

# ESTADÍSTICA

---

## **UNIDAD 3: Resumen de datos en forma gráfica**

Ciencias Económicas y Empresariales  
UCEL



# BASE DE DATOS

A continuación se presenta una parte de una base de datos de **EPH-Gran Rosario** que contiene información sobre un grupo de individuos, se mide la edad, el nivel educacional, ITF, IPCF, entre otras.

AÑO	TRIMESTRE	EDAD	EST- CIVIL	N°- hermanos	LEE- ESCRIBE	NIVEL_ED	ESTADO	CAT_OCUP	CAT_INAC	ITF	IPCF
2.019	1	38	1	2	1	4	1	3	0	64.800	12.960
2.019	1	49	2	1	1	2	1	3	0	24.000	6.000
2.019	1	45	2	1	1	6	1	3	0	146.000	29.200
2.019	1	66	2	1	1	2	1	3	0	51.000	7.286
2.019	1	23	5	1	1	5	3	0	3	65.000	13.000
2.019	1	15	5	1	1	3	3	0	3	20.000	4.000
2.019	1	72	5	4	1	1	3	0	1	9.800	9.800
2.019	1	29	1	1	1	3	1	2	0	16.800	4.200
2.019	1	39	2	1	1	4	1	3	0	40.000	10.000
2.019	1	34	3	4	1	4	1	3	0	6.400	3.200
2.019	1	17	5	4	1	3	3	0	3	19.000	4.750

“

”

**ESTADO CIVIL**

1= unido  
2=casado  
3=separado/a-  
divorciado/a  
4= viudo/a  
5= soltero/a

**ESTADO:  
condicion de la  
actividad**

1=ocupado  
2 = desocupado 3=inactivo  
4 =menor de 18

**NIVEL  
EDUCATIVO**

1= Primaria Incompleta  
2 = Primaria Completa  
3 = Secundaria Incompleta  
4= Secundaria Completa  
5= Superior Universitaria Incompleta  
6 = Superior Universitaria Completa  
7=Sin Instrucción 9

**CATEGORIA  
OCUPACIONAL**

1=Patrón  
2= Cuenta propia  
3 = Obrero o Empleado  
4= Trabajador sin  
Remuneración

**CATEGORIA  
DE INACTIVIDAD**

1= Jubilado/ pensionado  
2= Rentista  
3= Estudiante  
4= Ama de casa  
5= Menor de 18  
6= Discapacitado

**ITF: ingreso  
total fliar****IPCF: Ingreso per càpita**



# Definiciones

---

- **Unidad:** Elemento u objeto que observamos. Si es una persona, también se la llama “sujeto”.
- **Observación:** Información o característica registrada para cada unidad.
- **Variable:** Característica que varía de unidad a unidad.
- **Conjunto de datos:** Es un conjunto de observaciones sobre una o más variables.

# En el ejemplo...

---

- Las unidades o sujetos son los encuestados
- Una observación del encuestado 1 es su edad: 38 años
- Las variables estudiadas son: edad, estado civil, si sabe leer, nivel educacional, número de hermanos, estado, ingreso total familiar, etc.



# Estadística descriptiva

---

El objetivo de la estadística descriptiva es resumir la información contenida en un conjunto de datos, de la manera más concisa y completa posible. Esto puede hacerse mediante la construcción de **tablas** y **gráficos** y el cálculo de ciertas características numéricas llamadas "**estadísticas**".

[https://youtu.be/W1\\_eCwuYkAI](https://youtu.be/W1_eCwuYkAI)

# Tipos de variables

Podemos ver que cada una de las variables en la base toma distintos tipos de valores. Mientras que el **estado civil** presenta sólo 5 categorías: unido, casado, separado, viudo, soltero, para la edad tenemos muchos valores numéricos distintos, por ejemplo: 38, 48, 45, ... etc.

Se puede clasificar a las variables según los valores que tome en:

- Variables cualitativas
- Variables cuantitativas discretas
- Variables cuantitativas continuas



# Tipos de variables

**Cualitativas:** Son aquellas que clasifican a las unidades en categorías. Las categorías pueden tener o no un orden. También se las suele llamar **categorías**.

En el ejemplo, se pueden distinguir las siguientes variables cualitativas:

- Estado Civil: sus categorías (valores posibles) son unido, casado, separado, viudo y soltero. No hay orden entre las categorías.
- Nivel Educativo: sus categorías son primaria incompleta, primaria completa, secundaria incompleta, secundaria completa, etc. Hay un orden entre los niveles de la variables.



# Tipos de variables

**Cuantitativas:** Son aquellas cuyas observaciones provienen de procesos de medición o conteo se las puede clasificar en:

- Discretas: Si el conjunto de resultados posibles es finito o infinito numerable.
- Continuas: Si el conjunto de valores posibles es un intervalo o conjunto de intervalos de números reales.

En el ejemplo, las variables cuantitativas son: Edad, Número de hermanos, IPCF, ITF, etc.

# En el ejemplo...

La variable cuantitativa discreta es:

- Número de hermanos: los valores posibles son 0,1,2,3,....

Las variables cuantitativas continuas son:

- Edad: algunos de sus valores son 38, 45, 49, etc.
- IPCF (Ingreso per cápita) puede tomar los valores 12960, 6000, etc.



Ver distintos conjuntos de datos en el libro (págs. 125 y 126)



# MANOS A LA OBRA N°1

---

¿Qué tipo de variable es?

- El número de personas que ingresan al banco S el día viernes entre las 12.00 y las 13.00hs.
- Tiempo empleado para correr 100mts.
- Arrojar dos dados y registrar si las caras resultantes coinciden o no.
- Activos de supermercados de la ciudad de Rosario.

# Estadística descriptiva

---

- Descripción de un conjunto de observaciones de una **variable cualitativa**
- Descripción de un conjunto de observaciones de una variable cuantitativa
- Medidas descriptivas: de tendencia central y de dispersión.



# Formas de resumir una variable cualitativa

---

## Tabla de distribución de frecuencias

Agrupación de datos, en clases mutuamente excluyentes, que muestra el número de observaciones en cada clase.

## Gráficos

- De Sectores Circulares
- De Barras
- Pictogramas

# Distribución de una variable cualitativa

---

La **distribución de una variable** provee los **posibles valores** que una variable puede tomar y **cuán frecuentemente ocurren** estos posibles valores. La distribución de una variable muestra el modelo de variación de la variable.



# Distribución de una variable cualitativa

---

- Encuestados según estado o condición de la actividad

CATEGORÍA	Nº DE PERSONAS	PROPORCIÓN
OCUPADO	13	$0,42 = 13/31$
DESOCUPADO	3	$0,10 = 3/31$
INACTIVO	15	$0,48 = 15/31$
<b>TOTAL</b>	<b>31</b>	<b>1,00</b>

Fuente: EPH/2019

# Tabla de distribución de frecuencias

CATEGORÍA	Frecuencia Absoluta ( $f_j$ )	Frecuencia Relativa ( $h_j$ )	Frecuencia Relativa Porcentual ( $h_j\%$ )
OCUPADO	13	0.42	42
DESOCUPADO	3	0,10	10
INACTIVO	15	0,48	48
TOTAL	31	1,0	100

- $f_j$ : muestra el número de elementos correspondiente a cada una de las clases, es la **frecuencia absoluta**.
- $h_j$ : es el cociente entre la frecuencia absoluta de la clase “j” y el tamaño de muestra. Representa la proporción o **frecuencia relativa** de elementos correspondientes a esa clase:  $h_j = \frac{f_j}{n}$ .
- $h_j\%$ : la **frecuencia relativa porcentual** es la frecuencia relativa multiplicada por 100



# Tabla de distribución de frecuencias

Cada valor observado en la tabla tiene una interpretación. Veamos la interpretación de la primera fila:

- $f_j = 13$ . De los encuestados 13 se encuentran ocupados.
- $h_j = 0,42$ . La proporción de personas ocupadas es de 0,42.
- $h_j\% = 42$ . El 42% de las personas entrevistadas están ocupadas.



EJERCICIO PROPUESTO: Elegir otra fila de la tabla de frecuencias e interpretar cada uno de los valores.



# **Representación gráfica para distribuciones cualitativas**

---



# Gráfico de sectores circulares

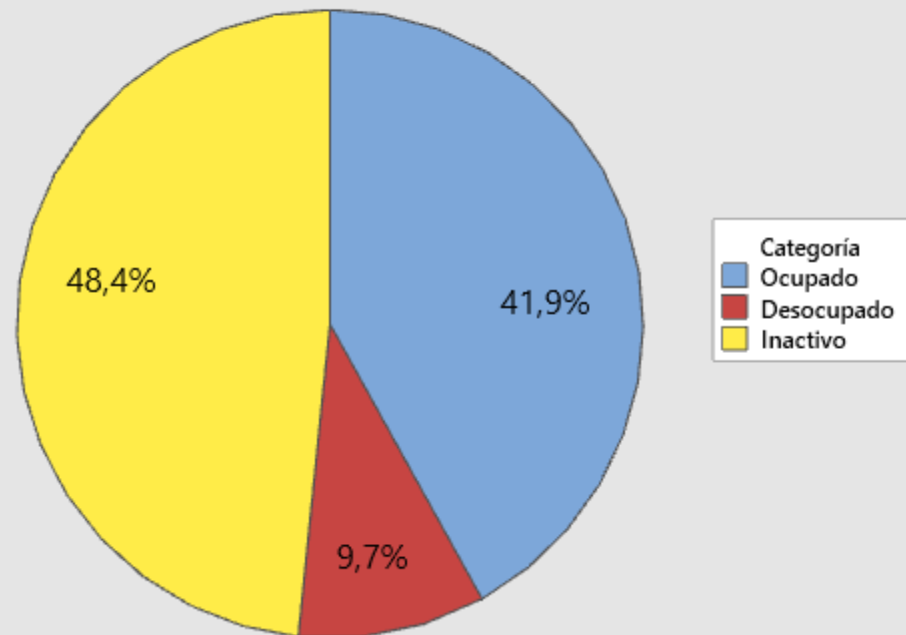
---

Un gráfico de sectores circulares muestra la distribución de una variable cualitativa.

La construcción de este gráfico consiste en diagramar un círculo que representa al 100 % de las unidades. El mismo se divide en tantos sectores como categorías existan. El área de cada sector es proporcional al porcentaje de unidades que pertenecen a la categoría que representa.

Retomando el ejemplo en el cual se estudiaba a los encuestados según estado o condición de la actividad.

Distribución de los encuestados clasificados según estado de la actividad



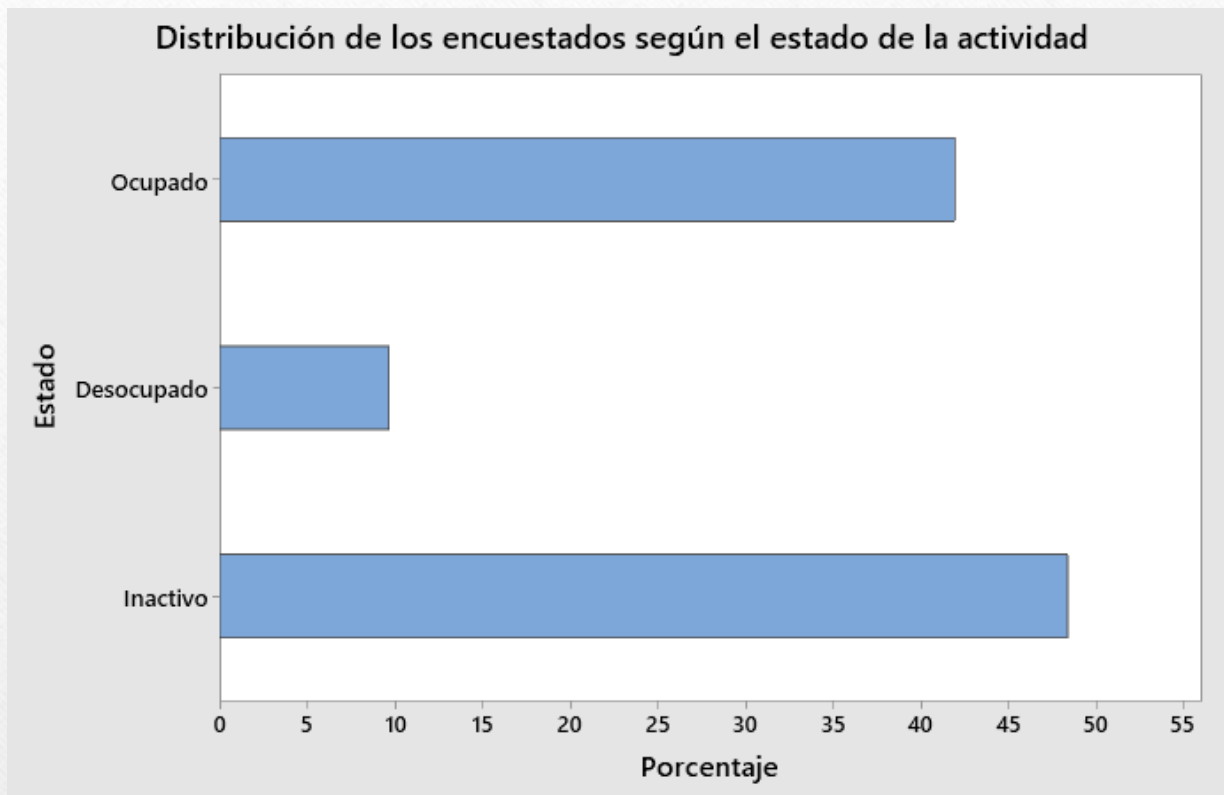
Equivale a representar las frecuencias relativas de la tabla mostrada con anterioridad.



# Gráfico de barras

---

Un gráfico de barras muestra la distribución de una variable cualitativa. Cuando se desean hacer comparaciones entre las distintas categorías de la variable analizada se utiliza el gráfico de barras horizontales. Se representa una barra para cada categoría y la longitud de cada una de ellas es igual a la cantidad (o proporción, o porcentaje) de unidades que pertenecen a la categoría y el ancho es el mismo para todas.



### IMPORTANTE

- Todas las barras deben tener el mismo ancho
- Las barras están separadas, nunca “pegadas”
- La distancia entre cada par de barras debe ser la misma
- Las barras pueden estar en sentido vertical u horizontal.





# CUIDADO!

---

Los gráficos de barras y de sectores circulares permiten representar la distribución de variables cualitativas.

# MANOS A LA OBRA N° 2

## ACTIVIDAD 1

Economistas que asistieron a un seminario de economía argentina, debieron completar un cuestionario. Una de las preguntas decía: "¿qué calificación le pondrías al seminario? Mediocre – Aceptable – Bueno – Muy Bueno – Excelente". Una muestra de médicos seleccionados al azar produjo las siguientes respuestas. Para facilitar el procesamiento informático de los resultados de los cuestionarios se utilizó una codificación numérica con:

1=Mediocre – 2=Aceptable – 3=Bueno – 4=Muy Bueno – 5=Excelente.

3	4	4	5	1	5	3	4	5	2	4	5	3	4	4	4	5	5	4	1	4	5	4	2	5	4	2	4	4	4
5	5	3	4	5	5	2	4	3	4	5	4	3	5	4	4	3	5	4	5	4	3	5	3	4	4	3	5	3	3

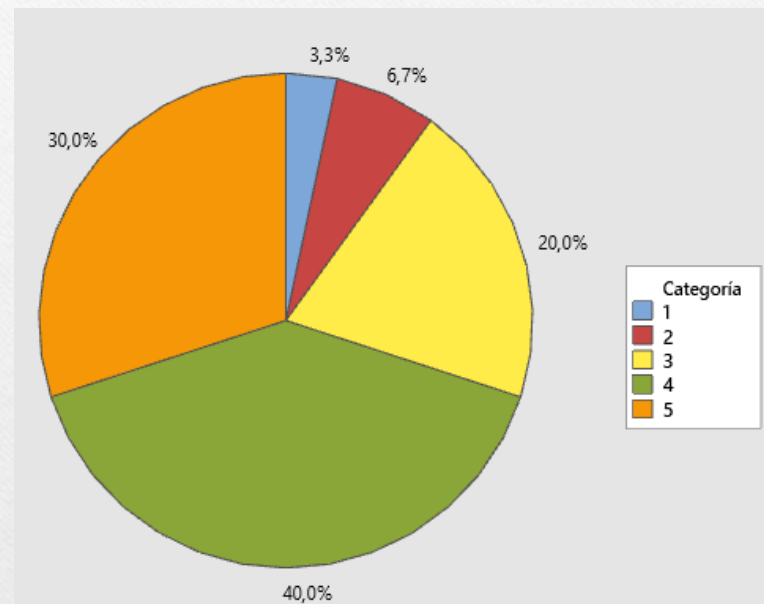
- La variable “calificación del seminario” es cualitativa o cuantitativa?
- Complete la tabla de frecuencias
- En base al gráfico provisto indique su nombre y comente sobre la evaluación general de los médicos acerca del seminario.



# MANOS A LA OBRA N° 2

## ACTIVIDAD 1

Calificación	frec. Abs
1	2
2	4
3	12
4	24
5	18
Total	60



# MANOS A LA OBRA N° 2

## ACTIVIDAD 2

La siguiente tabla muestra el número de ingresantes por carrera de cierta facultad en el año 2020

Carrera	Número de ingresantes
Contador Público	2168
Lic. en Administración de Empresas	632
Lic. en Comercialización	123
Lic. en Relaciones Laborales	79
Lic. en Economía	734

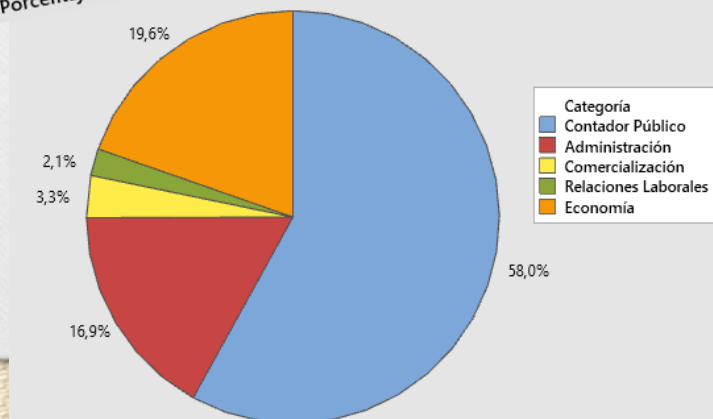
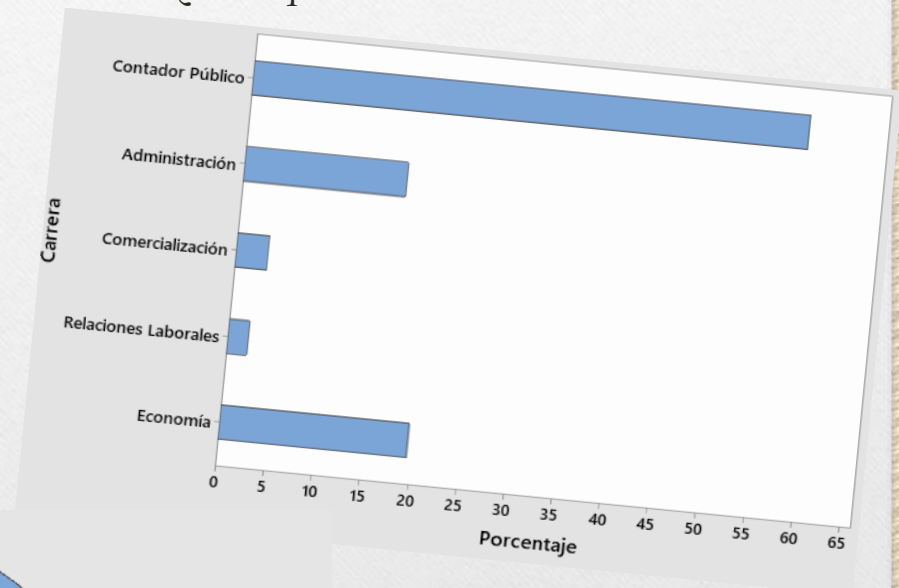
Complete la tabla de distribución de frecuencias.



# MANOS A LA OBRA N° 2

En base a la tabla de distribución de frecuencias responder:

- a- ¿Qué porcentaje de alumnos se inscribieron en la carrera de Lic. en Economía?
- b- ¿Qué porcentaje de alumnos lo hicieron en la carrera de Contador Público?
- c- ¿Qué gráfico representa a este conjunto de datos? ¿Por qué?



# Estadística descriptiva

---

- Descripción de un conjunto de observaciones de una variable cualitativa
- Descripción de un conjunto de observaciones de una **variable cuantitativa**
- Medidas descriptivas: de tendencia central y de dispersión.



# Formas de resumir una variable cuantitativa

---

Tabla de distribución de frecuencias

## Gráficos

- De bastones
- Tallo y hojas
- Histogramas
- Boxplot

## Descripción de un conjunto de observaciones de una variable cuantitativa

---

En las variables de naturaleza **cuantitativa** cada valor será un número que puede ser interpretado como tal.

Si los valores posibles de una variable cuantitativa provienen de un proceso de conteo, la variable se dice de tipo **discreta**.

Si los valores que puede asumir la variable cuantitativa corresponden potencialmente a cualquier número real, por supuesto en el rango de variación de la misma, la variable se dice **continua**.



Variable cuantitativa discreta


---

# Distribución de una variable cuantitativa discreta

---

Pensemos ahora que se toma una muestra de 15 alumnos y se les pregunta el número de hermanos, las respuestas resultan ser:

0-0-0-1-1-1-1-1-2-2-2-2-3-3-4

- ¿Cuál es la variable? ¿De qué tipo es?
- ¿Cuáles son los distintos valores posibles de esta variable?
- Construyamos una tabla de frecuencias 



# TABLA DE DISTRIBUCIÓN DE FRECUENCIAS

N° de hermanos	$f_j$	$h_j$	$h_j\%$	$F_j$	$H_j$	$H_j\%$
0	3	0,20				
1	5	0,33				
2	4	0,27				
3	2	0,13				
4	1	0,07				
Total	15	1				



Interpretemos la segunda fila

# TABLA DE DISTRIBUCIÓN DE FRECUENCIAS

En la tabla anterior, a las frecuencias absoluta, relativa y relativa porcentual agregamos:

**F<sub>j</sub>**: muestra el número de elementos con valor de la variable menor o igual al valor j. Se llama **frecuencia absoluta acumulada**.

**H<sub>j</sub>**: es la proporción de elementos con valor de la variable menor o igual al valor j. Se llama **frecuencia relativa acumulada** y se puede calcular:  $H_j = \frac{F_j}{n}$

**H<sub>j</sub>%**: es el porcentaje de elementos con valor de la variable menor o igual al valor j. Se llama **frecuencia relativa porcentual acumulada**.



# Gráfico de bastones

---

El gráfico de bastones permite representar gráficamente la distribución de una variable cuantitativa discreta, o la de una variable cuantitativa continua cuando se cuenta con muy pocos datos.

En el eje x se representa cada uno de los valores de la variable, para cada uno de estos valores se diagrama una línea cuya altura es igual a la frecuencia absoluta o relativa observada en la muestra.

Nº de hermanos	$f_j$	$h_j$
0	3	0,20
1	5	0,33
2	4	0,27
3	2	0,13
4	1	0,07
Total	15	1

Gráfico de bastones

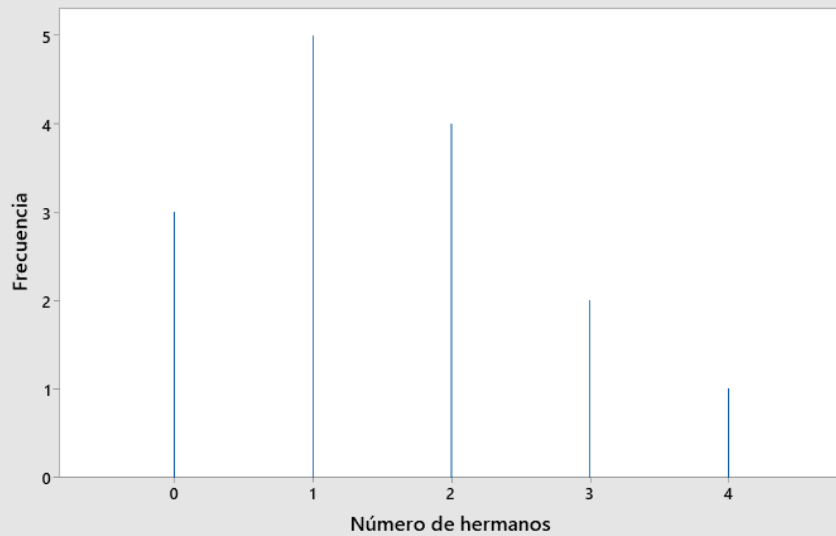
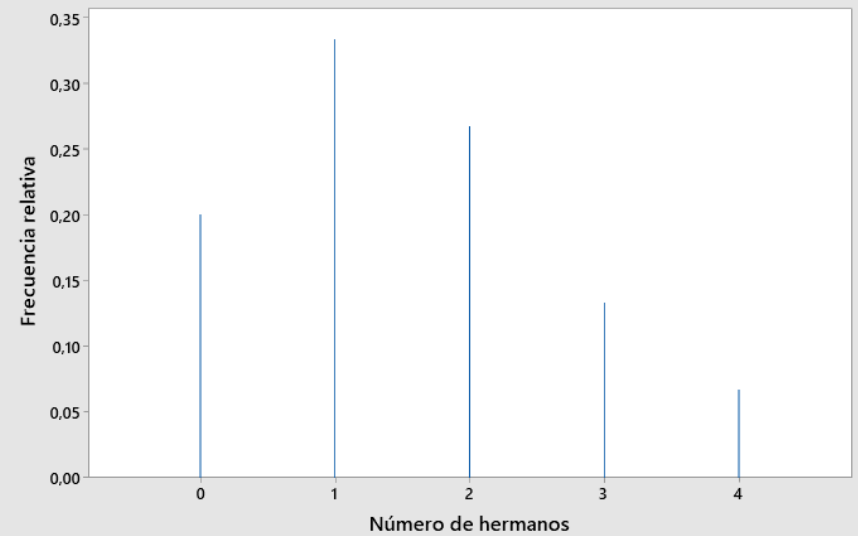


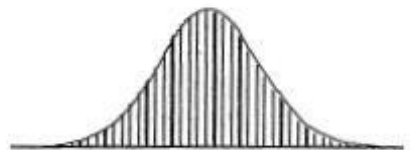
Gráfico de bastones



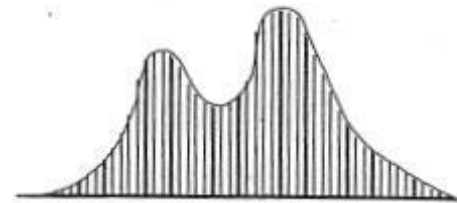


# Formas características de distribuciones

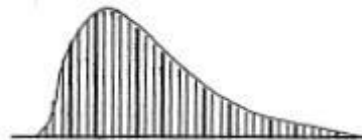
Ejemplos de las formas de distribuciones de algunas características:



Simétrica-Campanular-Unimodal



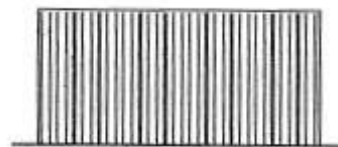
Bimodal



Asimétrica a la derecha  
(asimetría positiva)



Asimétrica a la izquierda  
(asimetría negativa)



Uniforme

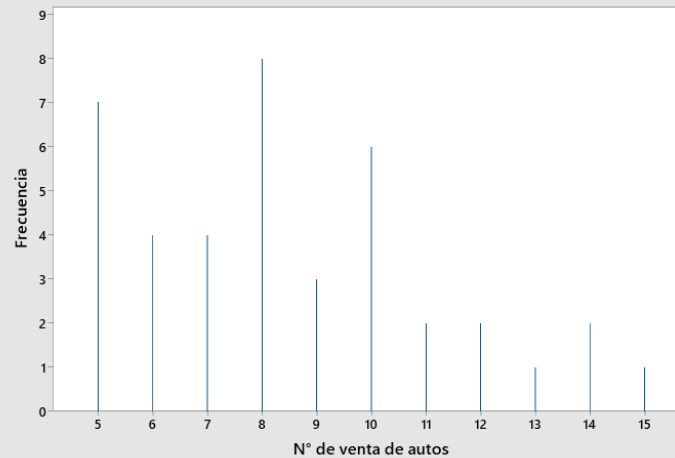
# MANOS A LA OBRA N° 3

Con el propósito de hacer una evaluación de desempeño y un ajuste de cuotas, Ralph Williams estuvo inspeccionando las ventas de automóviles de sus 40 vendedores. En un periodo de un mes, tuvieron las siguientes ventas de automóviles:

7	8	5	10	9	10	5	12	8	7	5	5	14	8	15	14	12	8	8	10
6	10	11	6	5	10	11	10	7	6	8	8	5	6	9	7	5	9	13	8

A- Construya la tabla de distribución de frecuencias.

B- ¿Qué puede decir sobre la forma de la distribución?





Variable cuantitativa continua

---

# Distribución de una variable cuantitativa continua

---

Consideremos la variable Ingreso Total Familiar (ITF) de la EPH.

- ¿Cuáles son los distintos valores posibles de esta variable?

La distribución de frecuencias de una variable continua es una tabla de resumen en la cual los datos se colocan en categorías ordenadas numéricamente (clases, intervalos) establecidas en forma conveniente.

La construcción de la tabla de frecuencias está íntimamente relacionada con el histograma. Este gráfico necesita de una tabla de frecuencias para poder ser construido.



# TABLA DE DISTRIBUCIÓN DE FRECUENCIAS

- Selección del número de clases: en general, se establecen entre 5 y 15 intervalos. Una forma de determinar  $m$ , el número de intervalos, es:  $m = \sqrt{n}$
- Determinación del ancho de un intervalo de clases: se recomienda que cada intervalo tenga la misma medida. Para determinar dicha medida, se divide el rango por el número de intervalos:  $Amplitud = \frac{valor\ máximo - valor\ mínimo}{m}$
- Establecimiento de los límites: de ser posible, los límites entre las clases deben ser de fácil lectura. Los formatos comunes son:  $[LI_j; LS_j)$  o  $(LI_j; LS_j]$  (es decir, intervalo que incluye el límite inferior de la clase o intervalo que incluye el límite superior de la clase).

# TABLA DE DISTRIBUCIÓN DE FRECUENCIAS

Supongamos que se extrajo una muestra de 170 personas de la EPH. Completemos la tabla de distribución de frecuencias para el ingreso total familiar en miles de pesos.

ITF	$f_j$	$h_j$	$h_j\%$	$F_j$	$H_j$	$H_j\%$
[50-60)	30					
[60-70)	55					
[70-80)	43					
[80-90)	25					
[90-100)	17					
Total	170					



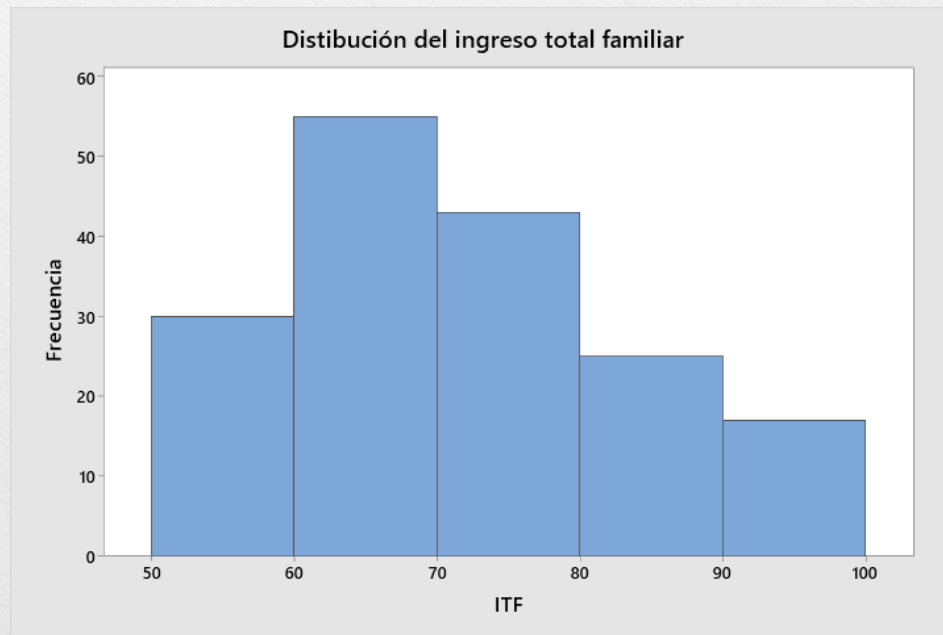
Interpretemos la tercer fila



# Histograma

Para describir una variable cuantitativa continua cuando hay una gran número de observaciones, se utilizan los histogramas.

La tabla de frecuencias realizada anteriormente nos permitirá construir este gráfico.



# MANOS A LA OBRA N° 4

Se cuenta con información de la altura de un grupo de 50 mujeres. Las alturas varían entre 1,46 m y 1,71 m.

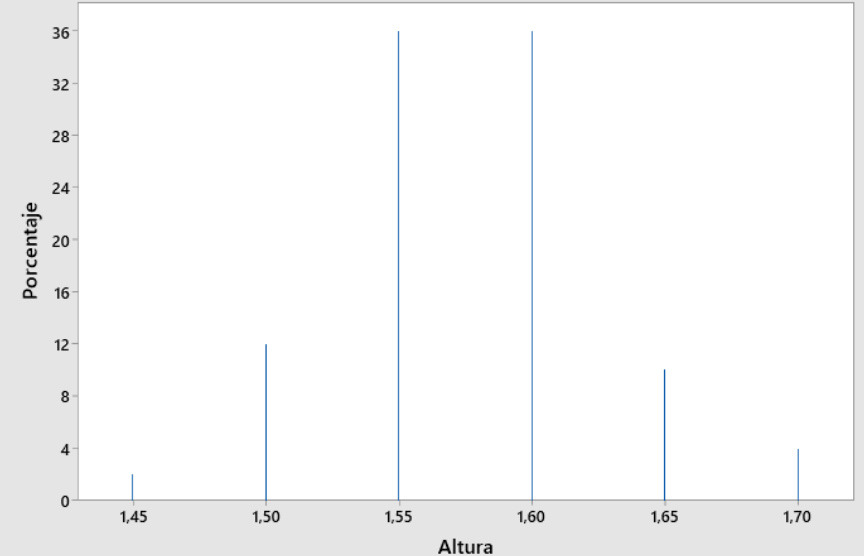
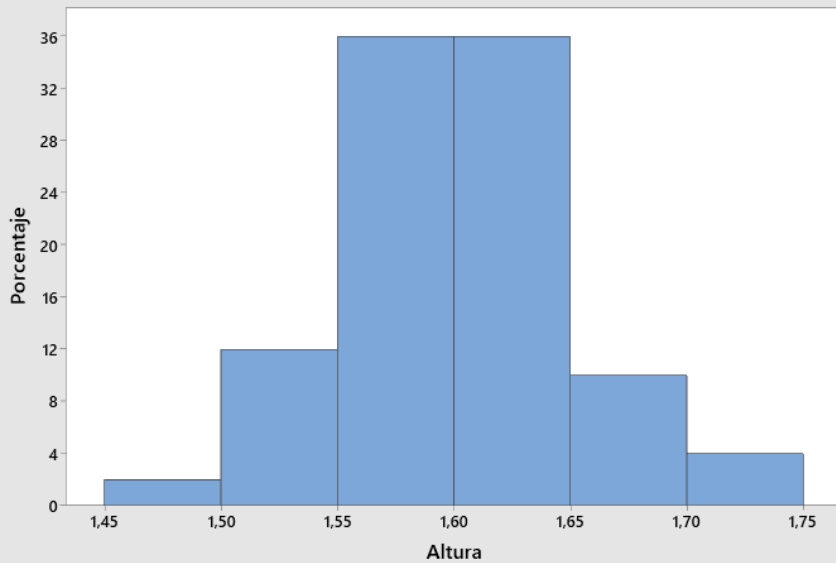
A- Complete la tabla de distribución de frecuencias.

ALTURA	$f_j$	$h_j$	$h_j\%$	$F_j$	$H_j$
[1,45-1,50)	1				
[1,50-1,55)	6				
[1,55-1,60)	18				
[1,60-1,65)	18				
[1,65-1,70)	5				
[1,70-1,75)	2				
Total	50				



# MANOS A LA OBRA N° 4

B- Indique que gráfico representa a este conjuntos de datos. ¿Qué puede decir sobre la forma de la distribución?



# MANOS A LA OBRA N° 5

## HISTOGRAMA DEL PESO DE LAS MUJERES

¿Qué información puede obtener de este gráfico? Describa la distribución de la variable.

