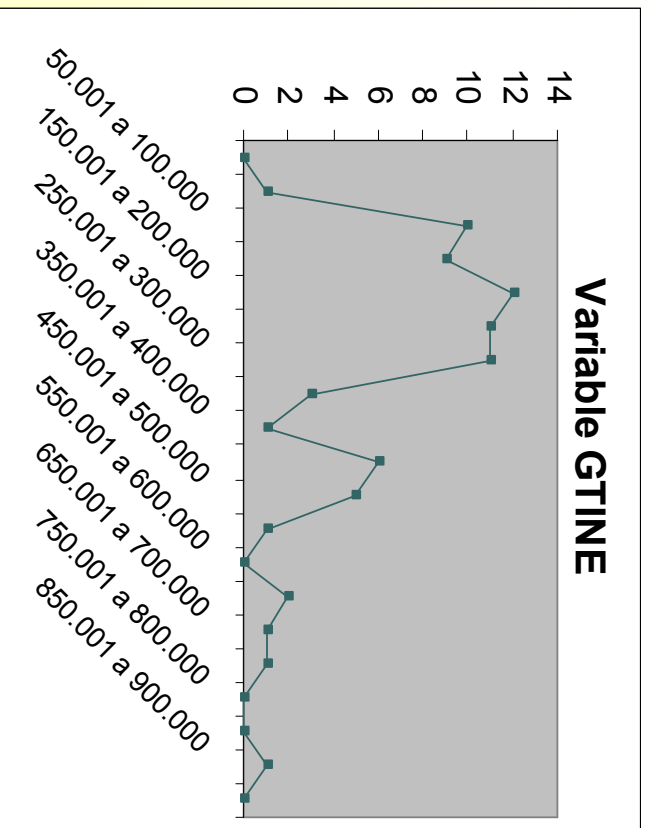


Tema 4- Tablas y gráficos

63

2. Representaciones gráficas

2.5. El polígono de frecuencias



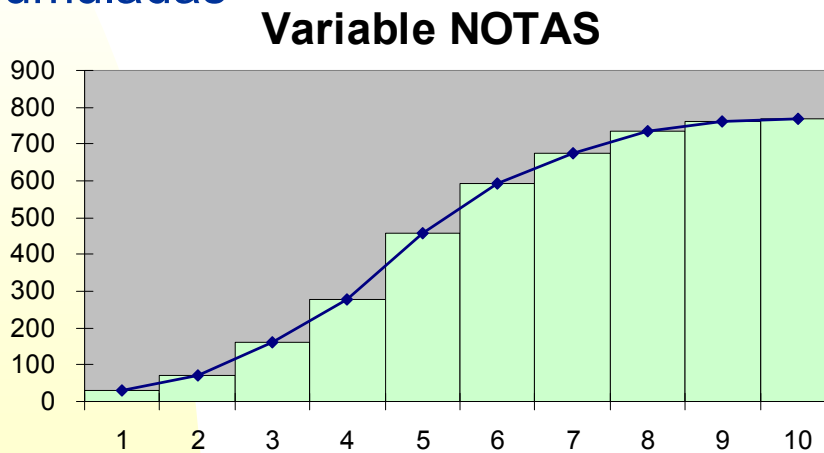
Tema 4- Tablas y gráficos

64

2. Representaciones gráficas

2.5. El polígono de frecuencias

- También puede hacerse con frecuencias acumuladas



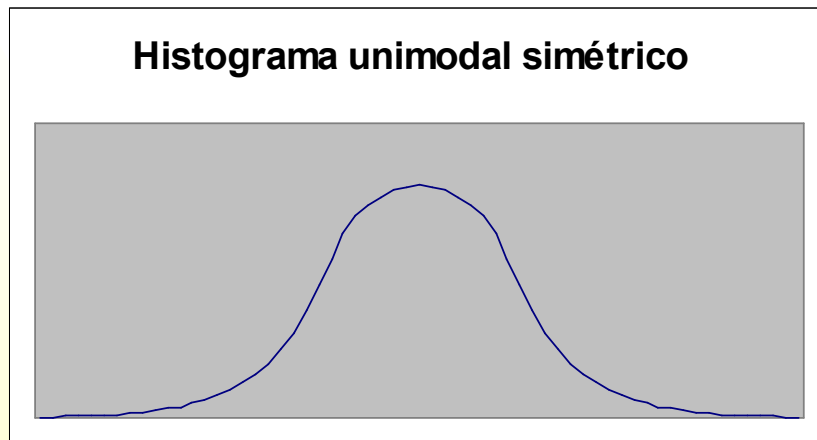
2. Representaciones gráficas

2.6. Interpretar los gráficos

- Histogramas y diagramas de barras son los más usados
- Nos dicen mucho sobre la distribución de la variable
- Datos dispersos en muchos valores, o concentrados en pocos valores
- Cuál es el valor más frecuente (la moda)
- Hay o no valores muy alejados del valor más frecuente
- Distribución es más o menos “simétrica”: igual número de casos con valores mayores y menores que el valor más frecuente...

2. Representaciones gráficas

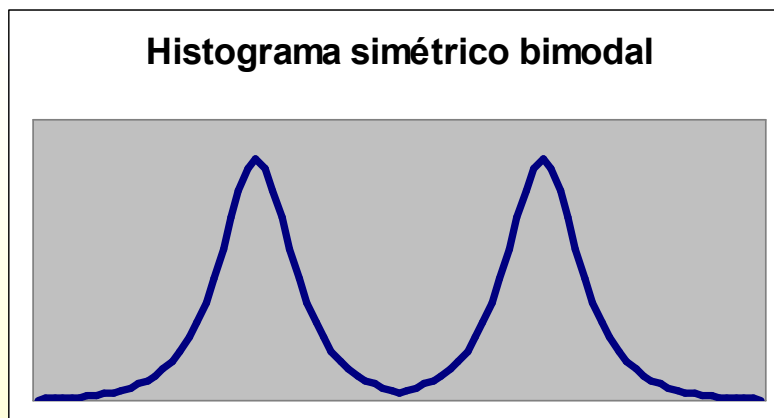
2.6. Interpretar los gráficos



Un valor más común (moda) aproximadamente en el centro
Simétrico
Ejemplos: notas, altura, peso...

2. Representaciones gráficas

2.6. Interpretar los gráficos

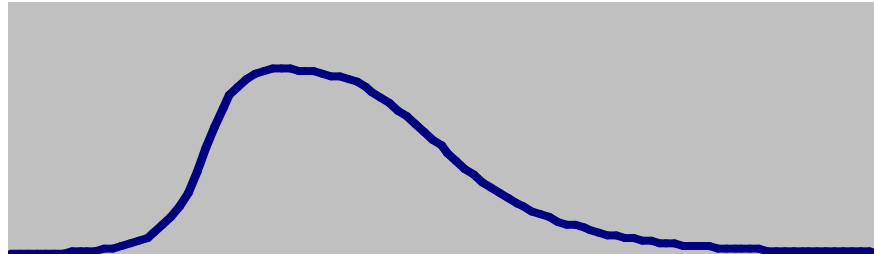


Dos valores con más casos (modas)
Simétrico en torno a valor con pocos casos
Extraño: ¿2 poblaciones?

2. Representaciones gráficas

2.6. Interpretar los gráficos

Histograma unimodal asimétrico a la derecha

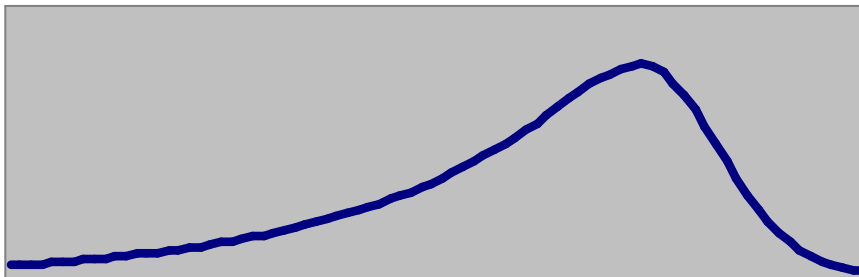


Un valor con más casos (moda)
“Cola” hacia la derecha
Ejemplos: ingresos o gastos

2. Representaciones gráficas

2.6. Interpretar los gráficos

Histograma unimodal asimétrico a la izquierda



Un valor con más casos (moda)
“Cola” hacia la izquierda
Ejemplo: esperanza de vida

Resumen: conceptos principales (1)

- Distribución de frecuencias
 - ◆ Frecuencias absolutas y relativas
 - ◆ Frecuencias acumuladas (absolutas y relativas)
 - ◆ Con EXCEL:
 - ★ Variables discretas:
 - =CONTAR.SI(rango;valores)
 - ★ Variables continuas:
 - =FRECUENCIA(datos;grupos)
 - Herramientas-Análisis de datos-Histograma

Resumen: conceptos principales (2)

- Representaciones gráficas
 - ◆ Diagrama de barras
 - ◆ Diagrama de Pareto
 - ◆ Pictograma
 - ◆ Histograma
 - ◆ Polígono de frecuencias

Ejercicios recomendados

- Tema 2 del libro
 - ◆ 2.7 (con Excel o a mano)
 - ◆ 2.9
- Tema 3 del libro
 - ◆ 3.1 (b) sólo con Excel)
 - ◆ 3.6 a) y b) (Histogramas sólo con Excel)
 - ◆ 3.9 (Con Excel)
 - ◆ 3.19 (Con Excel)

Ejercicios recomendados

- Ejercicios de exámenes
 - ◆ Feb02, Jun02: 4
 - ◆ Feb03, Sep03: 5
 - ◆ Feb04, Jul04: 4
 - ◆ Feb05, Jun05: 3
 - ◆ Feb06, Jun06, Ene07, Jul07, Ene08, Jul08: 2