

TEMA 3: VARIABLES ALEATORIAS

Clasificación. Escalas de medición.

Series de frecuencias. Concepto.

Amplitud de variación. Intervalos de clase.

Clases. Frecuencias absolutas, relativas y acumuladas.

Histograma. Curvas de frecuencia.

LOS DATOS

Para el estadístico o investigador, la información requerida proviene de los datos.

Son el recurso básico del investigador, su trabajo consiste en convertir esos datos en información útil para la toma de decisiones.

La calidad del trabajo depende de la calidad de los datos.

¿QUE ES UNA VARIABLE?

Medir una característica de un individuo (persona, objeto, animal, etc.) significa asignarle un número expresable en distintas unidades que la represente. El resultado de esa medición es **variable** y toma diferentes valores dependiendo del individuo a quien se le esta realizando la medición.

Una variable es una medida válida de un concepto si lo representa adecuadamente.

Muchas veces disponemos de un **instrumento** para realizar la medición. Para obtener la longitud de una mesa utilizamos una cinta métrica; estará expresada en centímetros para la UE y Argentina y en pulgadas para Estados Unidos, es decir, para expresar mediciones se pueden utilizar distintas unidades.

La cantidad de piezas defectuosas o la cantidad de días con saldos negativos en una empresa se pueden medir o contar sin demasiadas dificultades.

Ahora bien,

¿Como se mide la seguridad de los empleados o de un fumigador?

¿Cómo se mide el nivel de tecnología que utiliza un empresario?

¿Cómo se mide la eficiencia?



Para medir la seguridad se puede determinar una tasa de accidentes por ejemplo.

Para medir el nivel de tecnología se puede utilizar una escala (nulo, bajo, moderado, alto)

ESCALAS DE MEDICIÓN

- 1. Escala Nominal**
- 2. Escala Ordinal o de Rangos**
- 3. Escala de Intervalos**
- 4. Escala de Razones o Proporciones**

En la medida que avanzamos en la escala hay disponibles mas **TÉCNICAS DE ANÁLISIS DE LOS DATOS**

ESCALAS DE MEDICIÓN

1. Escala Nominal

Se dice que una variable se mide en una escala nominal, si sus valores son etiquetas o atributos y no existe un orden entre ellos.

2. Escala Ordinal o de Rangos

Se dice que una variable se mide en una escala ordinal, si sus valores son etiquetas o atributos pero existe un cierto orden entre ellos.

ESCALAS DE MEDICIÓN

3. Escala de Intervalos

Se dice que una variable cuantitativa se mide en una escala de intervalo si existe una noción de distancia entre los valores de la variable, aunque no existe necesariamente el valor natural cero.

4. Escala de Razones o Proporciones

Se dice que una variable cuantitativa se mide en una escala de razón si los valores de la variable tienen un sentido físico y existe el cero absoluto.

EJEMPLOS DE ESCALAS DE MEDICIÓN

Nominal o Clasificatoria:

Variable: Título de grado que otorga la UNNE

Categorías: Licenciado en RRLL; Médico Veterinario;

Abogado; Licenciado en Administración de Empresas

Ordinal o de Rangos:

Variable: Nivel de Educación

Categorías: Primario, Secundario, Terciario, Universitario.

Variable: Opinión acerca de una determinada política de empleo.

Categorías: Muy de acuerdo, De acuerdo, En desacuerdo,

Muy en desacuerdo

De Intervalos:

Variable: Escalas termométricas. Valores: 14, 25,..., 47 °C.

De Razones o Proporciones:

Variable: Número de alumnos que llegan tarde a la clase, durante un período de tres semanas. Valores: 0,1,2,3, 15.

TIPOS DE VARIABLES

- **CUALITATIVAS** : Expresan una cualidad o característica.
- **CUANTITATIVAS**: se expresan con números.

Discretas: números enteros.

Contínuas: números reales.

Ejemplos de variables de Cualitativas:

- a) Variable: Funciones empresariales;
Categorías: Marketing, Operaciones de producción, Administración de los RRHH, Contabilidad, Finanzas
- b) Variable: Dominio de los automotores registrados en la ciudad de Corrientes en el año 2002.
Categorías: ABC 123; BRJ 341; CDL 485;...
- c) Variable: Tamaño de los comercios de venta de comida de la ciudad de Resistencia.
Categorías: Pequeños, Medianos, Grandes

Ejemplos de variables Cuantitativas:

Discretas:

- Variable: Número de hojas sanas por árbol en un lote.
Valores: 0, 1, 2, 3, 4....25, 100,.....
- Variable: Número de alumnos que llegan tarde a la clase, durante un período de tres semanas.

Continuas:

- Variable: Ingresos diarios de un Maxi quiosco:
Valores: 45,56 - 52,08 - 50,16 kilogramos.
- Variable: Longitud de las piezas fabricadas
Valores: 2,4 – 2,45- 2,39 - 2,42 centímetros.

Cuando se define una variable hay que pensar en los posibles valores que tomaría.

ANÁLISIS DE DATOS

ANÁLISIS EXPLORATORIO DE DATOS

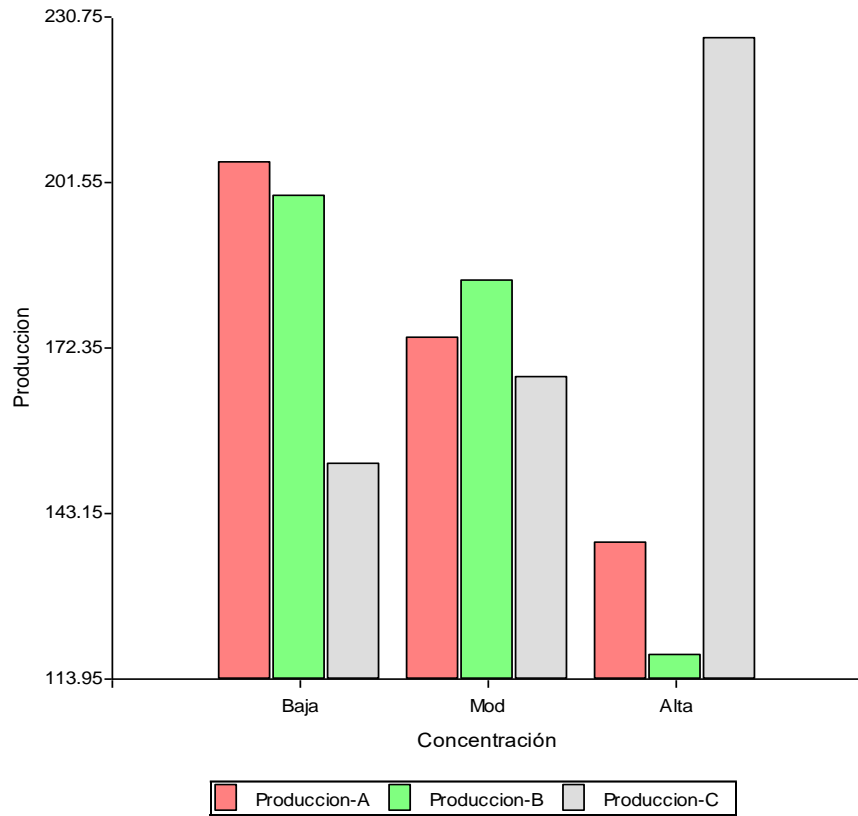
Se analizan los datos para entender la realidad, ellos nos darán indicios de lo que pasa. Por lo tanto debemos resumirlos, sintetizarlos y ser capaces de sacar información de ellos.

Los gráficos son una forma de resumir los datos.

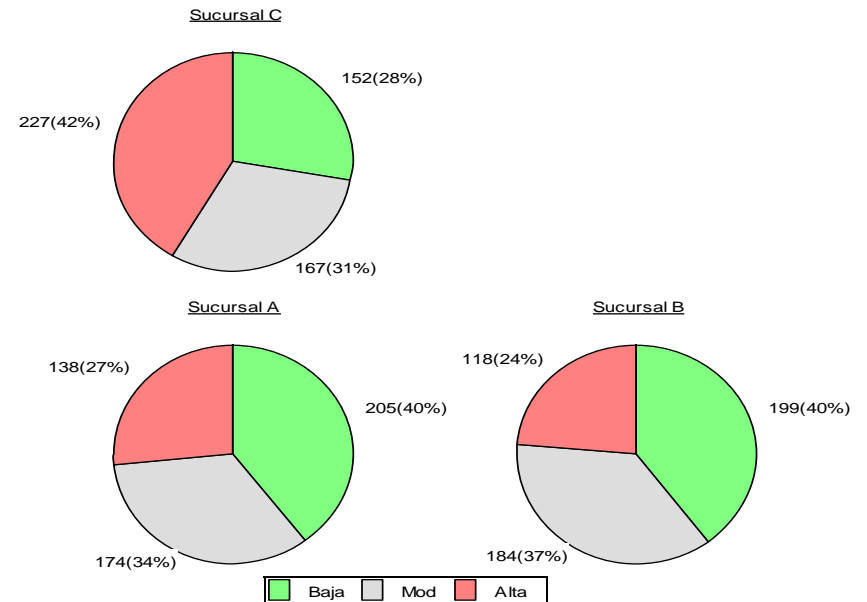
- Gráficos de barra.
- Gráficos de torta.
- Gráficos de caja (boxplots).
- Gráficos de tallo y hoja.

Se mencionan algunos gráficos aunque la actualidad hay tantos gráficos como personas.

Gráficos de barra



Gráficos de torta



Gráficos

Gráfico de tiempo: Este tipo de gráfico se utiliza para examinar la evolución a lo largo del tiempo de alguna variable. Tiene una unidad de tiempo en el eje horizontal (como meses o años) y en el eje vertical alguna cantidad (ingresos de los hogares, tasa de natalidad, ventas totales, porcentaje de la gente en favor del presidente, y así sucesivamente). En cada período de tiempo, la cantidad está representada por un punto, y los puntos están conectados por líneas.

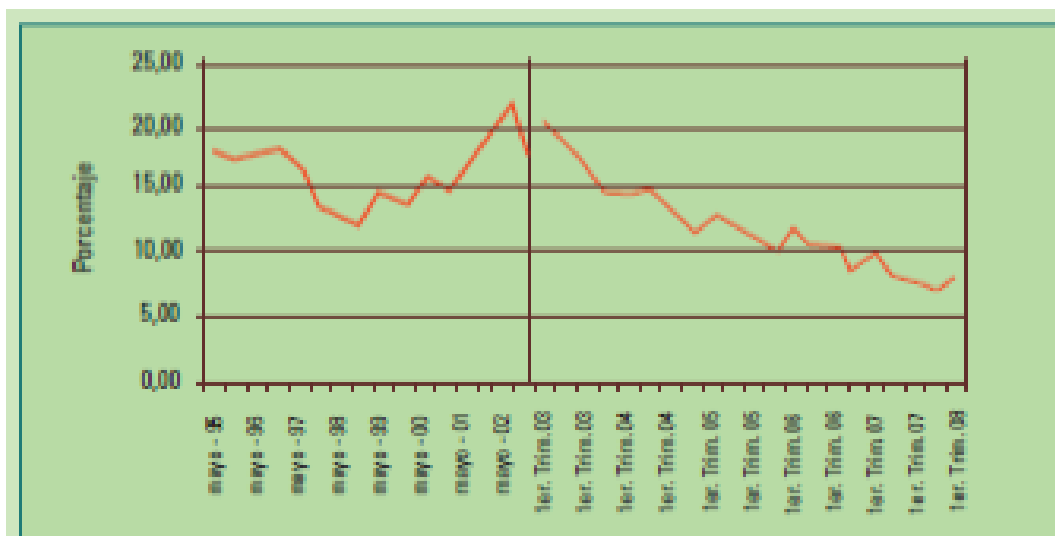
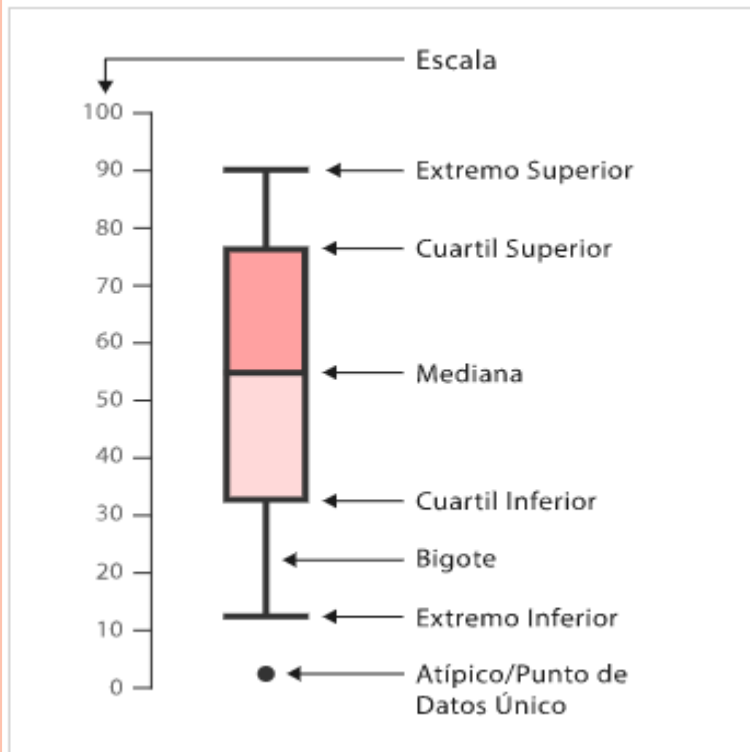


Figura 15.1. Tasa de desempleo según la EPH desde mayo de 1995 hasta octubre de 2002 (con periodicidad bianual, mayo, octubre) y desde el 1er trimestre del 2003 hasta el 1er trimestre de 2008 (con periodicidad trimestral).

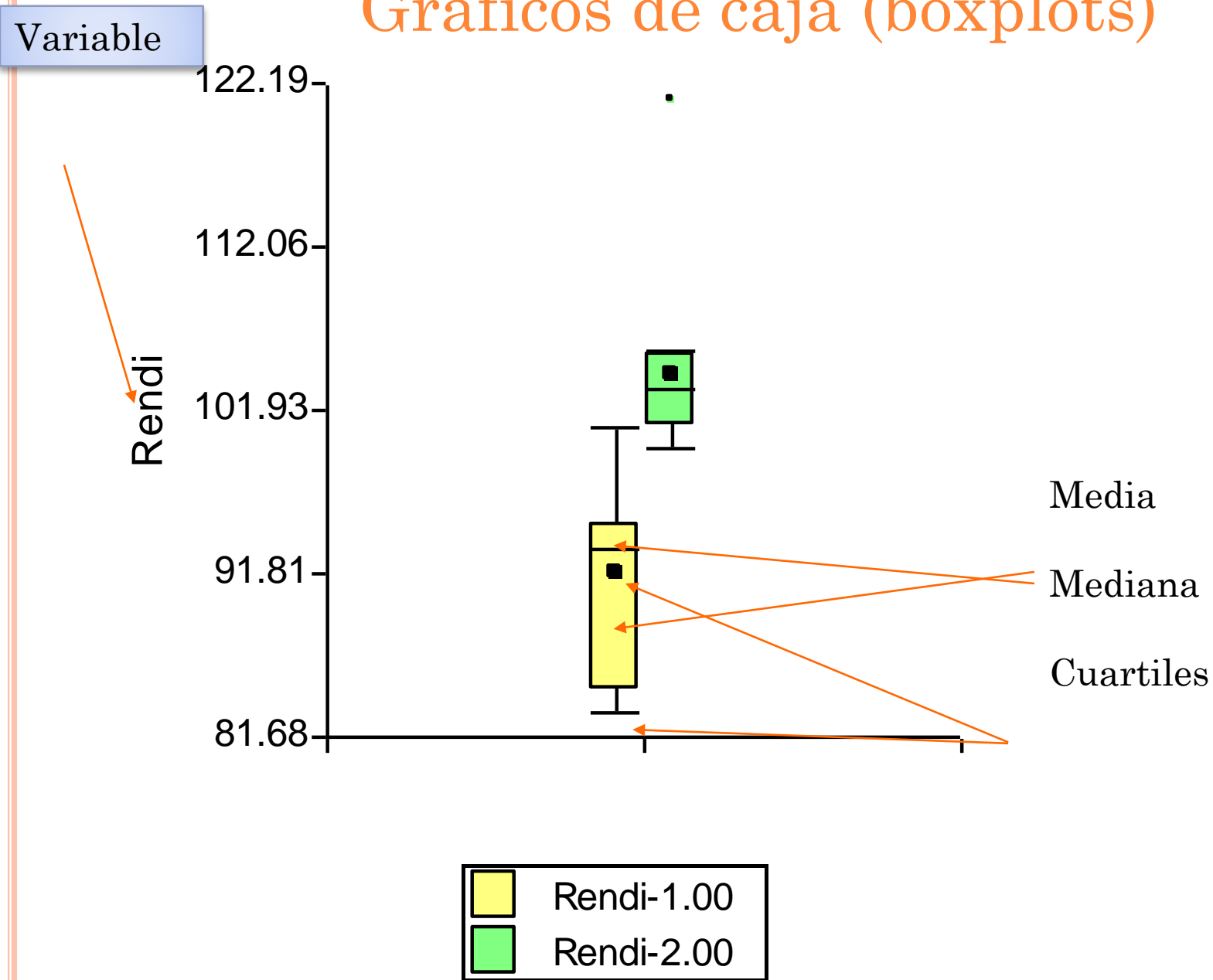
Gráficos de caja (boxplots)



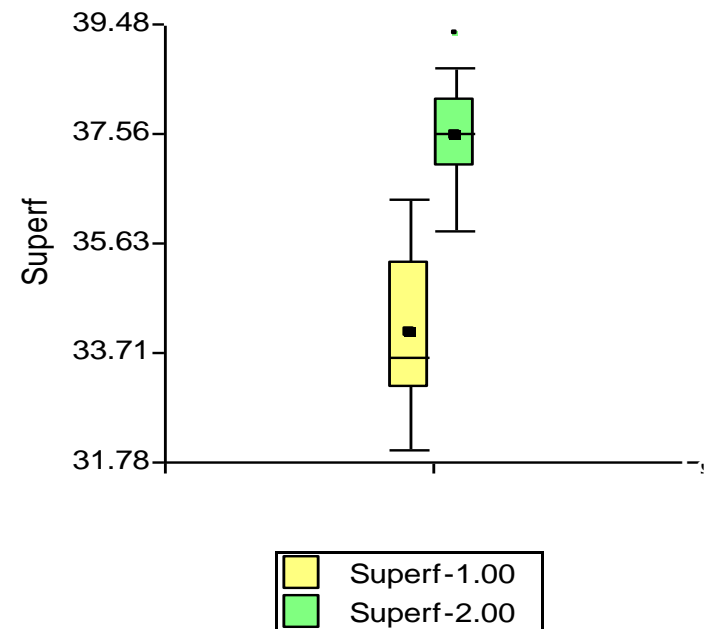
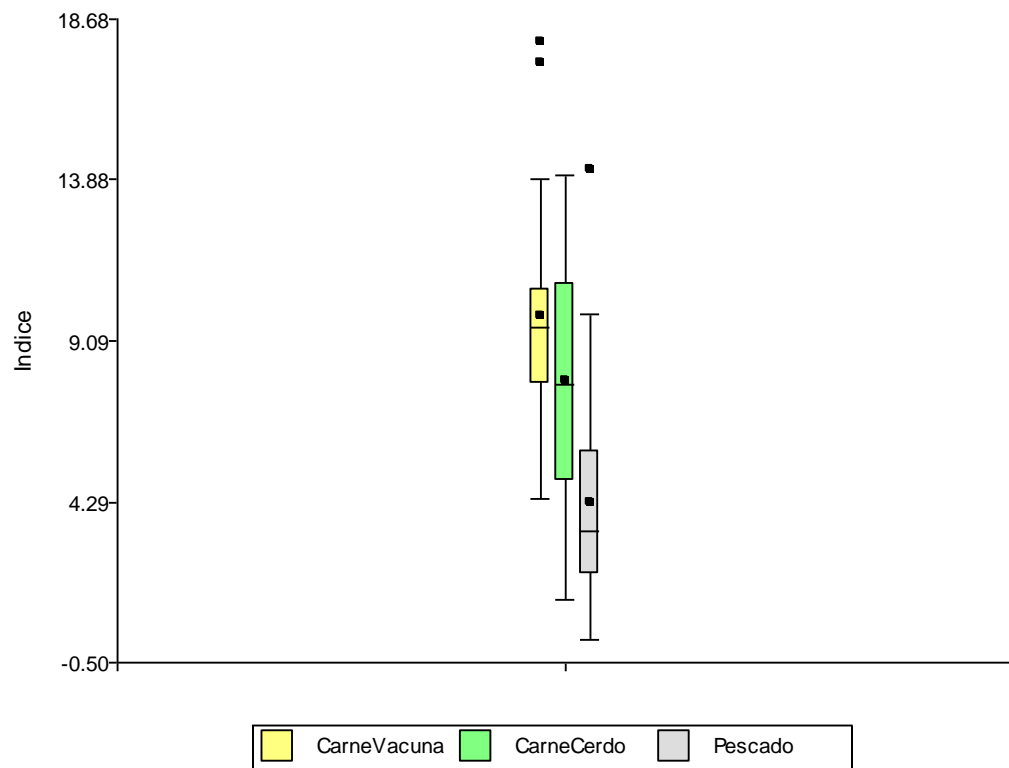
Un diagrama de cajas y bigotes es una manera conveniente de mostrar visualmente grupos de datos numéricos a través de sus cuartiles.

Las líneas que se extienden paralelas a las cajas se conocen como «bigotes», y se usan para indicar variabilidad fuera de los cuartiles superior e inferior. Los valores atípicos se representan a veces como puntos individuales que están en línea con los bigotes. Los diagramas de cajas y bigotes se pueden dibujar vertical u horizontalmente.

Gráficos de caja (boxplots)



Se visualizar 2 cajas, rendimientos de dos grupos 1 y 2. Se observa un dato atípico en el Rendi-2



*Ejemplo de consumo de carne y pescado.
Y el otro ejemplo se registraron datos de
superficie en metro cuadrado de aulas en
2 colegios.*

TABLAS DE FRECUENCIAS

TABLAS DE FRECUENCIAS

- Constituyen un primer análisis de los datos.
- Se construyen con la finalidad de sintetizar información.
- El objetivo de confeccionar tablas de frecuencias es poder analizar las frecuencias con que se presentan los diferentes valores de las variables y su distribución.

TABLAS DE FRECUENCIAS

Una tabla es un Arreglo tabular de las frecuencias con que ocurre cada característica.

Está compuesta por columnas y filas.

Si la variable es cualitativa, los datos se agrupan en **Categorías o Clases.**

Si la variable es cuantitativa los datos se agrupan **Intervalos de Clase y Clases.**

TABLAS DE FRECUENCIAS

La primer columna de la tabla contiene a las **Categorías o Intervalos de clase** y cada fila corresponderá a una categoría o a un intervalo.

Frecuencia absoluta (f_i o FA): Número de observaciones que pertenecen a cada categoría o intervalo.

Frecuencia relativa (f_x o FR): Cociente entre la frecuencia absoluta y el total de observaciones (n).

TABLAS DE FRECUENCIAS

Ejemplo de tablas de frecuencias una encuesta de preferencia que tipo de autos prefieren los norteamericanos.

Origen	FA	FR
American	128	0,38
Europeo	45	0,13
Japones	166	0,49
Total	339	1

Tipo	FA	FR
Familiar	176	0,52
Sport	109	0,32
Trabajo	54	0,16
Total	339	1

Tamaño	FA	FR
Chico	153	0,45
Grande	43	0,13
Medio	142	0,42
Total	338	1

0	1	2
2	3	3
2	5	5
3	3	2
3	3	1
1	4	3
3	5	0
4	2	4
5	3	3
3	4	4

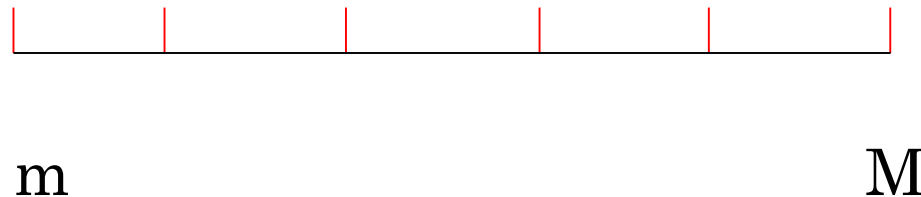
Se consultó a 30 alumnos de un curso sobre la cantidad de llaves que portaban.

X: número de llaves.

Categorías (x)	Frecuencias Absolutas (fi)	Frecuencias Relativas (fx)	Frecuencias Acumuladas Absolutas (Fi)	Frecuencias Acumuladas Relativas (Fx)
0	2	0,067	2	0,067
1	3	0,100	5	0,167
2	5	0,167	10	0,333
3	11	0,367	21	0,700
4	5	0,167	26	0,867
5	4	0,133	30	1
Total	30	1,00		

INTERVALOS

Cuando se trabaja con variables discretas con muchos valores o con variables continuas, es necesario dividir la amplitud total de variación en intervalos de clases.



Para luego determinar las frecuencias dentro de cada intervalo.

CÁLCULO DE INTERVALOS

Cada intervalo esta comprendido entre un **límite inferior** y uno **superior** (los intervalos extremos pueden ser abiertos).

Todos los intervalos deben tener **la misma amplitud**.

El **número de intervalos** de clase será aproximado a la cantidad: $N.I. = 1 + 3,3 \log n$

La **amplitud de cada intervalo**

$A.I. = \text{Amplitud de Variación (A.V.)} / N.I.$ donde

$A.V. = \text{Obs.mayor} - \text{Obs.menor} = M - m$

Clases (x), que se obtienen sumando los límites inferior y superior de cada intervalo y dividiendo el resultado por 2.

DETERMINACIÓN DE LAS FRECUENCIAS

Finalmente se calculan las **Frecuencias Acumuladas Absolutas (F_i)**, de la siguiente manera: la primer frecuencia acumulada es igual a la primer frecuencia absoluta, la segunda frecuencia acumulada será la primer frecuencia acumulada mas la frecuencia absoluta correspondiente a ese intervalo, la tercer frecuencia acumulada será la segunda frecuencia acumulada mas la frecuencia absoluta correspondiente a ese intervalo y así sucesivamente.

Las **Frecuencias Acumuladas Relativas (F_x)**, operando de la misma manera con las frecuencias relativas.

Y se confeccionan los gráficos.

EJEMPLO

Los datos corresponden a los registros de compra de un aserradero, registrándose para las compras el peso neto en toneladas y el número de tablones por camión, y las ventas de los tablones son por pie cuadrado.

El archivo tiene las siguientes variables:

Número= cantidad de troncos por camión

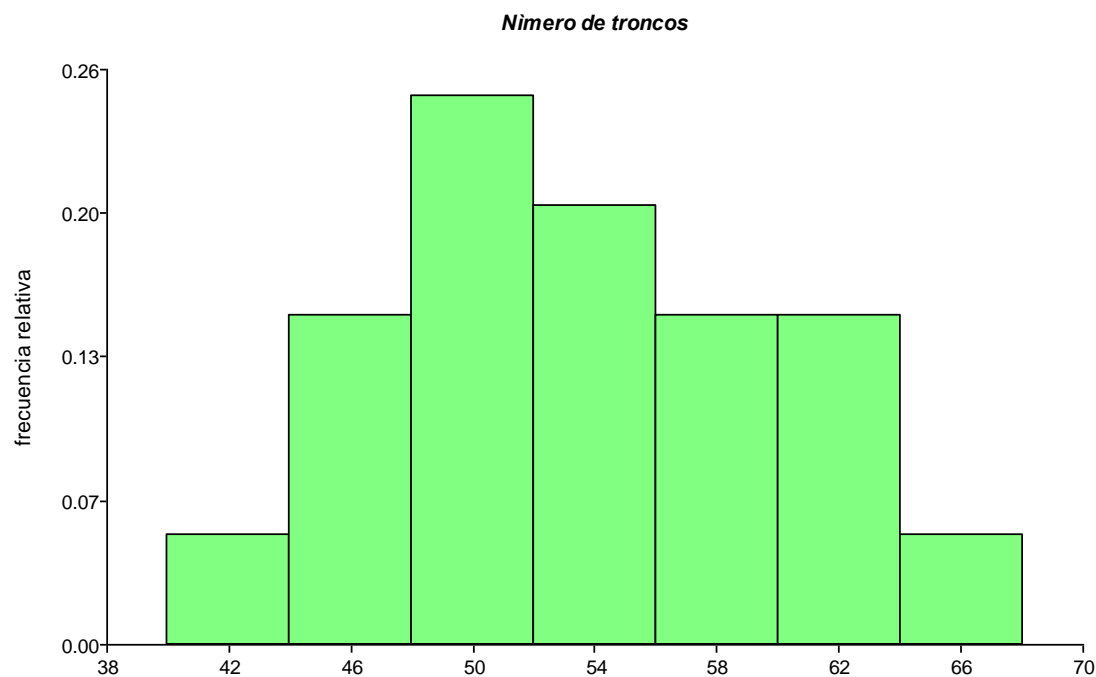
Peso total= toneladas por camión

Peso/numero= relación entre el peso y el numero de troncos

Rendimiento= rendimiento en pies cuadrados por camión

LI	LS	MC	FA	FR	FAA	FRA
40	44	42	1	0,05	1	0,05
44	48	46	3	0,15	4	0,2
48	52	50	5	0,25	9	0,45
52	56	54	3	0,15	12	0,6
56	60	58	2	0,1	14	0,7
60	64	62	4	0,2	18	0,9
64	68	66	2	0,1	20	1

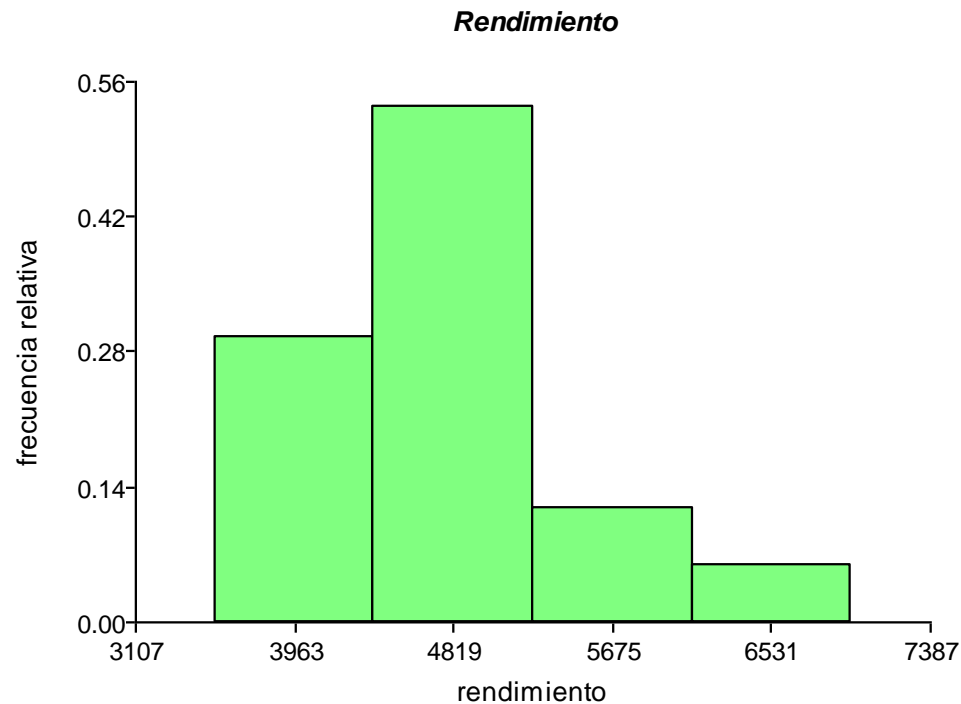
LI: Límite Inferior, LS : Límite Superior; MC: Marca de clase, FA: Frecuecia absoluta, FR Frecuencia Relativa, FAA: frecuencia acumulada absolula y FRA: Acumulada relativa



Tablas de frecuencias RENDIMIENTO

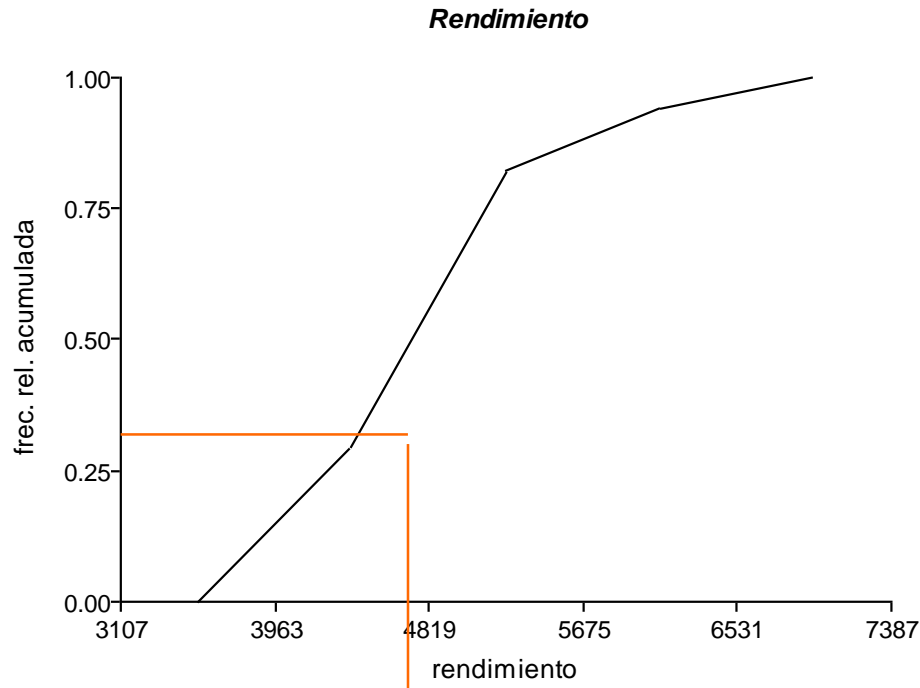
LI	LS	MC	FA	FR
[3535,00	4391,00)	3963,00	5	0,29
[4391,00	5247,00)	4819,00	9	0,53
[5247,00	6103,00)	5675,00	2	0,12
[6103,00	6959,00]	6531,00	1	0,06

HISTOGRAMA :es un gráfico de barras donde el eje “x” es la variable analizada y el eje “y” las frecuencias relativas.



OJIVA: GRÁFICO DE FRECUENCIAS ACUMULADAS

LI	LS	MC	FA	FR	FRA
[3535,00	4391,00)	3963,00	5	0,29	0.29
[4391,00	5247,00)	4819,00	9	0,53	0.82
[5247,00	6103,00)	5675,00	2	0,12	0.94
[6103,00	6959,00]	6531,00	1	0,06	1.00



Ojiva es un gráfico de frecuencias acumuladas, si trazamos una línea desde 0.50 hasta que toque al grafico y allí buscamos su correspondiente x , eso quiere decir que el 50% de los valores de la variable es menor que ese valor de x (4700 aprox .)