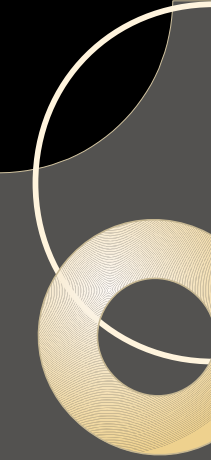


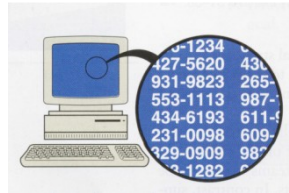
TEMA 4 DE ESTADÍSTICA



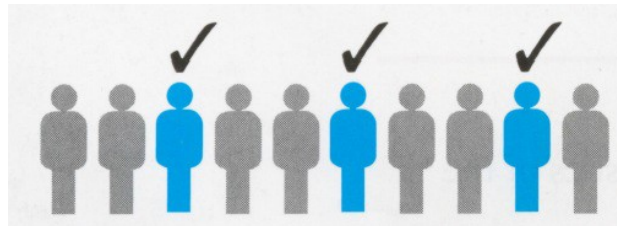
- 
- Muestreo :
 - **Muestreo Probabilísticos:** son aquellos métodos para los que se puede calcular la probabilidad de extracción de cualquiera de las muestras posibles. Los casos seleccionados son elegidos con criterios tales que permitan la generalización a toda la población de los resultados obtenidos al estudiar la muestra.
 - **No probabilísticos:** son aquellos métodos en los que no se puede calcular la probabilidad de extracción de una determinada muestra.

- Tipos de Muestreo:

- **Muestreo Aleatorio Simple:** Cada muestra de un mismo tamaño tiene las mismas posibilidades de ser seleccionada.



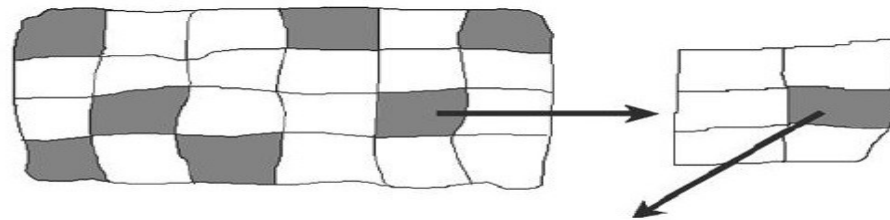
- **Muestreo Sistemático:** Selecciona uno de cada k elementos de la población, al ordenarlos según algún criterio no relacionado con las características que se desean estudiar.



- **Muestreo estratificado:** Divide la población en grupos o estratos homogéneos con respecto a la característica de interés, y toma una muestra aleatoria simple dentro de cada uno de estos estratos.

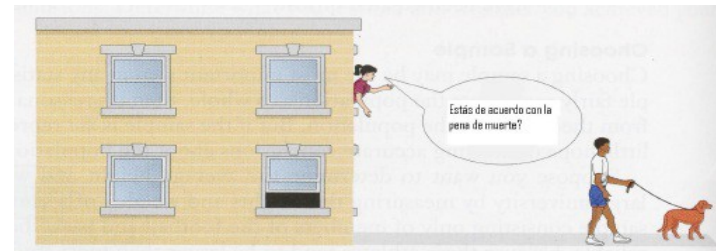


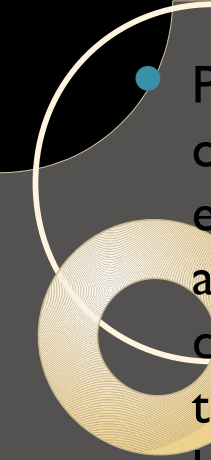
- **Muestreo por conglomerados:** Divide la población en grupos o conglomerados heterogéneos con respecto a la característica de interés, donde se asume que esta heterogeneidad es similar a la de la población completa. Toma una muestra aleatoria simple de conglomerados observando luego todos los elementos contenidos en este.



Se seleccionan todas las casas de la mansana.

- **Muestreo por conveniencia:** Muestreo no probabilístico, usa las unidades mas fáciles de obtener.





● Problema I: Se desea estudiar la composición del agua de un Lago para lo cual se divide en dos zonas: Costa y Centro del Lago. Además, se divide en 2800 cuadrículas de igual tamaño de las cuales se escogen aleatoriamente 80 en la zona Costa y 60 en la zona Centro. En cada una de las cuadrículas seleccionadas se mide el contenido de cierta sustancia toxica (en ppm), el contenido de oxígeno (en ml/100ml) y el numero de larvas de peces. Además se anota el color y calidad de transparencia del agua en la cuadrícula.

- a) Cual es la población en estudio?
- b) Que tipo de muestreo se utilizo?
- c) Cual(es) es (son) la(s) variable(s) en estudio, de que tipo es (son) y cual es el nivel de medición apropiado? (cualitativa o cuantitativa, nominal, ordinal, discreta o continua).
- RESPUESTAS: la población es el lago. El muestreo es un muestreo por conglomerado y aleatorio simple probabilistico. La variable independientes son consta y centro que es dicotomica y las variables dependientes son el oxígeno , la toxicidad , y el numero de larvas de peces



Datos perdidos o missing data:

- a) Dejar en blanco
- b) Dar valor raro para poder diferenciar ej (diferenciar no examen por enfermedad o perder la evaluación continua)
- c) Importante por formulas los missing data no han de contarse



- **Depuración de datos :**

- A. Depuración de datos : tomar encuesta y comprobar

- B. Control de calidad mirar que todo este bien

-  **Distribución de frecuencias:**

Forma de controlar y que todo este en orden explicación oficial:

Consideremos una población estadística de n individuos, descrita según un carácter o variable C cuyas modalidades han sido agrupadas en un número k de clases, que denotamos mediante c_1, c_2, \dots, c_k . Para cada una de las clases $c_i, i = 1, \dots, k$, introducimos las siguientes magnitudes:

-  **Frecuencia absoluta**

-  **Frecuencia relativa**

-  **Frecuencia absoluta acumulada**

-  **Frecuencia relativa acumulada**



□ porcentaje:

valido

acumulado

sirve para percentil y quartil

Ejercicios:

- La tabla muestra una distribución de frecuencias de la duración de 400 contratos temporales realizados durante el año pasado:

- a) Límite superior de la quinta clase.

- b) Límite inferior de la octava clase.

- c) Marca de clase de la séptima clase.

- d) Tamaño del intervalo de clase.

- e) Frecuencia de la cuarta clase.

- f) Frecuencia relativa de la sexta clase.

- g) Porcentaje de contratos cuya duración es menor a las 600 horas.

- h) Porcentaje de contratos cuya duración es mayor o igual a 900 horas.

- i) Porcentaje de contratos cuya duración es al menos de 500 horas pero menor de 1000 horas.

Duración (horas)	Número de Contratos
300 – 400	14
400 – 500	46
500 – 600	58
600 – 700	76
700 – 800	68
800 – 900	62
900 – 1000	48
1000 – 1100	22
1100 – 1200	6

- Decir cuáles de los que siguen representan variable discreta o variable continua:
 - a) Número de accidentes producidos en un determinado mes.
 - b) Número de incendios producidos en una localidad.
 - c) Temperaturas medidas en un observatorio.
 - d) Centímetros de lluvia en una ciudad.
 - e) Medición sismológica a la semana.
 - f) El tiempo que le tomará atender un paciente.
 - g) El número de rescates realizados en una determinada ciudad.

210	212	212	200	220	210	208	204	203
203	220	212	220	205	204	203	205	217
218	215	205	220	204	202	201	209	213
218	214	213	200	207	202	208	214	214
216	218	217	213	211	212	215	214	217
217	204	207	216	214	214	217	212	210