

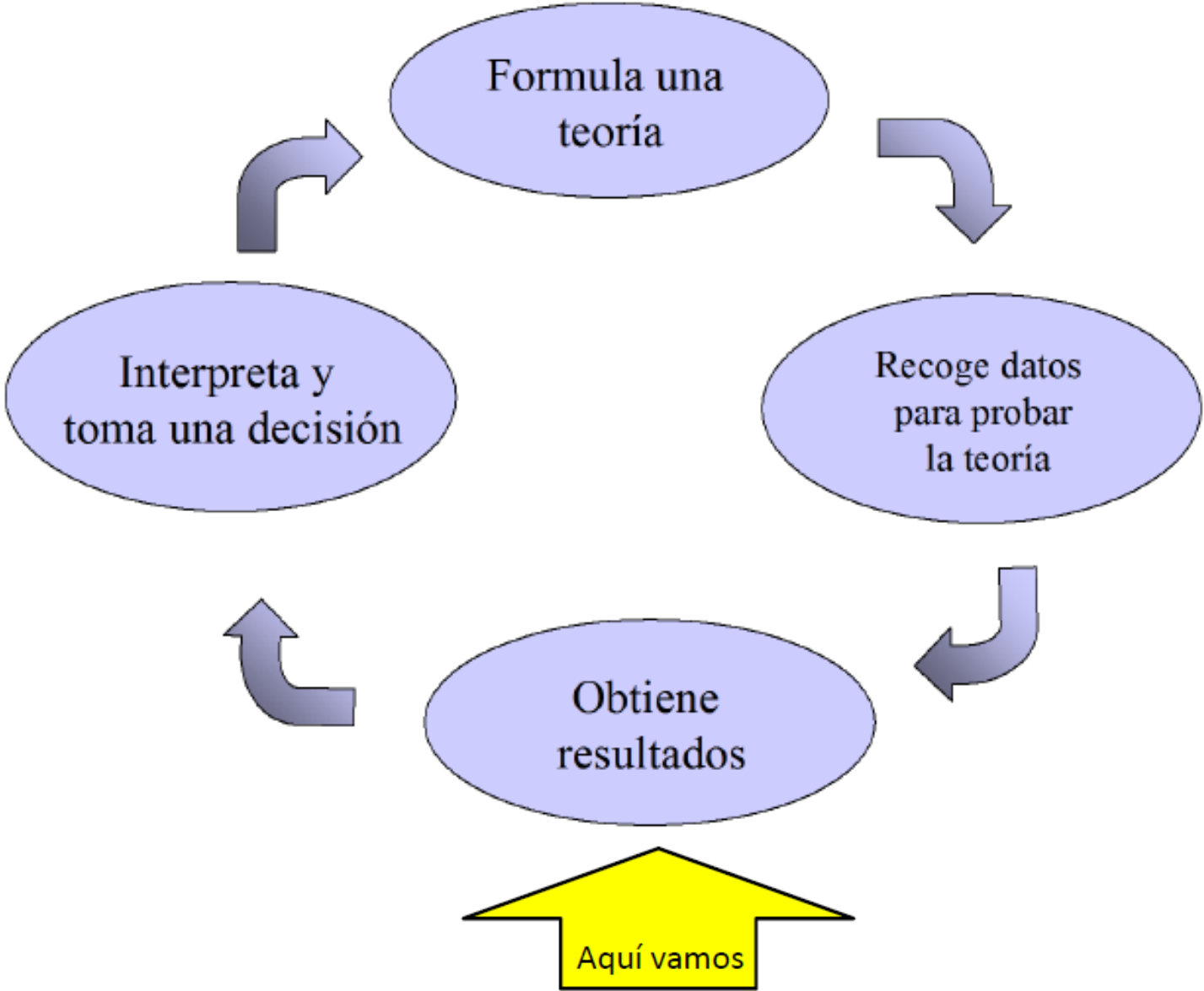


INSTITUTO DE MATEMÁTICA Y FÍSICA

Análisis Exploratorio de Datos:
Métodos tabulares y Representaciones Gráficas

DOCENTE Gloria Correa Beltrán

Etapas del Método Científico



Conceptos Básicos

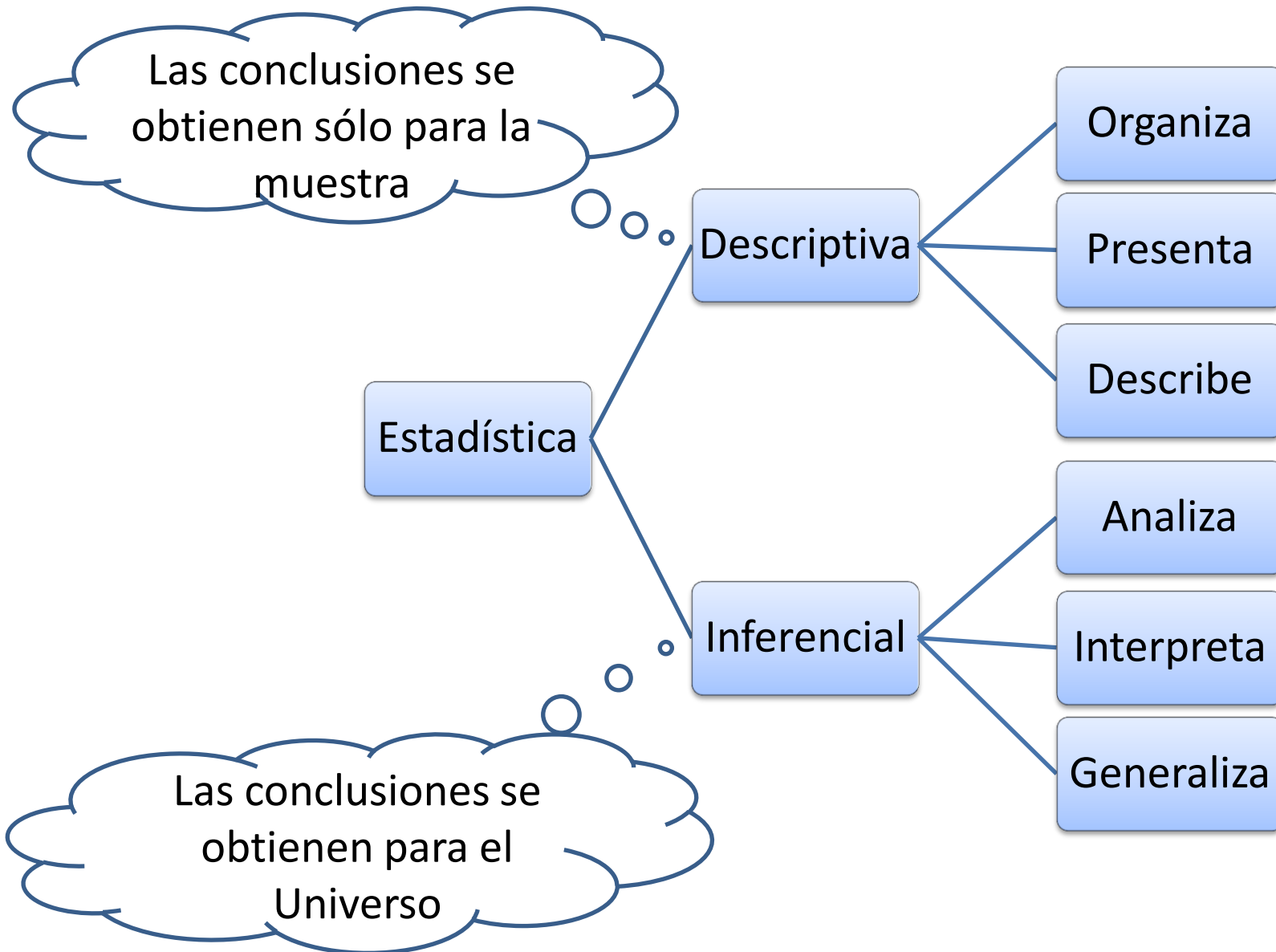
- **Unidad** es el objeto que observamos. Cuando el objeto es una persona, lo referimos como *sujeto*.
- **Variable:** Es la característica en estudio que puede variar de unidad en unidad.
- **Dato u Observación** es la información o característica que registramos de cada unidad. Son valores o categorías (número o palabra) asociados a una Variable dentro de una Población o Muestra.
- **Experimento:** Procedimiento cuyos resultados producen el conjunto de datos.
- **Base de datos:** Es una colección de observaciones con una o más variables se llama.



Veremos ejemplos de bases de datos en Excel

Conceptos Básicos

Clasificación de la Estadística



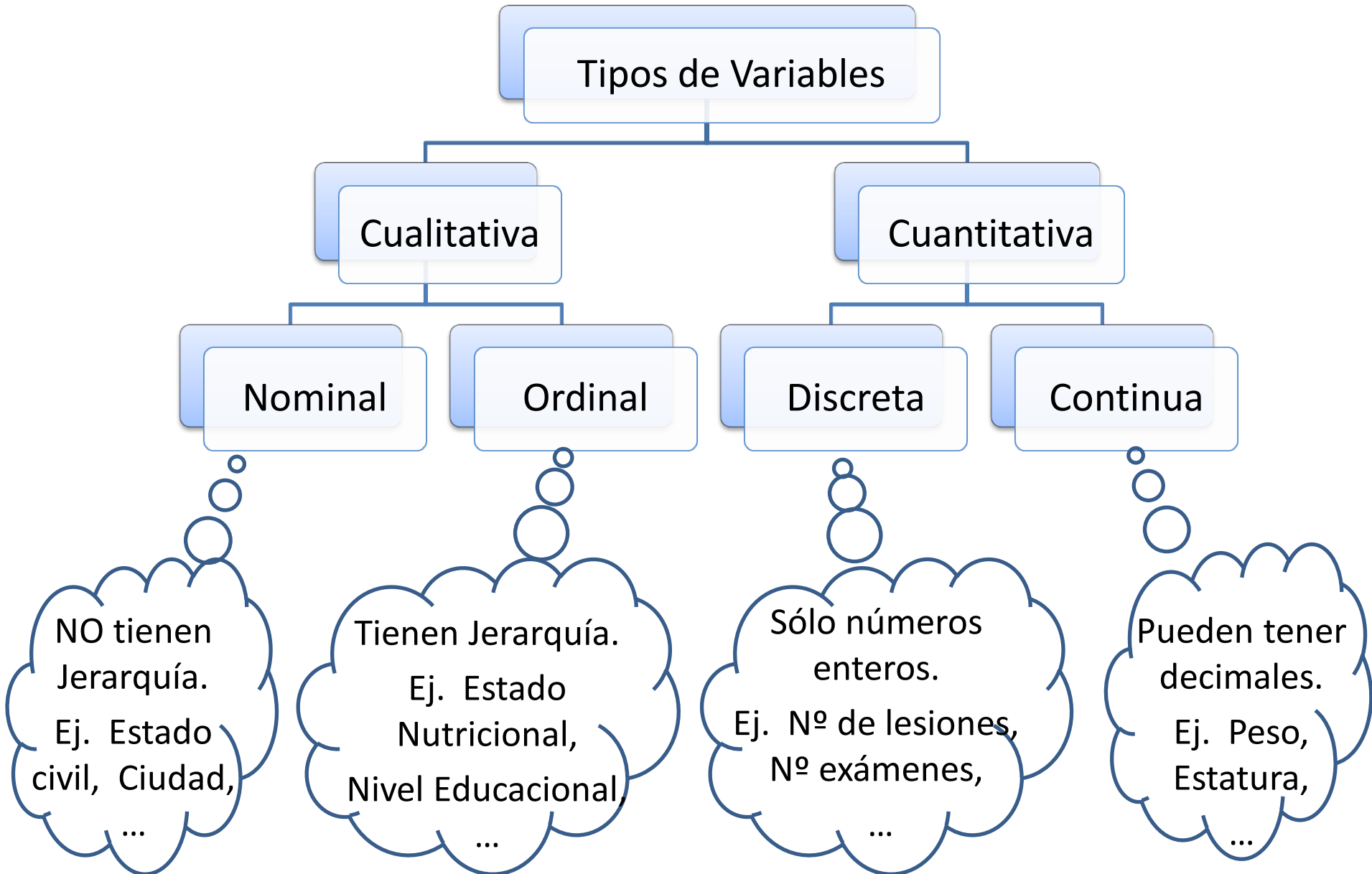
Ejercicio

En un estudio a nivel nacional, se desea saber cuál es el tiempo de espera promedio en ser atendido un paciente en los servicios médicos de los consultorios municipales. Talca, 2015

POBLACION	Pacientes de los servicios médicos de los consultorios municipales en Talca en el año 2015
MUESTRA	Pacientes seleccionados aleatoriamente
VARIABLE	Tiempos de espera
DATOS	20 min, 45 min, 2 horas, 10 min, 3 sem, ...
EXPERIMENTO	Observación, encuesta, ...
ESTADISTICA	Tiempo de espera promedio de la muestra (ej. 45 min)
PARAMETRO	Tiempo de espera promedio de la población (ej. 45 ± 10 min)

Conceptos Básicos

Clasificación de las Variables



Variable Cualitativa o Atributo

Los datos no se pueden asociar a un número, es decir, no son medibles.

- **Variable nominal:** no EXISTE UN ORDEN preestablecido entre las categorías de la variable. sitúa a cada dato en una u otra clase dada, que sólo indica que son distintos, no superior ni inferior.
- **Variable Ordinal:** EXISTE UN ORDEN O JERARQUIA, es decir, se puede distinguir entre superior e inferior.

Variable Cuantitativa o Numérica

Los datos se pueden asociar a un número, es decir, son medibles.

- **Variable Discreta:**
Los datos en teoría y práctica, corresponden a números enteros que en general se refieren a conteo.
- **Variable Continua:**
Los datos pueden tomar cualquier valor, al menos teóricamente.

Escalas de Medición

- **Nominal:** NO EXISTE ORDEN preestablecido entre las categorías. Sitúa a cada dato en una u otra clase dada, que sólo indica que son distintos.
- **Ordinal:** EXISTE UN ORDEN jerárquico. Sitúa los datos en orden, es decir, se puede distinguir entre mayor y menor, superior e inferior.
- **Intervalar:** No tiene cero absoluto.
- **de Razón:** Tiene un cero absoluto.

Observaciones:

- **Variables dicotómicas** o binarias tienen sólo dos respuestas posibles. Ej. Sexo (mujer, hombre), Asistencia (presente, ausente), etc.
- Las escalas nominal y ordinal son apropiadas para variables cualitativas.
- Las escalas intervalar y de razón se usan para variables cuantitativas.
- Es importante reconocer la escala de medición empleada, puesto que las operaciones aritméticas no se pueden utilizar en los niveles más bajos.

Ejercicio

En un estudio realizado en el primer semestre del año 2014, por la Facultad de Ciencias de la Salud de la Universidad de Talca, se hizo una evaluación del estado de salud de los niños de Educación Básica de las escuelas municipales de la comuna. Para ello se seleccionó aleatoriamente tres establecimientos, en cada uno de los tres establecimientos se seleccionó aleatoriamente un curso de los distintos niveles y luego se evaluó a todos los niños de los cursos seleccionados.

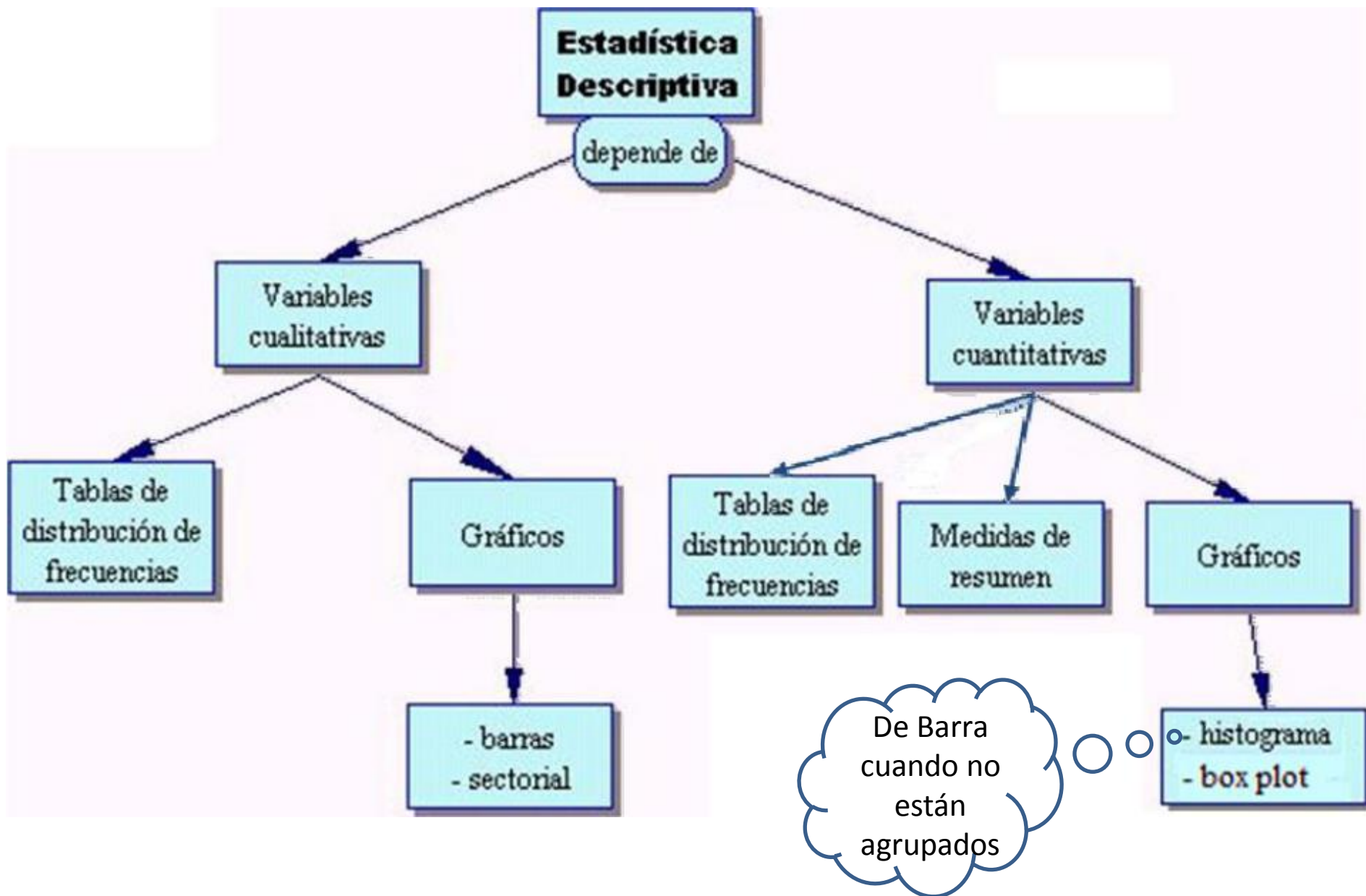
POBLACION	Niños de las escuelas municipales de la comuna de Talca en el primer semestre del años 2014.
MUESTRA	$\approx 3 * 8 * 40 = 960$

Ejercicio

Algunas de las variables evaluadas fueron:

Variable	Datos	Clasificación	Escala
Edad	6, 7, .. , 13, 14, ... años	cuantitativa continua	de razón
Curso	1°, 2°, ..., 8°	cualitativa ordinal	ordinal
Sexo	hombre, mujer	cualitativa nominal	nominal
Escuela	A,B, C, ...	cualitativa nominal	nominal
Peso	30, 45, ... kg	cuantitativa continua	de razón
Estatura	1,20; 1,55; 1,62; ... m	cuantitativa continua	de razón
Estatura	bajo, normal, alto	cualitativa ordinal	ordinal
Temperatura	35; 35,5; 36; ... °C	cuantitativa continua	intervalar
Nº de hermanos	0, 1, 2, 3, ...	cuantitativa discreta	de razón
Escolaridad madre	básico , medio, técnico, profesional	cualitativa ordinal	ordinal
Grupo sanguíneo	A, B, AB, O	cualitativa nominal	nominal

Pasos a seguir en el Análisis Exploratorio de Datos



Tablas de Distribución de Frecuencias para Variables Cualitativas

Ejemplo

En una Sede Social, al finalizar un periodo de educación en Vida Sana, realizado por alumnos tesistas del área de la salud, se hizo una evaluación para medir los logros alcanzados en todas las variables consideradas.

Los datos que se presentan se refieren al Estado Nutricional inicial.

Bajo Peso	Bajo Peso	Normal	Normal	Normal	Normal
Normal	Normal	Normal	Normal	Normal	Normal
Normal	Normal	Normal	Normal	Normal	Normal
Normal	Normal	Normal	Normal	Normal	Sobrepeso
Sobrepeso	Sobrepeso	Sobrepeso	Sobrepeso	Sobrepeso	Sobrepeso

Tablas de Distribución de Frecuencias para Variables Cualitativas

Ejemplo

Bajo Peso	Bajo Peso	Normal	Normal	Normal	Normal
Normal	Normal	Normal	Normal	Normal	Normal
Normal	Normal	Normal	Normal	Normal	Normal
Normal	Normal	Normal	Normal	Normal	Sobrepeso
Sobrepeso	Sobrepeso	Sobrepeso	Sobrepeso	Sobrepeso	Sobrepeso

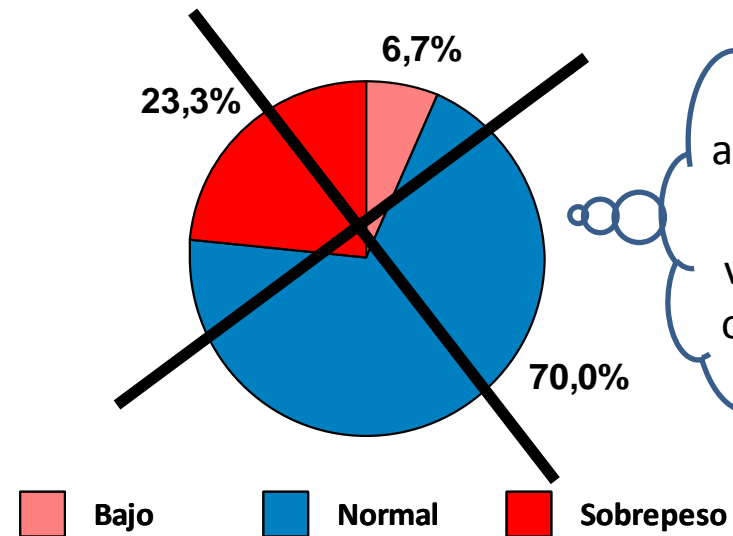
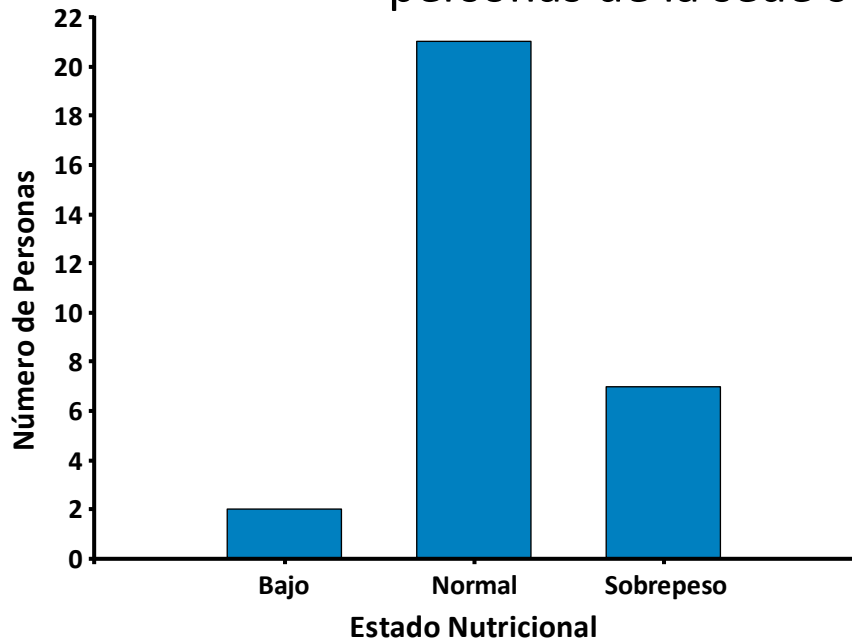
Estado Nutricional	Número de personas	Porcentaje de personas
Bajo Peso	2	6,7
Normal	21	70,0
Sobrepeso	7	23,3
Total	30	100,0

Gráficos para Variables Cualitativas: Ejemplo

Tabla N° 1: Distribución de Frecuencias del Estado Nutricional de las personas de la sede social XX, de la ciudad XX en el año XX.

Estado Nutricional	Número de personas	Porcentaje de personas
Bajo Peso	2	6,7
Normal	21	70,0
Sobrepeso	7	23,3
Total	30	100,0

Figura N° 1: Distribución de Frecuencias del Estado Nutricional de las personas de la sede social XX, de la ciudad XX en el año XX.



No es apropiado para variables ordinales

Tablas de Distribución de Frecuencias para Variables Cuantitativas con datos sin agrupar

Ejemplo

Para el ejemplo anterior, otra de las variables fue el número de veces que realiza actividad física en la semana. Los datos se presentan a continuación.

1	1	1	1	1	2
2	2	2	2	2	2
2	3	3	3	3	3
3	3	3	3	3	3
3	3	4	4	4	4

Tablas de Distribución de Frecuencias para Variables Cuantitativas con datos sin agrupar

Ejemplo

1	1	1	1	1	2
2	2	2	2	2	2
2	3	3	3	3	3
3	3	3	3	3	3
3	3	4	4	4	4

Actividad Física Semanal	Número de personas	Porcentaje de personas	Número de personas acumulado	Porcentaje de personas acumulado
1	5	16,7	5	16,7
2	8	26,7	13	43,3
3	13	43,3	26	86,7
4	4	13,3	30	100,0
Total	30	100,0		

Gráficos para Variables Cuantitativas con datos sin agrupar: Ejemplo

Tabla N° 2: Distribución de Frecuencias de la actividad física semanal de las personas de la sede social XX, de la ciudad XX en el año XX.

Actividad Física Semanal	Número de personas	Porcentaje de personas	Número de personas acumulado	Porcentaje de personas acumulado
1	5	16,7	5	16,7
2	8	26,7	13	43,3
3	13	43,3	26	86,7
4	4	13,3	30	100,0
Total	30	100,0		

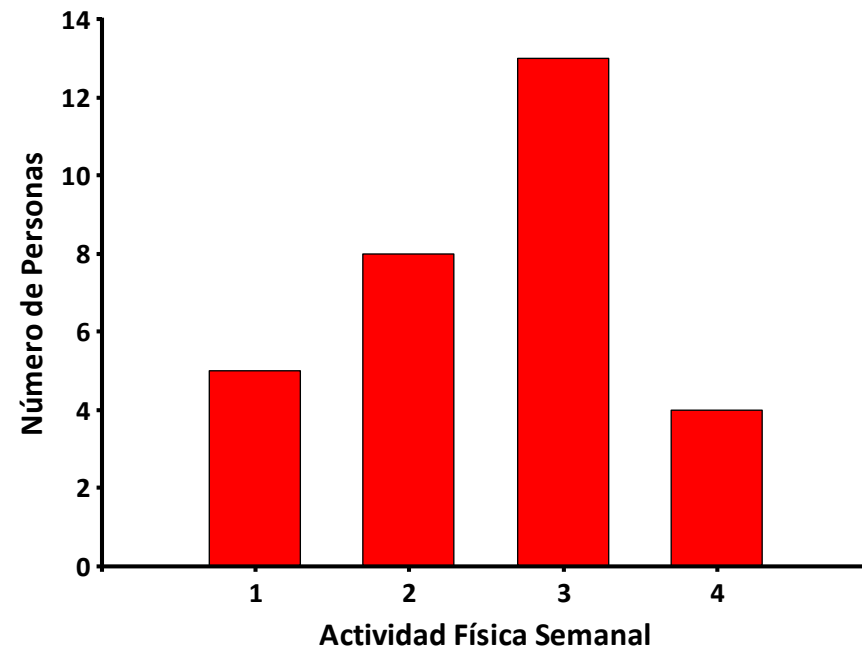
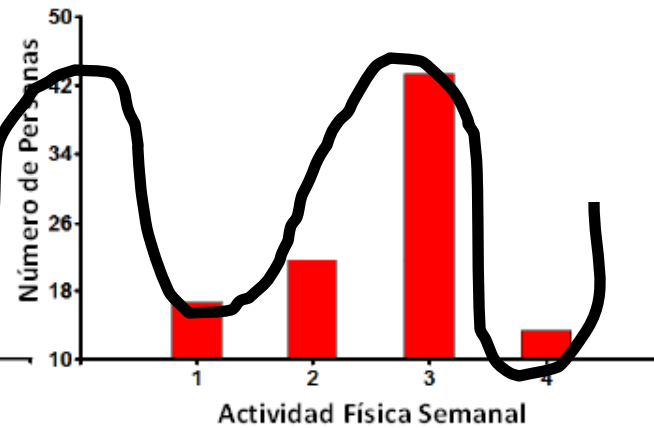
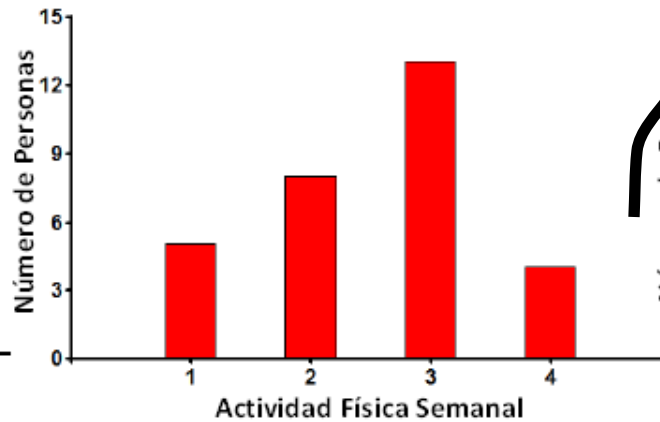
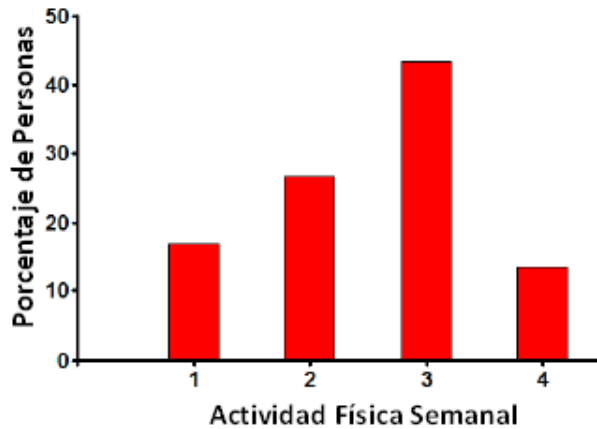
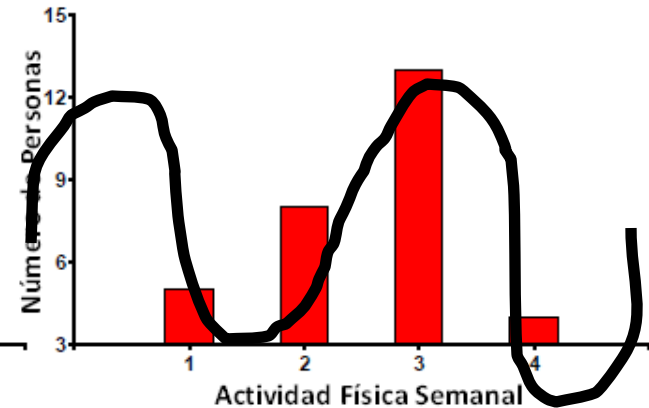
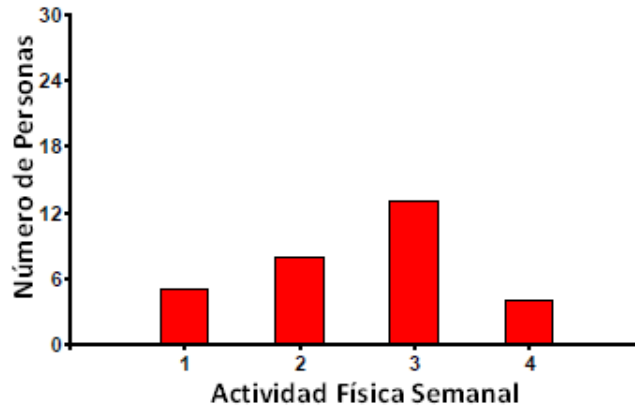
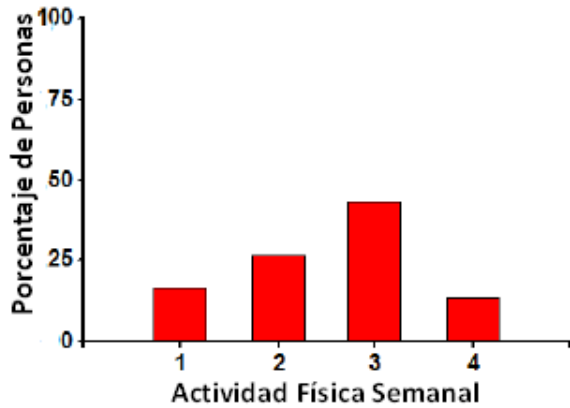


Figura N° 2: Distribución de Frecuencias de la actividad física semanal de las personas de la sede social XX, de la ciudad XX en el año XX.

Ejercicio

Compare los siguientes gráficos. Todos representan la situación planteada en el ejercicio anterior.



Tablas de Distribución de Frecuencias para Variables Cuantitativas con datos agrupados

Ejemplo

Para el ejemplo anterior, otra de las variables fue la edad (años) de las personas. Los datos se presentan a continuación.

60	62	62	67	68	68
68	71	71	72	73	73
75	76	76	76	77	77
77	78	79	79	81	82
85	86	87	87	88	89

Tablas de Distribución de Frecuencias para Variables Cuantitativas con datos agrupados

Ejemplo

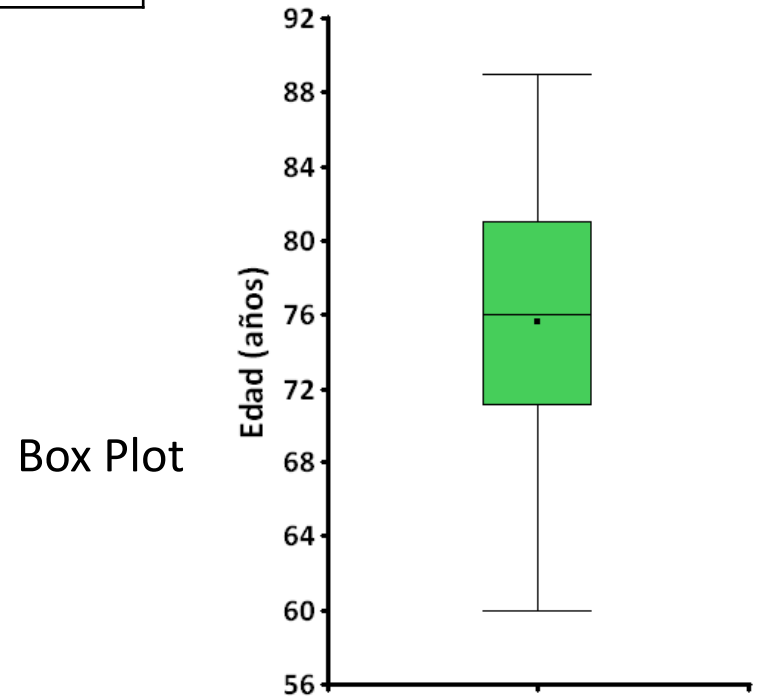
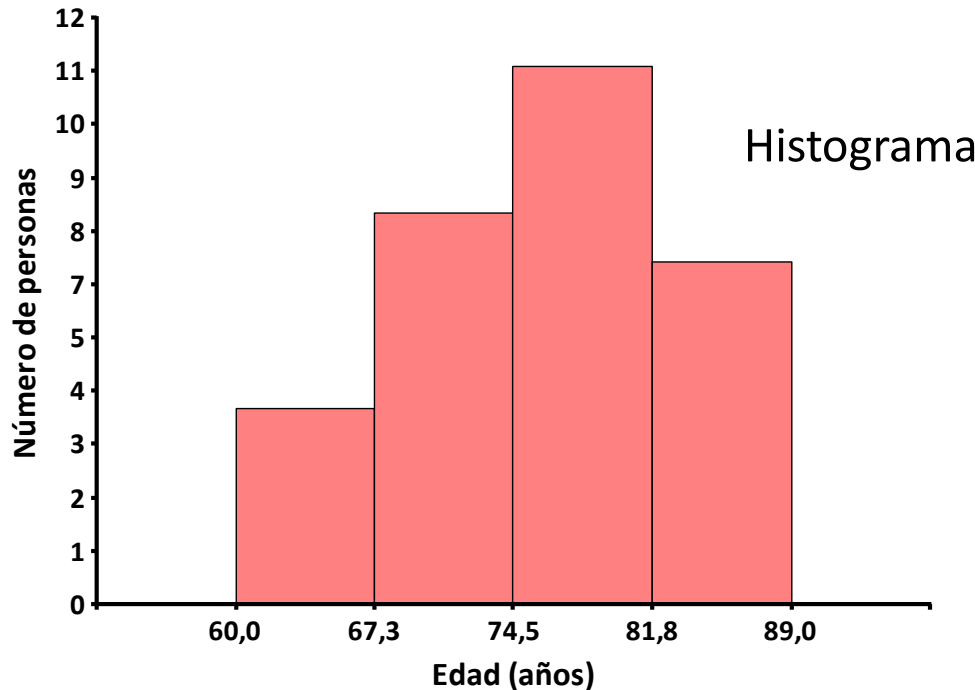
60	62	62	67	68	68
68	71	71	72	73	73
75	76	76	76	77	77
77	78	79	79	81	82
85	86	87	87	88	89

Edad	Marca de clase	Número de personas	Porcentaje de personas	Número de personas acumulado	Porcentaje de personas acumulado
[60,0 ; 67,3]	63,65	4	13,3	4	13,3
(67,3 ; 74,5]	70,90	8	26,7	12	40,0
(74,5 ; 81,8]	78,15	11	36,7	23	76,7
(81,8 ; 89,0]	85,40	7	23,3	30	100,0
Total		30	100,0		

Gráficos para Variables Cuantitativas con datos agrupados: Ejemplo

Tabla N° 2: Distribución de Frecuencias de la edad en años de las personas de la sede social XX, de la ciudad XX en el año XX.

Edad	Marca de clase	Número de personas	Porcentaje de personas	Número de personas acumulado	Porcentaje de personas acumulado
[60,0 ; 67,3]	63,65	4	13,3	4	13,3
(67,3 ; 74,5]	70,90	8	26,7	12	40,0
(74,5 ; 81,8]	78,15	11	36,7	23	76,7
(81,8 ; 89,0]	85,40	7	23,3	30	100,0
Total		30	100,0		



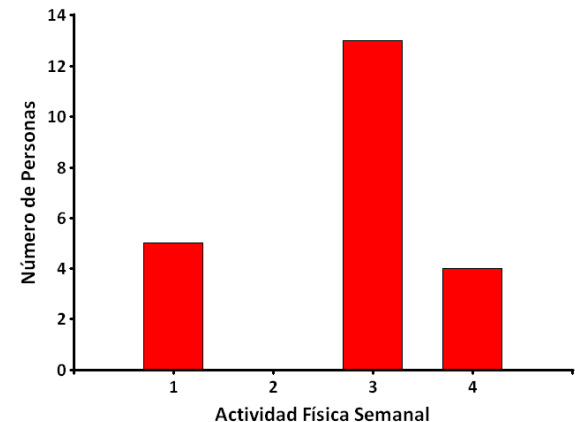
Observaciones:

- En los gráficos las frecuencias pueden ser reemplazadas por sus respectivos porcentajes. Al hacer esto la forma del gráfico se mantiene.
- Los porcentajes son recomendados para muestras grandes (>30) y las frecuencias para muestras pequeñas (<30).
- Al presentar un gráfico o una tabla siempre deben llevar un título apropiado y completo (describir el universo: tiempo y espacio).
- Tablas y gráficos deben estar enumerados (tabla N° 1, tabla N° 2, ..., gráfico N° 1, gráfico N° 2, ...).
- En un informe, sólo se justifica la presencia de un gráfico o de una tabla, si se hace mención de ellos en algún pasaje del informe.

Observaciones:

- El gráfico circular es apropiado para variables nominales y el de barra para variables ordinales.
- En los gráficos de barra se debe respetar la jerarquía de la variable cuando es ordinal y cuando es cuantitativa (sin agrupar) los datos de la variable deben figurar consecutivamente en el eje horizontal.

- En el gráfico de barra si alguno de los datos tiene frecuencia "0" no se debe omitir.



- En un gráfico de barras, estas deben quedar separadas y en un histograma deben quedar unidas.

Resumen

Edad	Marca de clase	Número de personas	Porcentaje de personas	Número de personas acumulado	Porcentaje de personas acumulado
[60,0 ; 67,3]					
(67,3 ; 74,5]	Actividad Física Semanal	Número de personas	Porcentaje de personas	Número de personas acumulado	Porcentaje de personas acumulado
(74,5 ; 81,8]					
(81,8 ; 89,0]					
Total					
	1	Estado Nutricional	Número de personas	Porcentaje de personas	
	2				
	3				
	4				
	Total				
	Total				

Cualitativa

Cuantitativa con datos sin agrupar

Cuantitativa con datos agrupados

