

BIOESTADÍSTICA
Sin Lágrimas



¿Ramras de la estadística?

La Estadística es la ciencia, con base matemática, referente a la recolección, análisis e interpretación de datos, que busca explicar condiciones regulares en fenómenos de tipo aleatorio.

Descriptiva

• **Sistematización, recogida, ordenación y presentación** de los datos referentes a un fenómeno que presenta variabilidad o incertidumbre para su estudio metódico, con objeto de

Probabilidad

deducir las leyes que rigen esos fenómenos,

Inferencia

• y poder de esa forma hacer previsiones sobre los mismos, tomar **decisiones** u obtener **conclusiones**.

¿Para qué sirve la estadística?

La Ciencia se ocupa en general de fenómenos observables

Ciencia se desarrolla observando hechos, formulando leyes que los explican y realizando experimentos para validar o rechazar dichas leyes

Los modelos que crea la ciencia son de tipo determinista o aleatorio (estocástico)

La Estadística se utiliza como tecnología al servicio de las ciencias donde la variabilidad y la incertidumbre forman parte de su naturaleza

“La Bioestadística[...] enseña y ayuda a investigar en todas las áreas de las Ciencias de la Vida donde la variabilidad no es la excepción sino la regla” Carrasco de la Peña (1982)

La Estadística para Ciencias de la Salud

- Sistematización, recogida, ordenación y presentación de los datos referentes a un fenómeno que presenta variabilidad o incertidumbre para su estudio metódico, con objeto de

deducir las leyes que rigen esos fenómenos y poder de esa forma hacer previsiones sobre los mismos, tomar decisionesu obtener conclusiones.

Método científico y estadística



PLANTEAR
HIPÓTESIS



DISEÑO
EXPERIMENTO



RECOGER
DATOS



ANÁLISIS



CONCLUSIÓN

Pasos en un estudio estadístico

A. Plantear hipótesis sobre una población

- Los fumadores tienen “más bajas” laborales que los no fumadores
- ¿En qué sentido? ¿Mayor número? ¿Tiempo medio?

B. Decidir qué datos recoger (diseño de experimentos)

- Qué individuos pertenecerán al estudio (muestras)
 - Fumadores y no fumadores en edad laboral.
 - Criterios de exclusión ¿Cómo se eligen? ¿Descartamos los que padecen enfermedades crónicas?
- Qué datos recoger de los mismos (variables)
 - Número de bajas
 - Tiempo de duración de cada baja
 - ¿Sexo? ¿Sector laboral? ¿Otros factores?

- Recoger los datos (muestreo)
 - ¿Estratificado? ¿Sistemáticamente?
- Describir (resumir) los datos obtenidos
 - tiempo medio de baja en fumadores y no (estadísticos)
 - % de bajas por fumadores y sexo (frecuencias), gráficos,...
- Realizar una inferencia sobre la población
 - Los fumadores están de baja al menos 10 días/año más de media que los no fumadores.
- Cuantificar la confianza en la inferencia
 - Nivel de confianza del 95%
 - Significación del contraste: $p=2$ al 5%

Clasificación

- **Estadística descriptiva**

La Estadística Descriptiva es aquella rama de la Estadística mediante la cual se trata de Presentar, Describir y Resumir datos obtenidos en una investigación.

- **Estadística inferencial**

La Inferencia Estadística es aquella rama de la Estadística mediante la cual se trata de sacar conclusiones de una población en estudio, a partir de la información de la muestra

Conceptos Básicos

- **Unidad de Análisis:** Personas, u objetos que contienen cierta información que se desea estudiar. Por ejemplo: un paciente, un establecimiento de salud, un equipo etc.
- **Población:** Es la totalidad de unidades de análisis que poseen algunas características comunes, susceptibles de observación. Una población se define en términos de inclusión y en términos de exclusión.

Población y Muestra



- **Población** (*‘population’*) es el conjunto sobre el que estamos interesados en obtener conclusiones (hacer inferencia).
 - Normalmente es demasiado grande para poder abarcarlo
- **Muestra** (*‘sample’*) es un subconjunto suyo al que tenemos acceso y sobre el que realmente hacemos las observaciones (mediciones)
 - Debería ser “representativo”
 - Está formado por miembros “seleccionados” de la población (individuos, unidades experimentales).

CONCEPTOS FUNDAMENTALES: Población

Conjunto definido de TODOS los INDIVIDUOS (personas, animales, objetos, plantas u organismos), de donde se observa cierta característica. Se representa con la letra **N**.

Finita

incluye una cantidad limitada contable de observaciones, individuos o medidas. Siempre que sea posible alcanzar (contar) el número total de todas las posibles mediciones, se considera como finita la población.



Infinita

incluye un gran conjunto de observaciones o mediciones que no pueden alcanzarse por conteo. Al menos, hipotéticamente, no existe límite en cuanto al número de observaciones que el experimento puede generar.



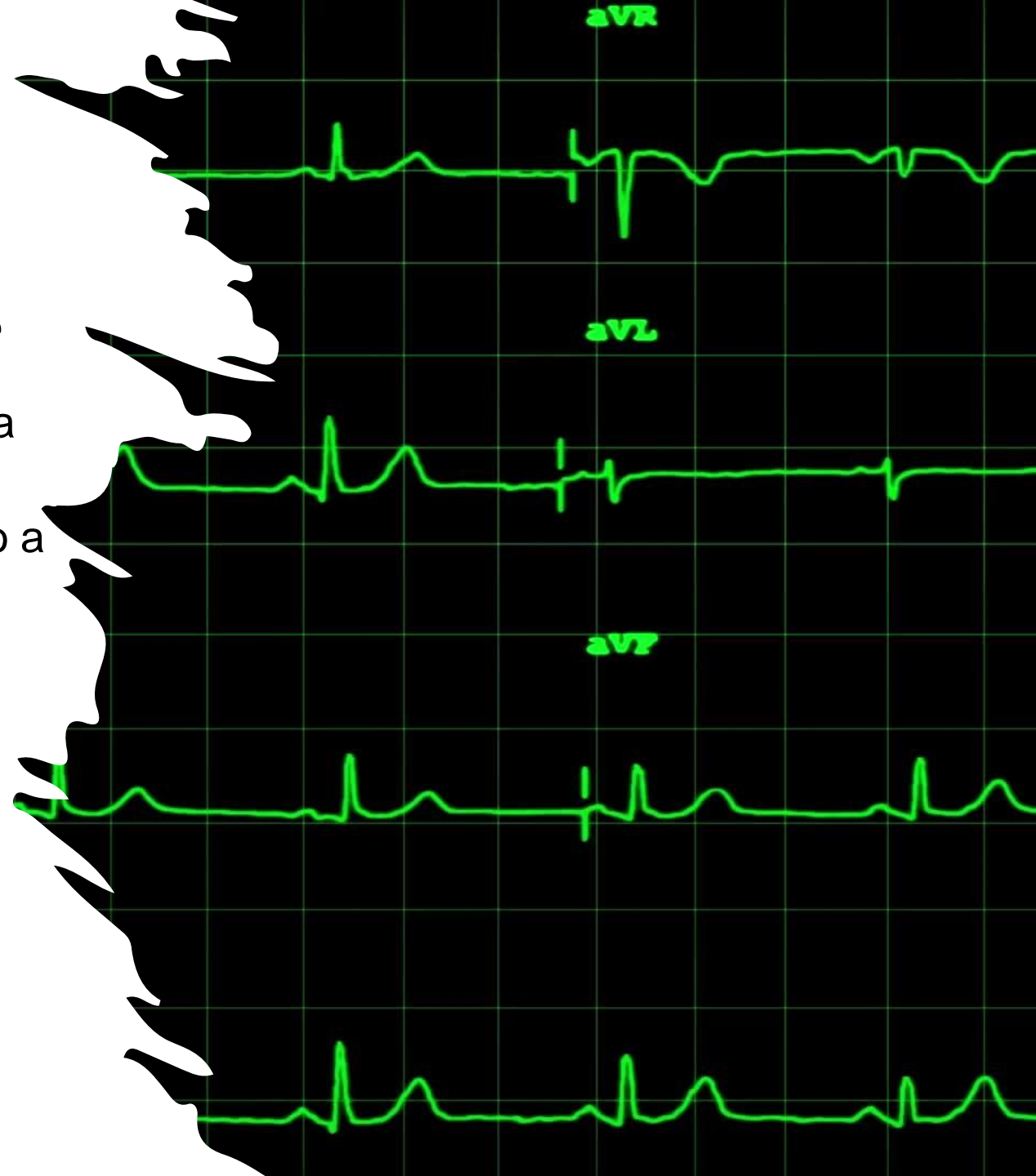
Conceptos Básicos

- **Parámetro:** es una medida de resumen que describe a una población.
- **Estadístico:** es una medida de resumen que describe a una muestra.

PARÁMETRO	ESTADÍSTICO
μ : Media poblacional	\bar{X} : Media muestral
σ^2 : Varianza poblacional	S^2 : Varianza muestral
σ : Desviación poblacional	S : Desviación muestral
P : Proporción poblacional	p : Proporción muestral

Variables

- Una **variable** es una característica observable *que varía entre los diferentes individuos* de una población. La información que disponemos de cada individuo es resumida en **variables**.
- En los individuos de la *población* española, de uno a otro **es variable**:
 - El grupo sanguíneo
 - {A, B, AB, O} Var. Cualitativa
 - Su nivel de felicidad “declarado”
 - {Deprimido, Ni fu ni fa, Muy Feliz} Var. Ordinal
 - El número de hijos
 - {0,1,2,3,...} Var. Numérica discreta
 - La altura
 - {1'62 ; 1'74; ...} Var. Numérica continua





- **Variable es una cosa que varía**
- Esta variación es susceptible de medirse.
- Ejemplos: Presión sanguínea, frecuencia cardíaca, peso de niños, edad de los pacientes que acuden a consulta

TIPOS DE VARIABLES

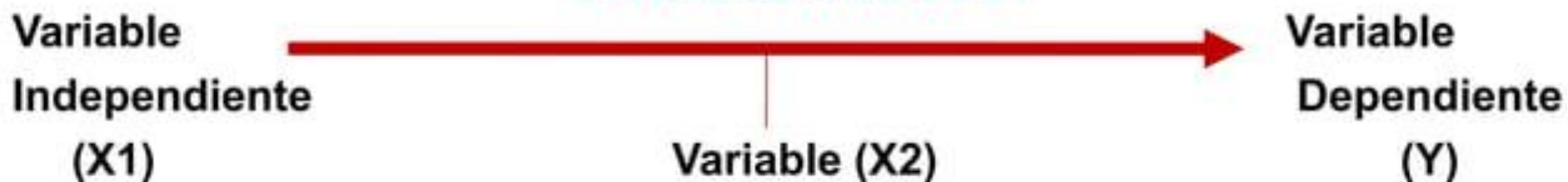
- POR SU INTERRELACION:

1. **INDEPENDIENTES:** determina cambio de valores en dependiente (es la causa).

2. **DEPENDIENTE:** está en función de la independiente (es el efecto).

3. **INTERVINIENTE:** influye ampliando o amortiguando efecto de la independiente.

VARIABLES INDEPENDIENTE Y DEPENDIENTE

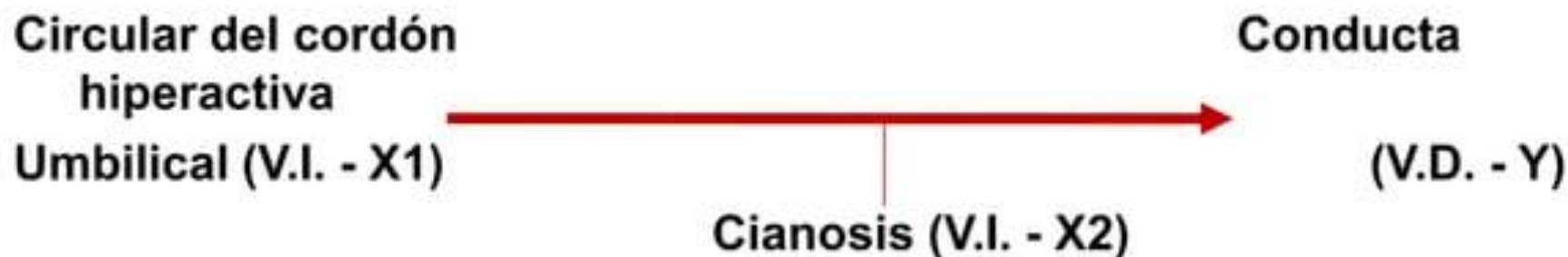


Ejemplo:

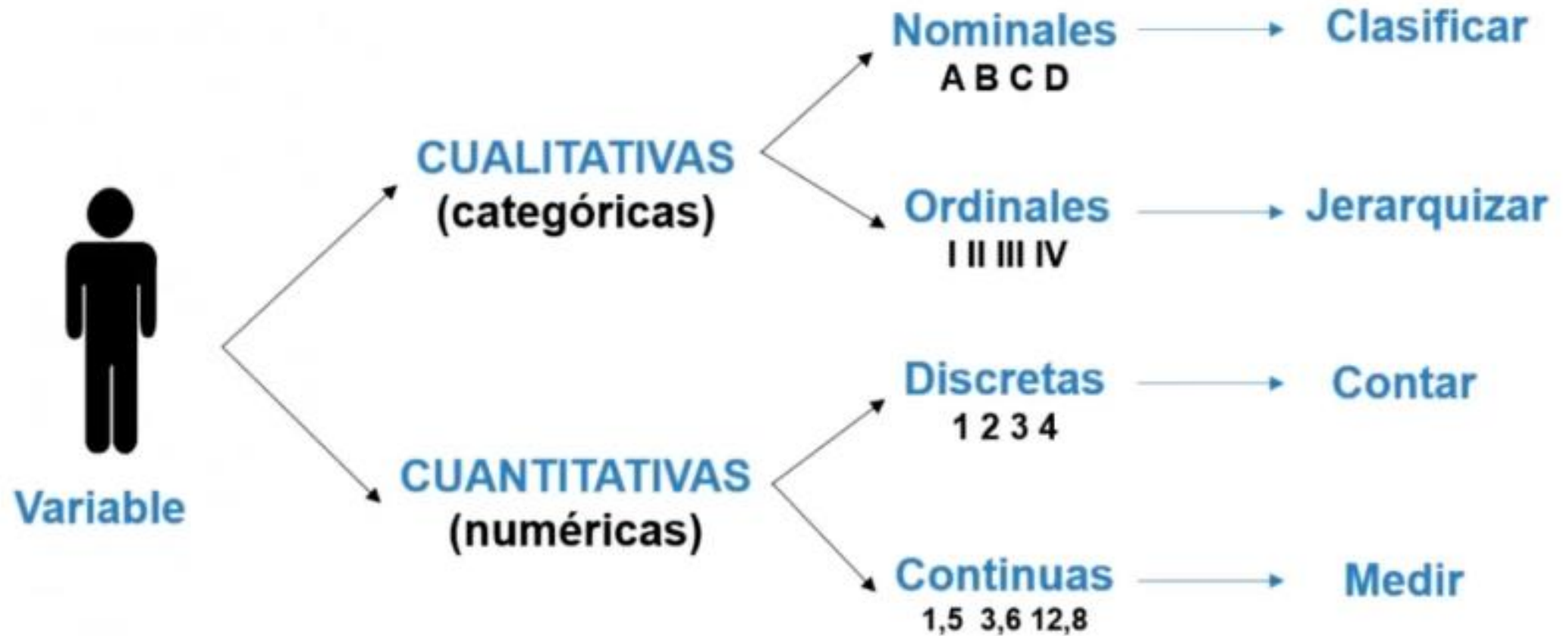
“Los alumnos que estudian en grupo obtienen mejores calificaciones que los que estudian individualmente”



“Los niños nacidos con circulación del cordón umbilical, desarrollan una conducta hiperactiva e irritable, siempre y cuando presenten cianosis al momento de nacer”



Tipos de variables



• Cualitativas

Si sus valores (*modalidades*) no se pueden asociar naturalmente a un número (**no se pueden hacer operaciones algebraicas con ellos**)

- **Nominales**: Si sus valores no se pueden ordenar
 - Sexo, Grupo Sanguíneo, Religión, Nacionalidad, Fumar (Sí/No)
- **Ordinales**: Si sus valores se pueden ordenar
 - Mejoría a un tratamiento, Grado de satisfacción, Intensidad del dolor



• **Cuantitativas o Numéricas**
Si sus valores son numéricos
(tiene sentido hacer
operaciones algebraicas con
ellos)

• **Discretas:** Si toma valores enteros

- Número de hijos, Número de cigarrillos, Num. de "cumpleaños"

• **Continuas:** Si entre dos valores, son posibles infinitos valores intermedios.

- Altura, Presión intraocular, Dosis de medicamento administrado, edad

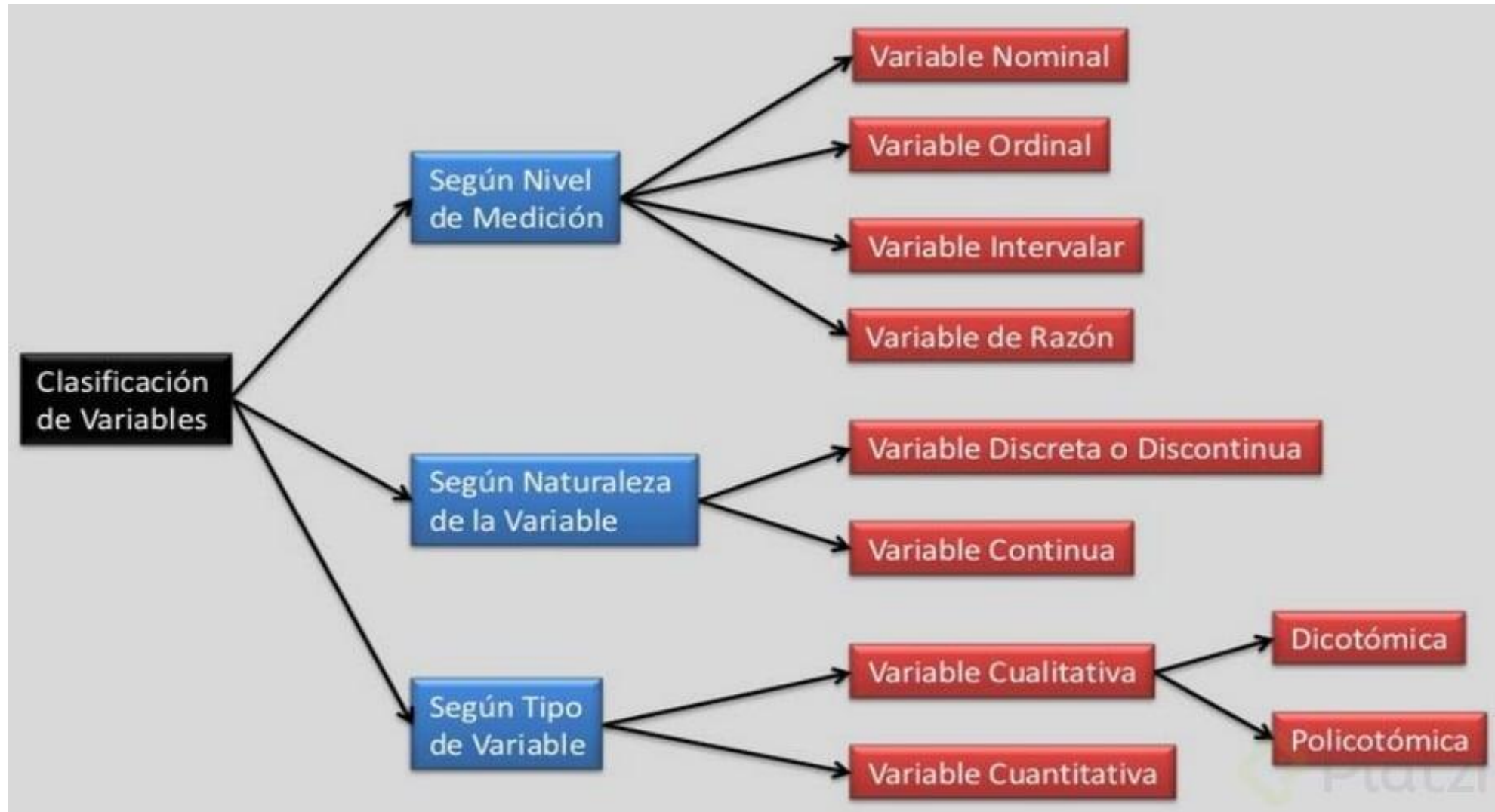


TIPOS DE VARIABLES

ESCALAS DE MEDICION

1. **NOMINAL**: se clasifican según categorías, se utiliza para v. cualitativas.
2. **ORDINAL**: clasifica en forma jerárquica.
3. **INTERVALO**: punto cero es arbitrario y convencional.
4. **PROPORCION (RAZON)**: tiene cero absoluto, existen proporciones.

Tipos de variables





Evaluación de Variables

- <https://forms.office.com/r/PDMyJ8uL6L?origin=lprLink>

OPERACIONALIZACION DE VARIABLES

- Es el proceso de llevar una variable del nivel abstracto a un plano concreto.
- Operacionalizar una variable es hacerla medible.

ESQUEMA DE OPERACIONALIZACION DE VARIABLES

Existen numerosos modelos y esquemas de Operacionalización. Uno práctico que se sugiere es el siguiente:

1. Variable
2. Definición
3. Dimensiones
4. Indicadores
5. Sub indicadores
6. Escala de Medición

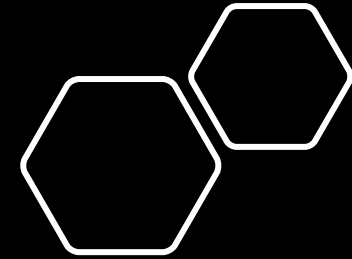
Variables	Definición	Dimensiones	Indicadores	Sub indicadores	Escala de Medición
Accesibilidad a los servicios de salud	Mayor o menor posibilidad de tomar contacto con los S.S. para recibir asistencia médica	Accesibilidad Geográfica	Tiempo medido en horas y minutos que tarda la persona en trasladarse de su domicilio al SS	Horas y minutos	Razón
		Accesibilidad Económica	Cantidad de dinero que gasta para recibir la atención	Monto en soles	Razón
		Accesibilidad Cultural	Percepción del problema de salud	Considera problema No considera problema	Nominal
Nivel de Glicemia	Nivel de glucosa en sangre cuantificado en ayunas en la mañana.		Hipoglicemia Normoglicemia Hiperglicemia	<70 mg/dl 70-110 mg/dl >110 mg/dl	Ordinal

EJERCICIO 1

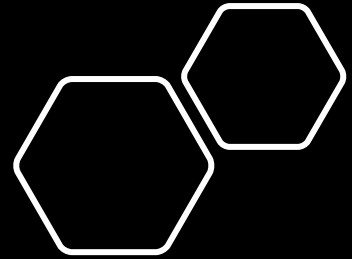
- III. Una especialista en nutrición está interesada en comparar la efectividad de tres dietas en la reducción del peso en señoras de 18 a 20 años de edad del nivel socioeconómico alto. Para tal efecto, se recurre a una muestra de 45 señoras, de tal manera que a cada dieta se le asigna 15 señoras. Se pide determinar:
- a. Unidad de análisis.
 - b. Población objeto de estudios.
 - c. Muestra
 - d. Variables
 - e. Tipos de variables
 - f. Escala de medición de las variables respectivas
 - g. En el estudio identifique variable independiente y dependiente

FUENTES DE INFORMACION

- Cuando se recoge información es posible que los datos ya hayan sido obtenidos de las unidades de observación y publicados o registrados por otros; o que los datos se tengan que obtener directamente de las unidades de observación en forma personal.
 - Fuente Primaria
 - Fuente Secundaria



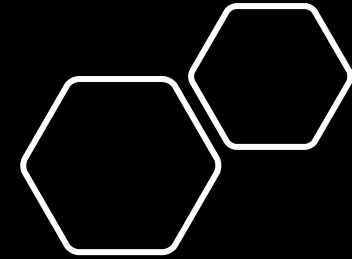
FUENTE PRIMARIA



- Si la información se obtiene directamente de la misma persona o entidad utilizando ciertas técnicas (encuesta, entrevistas, discusión grupal, etc..)
- Ejemplo: Observar la reacción de un grupo de pacientes con determinada enfermedad.

FUENTE SECUNDARIA

- Si la información a obtener ya ha sido recopilada y elaborada por otras personas o instituciones.
- Este tipo de información la encontramos en los informes estadísticos de las instituciones públicas y privadas. En el Perú, el Instituto Nacional de Estadística e Informática publica boletines estadísticos con respecto a la Producción, Inversión, población, migración, salud, etc..



Técnicas e instrumentos de Recolección de Datos

TÉCNICAS: corresponde a las distintas maneras de obtener los datos que luego de ser procesados, se convertirán en información. Entre estas se tiene: la observación (participante y no participante), la entrevista, la encuesta, y la discusión grupal.

INSTRUMENTOS: medios materiales que se emplean para la recolección de datos. Entre estos se tiene: el guión de observación, la lista de cotejo, el cuestionario, la guía de entrevista o guión de entrevista, el guión de discusión grupal.

TECNICAS DE RECOLECCIÓN DE DATOS



Observación

- Capta realidad sin distorsionar información
- Verdadera Realidad



Encuesta

- Información relativa a características predominantes de una población



Entrevista

- Comunicación verbal
- Entrevistador bien informado

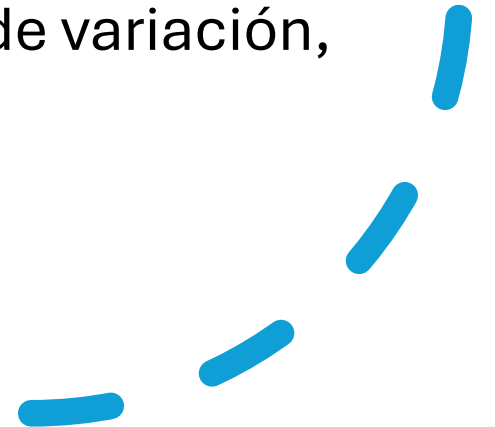


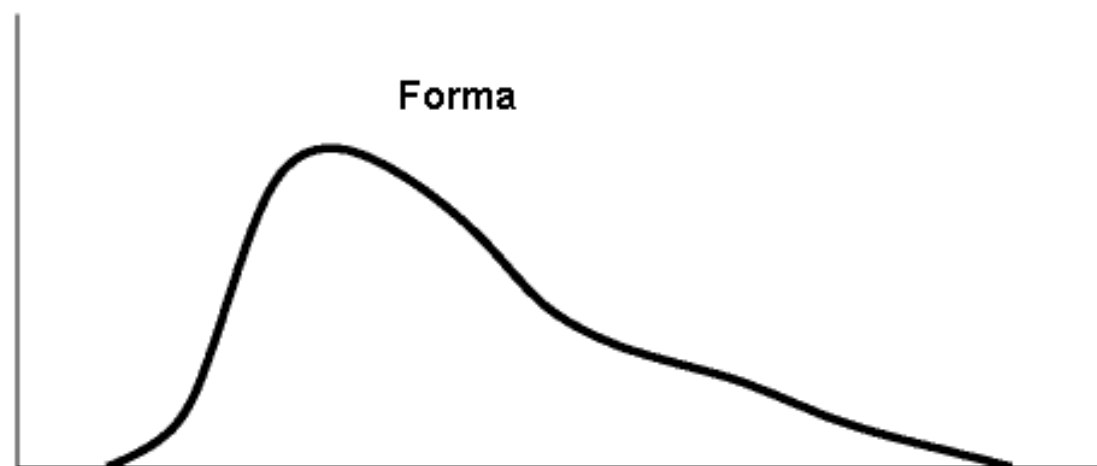
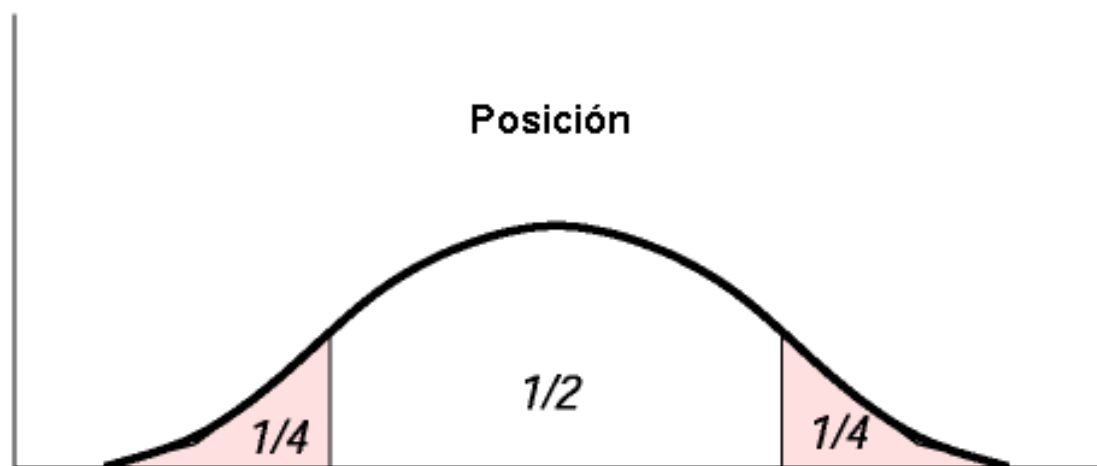
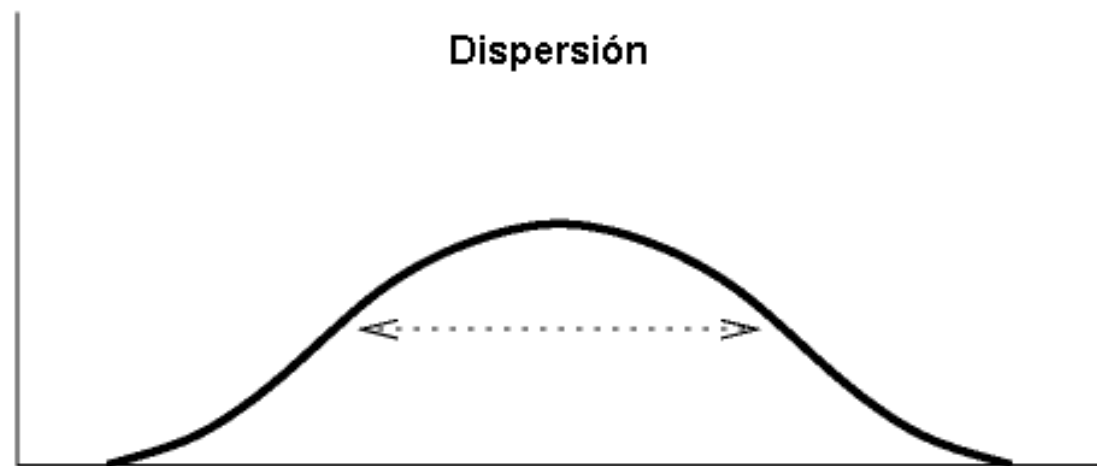
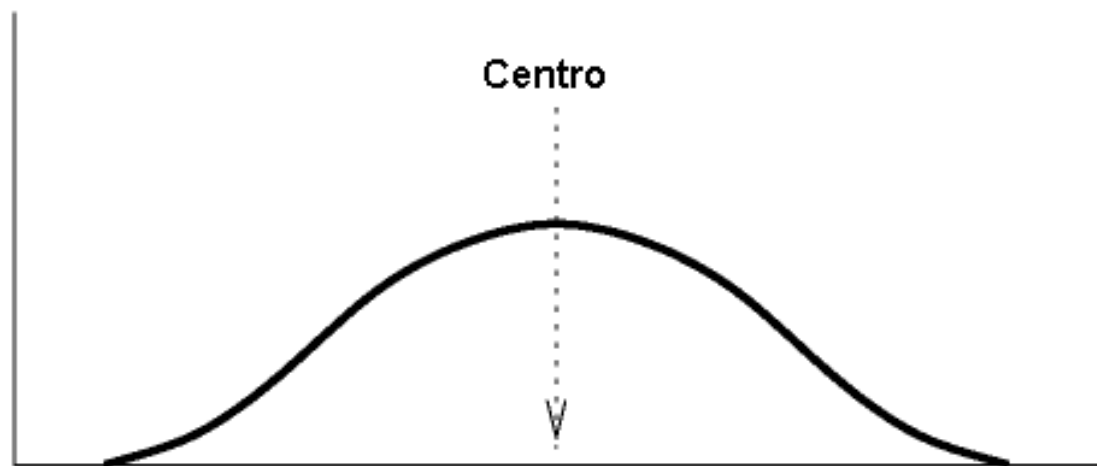
Conversación Grupal

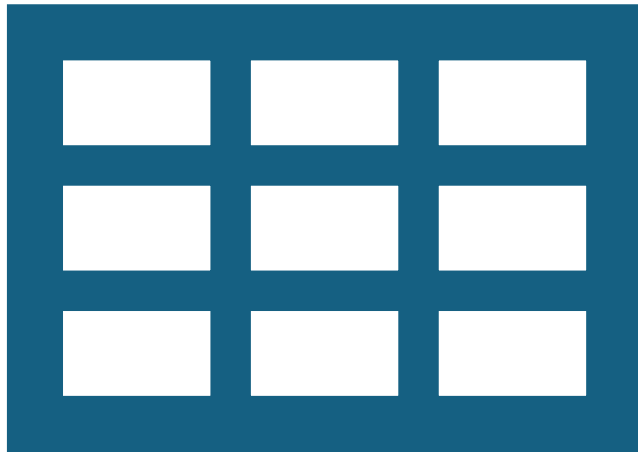
- Captar representación afectiva, valores e ideológica de una sociedad

Un brevísimo resumen sobre estadísticos

- Posición
 - Dividen un conjunto ordenado de datos en grupos con la misma cantidad de individuos.
 - Cuantiles, percentiles, cuartiles, deciles,...
- Centralización
 - Indican valores con respecto a los que los datos parecen agruparse.
 - Media, mediana y moda
- Dispersión
 - Indican la mayor o menor concentración de los datos con respecto a las medidas de centralización.
 - Desviación típica, coeficiente de variación, rango, varianza
- Forma
 - Asimetría
 - Apuntamiento o curtosis







Tablas de frecuencia

- Tabla que resume información de los datos obtenidos, **agrupándolos en categorías o clases**.
- **Frecuencia es el número de datos que hay en cada categoría.**
 - Σ = La suma de todos los datos debe ser el total de los datos originales.
 - **fi** = número de individuos en la categoría (frecuencia absoluta)
 - **fri** = frecuencia relativa f_i/n
 - **pi** = % de individuos en la categoría ($f_{ri} \cdot 100$)
 - **Fi** = suma de individuos en las categorías anteriores y la que estoy (frecuencia acumulada)
 - **Fri** = F_i/n (frecuencia relativa acumulada)

Para datos numéricos, se hacen intervalos que representan una categoría.

Ejemplo

Tabla de frecuencia: Interes profesional

Categoría	fi	fri	pi %	Fi	Fri
Silvestres	10	0,30	30,3	10	0,30
Aves	3	0,09	9,1	13	0,39
Menores	14	0,42	42,4	27	0,82
Equinos	2	0,06	6,1	29	0,88
Mayores	1	0,03	3,0	30	0,91
Produccion	1	0,03	3,0	31	0,94
Reproduccion	2	0,06	6,1	33	1,00
n	33				

**Medidas
descriptivas**

**Medidas de
tendencia
central**

**Medidas de
dispersión**

Media

Mediana

Moda

Rango

**Desviación
standard,
varianza**

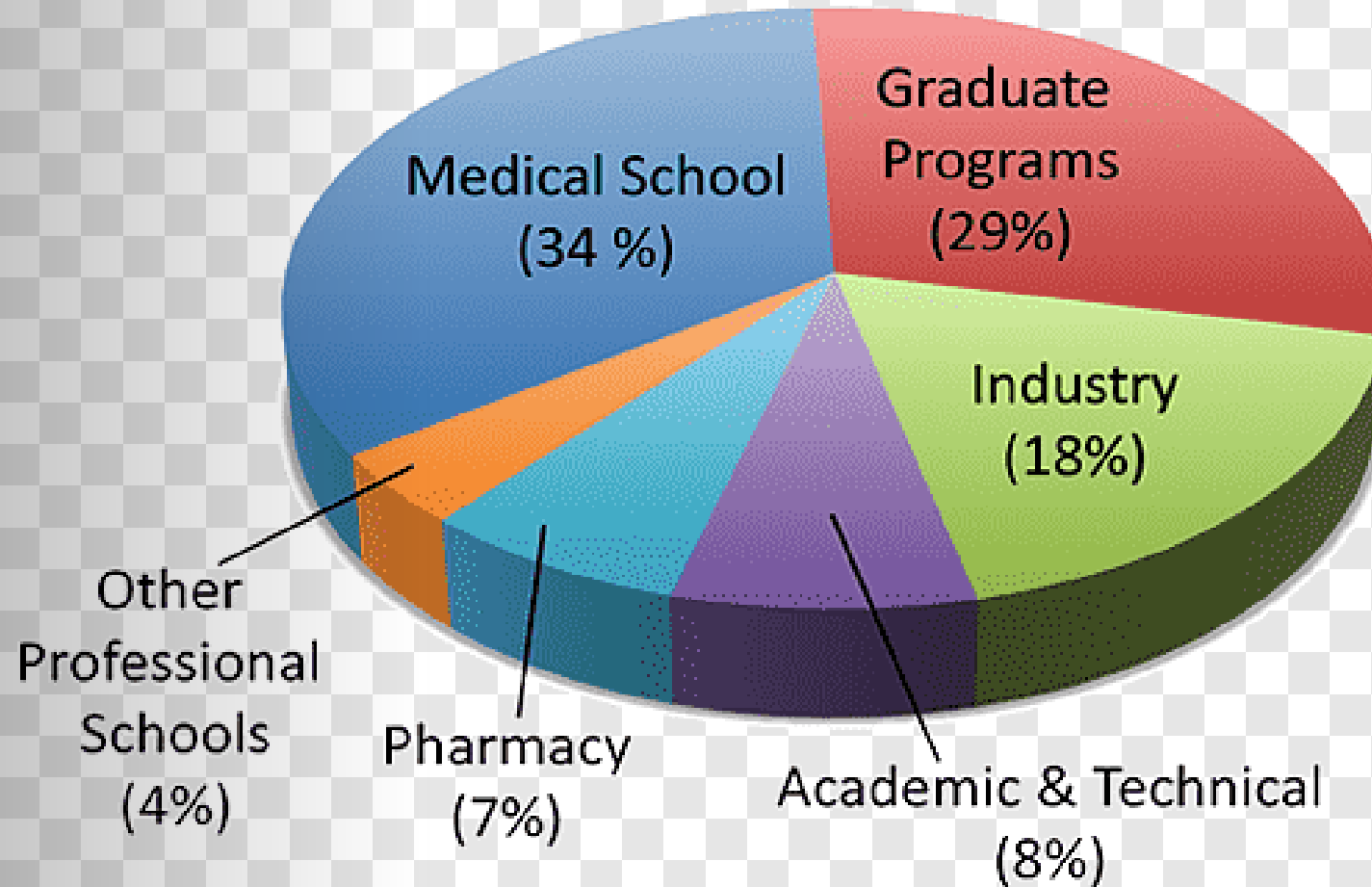
**Rango
intercuartílico**

Gráficos estadísticos

- Gráfico de pastel
 - Resumir datos **cualitativos**.
 - Círculo dividido en todas las categorías de la variable.
 - Mejor con pocas categorías.
 - Número de individuos que corresponde a cada una de las categorías brinda el tamaño de cada pedazo.
 - Debe llevar: nombre de la categoría, número de individuos o elementos y porcentaje.



Ejemplo Career Paths of Our Graduates



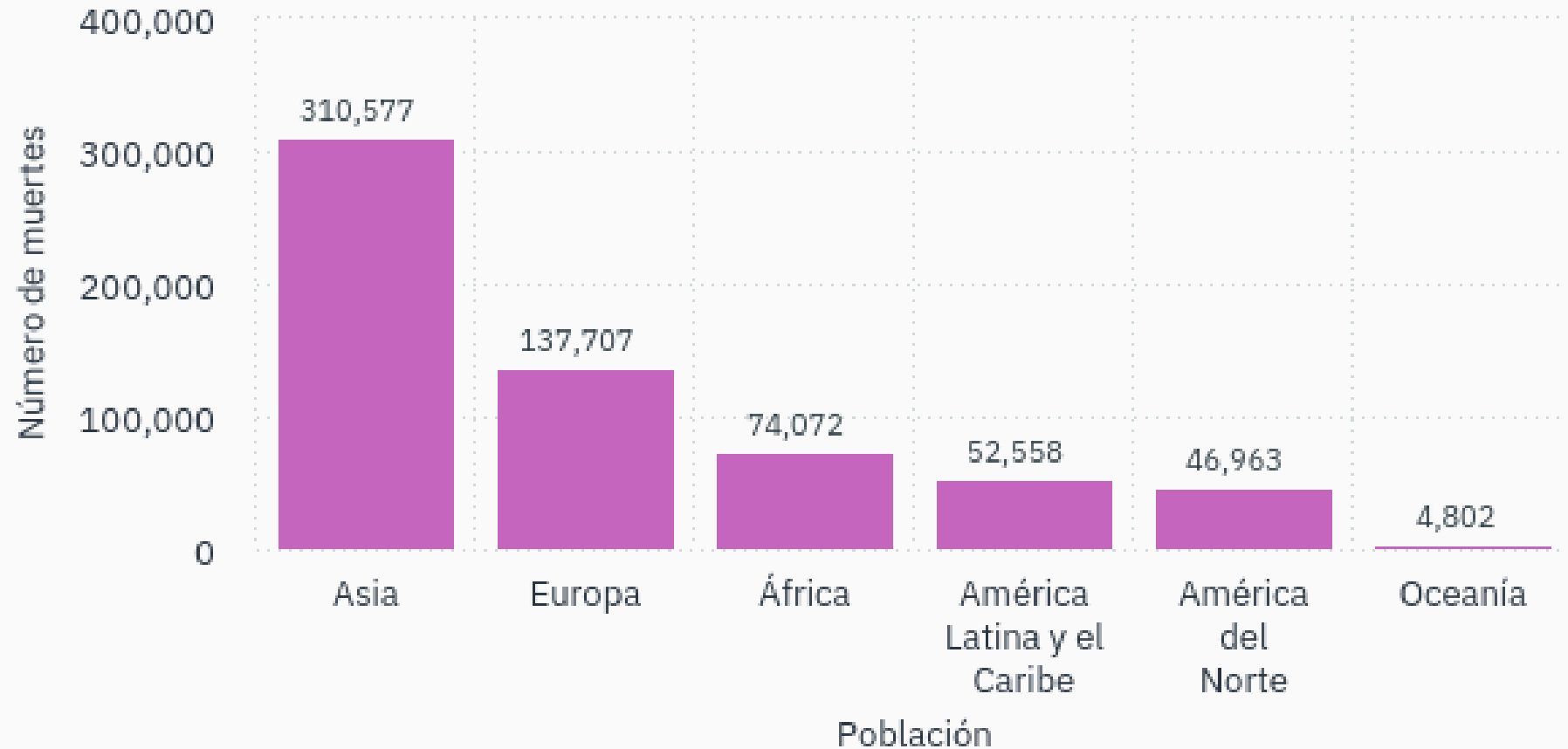
Gráficos estadísticos

- Gráfico de barras
 - Para datos cualitativos nominales y **ordinales**.
 - Rectángulos que representan una categoría.
 - Altura del rectángulo es proporcional al número de datos que hay en esa categoría.
 - Pueden representarse vertical u horizontalmente.
 - Debe llevar nombre de la categoría y número de elementos o individuos.

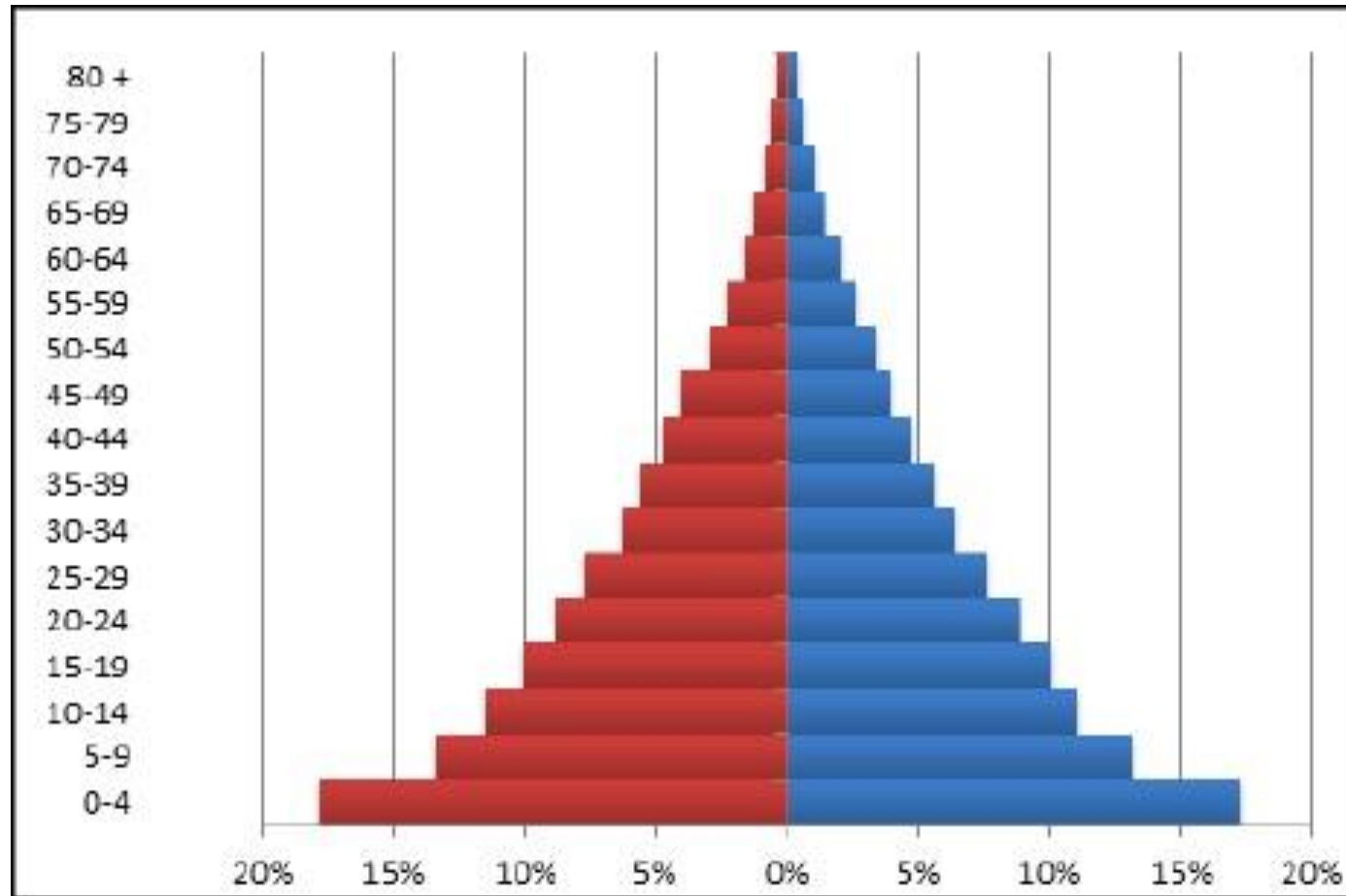


Ejemplo

Muertes estimadas de mujeres por cáncer de seno en 2018



Ejemplo

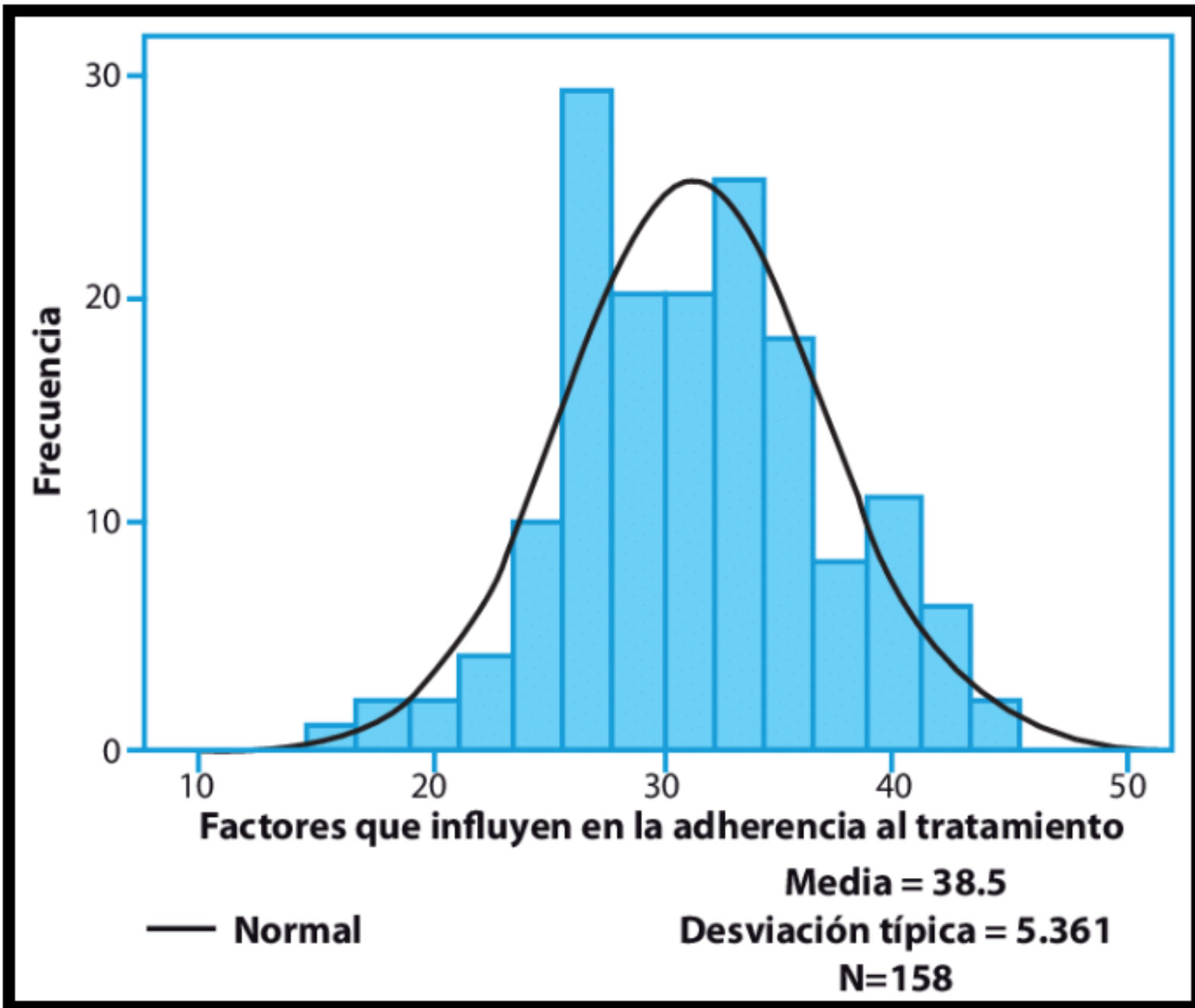


Gráficos estadísticos

- Histograma
 - Resume grandes cantidades de datos.
 - Para variables **numéricas**.
 - El eje horizontal es una recta con los valores que toma la variable.
 - El eje vertical es el número de veces que ocurre cada valor.

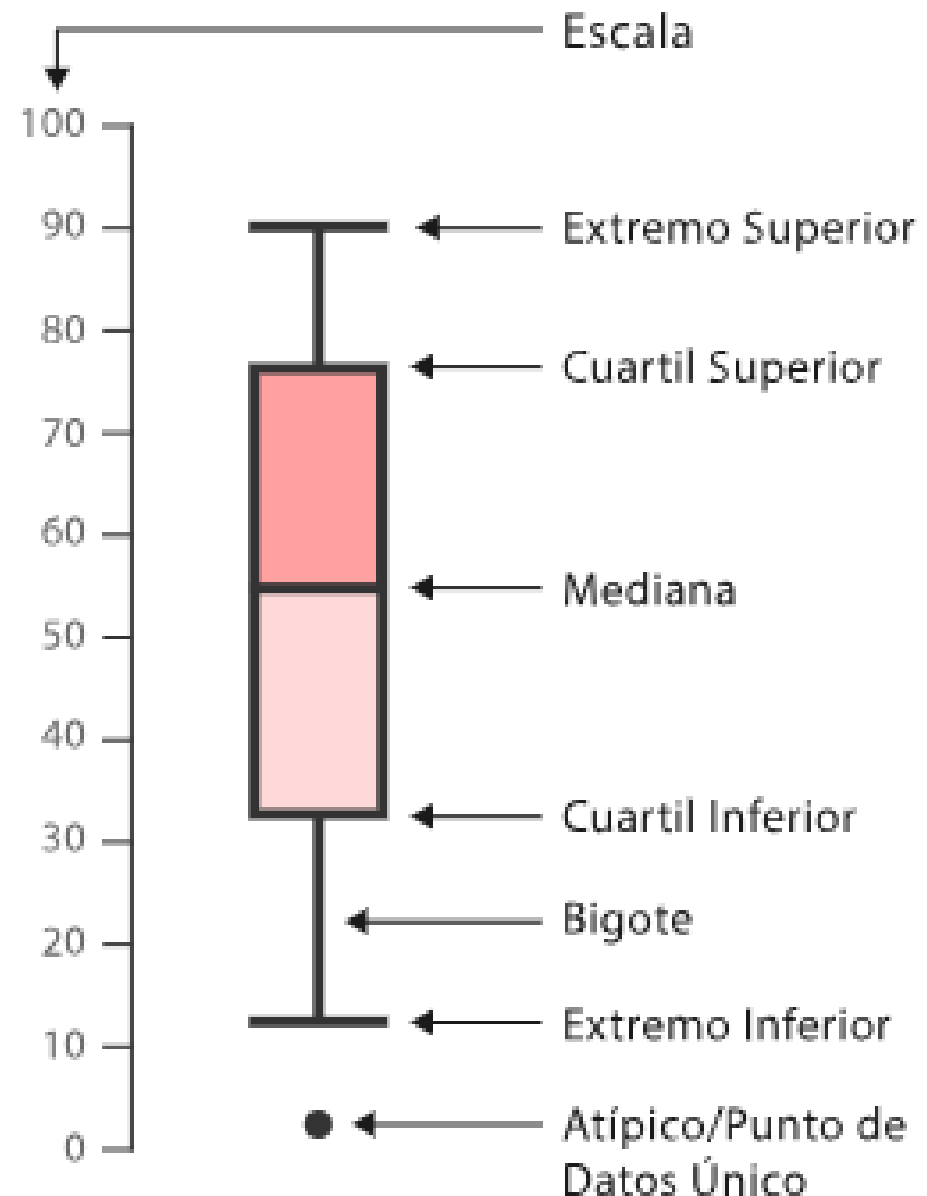


Ejemplo

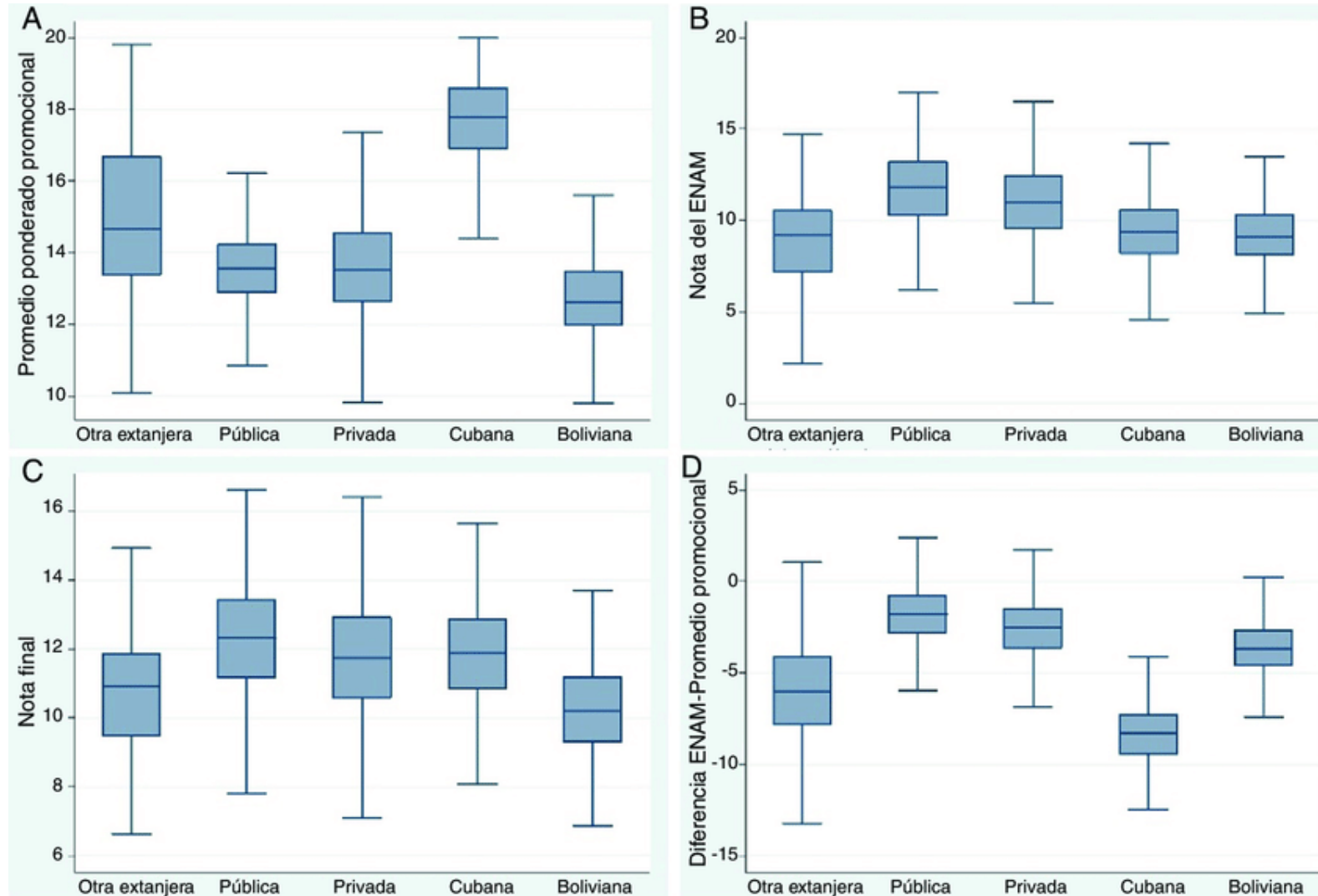


Gráficos estadísticos

- Diagrama de cajas y bigotes (Boxplot)
 - Se usa para variables **numéricas**.
 - La raya en la mitad de la caja es el valor de la mediana.
 - La caja representa el 50% de los individuos (percentil 25 a percentil 75).
 - Permite ver la localización de los datos, su dispersión e identificar valores extremos (“*outliers*”).



Ejemplo: análisis y conclusiones



EVALUACIONES

- Valorar los conocimientos de la estadística descriptiva.
 - <https://cibertest.com/examen-online/1040/examen-de-estadistica-descriptiva>
- Media, moda y mediana (actividad)
 - <https://www.liveworksheets.com/w/es/maticas/807837>
- Prueba de Diseño Experimental
 - <https://es.khanacademy.org/math/probability/study-design-a1/observational-studies-experiments/test/study-design-a1-unit-test?referrer=upsell>

Ejercicio

- Con los datos en Excel del archivo “Tabla de datos” realiza las siguientes actividades:
 - Genera una tabla de frecuencias con la variable “Interés”.
 - Genera un gráfico de tipo pastel para la variable “Interés”.
 - Genera un gráfico de tipo barras para la variable “Interés”.
 - Genera un gráfico de tipo histograma para la variable edad.
 - Genera un gráfico de tipo caja y bigotes para la variable edad.
 - Escribe una conclusión del gráfico de cajas y bigotes.