

REDACCIÓN de MONOGRAFÍAS, TESIS y ARTÍCULOS CIENTÍFICOS

Sergio Nesmachnow

Universidad de la República, Uruguay



Hoja de ruta

1. El proceso de investigación científica
2. Redacción de informes, monografías y tesis
3. Presentación y recursos gráficos
4. Redacción de artículos científicos
5. Citas y referencias bibliográficas
6. Latex y reseñas de trabajos relacionados
7. *Evaluación y reporte de resultados*
8. Gráficos
9. Casos de estudio



7



Evaluación y reporte de resultados: tablas

Introducción

- Son elementos gráficos para reportar resultados numéricos o características específicas
- Regla general: usar una tabla solo si se cuenta con un conjunto de datos *repetitivos*
- Exponer grandes cantidades de datos no ayuda a la buena ciencia. No son bien vistos los artículos con tablas extensas o con muchas tablas
 - El lector espera evidencia para corroborar las formulaciones presentadas, no una cantidad abrumadora de datos
- La mejor estrategia es incluir casos representativos o de extrema relevancia

Introducción

- Las tablas deben ser diseñadas para ser auto contenidas
- Muchas veces son consultadas independientemente del texto que las referencia
- De la tabla misma, el lector debe poder reconocer:
 - El propósito de la tabla
 - El contexto de los datos
 - La fuente de los datos
 - Las unidades de medida
 - Significado de términos y abreviaciones

Anatomía de un tabla

Table 8. Induction of creatinine deiminase in *C. neoformans* and *C. bacillisporus*

| N source | Total enzyme ^b | Cuerpo Los datos que se quieren mostrar | Sp act (U/mg of protein) |
|------------------|---------------------------|--|--------------------------|
| Ammonium sulfate | 0.50 | 0.22 | 0.50 |
| Glutamic acid | 0.50 | 0.22 | 0.50 |
| Aspartic acid | 0.50 | 0.22 | 1.47 |
| Arginine | 0.50 | 0.22 | 0.06 |
| Creatinine | 0.50 | 0.22 | 0.06 |

Numeración

Debe ser numerada según el orden en el que aparece referenciada en el texto.

Se debe omitir el signo de puntuación final. Debe ir arriba de la tabla.

Cuerpo

Los datos que se quieren mostrar

Encabezados de columnas

Deben ser etiquetados con palabras que describan la columna.

Etiquetas de filas

Describen el concepto descrito en cada fila

Notas

Cualquier aclaración sobre la tabla que no sea auto

explicativa.

^a The inoculum was grown in glucose broth with ammonium sulfate, washed twice, and then transferred into the media with the N sources listed below.

^b Enzyme units in cell extract obtained from ca. 10^{10} cells.

El título

Debe...

- Reflejar el contenido de la tabla
 - Debe ser conciso y no debe estar dividido en varias oraciones
 - Deben omitirse todas las palabras innecesarias
- Nombrar cada uno de sus principales componentes
 - Esto no quiere decir que sea necesario describir cada variable
- Especificar el contexto de los datos
 - ¿Dónde? ¿Cuándo? Se obtuvieron los datos
 - Debe especificarse la fuente en caso de ser externa
- Especificar las unidades de medida de los datos (de ser posible)
 - En caso de no ser posible, detallar estos datos en los encabezados

El título

Ejemplos:

- **Estadística descriptiva de las muestras**

Muy poco informativo. Qué tipo de estadísticas descriptivas? ¿En qué variables? ¿Para qué muestra?

- **Medias y desviaciones estándar de los componentes del suelo**

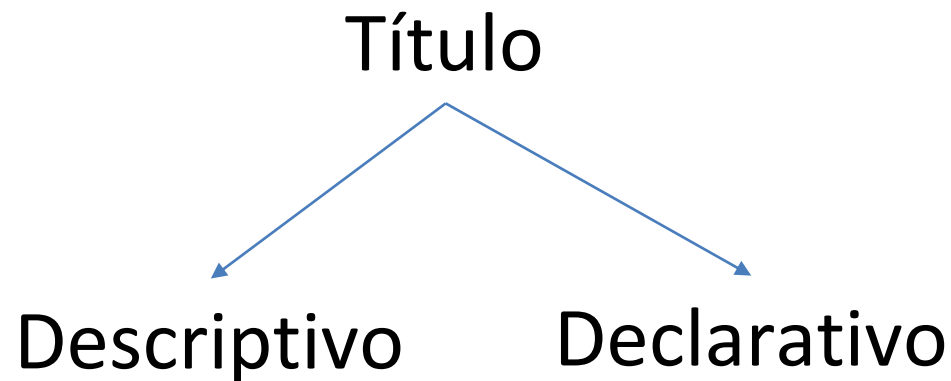
Deja claro el tipo de estadísticas y el tema de la tabla, pero aún no queda claro el contexto

- **Medias y desviaciones estándar de los componentes del suelo, 100 muestras, Cerro Chato, 1990**

Menciona el tema de la tabla, la fecha y el lugar donde fueron recolectados los datos y especifica el tamaño de la muestra

El título

- Un título puede ser **descriptivo** o **declarativo**
 - Puede describir el contenido o incluir un mensaje



- El título descriptivo se utiliza para las tablas cuyo objetivo es presentar información detallada

Table 8.1 Maternal age, gestational age, indication, size and type (avascular or vascular) of villi sampled, method used in processing the biopsy (direct preparation, 24 h culture, long-term culture), and karyotype in 80 diagnostic cases of first-trimester chorionic biopsy

| Case No. | Maternal age (y) | Gesta- tional age (wk) | Indication | Villi sampled | | | Cytogenic method | | |
|----------|------------------|------------------------------|---------------------------|---------------|----------------|---------------|------------------|-----------------|-------------------------|
| | | | | Weight (mg) | Avas- cular | Vas- cular | Direct | 24 h culture | Long term culture |
| 1 | 45 | 12 | Previous child Mb Down | 10 | | X | | X | 46,XY |
| 2 | 19 | 12 | Hemophilia | 2 | | | | X | 46,XY |
| 3 | 30 | 9 | Hemophilia | 10 | | X | | X | 92,XXYY |
| | | | | 5 | X | | | | |
| 4 | 41 | 11 | Age | 8 | | X | | X | 46,XY |
| [etc.] | | | | | | | | | |

- Un error común de los títulos descriptivos es repetir la información que figura en el encabezado

Table 8.1 Details of 80 diagnostic cases of first-trimester chorionic biopsy

| Case No. | Maternal age (y) | Gestational age (wk) | Indication | Villi sampled | | | Cytogenic method | | | |
|----------|------------------|----------------------|------------------------|---------------|-----------|----------|------------------|--------------|-------------------|-----------|
| | | | | Weight (mg) | Avascular | Vascular | Direct | 24 h culture | Long term culture | Karyotype |
| 1 | 45 | 12 | Previous child Mb Down | 10 | | X | | | X | 46,XY |
| 2 | 19 | 12 | Hemophilia | 2 | | | | | X | 46,XY |
| 3 | 30 | 9 | Hemophilia | 10 | | X | | | X | 92,XXYY |
| | | | | 5 | X | | | | | |
| 4 | 41 | 11 | Age | 8 | | X | | | X | 46,XY |
| [etc.] | | | | | | | | | | |

El título

- El título declarativo se usa cuando el objetivo de la tabla es mostrar una tendencia o una relación

Table 8.2 Increase in fractured clavicles with birth weight

| Birth weight, g | Deliveries | | Fractured clavicles | |
|-----------------|------------|-------|---------------------|-------|
| | <i>n</i> | % | <i>n</i> | % |
| -2500 | 434 | 8.5 | 9 | 2.07 |
| 2501-3000 | 1395 | 27.3 | 45 | 3.23 |
| 3001-3500 | 2047 | 40.0 | 108 | 5.28 |
| 3501-4000 | 1049 | 20.5 | 111 | 10.58 |
| 4001- | 193 | 3.7 | 24 | 12.44 |
| All | 5118 | 100.0 | 297 | 5.80 |

Table 8. Induction of creatinine deiminase in *C. neoformans* and *C. bacillisporus*

| N source ^a | <i>C. neoformans</i> NIH 12 | | <i>C. bacillisporus</i> NIH 191 | |
|-----------------------|--------------------------------|--------------------------|------------------------------------|--------------------------|
| | Total enzyme ^b | Sp act (U/mg of protein) | Total enzyme | Sp act (U/mg of protein) |
| Ammonia | 0.58 | 0.32 | 0.50 | 0.28 |
| Glutamic acid | 5.36 | 1.48 | 2.18 | 0.61 |
| Aspartic acid | 2.72 | 0.15 | 1.47 | 0.06 |
| Arginine | 3.58 | 2.18 | 3.38 | 2.19 |
| Creatinine | 97.30 | 58.40 | 104.00 | 58.30 |

→ Etiquetas de filas

← Encabezado de columnas

^a The inoculum was grown in glucose broth with ammonium sulfate, washed twice, and then transferred into the media with the N sources listed below.

^b Enzyme units in cell extract obtained from ca. 10¹⁰ cells.

Columnas y filas

- El encabezado de cada columna y la etiqueta de cada fila debe identificar su contenido
- Especificar la unidad de medida en los encabezados o las etiquetas
 - Nunca asumir que las unidades serán evidentes
- Evitar (en lo posible) el uso de abreviaciones y acrónimos
 - Es recomendable usar palabras sueltas o pequeñas frases



Table X. Descriptive statistics on sample

| Variable | Mean | Standard deviation |
|----------|------|--------------------|
| YNDSR | ## | ## |
| Q201a | ## | ## |

Acrónimos crípticos



Table X. Means and standard deviations on socioeconomic and attitudinal variables

| Variable | Mean | Standard deviation |
|---|------|--------------------|
| Income-to-needs ratio | ## | ## |
| Extent of agreement with current welfare system | ## | ## |

Columnas y filas

- Filas y columnas conceptualmente relacionadas pueden organizarse en paneles
 - Separadas por líneas horizontales o una fila en blanco
- Los paneles pueden introducir una nueva dimensión a la tabla
 - p.ej.: Dos tablas pequeñas pueden combinarse usando paneles si comparten encabezados de columnas

Table 6.1. Anatomy of a table

Households by type, race and Hispanic origin (thousands), United States, 1997

| Characteristic | All households | Family households | | | | Nonfamily household | | |
|------------------------------|----------------|-------------------|----------------|--------------------|------------------|---------------------|--------------------|------------------|
| | | Total | Married couple | Other families | | Total | Female householder | Male householder |
| | | | | Female householder | Male householder | | | |
| <u>Race/ethnicity</u> | | | | | | | | |
| White | 86,106 | 59,511 | 48,066 | 8,308 | 3,137 | 26,596 | 14,871 | 11,725 |
| Non-Hispanic White | 77,936 | 52,871 | 43,423 | 6,826 | 2,622 | 25,065 | 14,164 | 10,901 |
| Black | 12,474 | 8,408 | 3,921 | 3,926 | 562 | 4,066 | 2,190 | 1,876 |
| All other ^a | 3,948 | 2,961 | 2,330 | 418 | 212 | 986 | 455 | 532 |
| <u>Origin</u> | | | | | | | | |
| Non-Hispanic | 93,938 | 63,919 | 49,513 | 11,040 | 3,366 | 30,018 | 16,762 | 13,258 |
| Hispanic ^b | 8,590 | 6,961 | 4,804 | 1,612 | 545 | 1,630 | 754 | 875 |
| Total | 102,528 | 70,880 | 54,317 | 12,652 | 3,911 | 31,648 | 17,516 | 14,133 |

Source: U.S. Census Bureau 1998.

^a“All other” races includes Asians, Pacific Islanders, Native Americans, and those of unspecified race.

^b People of Hispanic origin may be of any race.

Table 8. Induction of creatinine deiminase in *C. neoformans* and *C. bacillisporus*

N source^a

| | <i>C. neoformans</i> NIH 12 | | <i>C. bacillisporus</i> NIH 191 | |
|---------------|--------------------------------|--------------------------------|------------------------------------|--------------------------------|
| | Total enzyme ^b | Sp act (U/mg of protein) | Total enzyme | Sp act (U/mg of protein) |
| Ammonia | 0.58 | 0.32 | 0.50 | 0.28 |
| Glutamic acid | 5.36 | 1.48 | 2.18 | 0.61 |
| Aspartic acid | 2.72 | 0.15 | 1.47 | 0.06 |
| Arginine | 3.58 | 2.18 | 3.38 | 2.19 |
| Creatinine | 97.30 | 58.40 | 104.00 | 58.30 |

^a The inoculum was grown in glucose broth with ammonium sulfate, washed twice, and then transferred into the media with the N sources listed below.

^b Enzyme units in cell extract obtained from ca. 10^{10} cells.

Los datos

- ¿Qué datos incluir en una tabla?
 - Es una herramienta para presentar evidencia, no es una base de datos
- En general no se deben presentar datos crudos
- Tampoco se deben presentar resultados de pasos intermedios!
- Es importante diseñar una tabla antes de construirla
 - Una tabla debe tener un propósito
 - Incluir solo los datos necesarios para mostrar los puntos sustantivos
 - Dejar el resto de los datos afuera

¿Cómo organizar los datos?

- Supongamos que tenemos cuatro organismos:
 - *S. fluoricolor*, *S. griseus*, *S. coelicolor*, y *S. nocolor*
- Que queremos reportar:
 - *Optimal growth temp (°C), Color of mycelium, Antibiotic produced, y Yield of antibiotic (mg/ml)*

Table 7. Characteristics of antibiotic-producing *Streptomyces*

| Organism | Optimal growth temp (°C) | Color of mycelium | Antibiotic produced | Yield of antibiotic (mg/ml) |
|-----------------------|--------------------------|-------------------|---------------------------|-----------------------------|
| <i>S. fluoricolor</i> | -10 | Tan | Fluoricillinmycin | 4,108 |
| <i>S. griseus</i> | 24 | Gray | Streptomycin | 78 |
| <i>S. coelicolor</i> | 28 | Red | Rholmondelay ^a | 2 |
| <i>S. nocolor</i> | 92 | Purple | Nomycin | 0 |



VERTICAL
(hacia abajo)

Table 6. Characteristics of antibiotic-producing *Streptomyces*

| Determination | <i>S. fluoricolor</i> | <i>S. griseus</i> | <i>S. coelicolor</i> | <i>S. nocolor</i> |
|-----------------------------|-----------------------|-------------------|---------------------------|-------------------|
| Optimal growth temp (°C) | -10 | 24 | 28 | 92 |
| Color of mycelium | Tan | Gray | Red | Purple |
| Antibiotic produced | Fluoricillinmycin | Streptomycin | Rholmondelay ^a | Nomycin |
| Yield of antibiotic (mg/ml) | 4,108 | 78 | 2 | 0 |

**HORIZONTAL
(a lo ancho)**

¿Cómo presentar datos numéricos?

- Los datos deben ser precisos pero no deben saturar al lector

Text or Chart

Table

| Type of Statistic | Text or Chart | | | Table | | |
|----------------------|---------------------------|---|------------------|--------------|--|------------------------------------|
| | Total digits ^a | Decimal places | Examples | Total digits | Decimal places | Examples |
| Integer ^b | 3 to 4 | Not applicable | 7 million 388 | Up to 6 | Not applicable | 7,123 thousand ^c 388 |
| Rational number | 3 to 4 | 1 to 2 | 32.1 -0.71 | Up to 6 | Up to 3; enough to show 2 significant digits | 32.1 -0.71 0.0043 |
| Percentage | 3 to 4 | 1 if several #s would round to same value; otherwise none | 72% 6.1% | 3 to 4 | 2 if several #s would round to same value; otherwise 1 | 72.1% 6.12% |
| Proportion | Up to 3 | 3 if several #s would round to same value; otherwise 2 | .36 .002 | Up to 3 | 3 if several #s would round to same value; otherwise 2 | .36 .002 |

Los datos

- Usar una escala y una cantidad de números decimales consistentes dentro de cada columna

| | | | | | | |
|--|---------|--|--------------------------------|---------|---|--------------------------------|
| Monetary value | 3 to 4 | None for large denominations; 2 for small | \$2 million \$12.34 | 3 to 4 | None for large denominations; 2 for small | \$2 million \$12.34 |
| Ratio | 3 to 4 | 1 to 2 | 12.7 0.8 | 3 to 4 | 2 if one or more ratios <1.0; otherwise 1 | 12.7 0.83 |
| Test statistic (e.g., t-statistic; chi-square) | 3 to 4 | 2 | $\chi^2 = 12.19$ $t = 1.78$ | 3 to 4 | 2 | $\chi^2 = 12.19$ $t = 1.78$ |
| <i>p</i> -value | Up to 3 | 2 | $p < 0.01$ $p = 0.06$ | Up to 3 | 2 for values ≥ 0.01 3 for values < 0.01 | $p = 0.06$ $p < 0.001$ |

Note: See Logan (1995) for considerations on appropriate number of digits for calculations, and Miller (forthcoming) for guidelines on digits and decimal places for results of multivariate statistical models.

^aIncluding decimal places. If number of digits exceeds this value, round or change scale.

^bIntegers include the positive and negative counting numbers and zero (Kornegay 1999). By definition, they have no decimal places.

^cThe word “thousand” (or other unit) would appear only in the column head, not in the table cells; see table 4.1b for an example.

También hay estándares sobre la alineación de los datos!

Encabezados de columnas centrados

| Characteristic | All households | Family households | | | | Nonfamily household | | |
|------------------------|----------------|-------------------|----------------|--------------------|------------------|---------------------|--------------------|------------------|
| | | Total | Married couple | Other families | | Total | Female householder | Male householder |
| | | | | Female householder | Male householder | | | |
| <i>Race/ethnicity</i> | | | | | | | | |
| White | 86,106 | 59,511 | 48,066 | 8,308 | 3,133 | | | 11,725 |
| Non-Hispanic White | 77,936 | 52,871 | 43,423 | 6,826 | 2,621 | | | 10,901 |
| Black | 12,474 | 8,408 | 3,921 | 3,926 | 562 | 4,066 | 2,190 | 1,876 |
| All other ^a | 3,948 | 2,961 | 2,330 | 418 | 212 | 986 | 455 | 532 |

Etiquetas de filas alineadas a la izquierda

Valores numéricos alineados a la derecha

Ejemplos de tablas poco útiles

Table 1. Effect of aeration on growth of *Streptomyces coelicolor*

| Temp (°C) | No. of expt | Aeration of growth medium | Growth ^a |
|-----------|-------------|---------------------------|---------------------|
| 24 | 5 | + ^b | 78 |
| 24 | 5 | – | 0 |

^a As determined by optical density (Klett units).

^b Symbols: +, 500-ml Erlenmeyer flasks were aerated by having a graduate student blow into the bottles for 15 min out of each hour; –, identical test conditions, except that the aeration was provided by an elderly professor.

- Dos columnas indican condiciones fijas, no variables del experimento
 - Si la temperatura es una variable, debe tener su propia columna y variar
 - En caso contrario, es un valor fijo que se puede describir en Materiales y Métodos, o en la descripción de los experimentos
- Los pocos datos presentados en la tabla pueden ser presentados en el texto mismo

Ejemplos de tablas poco útiles

Table 2. Effect of temperature on growth of oak (*Quercus*) seedlings

| Temp (°C) | Growth in 48 h (mm) |
|-----------|---------------------|
| -50 | 0 |
| -40 | 0 |
| -30 | 0 |
| -20 | 0 |
| -10 | 0 |
| 0 | 0 |
| 10 | 0 |
| 20 | 7 |
| 30 | 8 |
| 40 | 1 |
| 50 | 0 |
| 60 | 0 |
| 70 | 0 |

- La variable independiente está correctamente reportada
- Pero la variable dependiente prácticamente siempre reporta el mismo valor
- *“Las semillas de roble crecieron en temperaturas de entre 20 y 40°C; no se registró crecimiento en temperaturas menores a 20°C o mayores a 40°C.”*

Ejemplos de tablas poco útiles

Table 3. Oxygen requirements of various species of *Streptomyces*

| Organism | Growth under aerobic conditions | Growth under anaerobic conditions |
|-----------------------------|---------------------------------|-----------------------------------|
| <i>Streptomyces griseus</i> | + | – |
| <i>S. coelicolor</i> | + | – |
| <i>S. nocolor</i> | – | + |
| <i>S. everycolor</i> | + | – |
| <i>S. greenicus</i> | – | + |
| <i>S. rainbowenski</i> | + | – |

- La tabla no es informativa
- "*S. griseus*, *S. coelicolor*, *S. everycolor* y *S. rainbowenski* crecieron bajo condiciones aerobias, mientras que *S. nocolor* y *S. greenicus* requirieron condiciones anaerobias. “

Ejemplos de tablas poco útiles

Table 4. Bacteriological failure rates

| Nocillin | K Penicillin |
|------------------------|--------------|
| 5/35 (14) ^a | 9/34 (26) |

^a Results expressed as number of failures/total, which is then converted to a percentage (within parentheses). $P = 0.21$.

- Algunos autores creen que todos los resultados numéricos deben ser reportados en una tabla
- “La diferencia entre las tasas de fallo de la Nocilina (14%, 5 de 35) y de la Penicilina potásica (26%, 9 de 34) no fue significativa estadísticamente ($P = 0.21$)”
- El autor reporta valores no significativos estadísticamente ($P = 0.21$)

Table 5. Adverse effects of nicklecillin in 24 adult patients

| No. of patients | Side effect |
|-----------------|--------------------------------------|
| 14 | Diarrhea |
| 5 | Eosinophilia (35 eos/mm 3) |
| 2 | Metallic taste |
| 1 | Yeast vaginitis |
| 1 | Mild rise in urea nitrogen |
| 1 | Hematuria (8–10 rbc/hpf) |

- La “lista de palabras”: esta información puede presentarse en el texto
- Muchas veces la información ya figura naturalmente en el texto

Tablas poco útiles

Table 5

Simulation results for using full data, CRs only, and proposed method under four missing mechanisms

| Method | Bias ^a | | Variance ^b | | 95% CI ^c | |
|--|-------------------|-------------------|-----------------------|-------------------|---------------------|-------------------|
| | $(\hat{\beta}_W)$ | $(\hat{\beta}_X)$ | $(\hat{\beta}_W)$ | $(\hat{\beta}_X)$ | $(\hat{\beta}_W)$ | $(\hat{\beta}_X)$ |
| (M.1) $P(R = 1) = 0.66$ | | | | | | |
| Full | 0.01346 | 0.02229 | 0.04008 | 0.03685 | 0.955 | 0.950 |
| Comp | 0.03062 | -0.003561 | 0.1149 | 0.06732 | 0.960 | 0.955 |
| Impu | 0.01431 | 0.021 | 0.04088 | 0.05169 | 0.980 | 0.975 |
| (M.2) $\text{logit } P(R = 1) = 2Y$ | | | | | | |
| Full | 0.007908 | -0.02116 | 0.03838 | 0.03624 | 0.975 | 0.925 |
| Comp | 0.01945 | 0.07096 | 0.107 | 0.06581 | 0.960 | 0.950 |
| Impu | 0.006966 | 0.01597 | 0.04227 | 0.05226 | 0.975 | 0.985 |
| (M.3) $\text{logit } P(R = 1) = 2X$ | | | | | | |
| Full | 0.007908 | -0.02116 | 0.03838 | 0.03624 | 0.975 | 0.925 |
| Comp | 0.01225 | 0.0589 | 0.08856 | 0.06818 | 0.980 | 0.975 |
| Impu | 0.009563 | -0.04699 | 0.03865 | 0.04923 | 0.985 | 0.970 |
| (M.4) $\text{logit } P(R = 1) = X + Y$ | | | | | | |
| Full | 0.01346 | 0.02229 | 0.04008 | 0.03685 | 0.955 | 0.950 |
| Comp | 0.02404 | 1.613 | 0.1102 | 0.08202 | 0.955 | 0.580 |
| Impu | 0.01814 | 0.08289 | 0.0578 | 0.06075 | 0.955 | 0.970 |

^aBias = $(\hat{\beta} - \beta_0)/\beta_0$.

^bSimulation variance.

^cConfidence interval using jackknife standard error.

Paik, M. (2004). *Nonignorable Missingness in Matched Case-Control Data Analyses*. *Biometrics*, 60(2), 306-314, Tabla 5

- Se presentan demasiados dígitos!
- Falta consistencia! Si todos los dígitos en 0.04227 fueran significativos, entonces 0.107 debería ser 0.10700

Reglas básicas y buenas prácticas

- En la tabla mostrar solamente contenido relevante a la hipótesis del experimento
- La cantidad de contenido debe ser razonable
 - Si es muy poco, el lector puede no entender la tabla
 - Si es demasiado podemos llegar a confundir al lector
- Demasiadas columnas o filas dificultan su entendimiento
 - En estos casos es recomendable reducir la cantidad de datos o separarlos en diferentes tablas
- No usar una tabla si se tienen 2 (o menos) columnas o filas
 - En estos casos es suficiente una descripción textual
 - Es recomendable combinar tablas si se tienen columnas o filas idénticas en dos o más tablas

Reglas básicas y buenas prácticas

- Debe ser autocontenida y no debe depender de nada explicado en el texto para ser comprendida
- El significado de una buena tabla debe ser obvio para el lector
 - Si el lector debe perder tiempo descifrando una tabla, entonces no es efectiva
- La tabla y el texto deben complementarse, y no replicarse
 - En el texto se deben destacar y discutir solo los resultados más importantes de las tablas
- Debe estar ubicada cerca del texto donde es referenciada por primera vez
- Todas las tablas deben estar referenciadas en el texto
 - Si no es necesario referenciar una tabla, seguramente esa tabla no sea necesaria

Reglas básicas y buenas prácticas

- Las tablas deben referenciarse por su número
 - Nunca referenciar una tabla como «la siguiente» o «la que aparece a continuación»
- Se pueden usar expresiones del estilo de:
 - «Results are summarized/reported/presented/shown in Table 1»
 - «Table 2 summarizes/reports/presents/shows»

Reglas básicas y buenas prácticas

- Distinguir y resaltar los elementos principales de la tabla
 - Utilizar diferentes fuentes, estilos o colores para enfatizar los principales elementos

Table 6.1. Anatomy of a table

Households by type, race and Hispanic origin (thousands), United States, 1997

| Characteristic | All households | Family households | | | | Nonfamily household | | |
|------------------------|----------------|-------------------|----------------|--------------------|------------------|---------------------|--------------------|------------------|
| | | Total | Married couple | Other families | | Total | Female householder | Male householder |
| | | | | Female householder | Male householder | | | |
| <i>Race/ethnicity</i> | | | | | | | | |
| White | 86,106 | 59,511 | 48,066 | 8,308 | 3,137 | 26,596 | 14,871 | 11,725 |
| Non-Hispanic White | 77,936 | 52,871 | 43,423 | 6,826 | 2,622 | 25,065 | 14,164 | 10,901 |
| Black | 12,474 | 8,408 | 3,921 | 3,926 | 562 | 4,066 | 2,190 | 1,876 |
| All other ^a | 3,948 | 2,961 | 2,330 | 418 | 212 | 986 | 455 | 532 |
| <i>Origin</i> | | | | | | | | |
| Non-Hispanic | 93,938 | 63,919 | 49,513 | 11,040 | 3,366 | 30,018 | 16,762 | 13,258 |
| Hispanic ^b | 8,590 | 6,961 | 4,804 | 1,612 | 545 | 1,630 | 754 | 875 |
| Total | 102,528 | 70,880 | 54,317 | 12,652 | 3,911 | 31,648 | 17,516 | 14,133 |

Source: U.S. Census Bureau 1998.

^a“All other” races includes Asians, Pacific Islanders, Native Americans, and those of unspecified race.

^b People of Hispanic origin may be of any race.

¿Debería usar una tabla?

- Las tablas y el texto son buenas para presentar información detallada
- Las figuras son buenas para presentar los principales resultados, tendencias, una relación entre variables o conclusiones
- Si el texto está saturado de detalles, quizás sea más adecuado presentar esa información en una tabla
- Si una tabla tiene pocas filas o columnas, quizás sea más adecuado presentar sus datos en el texto
- Consolidar información relacionada en una tabla permite al lector hacer comparaciones fácilmente

La regla de oro

- Presentar los datos una única vez: en el texto, o en una tabla, o en una figura
- Nunca presentar los mismos datos varias veces!
- Aunque...
 - Algunos datos seleccionados pueden ser discutidos en el texto, por su relevancia, o pueden ser graficados para destacar su importancia

¿Cómo comparar resultados numéricos?

- El primer paso para comparar resultados numéricos es decidir qué datos comparar
- El valor de referencia puede ser arbitrario, pero debe ser consistente
- Se debe nombrar el valor de referencia en todas las comparaciones numéricas
- P. ej.: al comparar la región medio oeste con el resto de las regiones, se podría escribir: *“The Northeast is [measure of difference] larger (or smaller) than the Midwest”*
 - Sin la frase *“than the Midwest”* no queda claro con quien se hace la comparación

¿Cómo comparar resultados numéricos?

- Varias formas comúnmente usadas: **rango**, **diferencia absoluta**, **diferencia relativa**, **diferencia porcentual** y **z-score**
- **Rango** es la posición de un valor comparado en un conjunto ordenado de valores
- Al comparar el rango se debe describir la posición relativa y el concepto comparado
 - “más rápido”, “menos caro”, “segundo más denso”
- El rango es útil cuando solo importa el orden de los valores y no las distancias entre ellos

¿Cómo comparar resultados numéricos?

- La **diferencia absoluta** resta el valor de referencia (X) del valor de interés (Y)
- Es útil cuando la diferencia en si misma es de interés
 - ¿Cuánto más costará algo?
 - ¿Está aún dentro de mi presupuesto?
- Sin embargo, la diferencia absoluta no responde la pregunta: ¿Es importante una diferencia de \$0.50?
 - Depende del valor de X e Y. Para un caramelo que cuesta menos de \$1, un incremento de \$0.50 es sustancial. Para un auto que cuesta \$20 000, el mismo incremento de \$0.50 es despreciable

¿Cómo comparar resultados numéricos?

- La **diferencia relativa** es el ratio entre el número de interés (Y) y el número de referencia (X)
- Siempre describir en función de múltiplos
 - Evitar usar palabras como “ratio”, “relative difference”, etc.
- *“In 1990, the ratio between the South and the Midwest was 1.43”*
 - *Pobre*: No especifica que es lo que se está midiendo ni cual región tiene un valor superior
- *“In 1990, the South was 1.43 times as populous as the Midwest”*
 - *Mejor*: Es inmediatamente evidente que es lo que se está comparando, cual región es la mayor y por cuanto

¿Cómo comparar resultados numéricos?

- La **diferencia porcentual** expresa la diferencia entre un valor en relación con el valor de referencia, calculando $\left[\frac{(Y-X)}{X}\right] \times 100$
- Siempre identificar los casos comparados y la dirección del cambio
- *“The Western population changed 15.9%”*
 - *Pobre:* No especifica en qué período ocurrió ese cambio, ni su dirección
- *“During the 1990s, the population of the West region grew from 52.8 million to 61.2 million— an increase of 15.9%”*
 - *Mejor:* Esta versión reporta el valor, el cambio porcentual y su dirección

¿Cómo comparar resultados numéricos?

Mas ejemplos de **diferencia porcentual**...

- *“In 1990, the populations of the West and Midwest were 52.8 million and 59.7 million, respectively, so the percentage difference between the West and the Midwest is negative (-11.6%)”*
 - *Pobre:* Diferencias porcentuales negativas son difíciles de describir y poco intuitivas
- *“In 1990, the West had 11.6% fewer inhabitants than the Midwest (52.8 million persons and 59.7 million persons, respectively)”*
 - *Mejor:* “fewer inhabitants than the Midwest” claramente indica la dirección de la diferencia entre las poblaciones

¿Cómo comparar resultados numéricos?

- **Z-score** (o **Standardized scores**) compara un valor particular con el promedio, teniendo en cuenta la desviación estándar
- Se calcula $z = \frac{x - \mu}{\sigma}$, donde μ es el promedio y σ es la desviación estándar
 - Un z-score positivo indica un valor por encima de la media, mientras que uno negativo uno por debajo
- La unidad de z-score es múltiplos de la desviación estándar (ya no de la unidad original en la que fue medida la variable)
 - *“With a raw score of 47, Mia scored one standard deviation below the national average.”*

¿Cómo comparar resultados numéricos?

- **Regla de oro:** presentar valores relativos solo si son significativos
- Por ejemplo:
 - *«33,3% de la población de ratones fue curada por la droga testeada;*
 - *33,3% de la población no fue afectada por la droga y permaneció en estado moribunda;*
 - *el tercer ratón se escapó»*

¿Cómo presentar resultados estadísticos?

- **Media o mediana, esa es la cuestión**
- Si los datos siguen una distribución normal: usar media y desviación estándar
- Si NO siguen una distribución normal: usar mediana y rango intercuartílico
- Por ejemplo:
 - En un estudio sobre dolor de espalda durante el embarazo, se le pidió a un grupo de mujeres que intentara tocar el suelo con la punta de los dedos sin flexionar las rodillas y se midió la distancia desde la punta de los dedos hasta el piso.
 - Se reportó la media y desviación estándar: 12 ± 14 cm
 - ¡¿Se constataron mediciones entre -2 y 26 cm?!

¿Cómo presentar resultados estadísticos?

- La media es particularmente susceptible a muestras sesgadas (skewed) o valores no usuales (outliers). Por ejemplo:

| | | | | | | | | | | |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Empleado | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Sueldo | 15k | 18k | 16k | 14k | 15k | 15k | 12k | 17k | 90k | 95k |

- El salario medio es de \$30.7k
- Resulta claro que la mayoría de los trabajadores ganan entre \$12k y \$18k
 - La media está sesgada por los dos salarios más grandes
- Para este ejemplo la mediana es \$15k

¿Cómo presentar resultados estadísticos?

- Distinguir entre significancia estadística e importancia práctica
- Un resultado puede ser significativo estadísticamente pero resultar inconcluyente en la práctica
- Por ejemplo:
 - Una droga resultó bajar la presión arterial sistólica en una media de 8 mm Hg, de 100 a 92 mm Hg, de forma estadísticamente significativa ($P < 0.05$)
 - El intervalo de 95% de confianza se encuentra entre 2 y 14 mm Hg. Una baja en la presión arterial de 14 mm Hg es clínicamente importante, pero una baja de 2 mm Hg no lo es
 - *“La presión arterial sistólica disminuyó en una media de 8 mm Hg, desde 100 a 92 mm Hg (95% IC=2 a 14 mm Hg; $P=0.05$)”*

¿Cómo presentar resultados estadísticos?

- ¿ $P < 0.05$ es la verdad?
- Comúnmente aceptado que $P < 0.05 \Rightarrow$ la verdad
 - Mientras que $P > 0.05 \Rightarrow$ no publicable
- Sin embargo P valores de 0.04-0.06 deberían tener interpretaciones similares y no radicalmente diferentes
- Es más adecuado presentar P valores exactos (=)
 - Reportar utilizando el signo $<$ solo en casos extremos, $P < 0.001$