



International Center for Tropical Agriculture
Since 1967 *Science to cultivate change*

Análisis descriptivo de datos

Managua, Nicaragua

Marzo 2019

Autores

Hugo Andrés Dorado

Juan Camilo Rivera

h.a.dorado@cgiar.org , j.c.rivera@cgiar.org

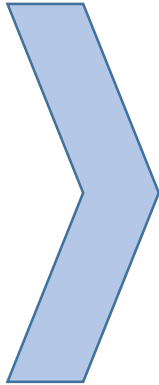


¿Análisis de datos, estadística?

“Es la ciencia que busca comprender al mundo a través de los datos”

Proporciona métodos para:

- Recolectar
- Describir
- Evaluar
- Interpretar



DATOS

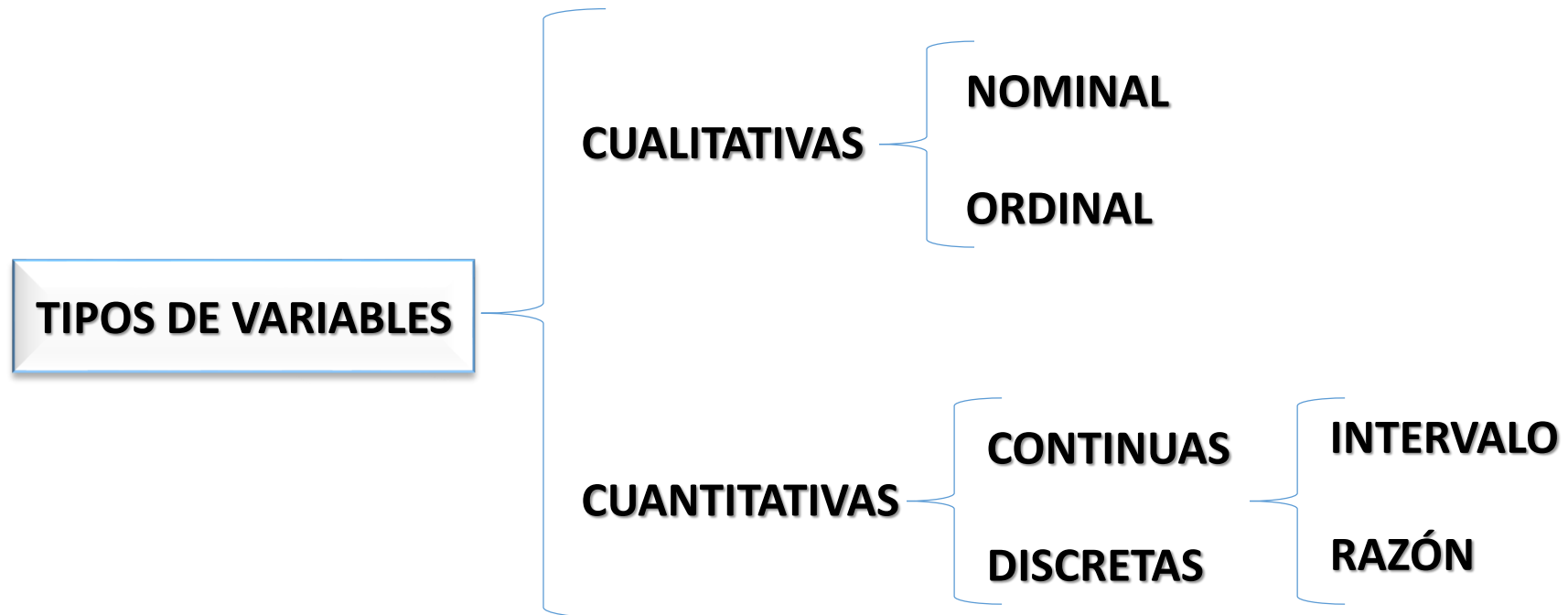
Transfórmalos en información útil para tomar decisiones.

Estadística descriptiva

Recolectar, ordenar y clasificar datos obtenidos por observaciones.

- Los datos del censo poblacional
- La cantidad de nacimientos en Managua la última semana
- La cantidad de goles anotados por Colombia en el último mundial.

VARIABLES



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, Religión, Nacionalidad, Fumar (Si/No)

Ordinales:

Si sus valores se pueden ordenar.
Grado de satisfacción, Intensidad de dolor, Mejoría a un tratamiento.

CUANTITATIVAS

- Si sus valores son numéricos.
- Tiene sentido hacer operaciones algebraicas con ellos.

Discretas:

Si toma valores enteros.
Número de hijos, Número de carros.

Continuas:

Si entre dos valores, son posibles infinitos valores intermedios.
Altura, Temperatura, Duración de una batería, Peso(kg).

ESCALAS DE MEDICIÓN

CUALITATIVAS

1. Escala Nominal:

No puede establecer un orden jerárquico entre las opciones de respuesta.

Color de Ojos (Verde, Azul, Gris, Negro, Café).

2. Escala Ordinal:

Existe un ordenamiento natural de las opciones de respuesta.

Calificación de un servicio (Excelente, Bueno, Regular, Malo).

CUANTITATIVAS

3. Escala de Intervalo:

El valor 0 es un valor arbitrario, no implica la no presencia de una característica.

Temperatura = 0°C ¿No hay temperatura?

4. Escala de Razón:

El valor 0 refleja ausencia de la característica.

Altura = 0 mts

Variables

¿Qué tipo de variables tenemos?

¿Cuál es su escala?

Nombre	#Hijos	Genero	Raza	Salario	Cargo
Diego Giraldo	0	Masculino	Blanca	62.100	Directivo
Diana Sánchez	2	Femenino	Blanca	47.350	Técnico
Julián Castro	1	Masculino	Asiática	18.250	Administrativo
Simón Valdés	1	Masculino	Negra	76.600	Directivo

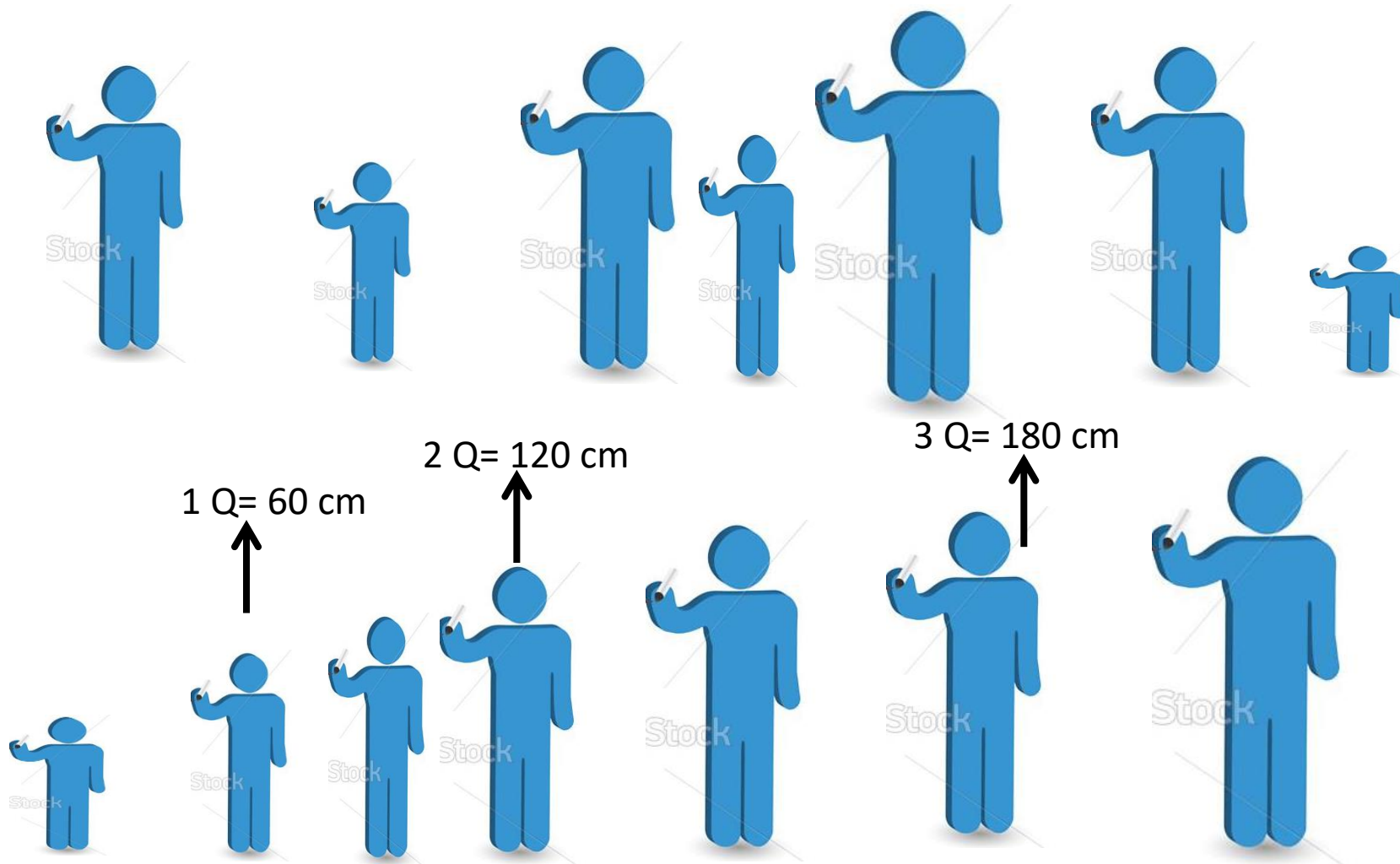
Algunas Medidas descriptivas

CUANTILES: Valores de la distribución que la dividen en partes iguales los mas usados son:

- **Percentiles:** son 99 valores que dividen en cien partes iguales el conjunto de datos ordenados.
- **Cuartiles:** son los tres valores que dividen al conjunto de datos ordenados en cuatro partes iguales.
- **Deciles:** son los nueve valores que dividen al conjunto de datos ordenados en diez partes iguales.

Medidas descriptivas

Cuartil

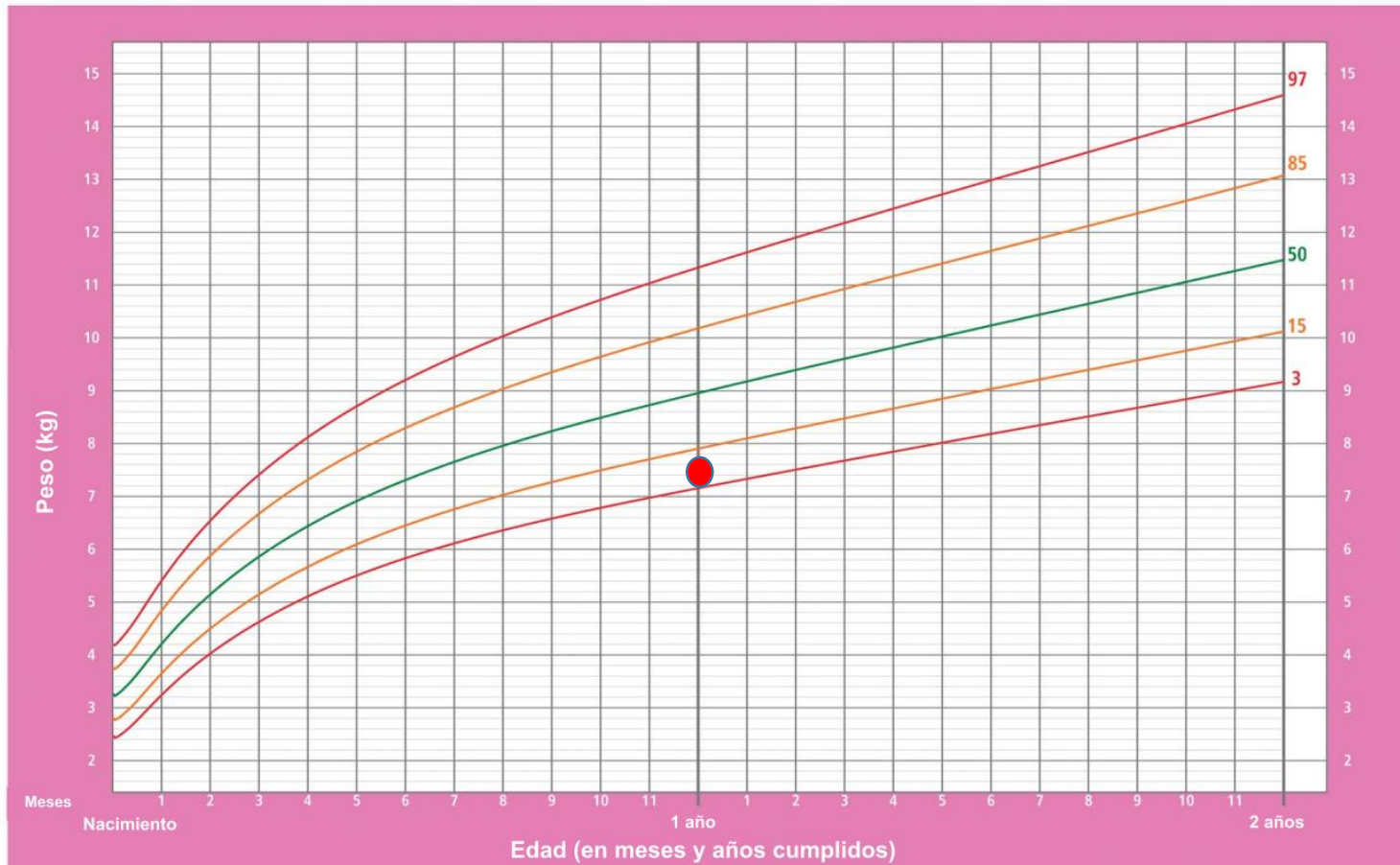


Algunas Medidas descriptivas

Percentil

Peso para la edad Niñas

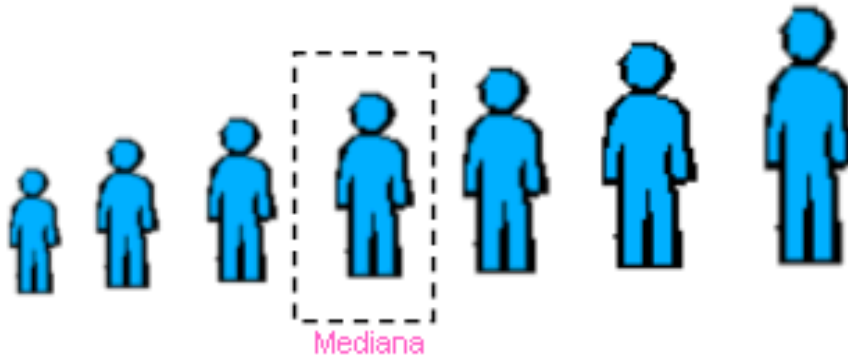
Percentiles (Nacimiento a 2 años)



Patrones de crecimiento infantil de la OMS

Medidas de Centralización

- **Mediana:** es el valor que separa por la mitad las observaciones ordenadas de menor a mayor.



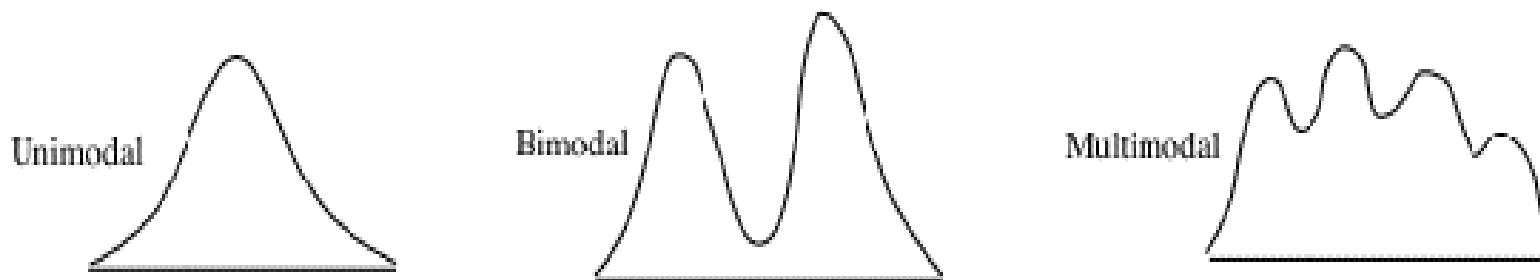
(18, 21, 24, 26, 53)

Me = 24

Mu= 30.8

Medidas Centralización

- **Moda:** es el valor de la variable que más veces se repite, es decir, aquella cuya frecuencia absoluta es mayor.



Medidas de dispersión

- **Varianza:** es el promedio del cuadrado de las distancias entre cada observación y la media aritmética del conjunto de observaciones.

$$V = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

Medidas de dispersión

- **Desviación típica:** La varianza viene dada por las mismas unidades que la variable pero al cuadrado.

$$S = \sqrt{S^2}$$

Medidas de dispersión

- **Recorrido o rango muestral** : Es la diferencia entre el valor de las observaciones mayor y el menor.

$$Re = x_{\max} - x_{\min}$$

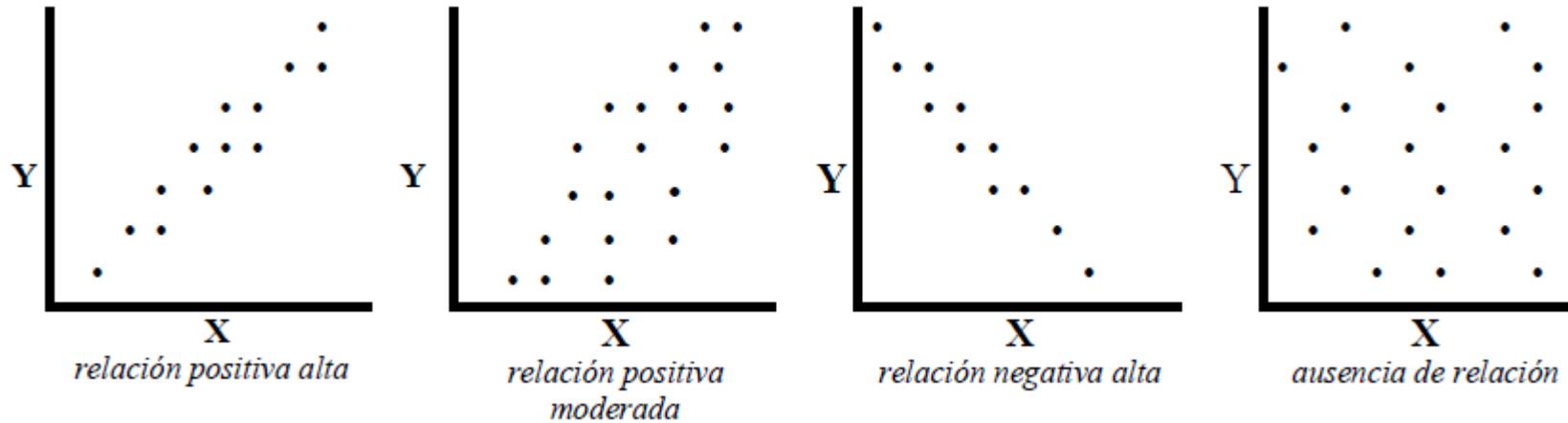
Medidas de dispersión

- **Coeficiente de variación de pearson:** Representa el número de veces que la desviación típica contiene a la media aritmética.

$$CV = \frac{s}{|\bar{x}|}$$

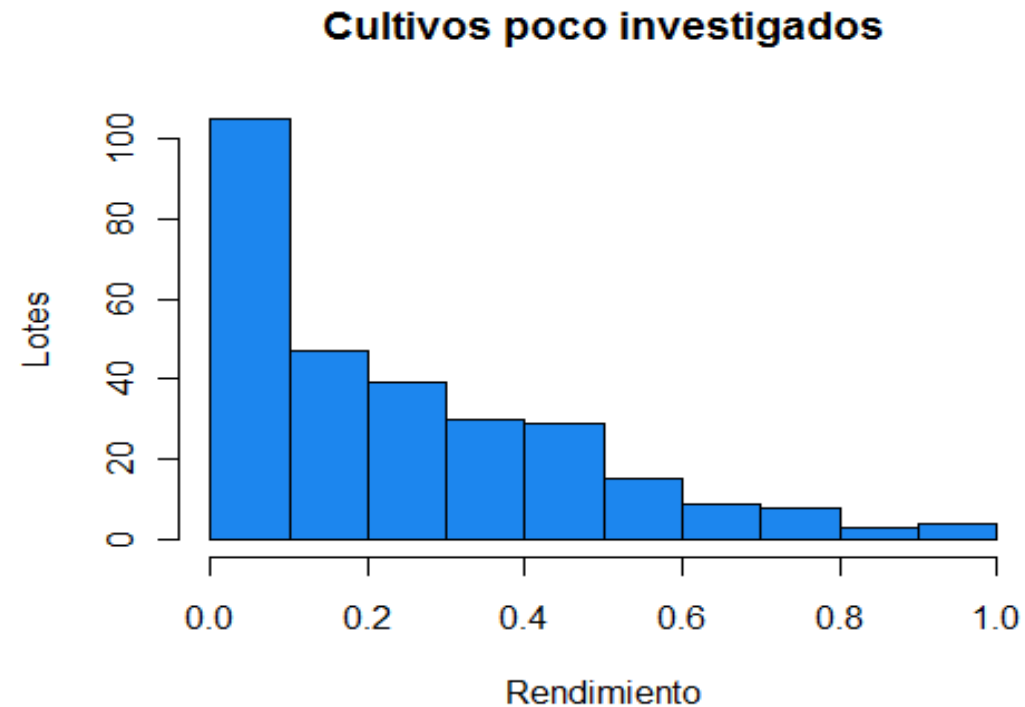
Medidas de covariación y correlación

Relación existente entre dos o mas variables cuantitativas.



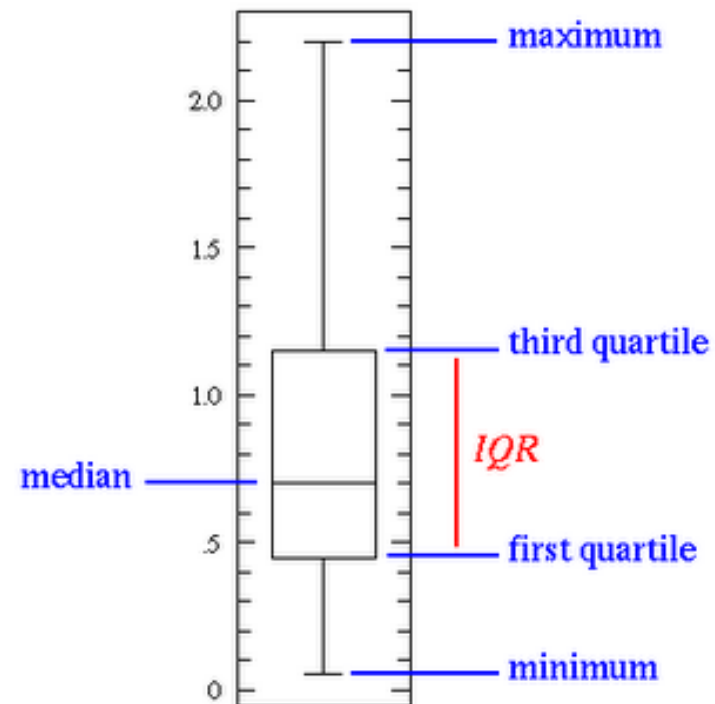
Gráficos descriptivos

- **Variables continuas: Histograma**



Gráficos descriptivos

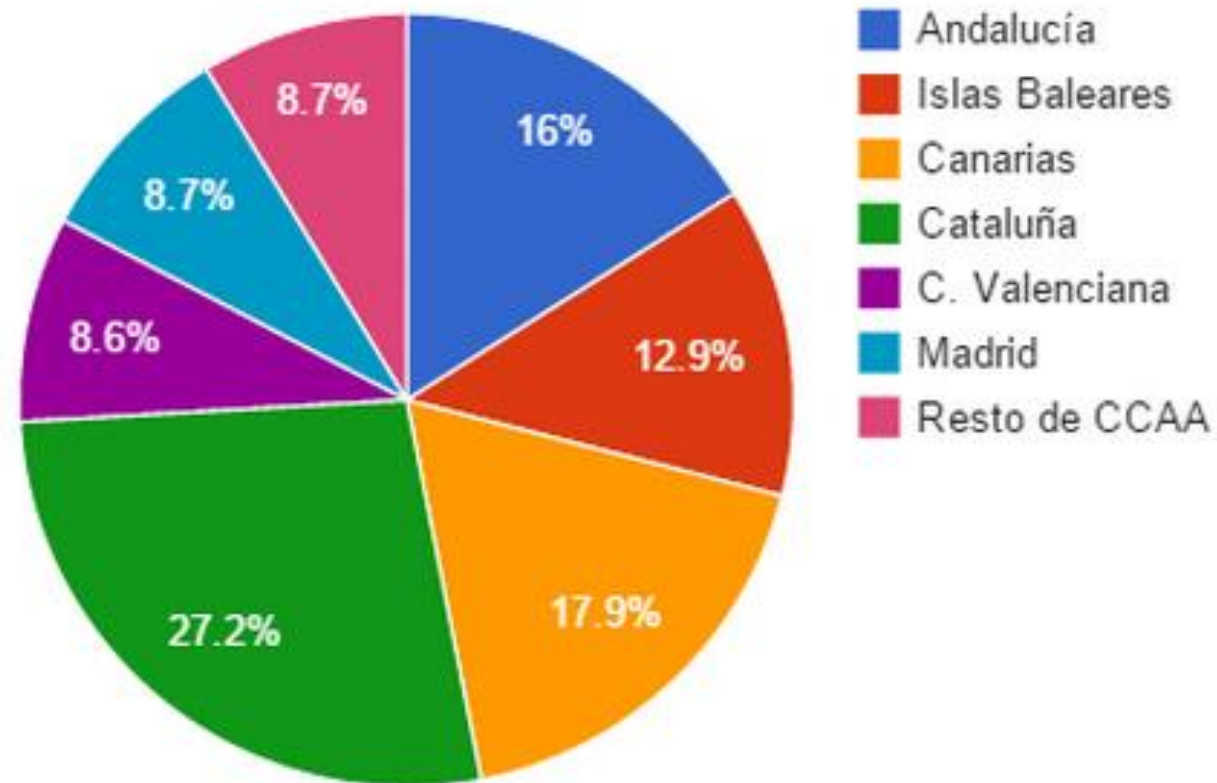
- **Variables continuas:** Boxplot – Diagrama de caja y alambres o de bigotes



Gráficos descriptivos Variables cualitativas

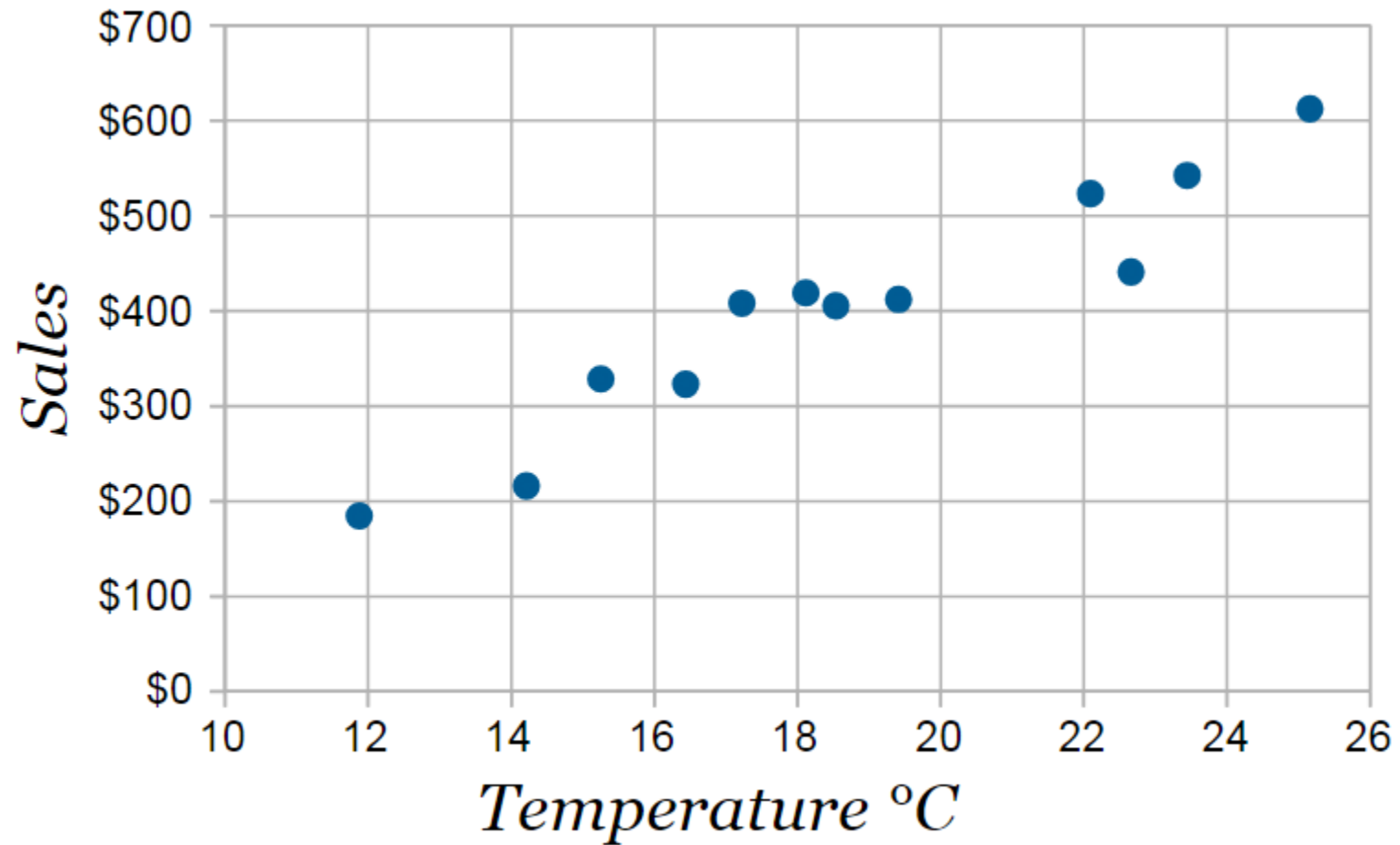
- Grafico circular – Diagrama torta

Recepcion de turistas



Gráficos descriptivos bi variados

- Gráficos de dispersión.



Gráficos descriptivos bivariados.

- Gráfico de barras

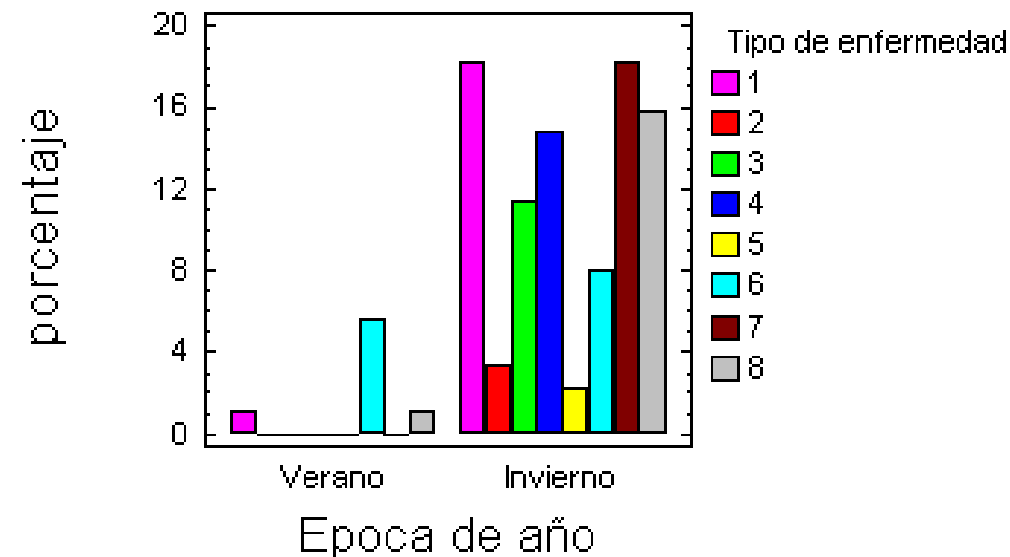


FIGURA 2. Gráfico del tipo de enfermedad y la época del año.

El presente y futuro

- Infografias

<https://www.visme.co/>

<https://www.edrawsoft.com/infographics/>

<https://www.easel.ly/>

