



Recursos y materiales de apoyo

“Muestreo y tablas de frecuencia”

Iniciemos con algunas preguntas para identificar nuestros conocimientos previos sobre el tema:

I. Utiliza un libro de probabilidad y estadística o bien recursos en la web para investigar los siguientes conceptos, después, relaciónalos con su definición.

Respuesta	Concepto	Definición
	Población	a) Es cuando en un estudio se eligen los elementos, en función de que sean representativos, según la opinión del investigador
	Muestreo	b) Es una forma de realizar un muestreo por donde las unidades muestrales no son elementos individuales de la población, sino grupos de elementos.
	Muestra	c) Forma de organizar un conjunto de datos, clasificándolos en categorías o clases. Ordenamiento o arreglo de los datos, clases o categorías; indicando para cada una de ellas, el número de elementos que contiene (frecuencia).
	Muestreo aleatorio	d) Se divide la población total en clases homogéneas, por ejemplo, por grupos de edades, por sexo. Hecho esto la muestra se escoge aleatoriamente en número proporcional al de los componentes de cada clase o estrato.
	Muestreo aleatorio simple	e) Número de observaciones que son menores que el límite superior de la clase, se obtiene sumando las frecuencias (absolutas o relativas) en sentido descendente.



	Muestreo aleatorio sistemático	f) Número de veces que aparece, sucede o se realiza una cosa durante un período o un espacio determinados.
	Muestreo aleatorio Estratificado	g) Este tipo de variables sólo nos permite establecer relaciones de igualdad/desigualdad entre los elementos de la variable. La asignación de los valores se realiza en forma aleatoria por lo que NO cuenta con un orden lógico.
	Muestreo aleatorio por conglomerados	h) Se ordenan previamente los individuos de la población; después se elige uno de ellos al azar, a continuación, a intervalos constantes, se eligen todos los demás hasta completar la muestra.
	Muestreo no aleatorio	i) Este tipo de variables nos permite establecer relaciones de igualdad/desigualdad y a su vez, podemos identificar si una categoría es mayor o menor que otra.
	Niveles de medición: nominal	j) Selección de un conjunto de personas o cosas que se consideran representativos del grupo al que pertenecen, con la finalidad de estudiar o determinar las características del grupo.
	Nivel de medición: ordinal	k) De Intervalo, De razón
	Distribución de frecuencias	l) Elegido el tamaño de la muestra, los elementos que la compongan se han de elegir aleatoriamente entre los N de la población.
	Nivel de medición métrico	m) Depende de la cantidad de datos u observaciones o de la amplitud en general. El número de clases recomendable está asociado con la cantidad de datos.
	Clases	n) Es el conjunto total de individuos susceptibles de poseer la información buscada.
	Frecuencia	o) Número de elementos u observaciones pertenecientes a una misma clase.

	Frecuencia absoluta	p) Se obtiene dividiendo la frecuencia absoluta por el número de observaciones.
	Frecuencia relativa	q) Se eligen los elementos, en función de que sean representativos, según la opinión del investigador.
	Frecuencia acumulada.	r) Es la parte de la población en la que se miden las características estudiadas.

II. Ahora, es tiempo de utilizar los conceptos anteriores en situaciones prácticas, para ello hagamos una situación hipotética sobre tu proyecto.

Situación hipotética

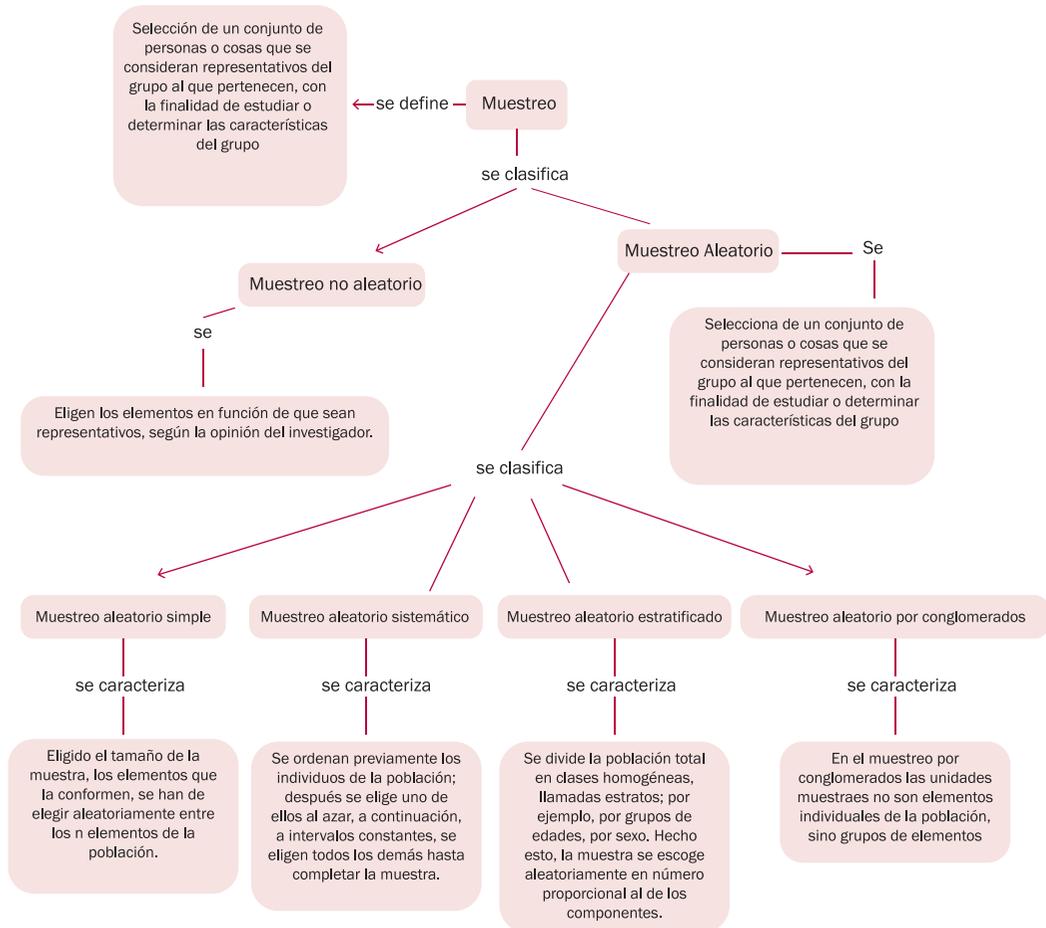
En la empresa que vas a presentar en el macroproyecto, necesitas empleados. Supongamos que para hacer funcionar tu empresa necesitas contratar 20 empleados. Para darlos de alta en el IMSS, te solicitan conocer la edad, género, nacionalidad.

¿Cómo puedes obtener estos datos?

¿Usarías el mismo proceso si tuvieras 100, 200 o 500 trabajadores?

En la situación anterior la población o Universo de estudio es de: 20, 100, 200 o 500 empleados, cuando se realiza un estudio no siempre es factible obtener los datos de cada individuo, en estos casos hacemos una selección, es decir, tomamos una parte de la población, la cual llamamos muestra. Para obtener el tamaño de la muestra y para definir la forma en cómo se van a obtener los datos, existen ciertas técnicas de muestreo, el más general es un muestreo aleatorio o al azar para dar cierto grado de confiabilidad al estudio que se va a realizar.

Analicemos algunos tipos de muestreo:



III. Para poner en práctica y verificar que diferenciamos los tipos de muestreo aleatorio, clasifica cada uno de las siguientes muestras como: aleatoria simple, aleatoria sistemática, estratificada o conglomerados (cuotas):

Muestra	Tipo de muestreo
En mi trabajo somos 10 empleados, el dueño quiere seleccionar una muestra aleatoria de tres empleados para ir a un viaje de trabajo y descanso a una de las sucursales de la playa. El dueño numera nuestros nombres y utiliza una tabla de dígitos aleatorios.	
En una escuela preparatoria hay 5 grupos de primer semestre, 5 de segundo semestre, ... ,5 de sexto semestre. Se eligen aleatoriamente a tres estudiantes de cada grupo para evaluar a sus maestros.	
En una escuela hay 30 grupos, se seleccionan al azar 5 grupos y encuestan a todos sus estudiantes.	
En una imprenta donde también trabajan la encuadernación sospecha que la máquina no está poniendo suficiente pegamento en el lomo del libro. Deciden inspeccionar en los últimos 50 libros para probar su teoría. Se etiquetan los libros con los números del 01 al 50 y se usa una tabla de dígitos al azar para seleccionar 5 libros y revisarlos.	
El profesor de la clase de matemáticas toma el número de lista que tiene asociado cada uno de sus estudiantes para seleccionar 10 y solicitarles que llenen una encuesta acerca de los servicios de la biblioteca.	
De una lista de 100 personas, se selecciona a todas las personas que son asociadas con un múltiplo de 5 para que contesten una encuesta sobre percepción de seguridad.	
En un evento masivo se eligen al azar la décima persona que se encuentra en una fila para el ingreso, a estas personas seleccionadas se les aplica una prueba rápida de COVID 19 para evitar contagios.	
Una franquicia internacional selecciona aleatoriamente algunas de sus ubicaciones para realizar un control de calidad. Los inspectores revisan cada área de las sucursales elegidas.	

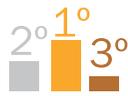
En el aeropuerto Internacional Miguel Hidalgo, se selecciona aleatoriamente 10 vuelos de procedencia internacional y se les aplica una prueba rápida de COVID 19.	
Del personal de una escuela se selecciona 5 personas del área administrativa, 5 personas del área de docentes, 5 personas del área de operativos para realizar una encuesta sobre su percepción de seguridad en el lugar de trabajo.	
En una rifa se asocia el nombre del comprador del boleto con un número de una bola en un recipiente, luego se sacan bolas sin mirar para seleccionar una muestra de ganadores.	

Una vez que se decidió el tipo de muestra mediante la cual vamos a obtener datos, necesitamos medirlos, para ello existen tres niveles de medición.

Niveles de medición



NOMINAL: Este tipo de variables sólo nos permite establecer relaciones de igualdad/desigualdad entre los elementos de la variable. La asignación de los valores se realiza en forma aleatoria por lo que NO cuenta con un orden lógico.



ORDINAL: Este tipo de variables nos permite establecer relaciones de igualdad/desigualdad y a su vez, podemos identificar si una categoría es mayor o menor que otra.



MÉTRICO: Que puede ser de Intervalo o de razón.

Es tiempo de pasar a aprender o reaprender formas de organizar la información, para ello están las tablas para representar información, que suele llamarse distribución de frecuencias.

Distribución de frecuencias

Se refiere a la forma de organizar un conjunto de datos, clasificándolos en categorías o clases. Es el ordenamiento o arreglo de los datos o clases o categorías indicando para cada una de ellas, el número de elementos que contiene (frecuencia).

Las clases

- Dependen de la cantidad de datos u observaciones o de la amplitud en general.
- El número de clases recomendable está asociado con la cantidad de datos.

Supongamos que después de preguntar a tus 20 empleados obtuviste las siguientes respuestas:

Edad: 21, 25, 40, 35, 35, 35, 33, 21, 23, 23, 23, 22, 21, 23, 24, 24, 23, 23, 22, 23

Género: F, F, F, M, M, M, F, F, F, M, M, M, M, M, M, F, M, M, F, M.
donde F=femenino y M=masculino

Nacionalidad: E, M, M, M, M, M, M, E, M, M.

IV. ¿Consideras que es la forma adecuada de presentar los datos?

Tu respuesta al punto anterior habrá sido negativa, una forma algo atractiva o elegante expresar los datos es mediante una tabla, que se le va llamar tabla de frecuencias, donde la frecuencia es el número de veces que se repite un dato.

V. Con la información anterior, completa las siguientes tablas de frecuencias:

Género	
	Frecuencia
Masculino	
Femenino	
Total	

Nacionalidad	
	Frecuencia
Mexicano	
Extranjero	
Total	

Las tablas de frecuencias que acabas de contestar se les conoce como tablas de frecuencias absolutas y representan el número de elementos y observaciones pertenecientes a una misma clase.

Existe otros tipos de frecuencias, la frecuencia relativa que se obtiene dividiendo la frecuencia absoluta por el número de observaciones, y denota la importancia de la clase al expresarla en términos porcentuales. La Frecuencia acumulada se obtiene sumando las frecuencias absolutas o relativas en sentido descendente.

VI. Completa las siguientes tablas de frecuencias, considera los ejemplos:

Género			
	Frecuencia	Frecuencia Relativa	Frecuencia acumulada
Masculino		$\frac{12}{20} = 0.6 = 60\%$	
Femenino			20
Total		$\frac{20}{20} = 1 = 100\%$	

Nacionalidad			
	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Acumulada
Mexicano			
Extranjero			
Total			

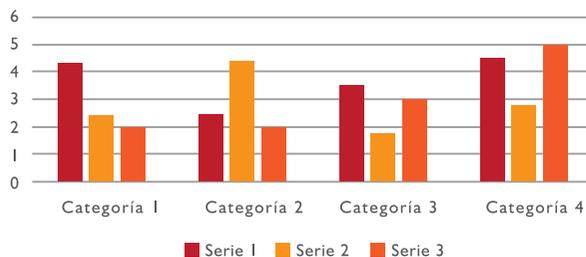
Problemario



Las tablas de frecuencia nos muestran una forma agradable de representar la información, existe otra forma de presentar la información y es mediante el uso de gráficas estas contribuyen a un mejor análisis de los datos y facilitan la comprensión del fenómeno considerado. Las representaciones gráficas:

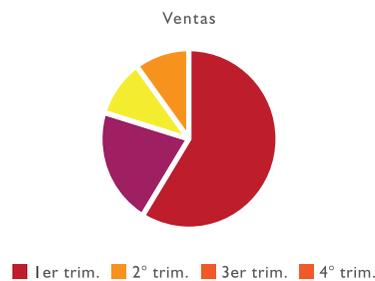
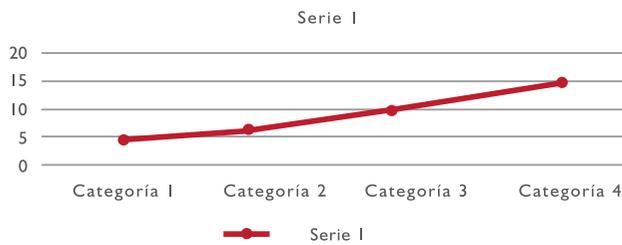
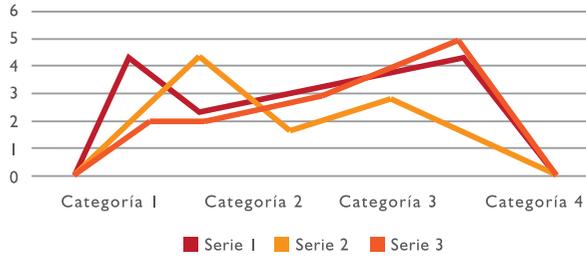
- Contribuye a un mejor análisis de los datos
- Facilita la comprensión del fenómeno considerado.
- Pierde detalle de información, pero se obtiene otro tipo de información.

Las gráficas más utilizadas son el histograma o gráfico de barras, el polígono de frecuencias, la ojiva y cuando se tienen pocos datos la gráfica circular o de pastel.



Histograma o gráfico de barras

- Gráfico de barras verticales.
- Definidas las escalas en el eje cartesiano se dibuja un rectángulo acorde a la frecuencia de la clase (altura).
- En la clase de igual amplitud las barras son proporcionales a la frecuencia de la clase.



Polígono de frecuencias

- La altura de cada punto la determina el punto medio (abscisa y la frecuencia simple (ordenada) de la clase.
- Los puntos se unen con secciones rectas.
- Se procede como si existiera una clase adicional al principio y al final, ambas frecuencias son cero.
- Útiles para representar dos distribuciones de frecuencia en un mismo gráfico.

Ojiva

- Para representar la frecuencia acumulada.
- La ordenada se levanta sobre el límite superior.
- Tiene forma de S alargada.

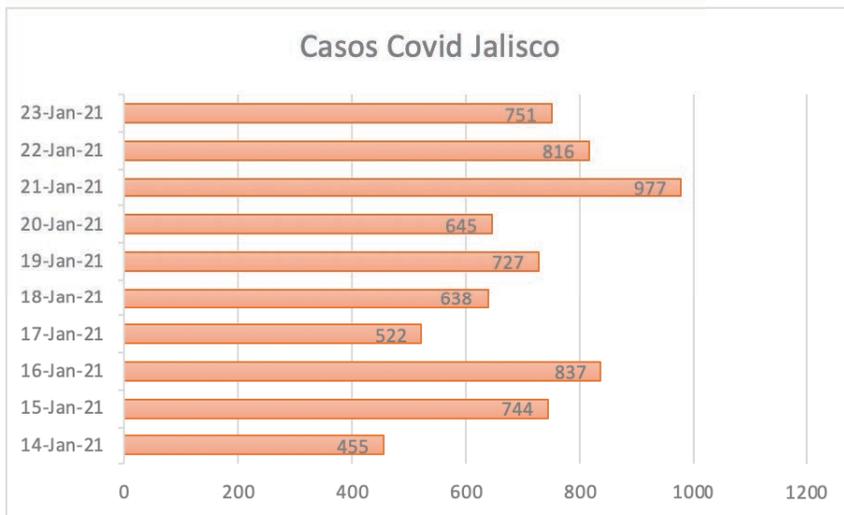
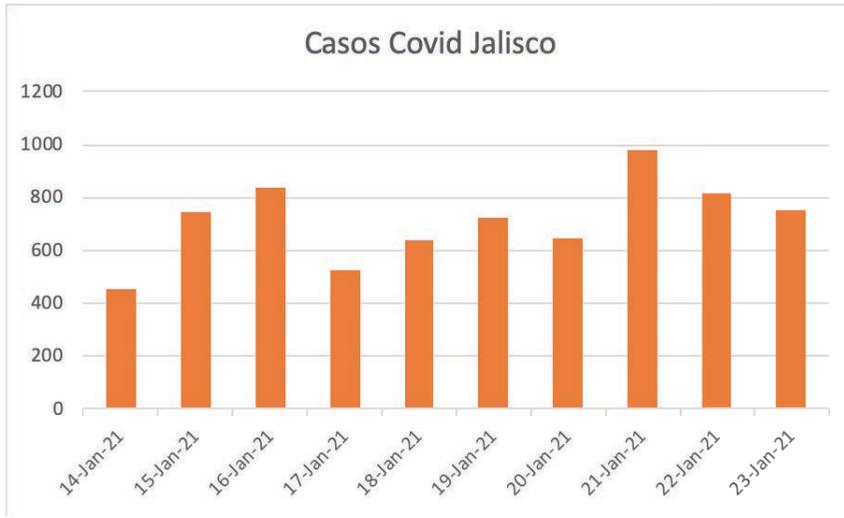
Circular o de pastel

- Se utiliza cuando hay pocos datos.
- Se basa en porcentajes.
- Se toma en consideración que la circunferencia mide 360°.

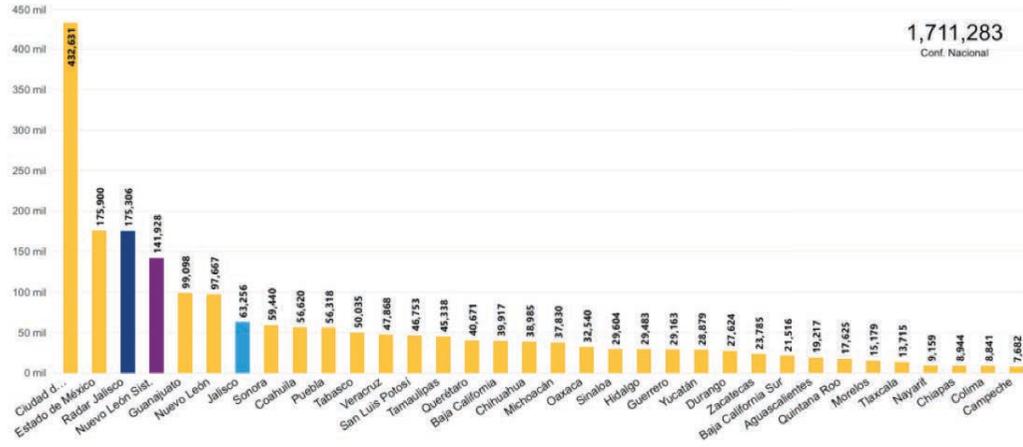
Te mostramos otros ejemplos con información relacionada con la pandemia que vivimos, a nivel internacional, durante el 2020 y 2021:



Histograma

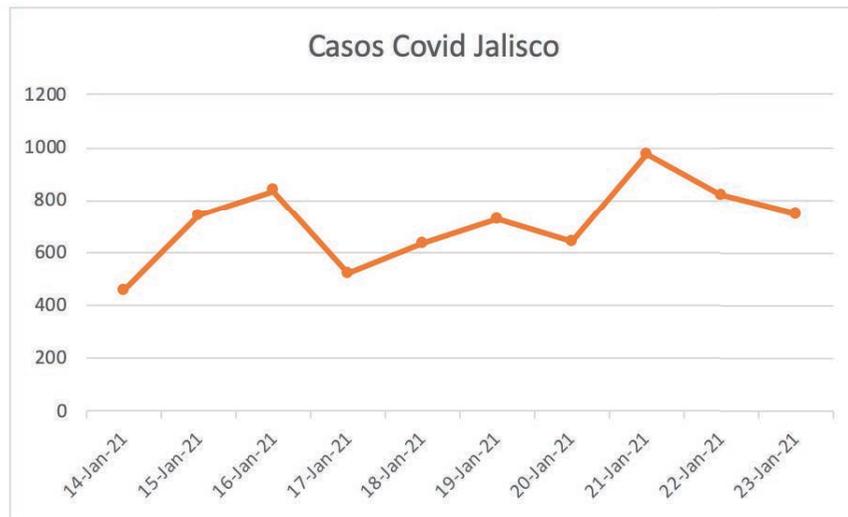


Datos consultados en:
(2021, Junio 30). Retrieved from
https://www.google.com/search?q=contagios+de+covid+en+jalisco&ei=kx7dYJn9MLG0qtsPx9eY4AQ&oq=contagios+de+covid+&gs_lcp=Cgdnd3Mtd2l6EAEYADICCAyAggAMgIIADICCAyAggAMgIIADICCAyAggAMgIIADIFCAAQyQM6BwgAELADEEM6BQgAELAD0gsIABCxAXCDARCwAzoECAAQZoGCAAQChBD0ggI



Fuente: SINAVE/DGE/InDRE Corte: 13:00 horas / 21 de enero de 2021

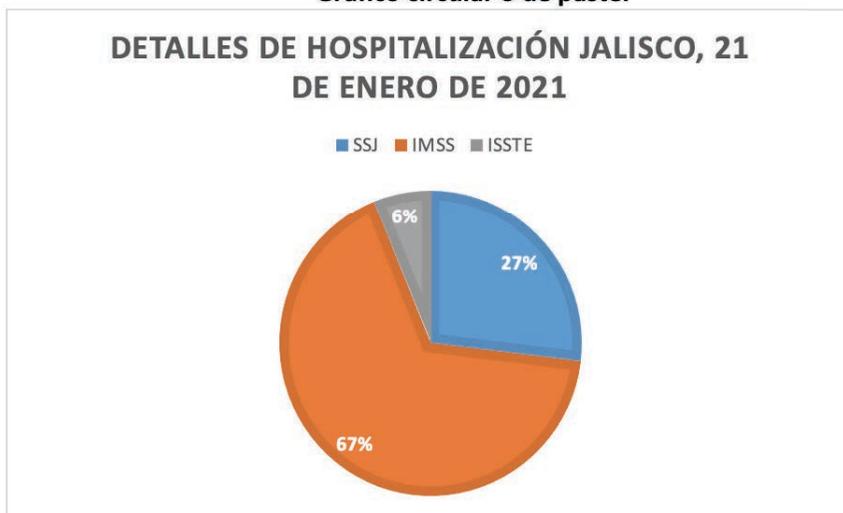
Polígono de frecuencias



Ojiva



Gráfico circular o de pastel





VII. De acuerdo a la información que acabas de revisar, relaciona cada gráfica con su tipo:





Para representar el dato referente a la edad mediante gráficos o tablas de frecuencias, aparecen algunos inconvenientes, pues, en un ejemplo, la edad mínima es de 21 años y la máxima es de 40, por lo que tendríamos muchas filas en la tabla de frecuencias y esto en lugar de facilitar la interpretación de información lo hará más complejo, por lo que es el momento de determinar clases.

El número recomendable de clases está asociado con el número total de datos, en nuestro ejemplo los datos varían de 21 a 40, por lo que tenemos una distancia de 20 datos.

¿Crees que sea conveniente utilizar 5 clases?

Veámoslo

VIII. Completa la siguiente tabla de frecuencias relativas a la edad considerando 4 clases:

Edad	Frecuencia absoluta	Frecuencia Relativa	Frecuencia acumulada
21 a 25			
26 a 30			
31 a 35			
36 a 40			
Total			