

Unidad

1 ESTADÍSTICA I 8

Repaso	10
Historia de la estadística	12
Conceptos básicos	13
Ordenando la información	14
Tabla de frecuencias de datos agrupados	15
Diagrama de tallo y hoja	16
Análisis de gráficos	17
Uso del computador	21
Ejercicios resueltos	24
Desafíos	26
Medios:	
Indicadores mensuales: INE	27
Síntesis	28
Evaluación	30
Ejercicios de refuerzo	32

Unidad

2 ESTADÍSTICA II 34

Repaso	36
Medidas de tendencia central	38
Medidas de dispersión	41
Desviación media	42
Desviación estándar o típica	43
Correlación	45
Medidas de localización:	
cuartiles, percentiles y deciles	46
Diagrama de cajas	49
Muestras al azar	50
Nivel de confianza	51
Tamaño de la muestra	52
Aplicaciones de la estadística	53
Distribución normal	56
Ejercicios resueltos	58
Desafíos	60
Medios:	
¿Cuántas personas tendrán un accidente mañana?	61
Síntesis	62
Evaluación	64
Ejercicios de refuerzo	66

Unidad

3 FUNCIÓN POTENCIA Y LOGARÍTMICA 68

Repaso	70
Funciones	72
Función inversa	73
Funciones periódicas	74
Gráficos con Javamath	75
Función potencia	76
Traslaciones verticales y horizontales	79
Concepto de logaritmo	80
Base de un logaritmo	81
Propiedades de los logaritmos	82
Función logarítmica	84
Distintas gráficas de la función logarítmica	86
Profundizando en los logaritmos	88
Logaritmo natural o neperiano	89
Ecuaciones logarítmicas con una incógnita	90
Aplicaciones de los logaritmos	92
Ejercicios resueltos	94
Desafíos	96
Medios: Modelación matemática	97
Síntesis	98
Evaluación	100
Ejercicios de refuerzo	102

EVALUACIÓN SEMESTRAL 1 104

Unidad

4 LA FUNCIÓN EXPONENCIAL 108

Repaso	110
Función exponencial	112
Aproximándonos al número e	116
Función exponencial natural	117
Función exponencial y función logarítmica	118
Ecuaciones exponenciales	120
Crecimiento exponencial	122
Decrecimiento exponencial	124

Aplicaciones de la función exponencial	126
Ejercicios resueltos	128
Desafíos	130
Medios: Ley de enfriamiento de Newton	131
Síntesis	132
Evaluación	134
Ejercicios de refuerzo	136

Unidad

5 VECTORES

138

Repaso	140
Rectas en el espacio	142
Planos en el espacio	144
Planos y sistemas de ecuaciones	145
Intersección de planos	146
Coordenadas cartesianas	148
Vectores	150
Operatoria con vectores	152
Producto de un número real por un vector	154
Propiedades del producto	155
Producto escalar	156
Producto cruz	157
Vectores en el plano cartesiano	158
Ecuación vectorial de la recta	160
Ecuación vectorial de la recta en el espacio	162
Ecuación vectorial de un plano en el espacio	164
Gráfico de rectas y planos	166
Intersección de rectas y planos en el espacio	168
Transformaciones geométricas	170
Composición de traslaciones	171
Homotecia	172
Composición de homotecias	173
Ejercicios resueltos	174
Desafíos	176
Medios:	
Ajedrez: un juego de razonamiento y concentración	177
Síntesis	178
Evaluación	180
Ejercicios de refuerzo	182

Unidad

6 GEOMETRÍA: ÁREAS Y VOLÚMENES

184

Repaso	186
Concepto de área	188
Concepto de volumen	189
Principio de Cavalieri	190
Teorema de Euler	191
Área y volumen de prismas	192
Área y volumen de pirámides	194
Área y volumen de cilindros	196
Área y volumen de conos	198
Área y volumen de la esfera	200
Secciones de una esfera	202
Proyecciones en el plano	204
Cuerpos generados mediante rotación	206
Problemas de aplicación I	208
Problemas de aplicación II	210
Ejercicios resueltos	212
Desafíos	214
Medios: Las latas de bebida	215
Síntesis	216
Evaluación	218
Ejercicios de refuerzo	220

EVALUACIÓN SEMESTRAL 2 222

Solucionario	226
Glosario	251
Bibliografía	255

Estadística I

Hoy en día no se entendería una campaña publicitaria sin los estudios previos basados en la información que aporta la estadística. En general, la mayoría de las empresas tienen su departamento de estudios estadísticos que se encarga de recopilar, organizar y analizar los datos referentes a un determinado producto.

Los avances tecnológicos de hoy, como la red de Internet, y los que vendrán, causarán efectos sobre la producción, ya que esta se orientará de acuerdo a la información que se obtenga sobre las necesidades, gustos e intereses de la población. De ahí, la importancia de conocer mecanismos para analizar la información que se tiene.



En esta unidad aprenderás a...

- ▶ Conocer algunos hitos importantes en el desarrollo de la estadística.
- ▶ Trabajar con algunos conceptos básicos de la estadística: muestra, población y tipos de variables.
- ▶ Ordenar y organizar la información.
- ▶ Analizar tablas y gráficos.
- ▶ Usar el computador para analizar y presentar la información.

Explora

Realiza el laboratorio 1 correspondiente a la unidad 1 que aparece en www.santillana.cl/mat4

¿Cuánto sabes?

1. La siguiente es una tabla que muestra el número de alumnos(as) que hay en 4° medio en un colegio, agrupados por curso y por sexo.

	Niñas	Niños
4° A	20	25
4° B	22	23

Escribe la razón entre:

- el número de niñas y el número de niños del 4° A.
 - el número de mujeres y el número de hombres.
 - los estudiantes del 4° A y del 4° B.
2. Completa la tabla.

Porcentaje	Fracción	Fracción irreductible	Expresión decimal
75%	$\frac{75}{100}$	$\frac{3}{4}$	0,75
	$\frac{62}{100}$		
		$\frac{1}{50}$	
			$0,\bar{3}$
90%			

3. Indica qué números enteros están contenidos en los siguientes intervalos.

a. $[2, 9]$ b. $]-3, 3[$ c. $]0, 1[$ d. $[-1, 10]$

4. Encuentra un intervalo de números reales que cumpla con lo pedido.

- Un intervalo abierto que contenga a $\frac{1}{2}$, 0 y $-\frac{1}{3}$.
- Un intervalo que contenga todos los números mayores que 5.
- Un intervalo que no contenga a los números positivos.
- Un intervalo semiabierto que no contenga ni al 8,3 ni al $\frac{7}{10}$.

5. En la siguiente tabla se muestran las edades de 6 niños pertenecientes a un taller de teatro.

Nombre	Edad
Pablo	8
Daniela	6
Enrique	10
Carolina	6
Angélica	8
Jaime	6

Construye un gráfico de barras que tenga como variables la **edad** y la **cantidad de personas** que tienen esa edad.

6. En la siguiente tabla se muestra la población por grupos de edad del censo de 1992. Completa los recuadros de la tabla con la frecuencia acumulada.

Grupos de edad	Habitantes	Frecuencia acumulada
0 – 14	3.929.468	
15 – 24	2.425.140	
25 – 39	3.286.011	
40 – 49	1.415.589	
50 – 64	1.415.149	
65 y más	877.044	

- 1 El $a\%$ de un número se puede representar con la fracción $\frac{a}{100}$.
Ejemplo: 34% se representa por $\frac{34}{100}$. Su fracción irreductible es $\frac{17}{50}$ y la expresión decimal equivalente es $0,34$.
- 2 $[a, b]$ es la representación del intervalo cerrado a, b ; por tanto, contiene a a y a b y a todos los números comprendidos entre ellos.
- 3 $]a, b[$ es la representación del intervalo abierto a, b ; por tanto, solo contiene a aquellos números que están comprendidos entre a y b .
- 4 $[a, b[$ o $]a, b]$ son representaciones de un intervalo semiabierto que contiene a a o b , según sea el caso, también contiene a los valores comprendidos entre a y b .
- 5 **Frecuencia:** es la cantidad de veces que ocurre un suceso.
- 6 **Frecuencia acumulada:** es la suma de las frecuencias observadas hasta un cierto punto.

¿Qué debes recordar?

Historia de la estadística

Los orígenes de la estadística, aunque no se sabe con exactitud cuándo se comenzó a utilizar, pueden estar ligados al antiguo Egipto como a los censos chinos que se realizaron hace unos **4.000** años, aproximadamente.

Sin duda, fueron los romanos, maestros de la organización política, quienes mejor supieron ocupar la estadística. Cada cinco años realizaban un censo de la población, cuyos datos de nacimientos, defunciones y matrimonios eran esenciales para estudiar los avances del imperio; sin olvidar los recuentos de ganado y las riquezas que dejaban las tierras.

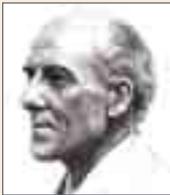
Desde esa época, diversos estados realizaron estudios sobre algunas características de sus poblaciones, sus riquezas, posesiones, etc.

En 1662, John Graunt (1620 – 1674), un mercader inglés, publicó un libro sobre los nacimientos y defunciones ocurridos en Londres; el libro contenía conclusiones acerca de ciertos aspectos relacionados con estos acontecimientos. Esta obra es considerada como el punto de partida de la estadística moderna.

La palabra estadística comenzó a usarse en el siglo XVIII, en Alemania, en relación a estudios donde los grandes números, que representaban datos, eran de importancia para el estado. Sin embargo, la estadística moderna se desarrolló en el siglo XX a partir de los estudios de Karl Pearson.

Hoy, la estadística tiene importancia no solo porque presenta información, sino que además permite inferir y predecir lo que va a ocurrir, y por lo tanto, es una herramienta fundamental a la hora de tomar decisiones de importancia.

HISTORIA



Karl Pearson
(1857 – 1936)

Matemático inglés, trabajó en la University College de Londres. Es considerado el padre de la Estadística Moderna.

EJERCICIOS

1. Averigua en qué parte del libro "Números", del Antiguo Testamento, se hace referencia a censos o recuentos estadísticos. ¿Qué semejanzas hay con los censos actuales?
2. ¿Qué importante acontecimiento relacionado con la estadística marcó el momento del nacimiento de Cristo?
3. En tu vida diaria, ¿cuándo usas la estadística para informarte? ¿Cuándo lo haces para tomar decisiones?
4. ¿Por qué crees tú que la estadística demoró tanto tiempo en desarrollarse?
5. Señala **4** áreas distintas en las cuales se utilice la estadística como herramienta de investigación.

Conceptos básicos

El Instituto Nacional de Estadísticas, (INE), es el organismo encargado de recoger, de forma fidedigna y oportuna, información relevante para la administración del Estado y para las actividades nacionales, con el objetivo de mejorar la calidad de vida de las personas.

En muchas ocasiones, para llevar a cabo una investigación se hacen encuestas, las cuales son dirigidas a una **muestra representativa** de la población. Para comprender mejor este tipo de estudio es importante que conozcas los siguientes términos básicos:

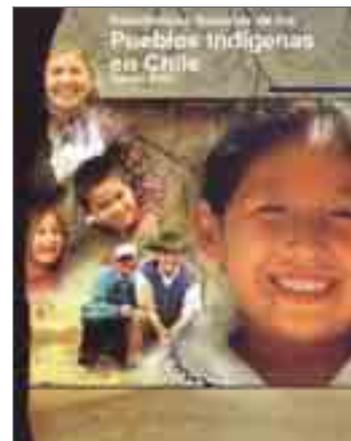
→ **Población:** es un conjunto de personas, eventos o cosas de las cuales se desea hacer un estudio, y tienen una característica en común.

Muestra: es un subconjunto cualesquiera de la población; es importante escoger la muestra en forma aleatoria (al azar), pues así se logra que sea representativa y se puedan obtener conclusiones más afines acerca de las características de la población.

Para estudiar alguna característica específica de la población se pueden definir los siguientes tipos de variables:

Variables cualitativas: relacionadas con características no numéricas de un individuo (por ejemplo: atributos de una persona).

Variables cuantitativas: relacionadas con características numéricas del individuo. Las variables cuantitativas se dividen en **discretas** (aquellas que no admiten otro valor entre dos valores distintos y consecutivos) o **continuas** (aquellas que pueden tomar una infinidad de valores entre dos de ellos).



Portada del estudio "Estadísticas Sociales de los Pueblos Indígenas en Chile" publicado por el Instituto Nacional de Estadística (INE) acerca de la información recopilada en el censo del año 2002.

ENLACES



En la página web www.ine.cl podrás encontrar más información relacionada con estudios estadísticos.

EJERCICIOS

1. Se desea saber si los dueños de automóviles catalíticos están dispuestos a pagar la conversión de sus motores a gas natural. Para ello se decide realizar una encuesta.
 - a. Determina cuál de las siguientes es la mejor muestra:
 - i. Escoger al azar a adultos que caminan por el centro de las principales ciudades del país.
 - ii. Escoger al azar a conductores de automóviles en las intersecciones más concurridas.
 - iii. Escoger al azar del registro de vehículos motorizados a dueños de automóviles catalíticos y enviarles un encuestador.
 - b. Explica la razón de tu elección, señala las ventajas y desventajas de cada alternativa.
 - c. ¿Cuáles son las variables utilizadas en la encuesta? ¿A qué tipo de variables corresponden? ¿Por qué?



Ordenando la información

Al ordenar datos muy numerosos, es usual agruparlos en **clases** o **categorías**. Al determinar cuántos pertenecen a cada clase, establecemos la **frecuencia**. Construimos así una tabla de datos llamada **tabla de frecuencia**.

Ejemplo

Los siguientes datos corresponden a las notas obtenidas por un curso de 24 alumnos en un trabajo de matemática:

3,2 4,2 5,6 6,0 2,8 3,9 4,2 4,2 5,0 5,0 3,9 3,9
3,2 3,2 4,2 5,6 6,0 6,0 3,2 6,0 4,2 5,0 5,6 5,0

Ordenemos estos datos en la siguiente tabla:

Nota	Frecuencia absoluta (fi)	Frecuencia relativa (hi)	Frecuencia relativa porcentual (%)
2,8	1	$\frac{1}{24}$	4,2
3,2	4	$\frac{4}{24}$	16,7
3,9	3	$\frac{3}{24}$	12,5
4,2	5	$\frac{5}{24}$	20,8
5,0	4	$\frac{4}{24}$	16,7
5,6	3	$\frac{3}{24}$	12,5
6,0	4	$\frac{4}{24}$	16,7

¿Qué conclusiones puedes obtener de la tabla anterior?

Solo un **16,7%** del curso obtuvo nota seis. El **33,4%** del curso obtuvo nota deficiente, etc.

TIPS



A veces, por efecto de las aproximaciones, es posible que la suma de las frecuencias relativas porcentuales no sea exactamente **100%**.

IR A LA WEB



Desarrolla el laboratorio 2.
www.santillana.cl/mat4



PARA ARCHIVAR

La **frecuencia absoluta** de una clase es el número de datos que forma dicha clase, mientras que la **frecuencia relativa** corresponde a la razón entre la frecuencia absoluta y el total de datos, la cual se puede expresar mediante el uso de porcentajes.

EJERCICIOS

- Los siguientes datos corresponden a los lugares favoritos de vacaciones de los empleados de una empresa.
Mar - Montaña - Campo - Mar - Mar - Montaña -
Campo - Mar - Mar - Montaña - Campo - Mar - Campo.
a. Completa la siguiente tabla y luego obtén al menos dos conclusiones.

Lugar \ Frecuencia	F. Absoluta	F. Relativa %
Campo		
Mar		
Montaña		
Total		

Tabla de frecuencia de datos agrupados

En ocasiones, el agrupar los datos en intervalos, nos puede ayudar para realizar un mejor análisis de ellos.

Ejemplo

Consideremos los siguientes datos, expresados en metros, correspondientes a las estaturas de **80** estudiantes de Cuarto año de Educación Media.

1,67	1,72	1,81	1,72	1,74	1,83	1,84	1,88	1,92	1,75
1,84	1,86	1,73	1,84	1,87	1,83	1,81	1,77	1,73	1,75
1,78	1,77	1,67	1,83	1,83	1,72	1,71	1,85	1,84	1,93
1,82	1,69	1,70	1,81	1,66	1,76	1,75	1,80	1,79	1,84
1,86	1,80	1,77	1,80	1,76	1,88	1,75	1,79	1,87	1,79
1,77	1,67	1,74	1,75	1,78	1,77	1,74	1,73	1,83	1,76
1,83	1,77	1,75	1,77	1,77	1,84	1,83	1,79	1,82	1,76
1,76	1,76	1,79	1,88	1,66	1,80	1,72	1,75	1,79	1,77

Estatura mayor: **1,93 m**; estatura menor: **1,66 m**; rango: **0,27 m = 27 cm**.
Formaremos **6** intervalos. Para calcular el tamaño de cada uno dividimos $27 : 6 = 4,5 \approx 5$.

Nos queda la siguiente tabla,

Intervalo	Marca de clase	Frecuencia absoluta
1,65 – 1,69	1,67	6
1,70 – 1,74	1,72	12
1,75 – 1,79	1,77	30
1,80 – 1,84	1,82	22
1,85 – 1,89	1,87	8
1,90 – 1,94	1,92	2
		Total: 80



PARA ARCHIVAR

Para construir una **tabla de frecuencias para datos agrupados**, determinamos el tamaño de cada intervalo, dividiendo el valor del rango por la cantidad de intervalos que se desea obtener.

AYUDA



El **rango**, está dado por la diferencia entre el máximo y el mínimo valor de una variable.

TIPS



El **tamaño del intervalo** se aproxima al impar más cercano.

AYUDA



La **marca de clase** es el representante de un intervalo, y corresponde al promedio entre los extremos de éste.

EJERCICIOS

- Utilizando los datos anteriores, haz una tabla de frecuencia para datos no agrupados. Luego responde:
 - ¿Cuántos alumnos miden entre **1,75 m** y **1,89 m**?
 - ¿Qué ventajas y desventajas tiene la utilización de cada tipo de tabla?
- Considera los siguientes datos: **1, 2, 5, 4, 7, 8, 9, 5, 6, 4, 7, 4, 1, 8, 5, 2, 3**, construye una tabla de datos agrupados determina la marca de clase de cada intervalo.

Diagrama de tallo y hoja

Otra forma de organizar la información, es la utilización del **diagrama de tallo y hoja**, este nos sirve para analizar la variabilidad de los datos, o bien para comparar dos grupos diferentes.

Ejemplo:

Los siguientes datos corresponden a la esperanza de vida de hombres y mujeres correspondientes a diversos países.

Hombre				Mujer			
68	62	62	56	75	66	66	67
42	43	47	63	46	47	50	69
68	62	80	53	71	78	73	52
74	76	71	66	77	82	75	72

Si observas los datos anteriores podrás apreciar que son similares, sin embargo, el siguiente diagrama de hoja nos permite apreciar algunas diferencias.

Esperanza de vida del hombre				Esperanza de vida de la mujer			
	7	3	2	4	6	7	
		3	6	5	0	2	
2	6	8	3	2	2	8	6
		4	6	1	7	5	1
				0	8	2	
							6
							6
							7
							9
							8
							3
							7
							5
							2

El diagrama anterior nos permite visualizar que la esperanza de vida de la mujer es mayor que la del hombre. Además podemos obtener otras conclusiones, como por ejemplo, que el intervalo $[62, 68]$, presenta la mayor frecuencia respecto a la esperanza de vida del hombre.

TIPS

La variabilidad de los datos se relaciona con cuán dispersos están estos.

ENLACES

Para mayor información acerca de datos estadísticos de diversos países ingresa a la página web: www.amstat.org/publications/jse/

AYUDA

En este caso el **tallo** representa la cifra de las decenas y las hojas, las unidades.

EJERCICIOS

1. Los siguientes datos corresponden a la tasa bruta de natalidad y mortalidad infantil de algunos países de Latinoamérica.

Natalidad (niños nacidos vivos en 1 año, por cada 1.000 habitantes):

21	47	29	27	23	33
28	29	35	33	18	28

Mortalidad (número de muertes al año por cada 1.000 habitantes, niños menores de 1 año):

26	51	63	40	17	63
56	43	42	109	22	23

a. Construye un diagrama de tallo y hoja para los datos anteriores.

b. Se afirma que la tasa de mortalidad infantil correspondiente a países africanos es de aproximadamente 96. ¿A qué crees que se debe la diferencia entre países latinoamericanos y africanos?

c. ¿A qué problemas puede conllevar la diferencia entre tasas de natalidad y mortalidad?

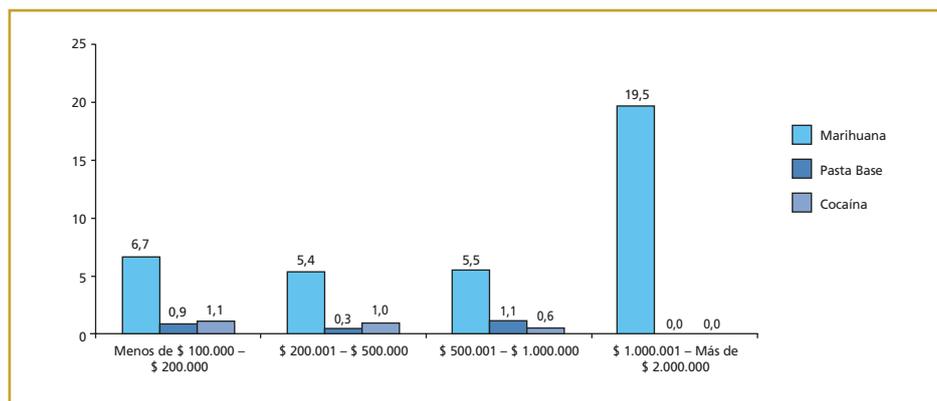
d. Averigua las tasas de natalidad y mortalidad correspondientes a otros grupos de países, por ejemplo, países de Oriente o Asia, y compáralos con las tasas de Latinoamérica. Comparte tus resultados con tus compañeros.

Análisis de gráficos

En mayo del 2005, el Consejo Nacional para el Control de Estupefacientes, (CONACE), publicó el Sexto Estudio Nacional de Drogas en Población General de Chile (realizado en el año 2004), relacionado con las tendencias en el uso de algunas drogas en el país.

En los siguientes gráficos se muestran las tendencias, de los adolescentes (entre 12 y 18 años) en el uso de ciertas drogas (lícitas e ilícitas), según el ingreso total, al mes, de la familia.

Ejemplo 1: Histograma

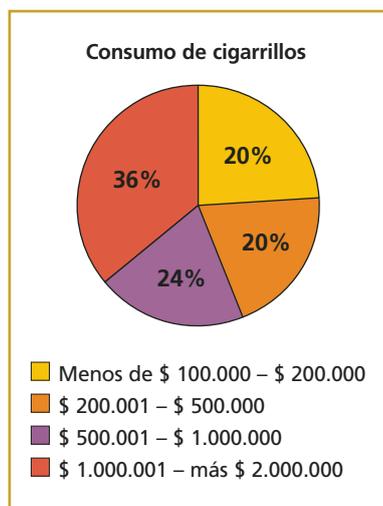


Fuente: Sexto Estudio Nacional de Drogas en Población General de Chile (2004), www.conace.cl, julio 2005.

La más alta frecuencia de consumo de marihuana se registra entre las personas cuyas familias tienen ingresos promedios mensuales sobre 1 millón de pesos, con una tasa cercana al 20%. Esta tasa, porcentual, de marihuana es tres veces más alta que en familias con los más bajos ingresos, con tasas de 6,7%.

El consumo de cocaína entre adolescentes está latente, con tasas que bordean el 1%, en familias de todos los niveles de ingresos, con la salvedad de las familias con los más altos recursos.

Ejemplo 2: Gráfico circular



De aquellos adolescentes que consumen cigarrillos, el 36% provienen de familias con los más altos ingresos mensuales. Dicha tasa es 16 puntos porcentuales más alta que en familias con los más bajos ingresos, con tasas de 20%.

Fuente: Sexto Estudio Nacional de Drogas en Población General de Chile (2004), www.conace.cl, julio 2005.

ENLACES



Para obtener más información visita el sitio www.conace.cl. Recuerda que el contenido de la página puede variar.

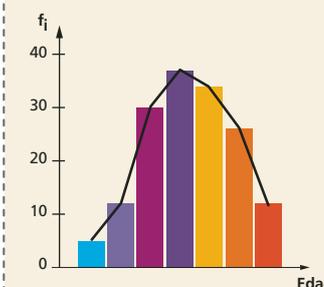
TIPS



El polígono de frecuencia se grafica a partir de un histograma. Se construye uniendo los puntos medios de cada barra (marca de clase).

Ejemplo:

Edad	Marca de clase	f_i
16 – 20	18	5
21 – 25	23	12
26 – 30	28	30
31 – 35	33	37
36 – 40	38	34
41 – 45	43	26
46 – 50	48	12

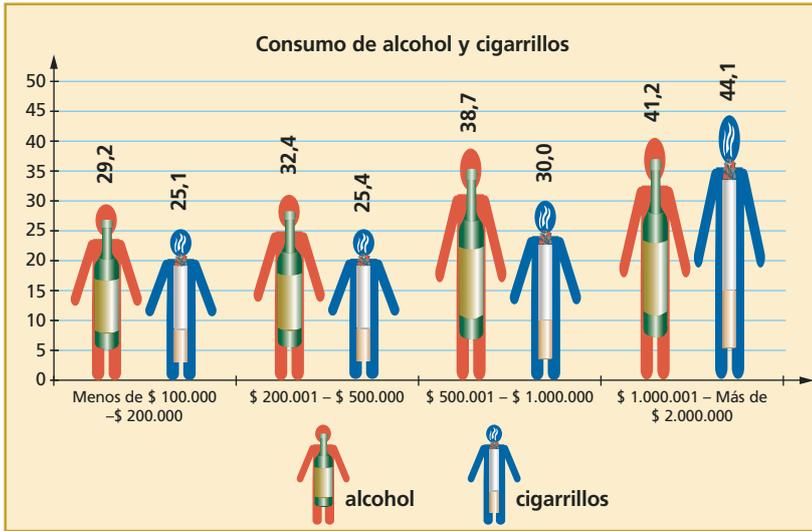


IR A LA WEB



Desarrolla el laboratorio 3. www.santillana.cl/mat4

Ejemplo 3: Pictograma



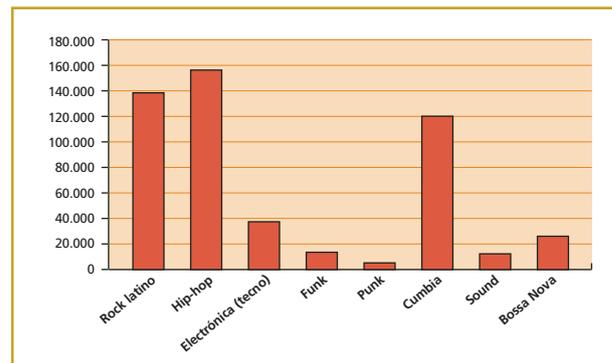
El consumo de cigarrillos en adolescentes de familias con el ingreso más alto, tiene una tasa que supera casi por 21 puntos porcentuales al consumo en familias con el ingreso más bajo.

Fuente: Sexto Estudio Nacional de Drogas en Población General de Chile (2004), www.conace.cl, julio 2005.

Ejemplo 4: Gráfico de barras

La encuesta Consumo de Cultura, realizada por el Instituto Nacional de Estadísticas, (INE), entre varias temáticas, arrojó la siguiente información relacionada con el tipo de música que escuchan hombres y mujeres.

Tipo de música	f _i
Rock latino	138.478
Hip-hop	156.305
Electrónica (techo)	37.436
Funk	13.529
Punk	5.205
Cumbia	120.129
Sound	12.308
Bossa Nova	26.074

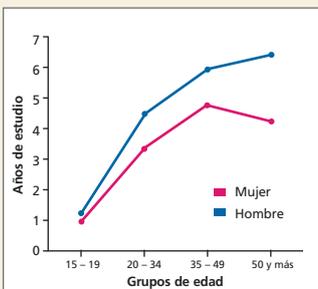


Fuente: Encuesta Consumo de Cultura, www.ine.cl, julio 2005.

TIPS

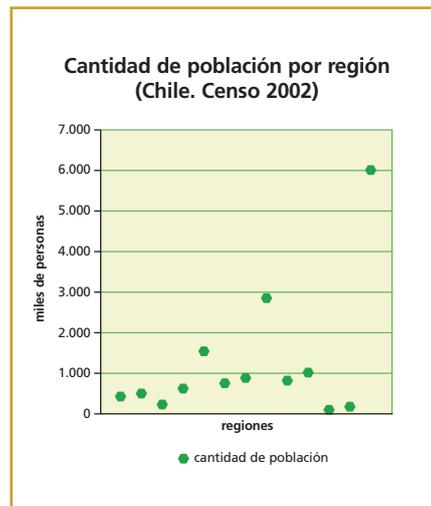


Observa otro tipo de gráfico que te permite un buen análisis de información.



Fuente: Estudio de la Mujer (2004), www.sernam.cl, julio 2005.

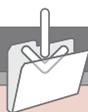
Ejemplo 5: Gráfico de dispersión



En la gráfica se observan datos obtenidos del Censo del año 2002, relacionado con la cantidad de población que hay en cada región del país.

Por millones de habitantes, está por sobre las demás una de las regiones, tiene poco más de seis millones de personas. Le siguen en tamaño, con más de un millón de habitantes, tres regiones más. De las restantes regiones, su cantidad de población es bastante homogénea.

Fuente: Censo 2002, www.ine.cl, julio 2005.



PARA ARCHIVAR

Utilidad de diversos tipos de gráficos:

Gráfico de barras: facilita la comparación entre las frecuencias de los valores.

Pictograma: se recomienda cuando la variable estudiada es una cualidad.

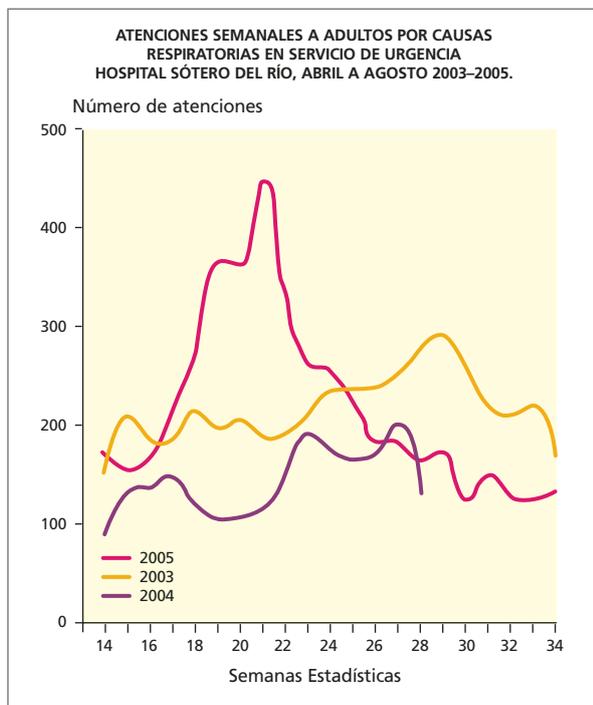
Gráfico circular: es útil cuando se necesita representar porcentajes.

Histograma: sirve para expresar información sobre datos que están agrupados.

Gráfico de dispersión: sirve para estudiar la homogeneidad o heterogeneidad de los datos.

EJERCICIOS

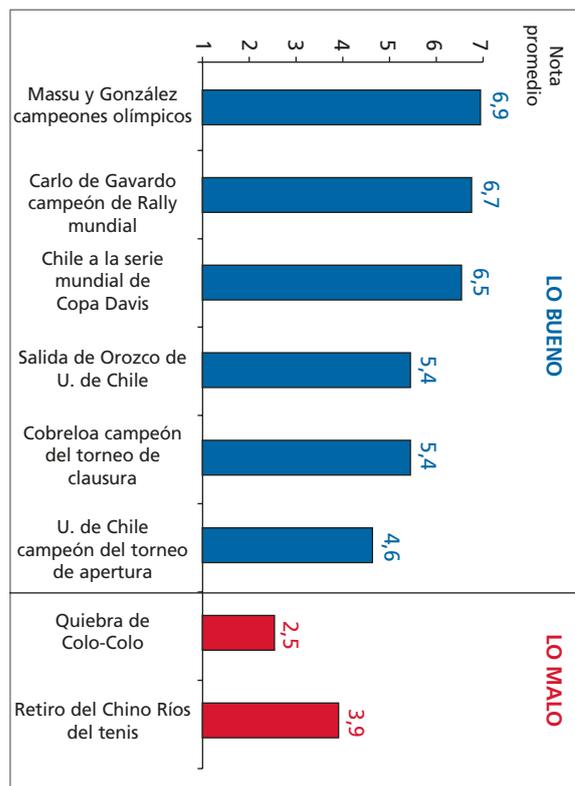
1. El gráfico muestra la cantidad de pacientes semanales que asistieron al hospital Sótero del Río, por motivos de enfermedades respiratorias.



Fuente: DEIS. Departamento de Estadísticas e Información de Salud, Ministerio de Salud.

- a. ¿Qué conclusiones puedes obtener a partir del gráfico?
- b. ¿En qué período las atenciones médicas fueron similares, en cantidad de pacientes?
- c. ¿En qué período se produjo mayor demanda en el hospital?

2. El siguiente gráfico nos presenta la información obtenida de 300 encuestados por la Fundación Futuro (2004), acerca de la pregunta: ¿Qué nota colocas a lo bueno y malo en el deporte chileno?



Fuente: Encuesta Lo bueno, lo malo y lo feo, www.fundacionfuturo.cl, julio 2005.

- a. ¿Cuál fue la categoría mejor evaluada? ¿Tú también la hubieras evaluado con esa nota? ¿Por qué?

EJERCICIOS

- b. De lo bueno, ¿qué área del deporte tiene el mejor promedio, el tenis o el fútbol?
 - c. Con la información obtenida en **b**, construye un gráfico circular que muestre la diferencia obtenida? ¿A qué atribuyes esta diferencia?
3. Dada la siguiente tabla, que muestra los resultados de la prueba SIMCE (Sistema de Medición de la Calidad de la Educación) año 2002 de 4° año Básico, responde:

Región	Matemática	Lenguaje
I Tarapacá	240	245
II Antofagasta	247	250
III Atacama	244	248
IV Coquimbo	242	249
V Valparaíso	249	254
VI L. Bdo. O'higgins	246	252
VII Maule	243	248
VIII Bío- Bío	243	247
IX Araucanía	235	243
X Los Lagos	242	249
XI Aysén	254	261
XII Magallanes	254	260
RM Región Metropolitana	254	257
Total	248	252

Fuente: Prueba SIMCE, 4° Año Educación Básica (2002), www.mineduc.cl, julio 2005.

- a. ¿Cuáles son las regiones que tienen menos de **246** puntos en Matemática?
- b. ¿Qué región obtuvo el puntaje más bajo en cada área? ¿Coinciden estos puntajes con la misma región?
- c. ¿Qué región obtuvo el mejor promedio en Lenguaje? Esta región, ¿también obtuvo el puntaje más alto en Matemática?
- d. ¿Qué tipo de variables son las consideradas en esta tabla?
- e. ¿Qué tipo de gráfico representa mejor la diferencia de puntajes totales en cada área?
- f. ¿Qué tipo de gráfico construirías para representar los puntajes de las mejores **5** regiones?

4. Uno de los problemas más complejos que debe abordar nuestra sociedad es la pobreza; un país que quiere surgir debe eliminar este problema. En la tabla se ven las comunas más pobres del país; en la mayoría de ellas vive población mayoritariamente mapuche que no ha podido salir del círculo de la pobreza.

	Comunas	Más pobres %
1	Mulchén	59,5
2	Angol	53,3
3	Carahue	50,8
4	Gorbea	50,6
5	Constitución	49,9
6	Coihueco	48,5
7	Curanilahue	47,8
8	Padre Las Casas	46,9
9	Nueva Imperial	46,4
10	Traiguén	45,4
11	Coronel	44,5
12	Lebu	44,4
13	Collipulli	44,1
14	Nacimiento	44,0
15	Cañete	43,8

Fuente: CASEN 1998, MIDEPLAN

- a. ¿Qué gráfico representaría mejor la información dada en la tabla? ¿Por qué?
- b. ¿Qué tipo de variable utilizaste para el gráfico anterior?
- c. De la tabla, determina los dos pueblos que presenten mayor porcentaje de pobreza y dos que tengan el menor porcentaje. Elige algún tipo de gráfico que te permita estudiar la comparación, ¿qué puedes concluir?
- d. ¿Qué factores culturales crees tú que afectan al pueblo mapuche y le impiden salir de la pobreza?
- e. ¿Qué factores de nuestra sociedad impiden a los mapuches vivir como ellos desean?
- f. ¿Qué soluciones ves tú al problema?
- g. Averigua en cuáles de las comunas del cuadro vive mayoritariamente gente mapuche.

Uso del computador

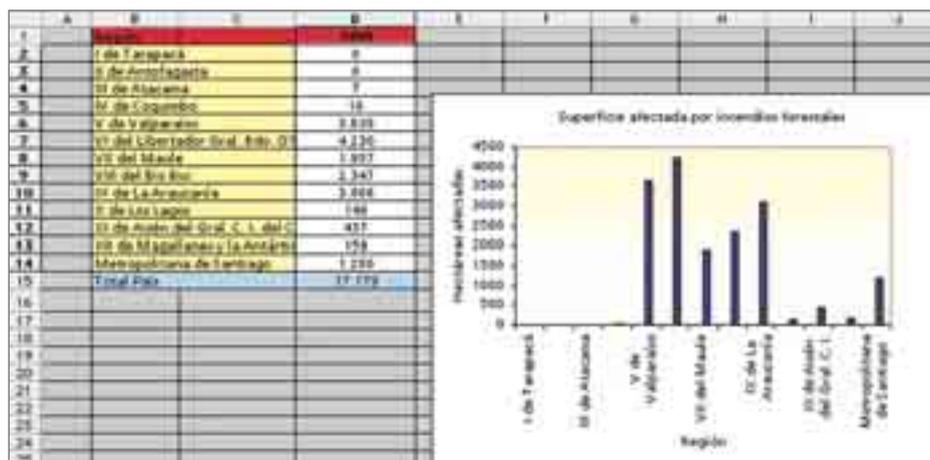
Las planillas de cálculo permiten ahorrar gran cantidad de tiempo al hacer trabajos estadísticos. A continuación, se presenta un ejemplo de cómo utilizar el programa Excel para graficar un conjunto de datos.

Lo primero que se debe hacer es construir una tabla de valores, luego seleccionarla y por último pulsar "Asistente de gráficos".

Ejemplo

La siguiente tabla muestra las hectáreas afectadas en 1999 por incendios forestales, para graficarla realizamos lo siguiente:

1. Seleccionamos presionando con el mouse, desde la columna **B2** hasta la columna **D14**.
2. En la barra de menú, selecciona "Insertar", luego selecciona "Gráfico" en el submenú.
3. Elegimos "Tipo de gráfico", en este caso seleccionamos un gráfico de barras.
4. Finalizamos nuestro gráfico en "Terminar".



TIPS

Puedes personalizar tu gráfico, haciendo clic sobre él, de esta manera puedes cambiar los colores. Además en "Título", puedes poner nombre a los ejes y al gráfico.

IR A LA WEB

Desarrolla el laboratorio 4.
www.santillana.cl/mat4

EJERCICIOS

1. Utilizando los datos anteriores realiza lo siguiente:
 - a. Ingresas la tabla anterior en una planilla Excel.
 - b. Realiza un gráfico de dispersión y otro circular. ¿Qué ventaja tiene la utilización de cada tipo de gráfico?
 - c. ¿En qué regiones se observa mayor cantidad de hectáreas afectadas por incendios forestales? ¿A qué crees que se debe?
 - d. Está comprobado que la mayor cantidad de incendios forestales es causada directa o indirectamente por el ser humano. ¿Qué medidas tomarías tú para preservar nuestros bosques?

EJERCICIOS

2. Las siguientes son las respuestas de un grupo de jóvenes a la pregunta: ¿Cuál es tu deporte favorito?

Fútbol - Tenis - Fútbol - Basquetbol - Fútbol - Automovilismo - Tenis - Fútbol - Natación - Fútbol - Tenis - Automovilismo - Gimnasia - Fútbol - Hockey - Fútbol - Tenis - Atletismo - Fútbol - Gimnasia - Tenis - Atletismo - Gimnasia

- a. Construye en Excel un gráfico circular e interpreta los resultados.
- b. ¿Qué deporte presenta mayor frecuencia?

3. Según la Empresa Metropolitana de Obras Sanitarias (EMOS), el consumo promedio de agua, en metros cúbicos, en una familia de 5 integrantes es:

Uso	Invierno	Verano
Duchas	250	350
Aseo en lavatorios	50	60
Descarga WC	300	300
Comida y lavado de vajilla	80	90
Lavado general	150	185
Riego	5	165
Total diario	835	1.150
Total mensual	25.050	34.500

Fuente: EMOS.

- a. Construye en Excel, un gráfico que permita comparar el consumo de una familia de 5 integrantes en invierno y verano.
- b. Construye un gráfico circular, para el consumo de invierno que muestre los porcentajes de agua destinados a cada fin.
- c. Repite el ejercicio anterior para mostrar el consumo de agua en verano.
- d. ¿A qué crees que se deba el incremento del consumo de agua en verano?
- e. Divide cada uno de los valores dados en la tabla por 5, luego construye un gráfico que muestre estos valores. ¿Qué resultados nos entrega este gráfico?

f. Discute con tus compañeros acerca de la escasez del agua y su mal uso.

4. La siguiente tabla de frecuencias muestra la cantidad de colesterol total de un grupo de pacientes cuya edad es de 50 a 60 años.

Colesterol total (mg/dl)	Frecuencia
170 – 179	4
180 – 189	7
190 – 199	12
200 – 209	16
210 – 219	35
220 – 229	37
230 – 239	11
240 – 249	8

- a. Calcula las frecuencias relativas para cada intervalo.
- b. Se considera un nivel normal de colesterol entre 200 y 239 (mg/dl). ¿Cuántos de los pacientes se encuentran dentro de los niveles normales?
- c. Construye en Excel un histograma para comparar la frecuencia de cada intervalo. ¿Qué puedes concluir?



EJERCICIOS

5. En septiembre del año 2003, la Fundación Futuro realizó un estudio en 34 comunas de Santiago, que arrojó los siguientes resultados, dada la siguiente pregunta:

¿En qué lugar se siente más seguro?

	Casa	Lugar de trabajo	Lugares públicos	Calle
Muy seguro	53%	41%	41%	13%
Muy inseguro	47%	30%	55%	86%
No responde	1%	29%	5%	1%

Fuente: Estudio Fundación Futuro, julio 2005.

- a. Construye un gráfico circular para cada uno de los lugares. ¿Qué puedes concluir?
- b. Construye un histograma que muestre las diferencias entre los cuatro lugares. ¿A qué crees que se deba esta diferencia?
- c. Si la muestra de la encuesta anterior fue de 402 personas, ¿cuántas personas corresponden a cada categoría?
- d. La encuesta fue realizada telefónicamente. ¿Cómo influye este hecho en los resultados de la encuesta? Discútelo con tus compañeros(as).
- e. ¿Qué medidas implementarías para mejorar los problemas relacionados con la seguridad?



6. La siguiente tabla muestra la disponibilidad de agua (en miles de metros cúbicos) por persona en el año 1950 y en el año 2000.

	1950	2000
África	17,8	4,8
Asia	7,6	2,9
Europa	5,9	4,5
América del Norte	32,4	17,6
América Latina	72,1	22,8
Ex URSS	24,1	14,8
Oceanía	159,5	65,6

Fuente: FAO (Food and Agriculture, Organization of the United Nations)

- a. Construye un gráfico de barras que permita comparar la disponibilidad de agua durante ambos períodos.
- b. Calcula el porcentaje de descenso para cada lugar.
- c. ¿Por qué crees que en algunos lugares el descenso de la cantidad de agua es mayor que en otras?
- d. Qué crees que sucederá con la disponibilidad de agua en 50 años más?
- e. Construye un gráfico circular que muestre la diferencia de disponibilidad de agua en el año 2000. ¿Qué puedes concluir? ¿A qué se debe la diferencia?



Ejercicio 1

Los siguientes gráficos piramidales, muestran la distribución poblacional de Chile en tres años diferentes. Observa y luego responde las siguientes preguntas.

- a. ¿Cuántos hombres aproximadamente comprende el intervalo [10, 24] en cada uno de los años mostrados en los gráficos?
- b. ¿En qué año la población masculina comprendida en el intervalo [10, 24] presentó una mayor diferencia por tramos de edad?
- c. ¿Qué consecuencias geográficas podrían derivarse de la pirámide poblacional proyectada para el año 2025?
- d. ¿En qué tipo de análisis es recomendable la utilización de gráficos piramidales?



Solución

- a. Para responder, debemos determinar la frecuencia de cada uno de los tramos comprendidos en el intervalo [10, 24], es decir, en los tramos [10, 14], [15, 19], [20, 24] para cada año.

1950			2000			2025		
Intervalo	Frecuencia absoluta	Frecuencia acumulada	Intervalo	Frecuencia absoluta	Frecuencia acumulada	Intervalo	Frecuencia absoluta	Frecuencia acumulada
[10, 14]	300.000	300.000	[10, 14]	700.000	700.000	[10, 14]	700.000	700.000
[15, 19]	280.000	580.000	[15, 19]	650.000	1.350.000	[15, 19]	700.000	1.400.000
[20, 24]	300.000	880.000	[20, 24]	600.000	1.950.000	[20, 24]	700.000	2.100.000

Cantidad de habitantes por kilómetro cuadrado.

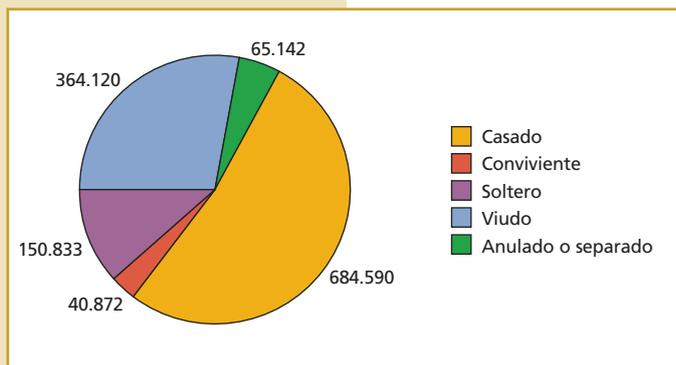
Cantidad de hombres comprendidos en el intervalo [10, 24], por cada año.

- b. Si observamos la frecuencia acumulada para cada año, podemos concluir que en el año 2025 la población masculina comprendida en el intervalo [10, 24] presentará una mayor diferencia por tramos de edad.
- c. Dado que en el año 2025 se observa un importante incremento de la población, uno de los principales problemas podrá estar dado por la densidad, y como consecuencia, el espacio disponible por individuo se verá disminuido.
- d. Se recomienda el uso de gráficos piramidales para realizar comparación de variables que presentan más de una categoría, por ejemplo, sexo.

Ejercicio 2

El siguiente gráfico circular muestra la distribución de personas de 60 años o mayores, según estado civil en Chile.

- Determina el porcentaje correspondiente a cada categoría.
- Determina el ángulo central aproximado correspondiente a cada uno de los grupos indicados en el gráfico.



Fuente: Instituto Nacional de Estadísticas, www.inec.cl, Julio 2005.

Solución

- Para calcular el porcentaje correspondiente a cada categoría, completaremos la siguiente tabla de frecuencias.

Categoría	Frecuencia absoluta	Frecuencia relativa	Frecuencia relativa porcentual
Casado	684.590	0,524	52,43 %
Conviviente	40.872	0,031	3,13 %
Soltero	150.833	0,115	11,55 %
Viudo	364.120	0,27	27,9 %
Anulado o separado	65.142	0,05	4,98 %
Total	1.305.557	0,99	99,99 %

Porcentaje correspondiente a cada categoría.

- Ahora que hemos calculado los porcentajes correspondientes, determinaremos el ángulo central correspondiente a cada grupo.

Sabemos que los 360° del círculo representan la frecuencia relativa porcentual acumulada, es decir 100% , por lo tanto cada 1% corresponderá a $3,6^\circ$. Luego para obtener el ángulo correspondiente, basta con multiplicar cada porcentaje por $3,6$. Nos queda:

Resolviendo la proporción

$$\frac{360^\circ}{x^\circ} = \frac{100\%}{1\%}$$

$$x = 3,6^\circ$$

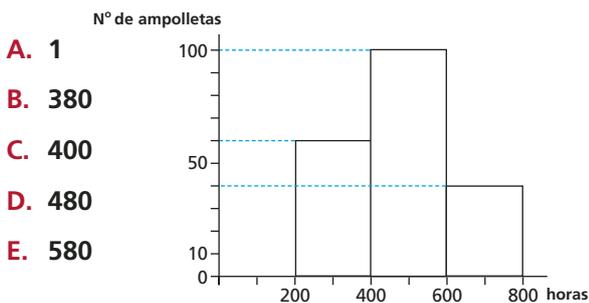
Categoría	%	Ángulo
Casado	52,43	188,75°
Conviviente	3,13	11,27°
Soltero	11,55	41,58°
Viudo	27,9	100,44°
Anulado o separado	4,98	17,93°
Total	99,99	359,9°

Ángulo correspondiente a cada categoría.

1. (Ensayo PSU, 2004) En un curso cada estudiante puede optar solamente por un actividad extraprogramática: las tres cuartas partes de los estudiantes elige deportes y una sexta parte del curso elige teatro. ¿Cuál de las siguientes es la mejor estimación del porcentaje de estudiantes que participa en alguna de estas dos actividades?

- A. Menos del 91%
- B. Entre el 91% y el 93%
- C. Entre el 93% y el 95%
- D. Entre el 95% y el 97%
- E. Más del 97%

2. (Ensayo PSU, 2004) La distribución del número de horas que duraron encendidas 200 ampolletas está dada en el gráfico siguiente. La duración promedio de una ampolleta en horas, aproximadamente, es:



- A. 1
- B. 380
- C. 400
- D. 480
- E. 580

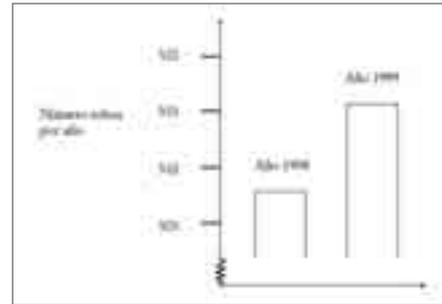
3. (Facsímil PSU, Demre, 2004) El estadio A de una ciudad tiene capacidad para 40.000 personas sentadas y otro estadio B para 18.000. Se hacen eventos simultáneos; el A se ocupa hasta el 25% de su capacidad y el B llena solo el 50%. ¿Cuál(es) de las siguientes afirmaciones es(son) verdadera(s)?

- I) El estadio A registró mayor asistencia de público que el B.
- II) Si se hubiese llevado a los asistentes de ambos estadios al A, habría quedado en éste, menos del 50% de sus asientos vacíos.
- III) Los espectadores que asistieron en conjunto a los dos estadios superan en 1.000 a la capacidad de B.

- A. Solo I
- B. Solo II
- C. Solo III
- D. I y II
- E. I y III

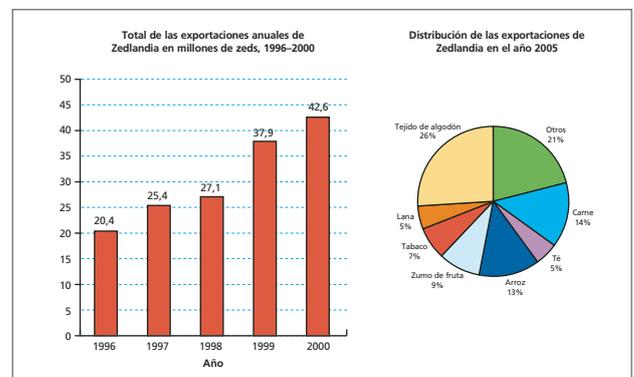
4. (Pisa, 2003) Un presentador de TV mostró este gráfico y dijo:

“El gráfico muestra que hay un enorme aumento del número de robos comparando 2001 con 2002”.



¿Consideras que la afirmación del presentador es una interpretación razonable del gráfico? Da una explicación que fundamente tu respuesta.

5. (Pisa, 2003) Los siguientes gráficos muestran información sobre las exportaciones de Zedlandia, un país cuya moneda es el zed.



¿Cuál fue el valor de las exportaciones de zumo de fruta en el año 2000?

- A. 1,8 millones de zeds.
- B. 2,3 millones de zeds.
- C. 2,4 millones de zeds.
- D. 3,4 millones de zeds.
- E. 3,8 millones de zeds.

Indicadores mensuales: INE

En la página web del Instituto Nacional de Estadísticas (INE) se publica mensualmente las variaciones que experimentan los precios de los productos con lo cual se obtiene el IPC del mes. El siguiente texto corresponde a las variaciones de julio del 2005.

Una variación mensual de **0,6%** experimentó el IPC en julio, con lo cual la inflación acumulada en el año es de **2,4%**. En doce meses se registra un alza de **3,1%**.

El grupo Transporte, con un aumento promedio de **1,5%**, muestra la más importante alza de precios. También se observaron aumentos en los grupos Vivienda (**0,8%**), Salud (**0,7%**), Alimentación (**0,6%**) y Equipamiento de la Vivienda (**0,1%**).

En tanto, los precios del grupo Vestuario cayeron en **0,8%**, fundamentalmente por las liquidaciones de temporada. Por otra parte, este grupo muestra una tendencia a la baja que se refleja en una caída de **16,4%** en los últimos cinco años.

Los grupos Educación y Recreación y Otros se mantuvieron sin variación respecto de junio.

Especial incidencia en el alza del grupo Vivienda tuvo el aumento del precio de la electricidad que alcanzó al **4,4%**. Éste obedeció al efecto rezagado del aumento de tarifas de mediados de junio.

Entre los veinte productos con mayor ponderación en el cálculo del IPC resaltan las alzas de la bencina y el gas licuado, frente a caídas en los precios medios del pasaje de micro, el agua potable y el dividendo hipotecario.



1. Actualiza esta información según la fecha en que te encuentres ingresando a la página web del INE.
2. ¿Cuál es la variación histórica del IPC del mes buscado?
3. ¿Cuáles son los productos con mayor variación? ¿Y cuáles son los que no tuvieron variación?
4. ¿Qué elementos importantes hacen que se produzcan variaciones importantes del IPC?

Mapa conceptual

Conceptos clave:

Población

Muestra

Variables

Clase

Frecuencia absoluta

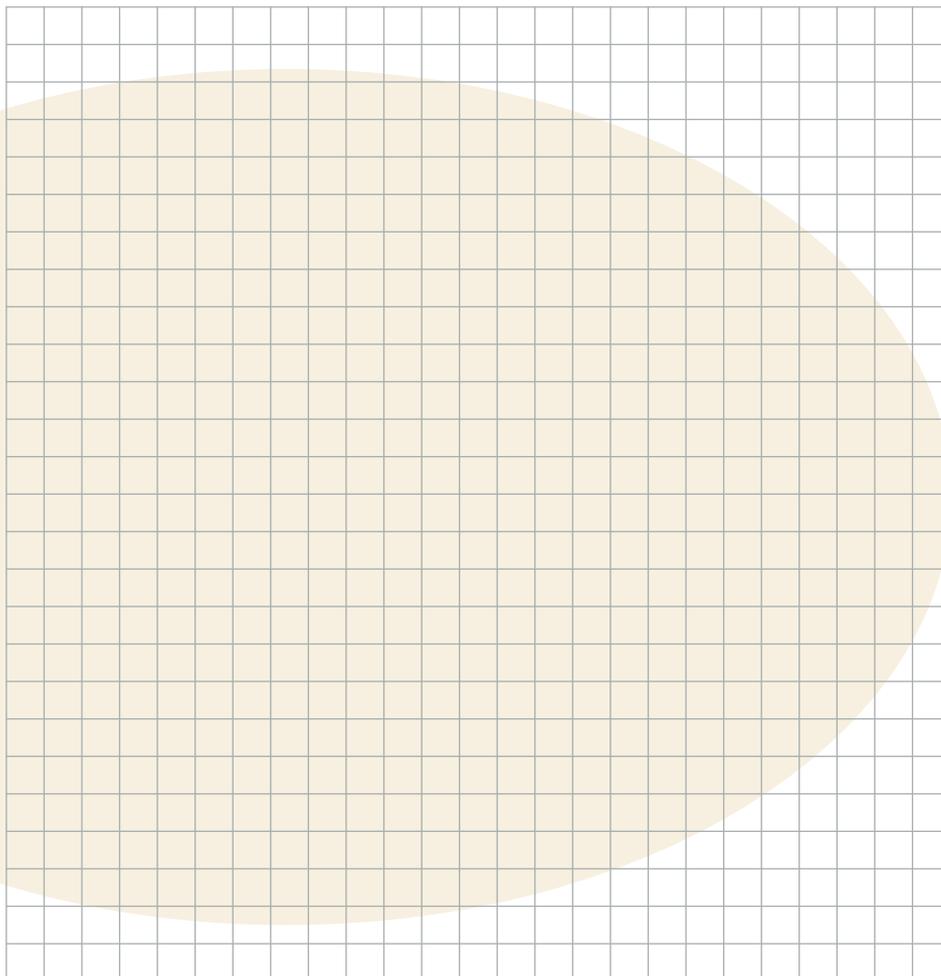
Frecuencia relativa

Intervalos

Diagrama de tallo y hoja

Gráficos

Construye tu mapa conceptual que relacione **al menos** los conceptos clave dados.



Resumen

- 1 Población:** conjunto completo de individuos u objetos a observar, que tienen una característica que se desea medir.
- 2 Muestra:** parte representativa de la población sobre la que se efectúa la medición.
- 3 Variable estadística:** característica o atributo que se observa en cada uno de los elementos de la población y que se mide en la muestra.
- 4 Variable cualitativa:** son aquellas que no se pueden expresar con números, pues representan una cualidad (color de pelo, comuna, deporte preferido, etc.).

- 5 **Variable cuantitativa:** son aquellas que se pueden expresar numéricamente, pues representan una cantidad (edad, peso, cantidad de habitantes, etc.).
- 6 **Frecuencia absoluta:** número de veces que se repite un valor de la variable en la muestra.
- 7 **Frecuencia relativa:** razón entre la frecuencia absoluta y el número total de elementos de la muestra.
- 8 **Frecuencia relativa porcentual:** corresponde a la frecuencia relativa expresada en porcentaje.
- 9 **Diagrama de tallo y hoja:** sirve para comparar la distribución de frecuencias, se puede realizar considerando una o dos variables.
- 10 **Tipos de gráficos:** los gráficos nos permiten representar la información de manera visual, algunos de ellos son:

Gráfico de barras

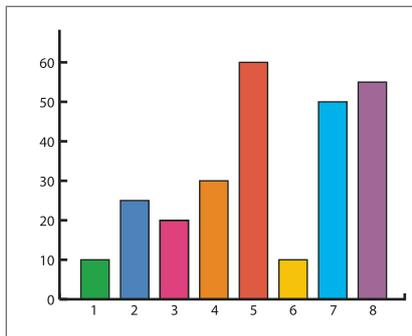


Gráfico circular

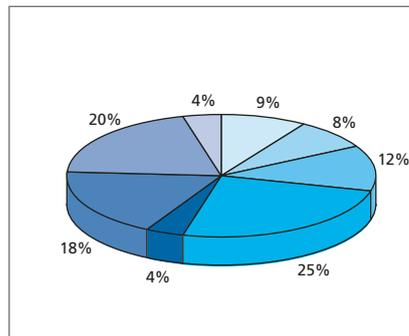
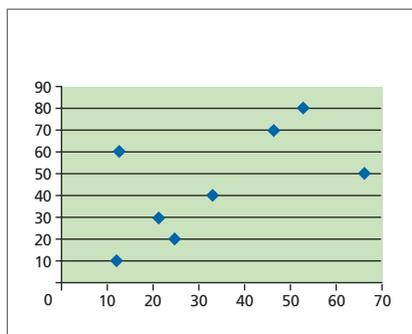
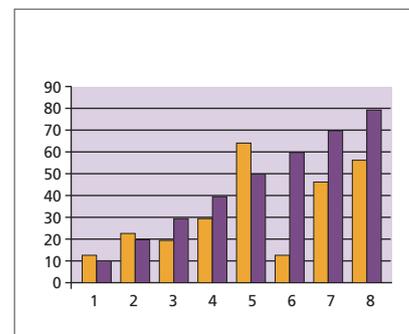


Gráfico de dispersión



Histograma



1. De las siguientes afirmaciones, son **correctas**:

- I) Los chinos hacían censos desde hace miles de años atrás.
- II) La palabra estadística comenzó a usarse en Alemania.
- III) Pearson es considerado el padre de la Estadística Moderna.

- A. Solo I
- B. I y II
- C. II y III
- D. I y III
- E. Todas.

2. En un análisis estadístico, el conjunto de todos los elementos que conforman el objeto de estudio se llama:

- A. rango.
- B. marca de clase.
- C. muestra.
- D. población.
- E. datos.

3. La estatura de un grupo de personas, empleada para un estudio estadístico, es una variable:

- I) cuantitativa.
- II) continua.
- III) discreta.

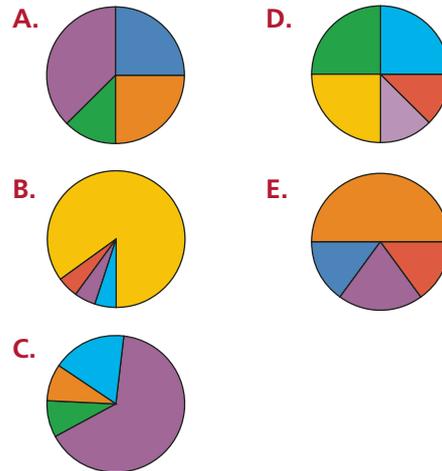
- A. Solo I
- B. Solo II
- C. Solo III
- D. Solo I y II
- E. Solo I y III

4. El tipo de muestra que es adecuado escoger para un estudio estadístico, es:

- A. una muy grande.
- B. una muy pequeña.
- C. una proporcional a la población.
- D. una representativa de la población.
- E. según sea el caso.

5. El gráfico que mejor representa la tabla es:

N° de semanas	f _i
0	2
1	4
2	15
3	2



6. Catalina quiere estudiar psicología. La tabla muestra sus resultados y las ponderaciones pedidas.

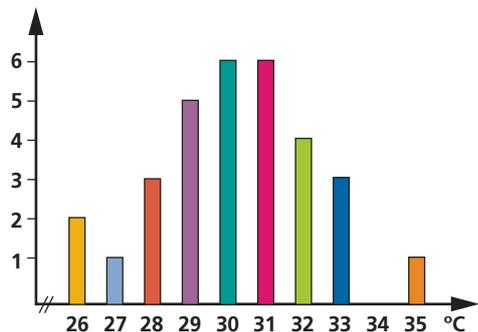
N.E.M.	PSU Leng.	PSU Matem.	PSU Hist. y Geog.	PSU Ciencia
740	712	770	605	610
20%	20%	30%	10%	20%

Con respecto a la tabla es verdadero que:

- I) El puntaje de postulación es levemente superior a **700**.
- II) La prueba de más valor es la de matemática.
- III) Si el **10%** del valor de la prueba de historia se va a la prueba de lenguaje, el puntaje de lenguaje aumenta unos **10** puntos.

- A. Solo I
- B. Solo II
- C. Solo III
- D. I y II
- E. Todas.

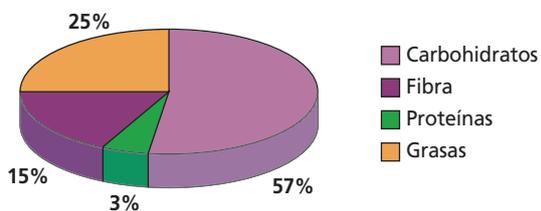
7. El gráfico muestra las temperaturas máximas del mes de enero en el Valle Central.



Con respecto a la información del gráfico es falso que:

- A. Más de la mitad del mes hubo entre 29° a 31°.
- B. Los días más calurosos tuvieron temperaturas de 30° y 31°.
- C. La menor frecuencia fue 34°.
- D. Ningún día la máxima fue 34°.
- E. 11 días hubo menos de 30°.

El gráfico circular nos muestra los porcentajes de los componentes alimenticios que el ser humano debiera consumir.



Fuente: RDA (Recommended Dietary Allowances)

Según el gráfico anterior contesta las siguientes preguntas:

8. ¿Qué porcentaje corresponde a aquellos componentes alimenticios que no sean carbohidratos?

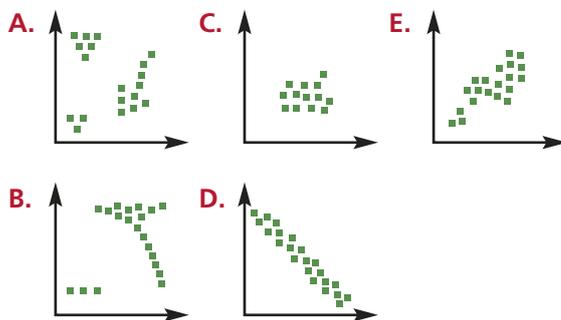
- A. $\frac{43}{100}$
- B. $\frac{3}{100}$
- C. $\frac{1}{4}$
- D. $\frac{3}{20}$
- E. $\frac{57}{100}$

9. El ser humano debe consumir mayormente:

- I) grasas.
- II) proteínas.
- III) carbohidratos.
- IV) fibra.

- A. Solo I
- B. Solo III
- C. Solo I y II
- D. Solo II y III
- E. Todas las anteriores.

10. De los siguientes gráficos el único que presenta una variabilidad homogénea es:



11. La siguiente tabla de frecuencias, muestra las calificaciones de un examen de matemática. ¿Cuál es la proposición falsa?

Calificaciones	Cantidad de alumnos
7.0	3
6.9 – 6.0	6
5.9 – 5.0	5
4.9 – 4.0	13
3.9 – 3.0	10
2.9 – 2.0	3

- A. Hay 6 alumnos que tienen una calificación entre 6.0 y 6.9.
- B. Hay 14 alumnos que tienen una calificación mayor a 4.9.
- C. El total de la muestra es de 40 alumnos.
- D. Hay 13 alumnos que obtuvieron nota insuficiente.
- E. Hay 11 alumnos que calificaron con nota inferior a 7.0 y superior a 6.0.

1. Determina cuál de las siguientes muestras son representativas. En el caso de que no lo sean, explica el por qué.
 - a. Se aplicó una encuesta durante la campaña para la elección de senadores de una región. El muestreo se realizó seleccionando **2.000** personas al azar, a las cuales se las llamó por teléfono. Para la selección se usó la guía de la región.
 - b. En un hospital se hace una encuesta acerca de los hábitos alimenticios de los pacientes, para ello cada médico debe encuestar a tres pacientes en una semana; la selección debe ser al azar.
 - c. En un club social y deportivo quieren saber qué deportes nuevos le interesan a sus asociados, para ello encuestaron a los asistentes a un bingo un día sábado.

2. La siguiente tabla presenta los gustos musicales de los alumnos(as) de dos cuartos año medios.

Música	f_i
Sound	5
Hip-Hop	7
Romántica	12
Rock	16
Reagee	10

a. Calcula la frecuencia relativa de cada tipo de música.

b. Construye un gráfico circular.

3. El siguiente diagrama de tallo y hoja, nos permite visualizar el porcentaje de atenciones respiratorias en niños, de abril a julio del 2005 (datos aproximados) www.minsal.cl.

Niños menores 1 año Niños entre 1 y 14 años

```

                9 5 4 6 9
          6 4 2 0 6 1 2 3 4 9
    9 8 3 3 2 7 2 2 5 7 7 8
  7 5 3 2 1 0 8 1 1
                1 9
    
```

- a. Construir una tabla de frecuencia para cada categoría.
 - b. Construir un gráfico de dispersión para cada categoría.
 - c. Compara ambas distribuciones ¿qué conclusiones puedes obtener?
4. La siguiente tabla de distribución de frecuencias agrupa las marcas, expresadas en metros, obtenidas por un grupo de estudiantes en el lanzamiento del disco.

Intervalo (m)	f_i
34,1 – 34,9	12
35,1 – 35,9	15
36,1 – 36,9	18
37,1 – 37,9	30
38,1 – 38,9	28
39,1 – 39,9	20
40,1 – 40,9	17
41,1 – 41,9	6
42,1 – 42,9	4

- a. ¿Qué porcentaje de estudiantes obtuvo marcas en el intervalo **39,1 – 39,9**?
- b. ¿Qué porcentaje de alumnos obtuvo una marca igual o superior a los **40 m**?

5. Completa la siguiente tabla de distribución de frecuencias correspondientes a las medidas de una pieza de motores, después de un año de uso. Expresa las frecuencias relativas aproximadas a las milésimas (tres decimales).

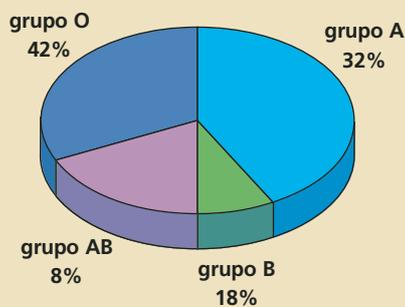
Intervalo (mm)	f_i	f_r	Frecuencia relativa porcentual
100 – 109	4		
110 – 119	17		
120 – 129	29		
130 – 139	18		
140 – 149	10		
150 – 159	5		
160 – 169	2		

6. Dibuja, en un solo gráfico, el histograma y el polígono de frecuencias correspondiente a la tabla del ejercicio anterior.
7. Los siguientes datos corresponden a la duración en horas, de uso continuo de 50 dispositivos electrónicos iguales, sometidos a un control de calidad.

480	496	724	780	801
570	802	795	886	714
775	712	683	830	560
826	560	794	676	760
890	590	750	489	725
666	746	668	880	570
830	452	810	720	680
680	660	490	895	660

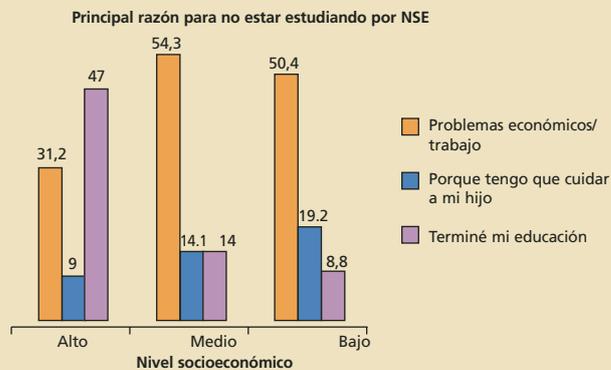
Construye una tabla de distribución de frecuencias agrupadas que considere las columnas: intervalo, frecuencia absoluta, frecuencia relativa.

8. A un curso de 40 estudiantes de cuarto medio se les preguntó su grupo sanguíneo.



- a. ¿A qué tipo de gráfico corresponde el representado?
- b. ¿Qué procedimiento utilizarías para encontrar la cantidad de personas por grupo sanguíneo?
- c. Realiza un gráfico de barras cuyas variables sean el grupo sanguíneo y su frecuencia absoluta.

9. Construye un diagrama de tallo y hoja con los datos del ejercicio 7.
10. El siguiente gráfico muestra la principal razón para no estar estudiando, según nivel socioeconómico (NSE).



- a. ¿Qué NSE presenta, en mayor medida, motivos para no poder terminar los estudios?
- b. ¿Cuál NSE es el que tiene más personas con sus estudios terminados?

11. Los datos que se indican a continuación corresponden a g/dl de hemoglobina en la sangre de pacientes hombres entre 25 y 35 años de edad.

14,3	15,1	15,3	15,5	13,0
15,0	14,5	15,2	14,2	15,9
15,2	15,7	15,4	15,8	17,5
13,2	15,4	16,1	17,1	15,2
15,4	16,2	14,2	15,4	13,3
15,2	15,3	16,7	15,5	16,9
15,1	15,2	14,2	13,2	15,3
14,3	14,6	13,3	15,2	14,3
15,5	14,1	15,5	14,8	13,6
13,9	15,0	16,2	15,2	14,9
14,7	14,7	15,0	14,9	15,9
15,8	16,4	17,3	14,7	16,3
14,8	14,8	16,4	16,8	15,0
15,7	16,5	14,8	15,6	14,8
14,6	14,9	15,6	16,0	14,7
16,3	16,5	16,9	17,3	15,8
17,2	15,8	16,3	15,9	16,9
16,0	17,1	16,8	16,7	17,3
17,5	16,8	16,4	17,4	16,0
15,7	15,9	16,1	15,8	16,4

Con ayuda de una planilla Excel, construye una tabla de frecuencias que agrupe estos datos.

Estadística II

A veces, cuando las poblaciones son muy grandes, es muy difícil, por problemas de tiempo y dinero, hacer un análisis que incluya a toda la población. Por este motivo, lo que se hace es estudiar una parte de ella, llamada muestra; cuando los individuos de la muestra han sido seleccionados de acuerdo a procedimientos estadísticos, se pueden sacar conclusiones que caracterizan a toda la población. A estos resultados los llamaremos inferencias.

Es común, hoy en día, recibir invitaciones a participar de encuestas en la mayoría de los sitios de Internet relacionados con las comunicaciones o empresas que necesitan saber lo que quieren sus clientes. Por ejemplo, en un diario electrónico se publicó una encuesta sobre la creencia en extraterrestres. En relación con ella, ¿se puede decir que el 67% de las personas cree en extraterrestres? ¿Bastará con encuestar a **316** personas para obtener conclusiones relevantes?

Encuesta

¿Crees en los extraterrestres?

Vota aquí Sí No

316 encuestados

Hasta ahora 67% Si – 33% No

Enviar

No quieres votar pero quieres [ver los resultados](#)



En esta unidad aprenderás a...

- Conocer las medidas de tendencia central: promedio, mediana y moda.
- Conocer las medidas de dispersión: rango, desviación media, desviación estándar.
- Trabajar con las medidas de localización: cuartiles, deciles, percentiles.
- Conocer y trabajar con muestras, identificando niveles de confianza y margen de errores.

Explora

Realiza el laboratorio 1 correspondiente a la unidad 2 que aparece en www.santillana.cl/mat4

¿Cuánto sabes?

1. Calcula el valor de x en las siguientes proporciones.

a. $\frac{6}{24} = \frac{15}{x}$

c. $\frac{1}{6} : \frac{15}{12} = \frac{1}{3} : x$

e. $\frac{0,7}{x} = \frac{2,5}{1,4}$

b. $\frac{7}{21} = \frac{x}{900}$

d. $x : 2,4 = 3 : 1,8$

f. $\frac{x}{6} = \frac{24}{x}$

2. Calcula los siguientes porcentajes.

a. 10% de 457

c. 99% de 1.246

e. 18% de 310.000

b. 25% de 398

d. 5,7% de 45.980

f. 60% de 94.327

3. Completa.

a. 281,49 representa el _____ % de 853

b. 38.000 representa el _____ % de 95.000

c. 13.891,5 representa el _____ % de 18.522

d. 2.809,8 representa el _____ % de 46.830

e. 652 representa el _____ % de 65.200

f. 55.928,95 representa el _____ % de 76.615

4. En una empresa se ha entregado la planilla de sueldos correspondiente al mes de julio. Completa la planilla para poder saber cuánto dinero recibe cada persona al cobrar su sueldo.

Nombre	Sueldo (imponible)	Fonasa o Isapre (7% del imponible)	AFP (13% del imponible)	Sueldo líquido
Daniel	\$ 165.249			
Carolina	\$ 237.860			
Andrea	\$ 551.925			
Sebastián	\$ 618.004			
Jorge	\$ 1.045.776			

5. Usando tu calculadora, evalúa cada expresión dados los siguientes valores (aproxima el resultado a tres decimales):

$$a = \frac{1}{9} \quad b = 0 \quad c = -5 \quad d = 80$$

a. $\sqrt{a} + \sqrt{bd} - c$

d. $(c - b : d^3) : (a^2 + \sqrt{d})$

g. $(b - c) : \sqrt{\frac{c^4}{2a}}$

b. $\frac{a}{c} : \sqrt{a^4} + b^2 \cdot acd$

e. $\sqrt{\frac{abc^3d}{c}}$

h. $(c - a) \cdot (b + d)$

c. $a \cdot d - \sqrt{c^2}$

f. $a - b + c - d^2$

i. $\sqrt[3]{(a + d)^6}$

6. Los siguientes datos corresponden a las estaturas, en metros, de los alumnos de IV Medio de un colegio.

Mujeres

1,56 – 1,49 – 1,63 – 1,71 – 1,56 – 1,55 – 1,61 – 1,74 – 1,68 – 1,52 – 1,57 – 1,48 – 1,54 – 1,60 – 1,55 – 1,54 – 1,49 – 1,50 – 1,56 – 1,53 – 1,72 – 1,66 – 1,53 – 1,62 – 1,59 – 1,63 – 1,71 – 1,69 – 1,73 – 1,67 – 1,59 – 1,63 – 1,65 – 1,76 – 1,61 – 1,57 – 1,58 – 1,71 – 1,51 – 1,66 – 1,64 – 1,63

Hombres

1,65 – 1,69 – 1,74 – 1,81 – 1,72 – 1,68 – 1,61 – 1,73 – 1,79 – 1,81 – 1,74 – 1,85 – 1,84 – 1,76 – 1,66 – 1,69 – 1,73 – 1,72 – 1,76 – 1,79 – 1,86 – 1,69 – 1,63 – 1,79 – 1,77 – 1,76 – 1,74 – 1,81 – 1,83 – 1,69 – 1,74 – 1,77 – 1,71 – 1,75 – 1,68 – 1,88 – 1,76 – 1,74 – 1,68 – 1,83 – 1,81 – 1,73 – 1,76 – 1,78 – 1,76 – 1,79 – 1,83 – 1,66

- Determina el valor máximo y mínimo en cada caso.
- Construye un gráfico de barras y un polígono de frecuencias con la estatura de todos los alumnos de IV Medio.
- Usando una planilla Excel, construye un gráfico de dispersión que permita comparar la estatura de los hombres y la estatura de las mujeres. ¿Qué conclusiones puedes obtener?

- 1 **Teorema fundamental de las proporciones:** “Dos razones forman una proporción si y solo si el producto de sus términos extremos es igual al producto de sus términos medios”.

$$\begin{array}{c}
 \text{términos extremos} \\
 \boxed{a : b = c : d} \\
 \text{términos medios}
 \end{array}
 \Leftrightarrow
 \frac{a}{b} = \frac{c}{d}
 \Leftrightarrow
 a \cdot d = b \cdot c, \quad b \neq 0, \quad d \neq 0$$

- 2 Para calcular el $a\%$ de un número b cualquiera, puedes aplicar el siguiente procedimiento:

$$\frac{a}{100} \cdot b$$

O puedes calcular $x \cdot b$, donde x es la expresión decimal que representa el $a\%$.

- 3 Sean a , b dos números reales cualesquiera. Para calcular a qué porcentaje corresponde a de b , puedes aplicar el siguiente procedimiento:

Si x es el porcentaje, entonces $x = \frac{a \cdot 100}{b}$.

- 4 **Población:** es un conjunto de personas, situaciones o cosas de las cuales se desea hacer un estudio y las cuales tienen una característica en común.

- 5 **Muestra:** es un subconjunto cualquiera de la población.

¿Qué debes recordar?

Medidas de tendencia central

Las medidas de tendencia central nos dan una idea acerca del comportamiento de los datos a los que se refieren. Se puede decir que expresan el grado de centralización de los datos que representan.

Antes de profundizar en las principales medidas de tendencia central, es necesario conocer la siguiente notación de sumatoria.

Una suma como $x_1 + x_2 + x_3 + x_4 + \dots + x_n$ se puede expresar de manera resumida mediante el uso del símbolo de sumatoria: Σ .

La suma de los términos de la forma x_k , donde k es un número natural que

varía desde 1 a n , se simboliza por $\sum_{k=1}^{k=n} x_k$.

Entonces, $\sum_{k=1}^{k=n} x_k = x_1 + x_2 + x_3 + x_4 + \dots + x_n$

TIPS



El símbolo Σ es la letra griega "sigma" mayúscula. Corresponde a la decimoctava letra del alfabeto griego, que equivale a la letra **S** de nuestro alfabeto.

AYUDA



$$1 + 2 + 3 + 4 + 5 = \sum_{k=1}^{k=5} k$$



PARA ARCHIVAR

El uso de **sumatoria** tiene evidentes ventajas, puesto que permite escribir fórmulas de manera reducida.

Media aritmética

La media aritmética de n datos numéricos que expresan cantidades, es el cociente entre la suma de todos los datos y la frecuencia total de ellos.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + x_4 + \dots + x_n}{n} = \frac{\sum_{k=1}^{k=n} x_k}{n}$$

Es decir,

TIPS



La media aritmética se designa por el signo \bar{x} .

TIPS



El promedio que se emplea en las calificaciones escolares, corresponde a la media aritmética de estas.

Ejemplo 1

La siguiente tabla muestra el precio de un cuaderno en diferentes tiendas comerciales (según un estudio del SERNAC).

Tienda 1	Tienda 2	Tienda 3	Tienda 4	Tienda 5
\$ 940	\$ 1.100	\$ 845	\$ 820	\$ 745

Calculemos la media aritmética de los precios anteriores:

$$\bar{x} = \frac{940 + 1.100 + 845 + 820 + 745}{5} = \frac{4.450}{5} = 890 \text{ pesos}$$

Observa que el promedio en este caso no coincide con ninguno de los valores dados en la tabla.

Ejemplo 2

La siguiente tabla muestra la distribución de frecuencias de los puntajes obtenidos por 50 alumnos en una prueba de matemática.

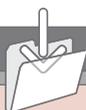
Intervalo	Frecuencia	Marca de clase	f · x
60 – 64	5	62	310
65 – 69	5	67	335
70 – 74	8	72	576
75 – 79	12	77	924
80 – 84	16	82	1.312
85 – 89	4	87	348
$\Sigma f = 50$			$\Sigma f \cdot x = 3.805$

AYUDA



Observa que, en el ejemplo 2, se calculó la media aritmética para datos agrupados (separados en intervalos, en una tabla de frecuencia).

El promedio de los valores está dado por $\bar{x} = \frac{\Sigma f \cdot x}{\Sigma f} = \frac{3.805}{50} = 76,1$ puntos.



PARA ARCHIVAR

Para calcular la media aritmética de datos no agrupados utilizamos la

fórmula $\frac{\sum_{k=1}^{k=n} x_k}{n}$, mientras que para datos agrupados utilizamos $\frac{\Sigma f \cdot x}{\Sigma f}$,

donde **f** es la frecuencia y **x** la marca de clase correspondiente a cada intervalo.

IR A LA WEB



Desarrolla el laboratorio 2.
www.santillana.cl/mat4

EJERCICIOS

- Un alumno obtuvo las siguientes notas parciales en Matemática: **4,8**; **2,5**; **6,0**; **3,9** y una quinta nota que no recuerda. Si su promedio fue **4,6**, calcula la nota que falta.
- Dos alumnos obtuvieron el mismo promedio semestral de notas. ¿Significa que tuvieron las mismas notas? Justifica numéricamente tu respuesta.
- En una oficina, el jefe gana **\$ 540.000** y tres empleados ganan **\$ 100.000**, **\$ 155.000** y **\$ 165.000**, respectivamente. La media aritmética de los sueldos, ¿es un valor representativo de esos sueldos?
- En una muestra de control se midieron **10** clavos de una bolsa, con los siguientes resultados: **5** de **2,00"**; **3** de **1,99"** y **2** de **2,05"**. Calcula la longitud media de la muestra.

Ejemplo 3

Un alumno que postula a la Universidad tiene los siguientes puntajes en las Pruebas de Selección Universitaria (PSU) y en sus notas de Educación Media.

	Puntaje	Ponderación (%)
Prueba Lenguaje	680	10
Prueba Matemática	752	20
Prueba Ciencias	640	10
Prueba Historia y Geografía	720	40
Notas E. Media	590	20

Calculemos su puntaje ponderado, es decir, la **media aritmética ponderada** de sus puntajes:

$$\bar{x} = \frac{10 \cdot 680 + 20 \cdot 752 + 10 \cdot 640 + 40 \cdot 720 + 20 \cdot 590}{10 + 20 + 10 + 40 + 20} = \frac{68.840}{100} = 688,4 \text{ puntos.}$$

Hemos calculado la media aritmética ponderada, la cual nos sirve para calcular el promedio de datos que no tienen igual ponderación.

AYUDA



La importancia de un dato se traduce en un número que corresponde a su ponderación.



PARA ARCHIVAR

Si p_k es la ponderación de un dato x_k , el promedio ponderado se obtiene

utilizando la siguiente expresión:
$$\bar{x} = \frac{\sum_{k=1}^{k=n} x_k \cdot p_k}{\sum_{k=1}^{k=n} p_k}$$

AYUDA



La mediana divide los datos en dos subconjuntos que contienen igual cantidad de elementos.

⇒ **Mediana**

La **mediana** de un conjunto de datos numéricos ordenados en forma creciente o decreciente, es el dato que se encuentra al centro de dicha ordenación, o la media aritmética de los datos centrales (en caso que la muestra tenga un número de datos pares).

Ejemplos

1 3 **6** 7 9
Mediana: **6**

1 3 **4** **5** 7 9
Mediana: **4,5** (promedio entre 4 y 5)

EN EQUIPO



Averiguen el promedio de notas por cada alumno del curso. Ordenen la información en una tabla de frecuencia, luego determinen media, mediana y moda. ¿Qué pueden concluir?

⇒ **Moda**

La **moda** de un conjunto de datos, es aquel que tiene la mayor frecuencia.

Medidas de dispersión

Las medidas de dispersión determinan cuán cercanos o lejanos están los datos de un valor central, respecto a la media aritmética. También indican el grado de variabilidad de los datos. En estas páginas estudiaremos algunas de ellas: **rango**, **desviación con respecto a la media** y la **desviación estándar o típica**.

Ejemplo

El colegio otorgará una beca de matrícula para la universidad, al alumno cuyo buen rendimiento se haya mantenido por mayor tiempo, en el último trimestre de 4° medio. Para calcular el mejor promedio solo consideraron algunas asignaturas. Los mejores alumnos de la promoción fueron Pablo y Soledad.

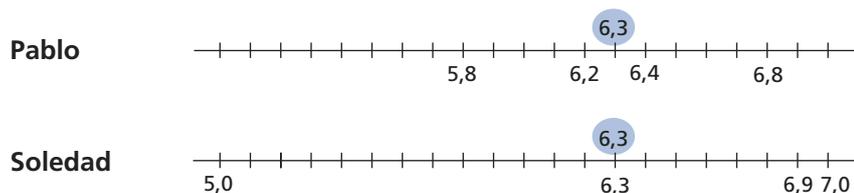
La media aritmética (promedio) de cada uno es **6,3**.

Si solo uno debe ser elegido ¿quién ganará la beca?

Las calificaciones son las siguientes:

	Lenguaje	Matemática	Historia	Ciencias
Pablo	6,2	6,8	5,8	6,4
Soledad	6,9	5,0	7,0	6,3

Observa la siguiente representación de las calificaciones,



Las calificaciones de Pablo se encuentran más cercanas a la media aritmética, que las notas de Soledad. Es decir, las calificaciones de Soledad se encuentran más **dispersas**.

¿Es suficiente este argumento para optar por Pablo, como un alumno que ha mantenido su buen rendimiento?

Las medidas de dispersión nos permitirán realizar un análisis más certero.

Rango

Anteriormente utilizamos el rango para determinar el tamaño de cada intervalo en una tabla de frecuencias. Simbolizaremos el **rango** por la letra **R**. Aunque no es una medida muy significativa, este nos indica cuán dispersos se encuentran los datos entre los valores de los extremos.

Pablo ▶ $R: 6,8 - 5,8 = 1$ **Soledad** ▶ $R: 7,0 - 5,0 = 2$

Como el valor del rango de las notas de Pablo es menor que el de Soledad, podemos decir que sus calificaciones son menos dispersas. Por lo tanto, sería el más apto para ganar el premio por mantener un buen rendimiento.



AYUDA



Recuerda que el **rango** de un conjunto de datos numéricos, se calcula como la diferencia entre el mayor y el menor valor de la distribución.

TIPS



¿Qué significado tiene un rango de notas **4,2** respecto de las notas de otro alumno cuyo rango es **2,1**?

En el primer caso las notas están más dispersas que en el segundo. Sin embargo, no sabemos en qué caso son mejores; para determinarlo debemos disponer de más información.

Muchos conjuntos de notas pueden tener rango **2,1** y sus respectivas medias aritméticas ser muy diferentes.

TIPS



La idea de desviación representa el mayor o menor alejamiento de un dato con respecto a \bar{x} .

Desviación media

La media aritmética de ambos alumnos es de **6,3**.

Si calculamos la diferencia de una nota con la media aritmética tendremos la desviación de la nota con respecto a \bar{x} . Las desviaciones de todas las notas, de Pablo y Soledad, con respecto a $\bar{x} = 6,3$ se indican a continuación:

Pablo

Nota	x	6,2	6,8	5,8	6,4
Desviación respecto de la medida	$x - \bar{x}$	-0,1	0,5	-0,5	0,1

Soledad

Nota	x	6,9	5,0	7,0	6,3
Desviación respecto de la medida	$x - \bar{x}$	0,6	-1,3	0,7	0

AYUDA



Comprobemos que la suma de las desviaciones es siempre **0**:

Pablo:
 $-0,1 + 0,5 - 0,5 + 0,1 = 0$

Soledad:
 $0,6 - 1,3 + 0,7 + 0 = 0$

Si sumamos las desviaciones medias de cada uno resulta **0**.



PARA ARCHIVAR

La **desviación de una variable x** con respecto a la media aritmética \bar{x} está dada por la diferencia: $d = x - \bar{x}$.

La suma de las desviaciones de todos los datos con respecto a su media aritmética es cero.

Para conocer quién presenta un valor de desviación, que nos indique cuán cercano o lejano está de la media aritmética, será necesario calcular el valor absoluto de la desviación.

Pablo

Nota	x	6,2	6,8	5,8	6,4
Desviación respecto de la medida	$ x - \bar{x} $	0,1	0,5	0,5	0,1

Soledad

Nota	x	6,9	5,0	7,0	6,3
Desviación respecto de la medida	$ x - \bar{x} $	0,6	1,3	0,7	0

TIPS



La desviación se puede calcular con respecto a cualquier valor, no solo respecto a la media aritmética. Esta puede ser positiva, cero o negativa.

Desviación absoluta de Pablo: $\sum_{k=1}^4 |x_k - \bar{x}| = 0,1 + 0,5 + 0,5 + 0,1 = 1,2$

Desviación absoluta de Soledad: $\sum_{k=1}^4 |x_k - \bar{x}| = 0,6 + 1,3 + 0,7 + 0 = 2,6$

Como el valor de la desviación absoluta de Soledad es mayor, entonces las notas de Pablo son las que se encuentran más cercanas de mantener un buen rendimiento durante un período de tiempo.

EN EQUIPO



Si las desviaciones se calculan en relación a un valor distinto de la media aritmética, ¿cuánto suman sus valores?, ¿por qué?



Se define **desviación media** como la media aritmética de las desviaciones absolutas respecto de la media. La designaremos como **DM**.

$$DM = \frac{\sum_{k=1}^n |x_k - \bar{x}|}{n}$$

La **DM** de Pablo es **0,3** y la de Soledad es **0,65**.

Desviación estándar o típica

Otra importante medida de dispersión es la desviación estándar.

→ La **desviación estándar o típica** expresa el grado de dispersión de los datos con respecto a \bar{x} . Se designará con la letra s , y se calculará de la forma:

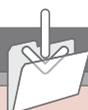
$$s = \sqrt{\frac{\sum_{k=1}^n (x_k - \bar{x})^2}{n}}$$

Continuando con el análisis de quien ganará la beca, obtendremos el valor de la desviación estándar de cada alumno:

Pablo	Nota (x)	6,2	6,8	5,8	6,4	$\sum_{k=1}^4 (x_k - \bar{x})^2$	$s = \sqrt{\frac{0,52}{4}} \approx 0,36$
	$(x_k - \bar{x})^2$	0,01	0,25	0,25	0,01	0,52	
Soledad	Nota (x)	6,9	5,0	7,0	6,3	$\sum_{k=1}^4 (x_k - \bar{x})^2$	$s = \sqrt{\frac{2,54}{4}} \approx 0,79$
	$(x_k - \bar{x})^2$	0,36	1,69	0,49	0	2,54	

Observamos que el valor de la desviación estándar de las notas de Pablo es menor que la de Soledad, entonces, podemos decir que las calificaciones de Pablo están más cercanas a la media, y son menos dispersas.

Por lo tanto, el más indicado para ganarse la beca que otorga el colegio es Pablo, ya que sus calificaciones cumplen con haber mantenido un buen rendimiento.



PARA ARCHIVAR

Mientras menor sea el valor del **coeficiente de la desviación estándar**, el grupo de observaciones es más "homogéneo" que si el valor de la desviación estándar fuera más grande. O sea, a menor dispersión mayor homogeneidad y a mayor dispersión, menor homogeneidad.

Desviaciones para datos agrupados

Recordemos que los datos agrupados pertenecientes a una clase se consideran iguales a la respectiva marca de clase.

→ Para **datos agrupados**, el cálculo de ambos tipos de desviaciones se puede aplicar al método abreviado, tal como se obtuvo en la media aritmética:

Desviación media:

$$DM = \frac{\sum f \cdot |x - \bar{x}|}{\sum f}$$

Desviación estándar:

$$s = \sqrt{\frac{\sum f \cdot (x - \bar{x})^2}{n}}$$

TIPS



La desviación estándar es muy inestable a pequeñas variaciones que se producen respecto de la media.

TIPS



Recuerda que s puede estar referida a otro valor que no sea \bar{x} . Si se emplea \bar{x} el valor de s que se obtiene es mínimo. En otros casos este valor sería mayor.

AYUDA



La desviación típica es un valor de la misma naturaleza que los datos con que se calcula. Si el valor de s en un conjunto de notas es $s = 1,8$, el número **1,8** se refiere a puntos de notas.

EN EQUIPO



Averigüen, ¿qué significa la palabra "homogéneo" y "heterogéneo"?

EJERCICIOS

1. El análisis de las notas de un curso señala que en ambos trimestres el promedio en matemática es **5,1**, al término del primer y segundo trimestre, la nota máxima es **7,0** y la mínima es **3,2**. Sin embargo, los alumnos tienen la sensación de mejores resultados en un trimestre que en otro.

Primer trimestre

7,0	5,6	4,3	7,0	5,4	4,3
6,9	5,4	4,3	6,8	5,2	4,1
6,5	5,2	4,1	6,3	5,2	4,1
5,8	4,8	4,1	5,7	4,8	3,2
5,6	4,8	3,2	5,6	4,5	3,2

Segundo trimestre

7,0	5,3	5,0	6,4	5,2	4,9
6,1	5,2	4,7	6,0	5,2	4,1
5,7	5,1	4,7	5,5	5,0	4,6
5,4	5,0	4,5	5,3	5,0	4,5
5,3	5,0	3,2	5,3	5,0	3,2

Responde las siguientes preguntas:

- ¿Cuánto es el coeficiente del rango en cada trimestre? ¿Qué trimestre tiene un coeficiente de rango menor?
- Según el coeficiente del rango, ¿qué trimestre presenta calificaciones más dispersas, en relación al promedio?
- ¿Cuánto es el valor del coeficiente de la desviación media en cada trimestre?
- Según la situación, ¿cómo interpretarías el coeficiente de desviación media? ¿Corroboras la "sensación" de los estudiantes?
- Calcula el coeficiente de la desviación estándar para cada trimestre.
- ¿Qué trimestre presenta calificaciones más homogéneas?
- ¿Cómo interpretarías el valor del coeficiente de desviación estándar?
- ¿Cuál fue el mejor trimestre?, ¿por qué lo consideras mejor?
- Construye la gráfica que mejor represente la situación.

2. Un grupo de alumnos obtuvo las siguientes marcas, en salto con garrocha, expresadas en metros: **2,50 ; 2,80 ; 2,60 ; 3,00 ; 2,90**.

- Comprueba que la suma de las desviaciones de estos datos respecto a \bar{x} es 0.
- Calcula la desviación media de los datos.

3. La tabla de distribución de frecuencias muestra la puntuación obtenida por **1.800** alumnos de 5° a 8° Básico en un cuestionario de cultura general.

Puntaje	Frecuencia
0 – 2	21
3 – 5	50
6 – 8	110
9 – 11	241
12 – 14	423
15 – 17	457
18 – 20	275
21 – 23	134
24 – 26	66
27 – 29	23

- Calcula la desviación estándar de la distribución.
 - ¿A qué cantidad de puntos corresponden los valores de $\bar{x} + s$ y $\bar{x} - s$?
4. En una misma prueba de Matemática dos cursos **A** y **B**, obtuvieron resultados cuyos datos estadísticos son los siguientes:

	Curso A	Curso B
\bar{x}	5,3	5,4
s	0,7	0,4

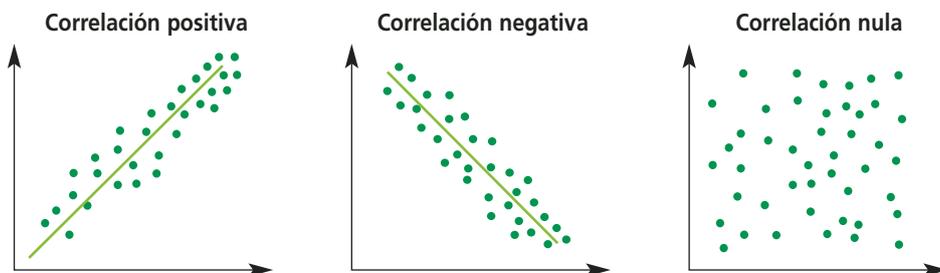
De acuerdo con estos datos:

- Un alumno del curso **A** obtuvo un **6,7** y uno del curso **B** un **6,6**. ¿A cuál de los alumnos le fue mejor en la prueba, en relación a su curso?
- Justifica la respuesta anterior y compártela con un compañero.

Correlación

La **correlación** indica el grado de asociación de dos variables; la influencia que pueda tener una sobre la otra, lo que a veces permite encontrar funciones que predicen ciertos comportamientos, como, por ejemplo, el modelo que se usa para aplicar la restricción vehicular.

Veamos algunos ejemplos gráficos.



PARA ARCHIVAR

El **grado de asociación** o **correlación** de dos variables puede ser:

- positiva: están directamente relacionadas.
- negativa: se relacionan de manera inversa.
- nula: no existe relación entre ellas.

TIPS



Para calcular el coeficiente de correlación en Excel se utiliza la función estadística COEF. DE CORREL.

AYUDA



La correlación se mide usando el coeficiente de correlación lineal de Pearson (r).

$$r = \frac{S_{xy}}{S_x \cdot S_y}$$

$$S_{xy} = \frac{\sum_{i=1}^n x_i \cdot y_i}{n} = \bar{x} \cdot \bar{y}$$

S_x = Desviación típica x .

S_y = Desviación típica y .

r cercano a **1**, indica correlación positiva.

r cercano a **-1**, indica correlación negativa.

r cercano a **cero**, indica correlación nula.

EJERCICIOS

1. En las siguientes situaciones señala si la correlación es positiva, negativa o nula.
 - a. Cantidad de hijos de una familia y dinero gastado por esa familia en el supermercado.
 - b. Edad de una persona y cantidad de libros que ha leído.
 - c. Promedio en matemática de cuarto medio y resultado de esa persona en la PSU de matemática.
2. Averigua los promedios que tus compañeros obtuvieron en el primer semestre en las asignaturas de Castellano, Historia, Matemática, Biología, Química y Física.
 - a. Calcula los coeficientes de correlación entre:
 - Matemática y Física
 - Matemática y Química
 - Lenguaje e Historia
 - Química y Biología
 - Lenguaje y Matemática
 - b. ¿Entre qué asignaturas existe mayor correlación?
 - c. ¿Son lógicos los resultados? Justifica.
3. Al estudiar la relación entre la masa y la edad de los niños de un jardín infantil, la directora obtuvo que el coeficiente de correlación de Pearson era de **0,85**, por lo que dedujo que había un alto grado de asociación entre ambas variables. Por otra parte, el director de una casa de reposo para ancianos hizo el mismo estudio, obteniendo como coeficiente de correlación **0,345**, por lo que determinó que la edad no tenía ninguna relación con la masa. ¿A qué se deben estas diferentes conjeturas?

Medidas de localización: cuartiles, percentiles y deciles



La prueba de tolerancia a la glucosa se realiza mediante muestras de sangre, determinando si los niveles de glicemia están dentro de los percentiles considerados normales.

Anteriormente aprendiste que la mediana de un conjunto de datos ordenados, de acuerdo a su magnitud, los separa en dos mitades.

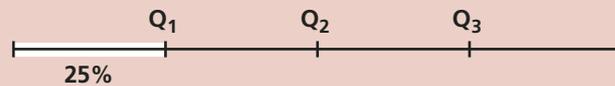
Ahora estudiaremos otros valores típicos que dividen a un conjunto de datos numéricos en cierta cantidad de partes iguales, como los cuartiles, deciles, percentiles.



Cuartil

Los cuartiles de una distribución de datos numéricos, corresponden a los 3 valores que dividen a estos en 4 partes iguales, es decir, al 25%, 50% y 75%.

Los cuartiles se designan por $Q_1(25\%)$, $Q_2(50\%)$ y $Q_3(75\%)$.



TIPS



Observa que, en el caso de los cuartiles, la mediana corresponde a Q_2 . En el caso de los deciles, corresponde a D_5 .

Ejemplo

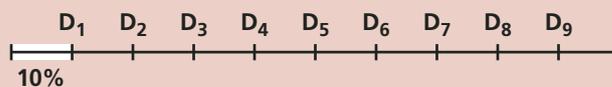
En la distribución de notas de un grupo de alumnos, el cuartil Q_2 es una nota de referencia que permite afirmar que el 50% de los alumnos obtuvo esa nota o una menor.



Deciles

Los deciles de una distribución de datos numéricos corresponden a los 9 valores que dividen a estos en 10 partes iguales.

Los deciles se designan por D_1, D_2, \dots, D_9



Ejemplo

En el caso anterior, el decil D_6 es una nota de referencia que nos permite afirmar que el 60% de los alumnos obtuvo esa nota o una menor.

IR A LA WEB



Desarrolla el laboratorio 3.
www.santillana.cl/mat4

Cuando queremos estudiar una muestra que contiene muchos datos, podemos subdividir esta en **percentiles**. Los percentiles de una distribución de datos numéricos, corresponden a los **99** valores que dividen a estos en **100** partes iguales.

Los percentiles se designan por P_1, P_2, \dots, P_{99}

Ejemplo

El percentil P_{70} de una distribución de frecuencias dadas en una competencia del lanzamiento de la jabalina, nos indica que el **70%** de los competidores alcanzó esa distancia o una menor.

AYUDA

Observa que: P_{50} equivale a la mediana.

TIPS

En Excel podemos calcular percentiles utilizando la función: **=PERCENTIL()**.

Por ejemplo, el percentil 5 de los datos **A1** hasta **A6** se ingresa: **=PERCENTIL(A1:A6,5)**.



PARA ARCHIVAR

Para calcular el n -ésimo percentil utilizamos la siguiente fórmula

$$P_n = l_i + (l_{i+1} - l_i) \cdot \frac{f_n - f_i}{f_{i+1} - f_i}, \text{ con}$$

l_i : extremo izquierdo del intervalo donde se ubica el percentil.

l_{i+1} : extremo derecho del intervalo donde se ubica el percentil.

f_i : frecuencia acumulada hasta l_i .

f_{i+1} : frecuencia acumulada hasta l_{i+1} .

f_n : frecuencia acumulada hasta el percentil buscado (P_n).

Ejemplo

Calculemos el percentil **45** considerando la distribución de frecuencias de **212** puntajes obtenidos en la PSU.

Puntaje	Frecuencia absoluta	Frecuencia acumulada
[400, 449[10	10
[450, 499[9	19
[500, 549[20	39
[550, 599[31	70
[600, 649[80	150
[650, 699[42	192
[700, 749[10	202
[750, 799[10	212
	212	

AYUDA

La frecuencia acumulada hasta el percentil P_n , se calcula de la siguiente manera:

$$f_n = \frac{n}{100} \cdot N$$

(N : tamaño de la muestra).

TIPS



Los percentiles, deciles y cuartiles reciben el nombre de **cuantiles**. Conocer estos valores nos proporciona una importante información acerca de los datos de una cierta distribución.

El **45%** de los datos es **95,4**, entonces $f_n = 95,4$, este valor se encuentra en el intervalo **[600, 649]**. Además $l_i = 600$; $l_{i+1} = 649$; $f_i = 70$; $f_{i+1} = 150$.

Remplazando en la fórmula tenemos:

$$P_{45} = 600 + (649 - 600) \cdot \frac{(95,4 - 70)}{150 - 70} = 615,5.$$

El resultado nos indica que el **45%** de los alumnos obtuvo puntajes menores o iguales a **615,5**.

Nota

La fórmula para encontrar un determinado percentil se puede generalizar para encontrar cuartiles y deciles, solo varía el cálculo de f_n .

EJERCICIOS

1. A partir de los datos dados en la tabla anterior:
 - a. Calcula D_3 .
 - b. Calcula Q_3 .
 - c. ¿Qué información nos entrega (a) y (b)?
 - d. ¿Qué porcentaje de los **212** alumnos obtuvieron resultados entre **620** y **680** puntos?

2. Dada la siguiente tabla de distribución de frecuencias, que muestra los puntajes obtenidos por **50** alumnos en un test, calcula

Intervalo	F. Absoluta	F. Acumulada
[60, 64[5	5
[65, 69[5	10
[70, 74[8	18
[75, 79[12	30
[80, 84[16	46
[85, 89[4	50

- a. P_3
- b. P_{90}
- c. Q_1
- d. Q_3
- e. Interpreta los resultados obtenidos.

3. ¿Qué significa que un alumno haya obtenido un puntaje superior al noveno decil D_9 en un cuestionario de intereses científicos?
4. Analiza el siguiente cuadro que muestra la evolución de la distribución del ingreso per cápita entre 1987 y 1998 según quintiles (divide a la muestra en 5 partes iguales).

Quintil	Años					
	1987	1990	1992	1994	1996	1998
I	4,3	4,4	4,6	4,3	4,1	4,1
II	7,9	8,2	8,5	8,2	8,2	8,2
III	11,7	12,3	12,2	12	11,9	11,8
IV	19	18,1	18,4	18,5	19,1	19,1
V	57,2	56,9	56,3	56,9	56,7	56,9
Total	100	100	100	100	100	100

Fuente: MIDEPLAN, encuesta CASEN.

- a. Investiga sobre el monto de ingresos per cápita en los años que indica el cuadro y establece los valores por año y quintil.
- b. Establece el significado de los quintiles y su aporte como complemento a la media aritmética que es el ingreso per cápita.

Diagrama de cajas

El **diagrama de cajas** consiste en un gráfico que muestra simultáneamente diferentes elementos de la distribución de una o más muestras, por ejemplo, mediana, rango, cuartiles, deciles, etc. La principal utilidad de este tipo de gráficos es la comparación de distribuciones de diferentes grupos.

TIPS

A este tipo de gráfico se le llama también "cajón con bigotes".

Ejemplo

Los siguientes datos corresponden a la masa (en kg) de 24 mujeres de 17 años.

44 48 48 48 48 50 50 51
 52 52 54 54 54 55 55 55
 55 57 57 57 57 58 60 61

Al analizar estos datos podemos obtener lo siguiente:

Tamaño muestra	Mediana	Cuartiles	Valor mínimo	Valor máximo	Rango
24	54	Q ₁ = 50 Q ₂ = 54 Q ₃ = 57	44	61	17

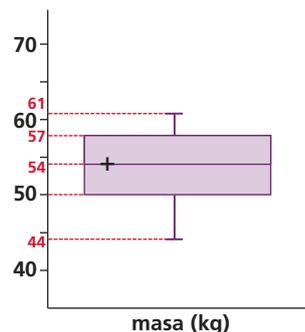
AYUDA

Observa que $Q_1 = 50$ indica que el 25% pesó menos de 50 kilos o igual; Q_2 que el 50% pesó menos o igual que 54 kilos y Q_3 que el 75% pesó menos de 57 kilos o igual.

Visualizaremos todos los elementos anteriores mediante el siguiente diagrama de caja.

Observa que en el gráfico, los extremos del rectángulo indican los cuartiles Q_1 y Q_3 , mientras que la línea que divide a este horizontalmente indica la mediana (Q_2).

Las líneas que sobresalen del rectángulo, indican el valor mínimo y máximo de la distribución, y el signo + indica la media aritmética.



AYUDA

En un gráfico de cajas se pueden expresar los datos de manera vertical u horizontal.

EJERCICIOS

1. La siguiente tabla muestra la tasa de desocupación, correspondiente a los meses de abril, mayo y junio del 2005, según el Boletín Informativo del Instituto Nacional de Estadísticas.

Actividad	Tasa desocupación
Agricultura, caza, pesca	685,01
Minas y canteras	73,07
Industria manufacturera	798,13
Electricidad, gas y agua	30,69
Construcción	451,36
Comercio	1.122,93

- Calcula Q_1 , Q_3 y la mediana.
- Construye el diagrama de caja correspondiente.
- ¿A qué crees que se debe la diferencia entre la tasa de desocupación de cada actividad?
- Si la tasa de desocupación de servicios financieros es de 510,32, ¿a qué cuartil corresponde?

Muestras al azar



En determinadas ocasiones se debe obtener el número de elementos que tiene una cierta población. Para este fin, se toma una muestra, se marca y se devuelve a la población originaria. Se vuelve a tomar una segunda muestra, y con los elementos marcados de esta muestra, se forma una razón con su total, entregando así un total aproximado del tamaño de la población.

Ejemplo

Un grupo de científicos llegó al parque nacional Pan de Azúcar a estudiar la fauna del lugar. Observaron una gran colonia de pingüinos Humboldt, para calcular la cantidad total siguieron el siguiente procedimiento: Durante 4 días, en diversos lugares del parque, capturaron 120 pingüinos, a los cuales marcaron con una cinta: a la semana, en diversos sitios del parque, capturaron 160 pingüinos, de los cuales 30 estaban marcados. Con esta simple proporción obtuvieron la cantidad aproximada de pingüinos en la isla.

$$\frac{30}{160} = \frac{120}{N} \Rightarrow 160 \cdot \frac{120}{30} = 640$$

TIPS



El término **muestreo** es el nombre que recibe la forma de seleccionar a un individuo de la población, para una muestra. Algunas técnicas de muestreo son: muestreo aleatorio, muestreo sistemático, muestreo estratificado, entre otros.



PARA ARCHIVAR

El **tamaño de la población** se calcula de la siguiente manera: $\frac{m}{n_2} = \frac{n_1}{N}$.

Donde, n_1 : tamaño de la primera muestra.

n_2 : tamaño de la segunda muestra.

m : número de individuos marcados en la segunda muestra.

N : tamaño de la población.

Muestras representativas

El cálculo anterior no es más que una estimación de la cantidad de población, ya que dependerá de lo **representativa** que sea la muestra escogida. La estimación en la práctica es muy difícil, por esta razón se toman varias muestras para mejorarla.

Para que la muestra sea representativa se deben considerar varios aspectos; uno de ellos es el **tamaño de la muestra**, mientras mayor sea su tamaño mayor será su confiabilidad, pero a su vez más costoso será el estudio. Otro aspecto se relaciona con que todos los integrantes de la población tengan la misma probabilidad de ser seleccionados en la muestra, por este motivo la selección debe ser al azar, es decir una **muestra aleatoria**.

Las muestras, al igual que las poblaciones, nos permiten calcular parámetros estadísticos como la media, la desviación estándar, etc.; para diferenciarlos usaremos \bar{x} y s , respectivamente, en el caso de la muestra, μ y σ en el caso de la población.

AYUDA



Recuerda:

\bar{x} : media muestral.

s : desviación estándar muestral.

μ : media poblacional.

σ : desviación estándar poblacional.

Nivel de confianza

Si se desea conocer la media aritmética de una población, se puede obtener un intervalo, que con cierto **nivel de confianza**, pueda asegurar que esta se encuentra dentro de un intervalo.



PARA ARCHIVAR

Llamaremos **intervalo de confianza** al intervalo que, con cierto nivel de confianza, contiene el valor del parámetro que se está estimando. El **nivel de confianza** es la "probabilidad" de que el intervalo calculado contenga al verdadero valor del parámetro.

La media poblacional está dada por el siguiente intervalo de confianza:

$$\bar{x} \pm k \frac{s}{\sqrt{n}}, \text{ donde } \begin{array}{l} \bar{x}: \text{media muestral} \\ k: \text{coeficiente asociado al nivel de confianza} \\ s: \text{desviación estándar de la muestra} \\ n: \text{número de elementos de la muestra} \end{array}$$

AYUDA



La estimación por intervalos, es más útil, ya que se calculan dos valores, entre los que se encontrará el parámetro, con un nivel de confianza fijado de antemano.

AYUDA



Un parámetro es una característica numérica de una población. Equivale a una constante fija para cada estudio particular.

Margen de error

El margen de error depende del nivel de confianza y del tamaño de la muestra.



PARA ARCHIVAR

Al estimar la media poblacional a partir de una muestra, el **margen de error** está dado por: $[\bar{x} - E, \bar{x} + E]$ siendo $E = k \cdot \frac{s}{\sqrt{n}}$ (error).

Ejemplo

Un grupo de médicos de distintos hospitales desea saber cuánto tiempo permanecen hospitalizados los pacientes con problemas cardíacos. Extraen una muestra de **80** pacientes obteniendo una media muestral de **2,5** días; ellos sabían que la desviación típica era de **4** días. Si el nivel de confianza es de un **95%**, ¿cuál es el intervalo?

$$2,5 \pm 1,96 \cdot \frac{4}{\sqrt{80}} \Rightarrow [1,62 ; 3,38] \text{ con un } 95\% \text{ de confianza.}$$

Por lo tanto, la cantidad de días que permanecerán los pacientes, será aproximadamente entre los valores dados en el intervalo.

AYUDA



El coeficiente k se obtiene de la siguiente tabla.

Nivel de confianza	Coficiente k
68%	0,99
75%	1,15
80%	1,28
90%	1,64
95%	1,96
96%	2,05
97%	2,17
98%	2,32
99%	2,58

Tamaño de la muestra

TIPS



El error porcentual está dado por $\frac{E}{\bar{x}} \cdot 100$.

El tamaño de la muestra está dado por el número de sujetos que componen la muestra extraída de una población.



PARA ARCHIVAR

El tamaño de la muestra se calcula de la siguiente forma:

$$n = \left(\frac{k \cdot \sigma}{E} \right)^2$$

donde, **k**: nivel de confianza
 σ : desviación estándar de la población
E: margen de error

Ejemplo

En un colegio de **1.600** alumnos se está estudiando la relación entre la estatura de los niños al nacer y otras variables. Se sabe que la desviación típica poblacional es de **1,5 cm** y se desea estimar la media con un **99%** de confianza y con un error máximo de **0,5 cm**.

$$E = k \cdot \frac{\sigma}{\sqrt{n}} \rightarrow 0,5 = 2,58 \cdot \frac{1,5}{\sqrt{n}}$$

$$\sqrt{n} = \frac{k \cdot \sigma}{E} \rightarrow \sqrt{n} = \frac{2,58 \cdot 1,5}{0,5} = 7,74$$

$$n = \left(\frac{k \cdot \sigma}{E} \right)^2 \rightarrow n = 59,9076 \approx 60$$

Se debe tomar al menos una muestra de **60** alumnos.

EJERCICIOS

1. Para estimar la cantidad de salmones en un lago se realizó lo siguiente:

- I. Se capturó una muestra al azar, se les marcó y fueron devueltos al agua.
 - II. Breve tiempo después, se capturó una nueva muestra, se registró la proporción de salmones marcados versus el total de salmones de la muestra.
- a. Si en el primer proceso se capturan y marcan **100** salmones. Posteriormente, **80** salmones, de los cuales **20** están marcados, ¿cuántos salmones hay aproximadamente en el lago?

2. Trabajo experimental: Se dispone de una bolsa con **100** fichas numeradas y distribuidas como lo indica la tabla.

Nº de fichas	1	2	3	4	5	6	7	8	9	10
Cantidad	10	10	10	10	10	10	10	10	10	10

- a. Obtén muestras al azar de tamaño **10**, **20** y **30**. Calcula para cada una de ellas la media de los valores de las fichas y su desviación estándar.
- b. Compara los valores de las medias y desviaciones estándar obtenidos para cada muestra de la pregunta **a**.
- c. ¿Qué inferencias puedes sacar a partir de las medias poblaciones anteriores?

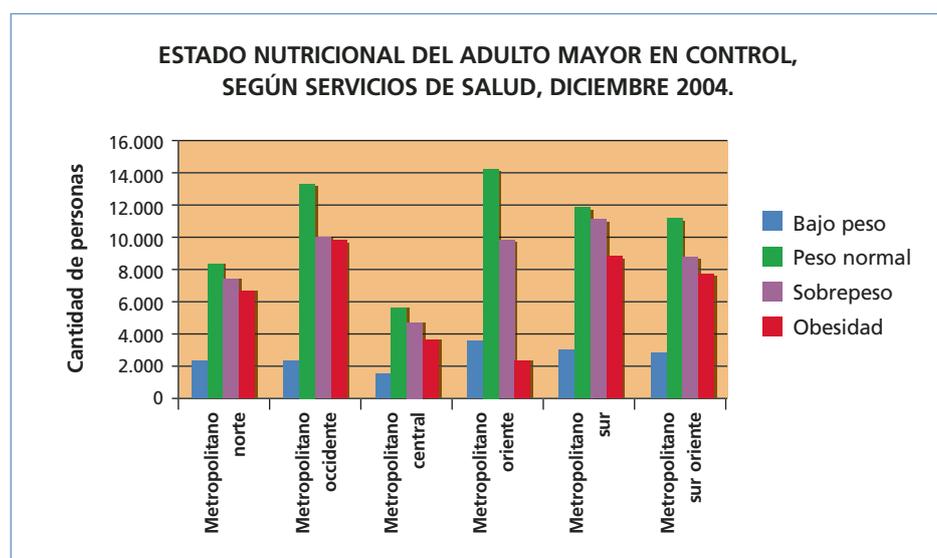
Aplicaciones de la estadística

Ciencias naturales

Los estudios estadísticos, realizados por el Ministerio de Salud, relacionados con el estado nutricional de las personas, tienen como uno de sus objetivos el “supervisar la situación alimentario-nutricional de la población chilena, detectando grupos en riesgo de sufrir alguna forma de malnutrición, y normar la implementación de acciones y programas orientados a prevenir el daño en dichos grupos y en la población general” (www.minsal.cl).

La siguiente tabla y gráfico nos muestran la cantidad de población adulto mayor, de la Región Metropolitana, que se encuentra en algún estado de normalidad o no, en relación a su masa.

	Bajo peso	Peso normal	Sobrepeso	Obesidad
Metropolitano norte	2.220	8.454	7.580	6.878
Metropolitano occidente	2.218	13.356	10.000	9.988
Metropolitano central	1.453	5.760	4.806	4.019
Metropolitano oriente	3.647	14.514	9.433	7.183
Metropolitano sur	2.727	11.827	11.017	8.894
Metropolitano oriente	2.527	10.898	8.926	7.816



Fuente: Estado nutricional del adulto mayor en control, según servicios de salud, (diciembre 2004), www.minsal.cl, julio 2005.

En los últimos 40 años, el grupo llamado “adulto mayor” ha crecido más de un 25%, llegando a representar más del 10% de la población total. Por este motivo, el país ha creado nuevas políticas, con el fin de mejorar su calidad de vida.

¿Qué puedes concluir?

De la población estudiada, observamos que en toda la Región Metropolitana la población se ordena en: personas con peso normal, sobrepeso, obesidad y bajo peso. Por otro lado, la cantidad de población con sobrepeso y obesidad, supera en demasía a la población con peso normal. Siendo, el rango de bajo peso, el que considera la menor porción de población adulta mayor.

Al igual que las ciencias naturales, la estadística es aplicable a otros campos, tales como las ciencias humanas, económicas, entre otras.

Ciencias humanas

Tasa de asistentes (Por cien mil habitantes) Cine, Teatro, Recitales, Conciertos y Espectáculos Deportivos 1989-1999 (promedio mensual)					
Año	Conciertos	Espectáculos deportivos	Cine	Teatro	Recitales
1989	110	6.957	9.258	195	246
1990	122	7.034	7.257	150	275
1991	113	7.524	6.242	158	191
1992	127	8.816	5.189	246	333
1993	203	8.140	4.856	179	273
1994	159	7.807	4.262	221	331
1995	145	5.467	4.403	175	403
1996	156	7.483	4.054	247	306
1997	225	7.145	5.039	354	440
1998	230	6.300	6.198	352	373
1999	308	5.885	7.739	407	512

Fuente: Anuario de Cultura y Medios de Comunicación 1989-1998 y Datos preliminares 1999.

En la tabla se observa que la asistencia de los chilenos a espectáculos masivos ha cambiado en los 10 últimos años. Se aprecia claramente un aumento en eventos de orden artístico cultural, como son el teatro, los conciertos y los recitales; un resurgimiento del cine, y una leve baja en los espectáculos deportivos. El cine bajó mucho durante la década de los 90 debido al auge de los video club y cuando parecía que este iba a ser un espectáculo cada vez menos masivo, hubo un resurgimiento producto de un cambio en el concepto del cine.

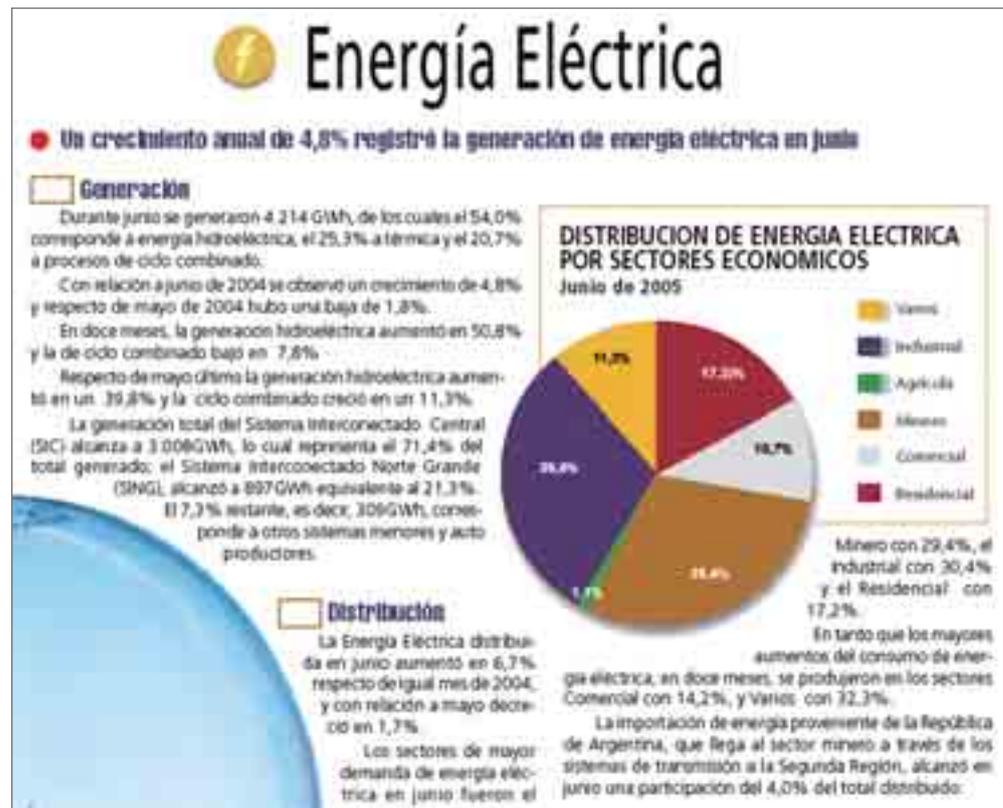
EN EQUIPO

Según la información de ciencias humanas, contesten las siguientes preguntas:

- ¿Cuál de los espectáculos tuvo mayor aumento en estos años?
- De mantenerse el ritmo de aumento de asistencia al cine, ¿en qué año se superará la asistencia de 1989? ¿la asistencia a este tipo de espectáculos, sería superior a los años señalados en la tabla? Fundamenten.
- Si hicieran una encuesta este año, ¿la asistencia a este tipo de espectáculos sería superior a los años señalados en la tabla? Fundamenten.

En vez de tener grandes salas para mostrar una gran película, ahora se tienen muchas salas pequeñas con gran variedad de películas, más un ambiente acogedor y venta de chocolates, cabritas, bebidas, etc. El cine volvió a ser atractivo, pues es una alternativa interesante y entretenida que permite desconectarse de los deberes del hogar, cosa que no se logra con el video.

Ciencias económicas



Fuente: indicadores del mes de INE, Empleo y sectoriales. Boletín N° 81 de junio de 2005
Distribución de energía eléctrica por sectores económicos. Junio 2005.

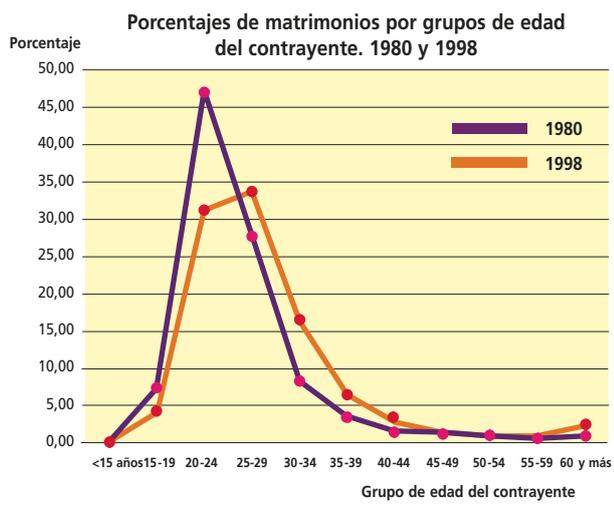
EN EQUIPO

¿A qué causas atribuyen ustedes que el área agrícola utilice menor cantidad de energía eléctrica?

EJERCICIOS

1. En la tabla y en el gráfico se aprecia la evolución de la relación edad/matrimonio en los últimos años.

Edad media al matrimonio por sexo de los contrayentes 1980-1998		
Año	Hombre	Mujer
1980	26,6	23,8
1981	26,7	23,9
1982	27,0	24,3
1983	26,9	24,2
1984	27,0	24,3
1985	27,0	24,3
1986	27,0	24,4
1987	27,1	24,6
1988	27,2	24,7
1989	27,2	24,7
1990	27,5	25,0
1991	27,8	25,2
1992	27,9	25,3
1993	27,7	25,2
1994	27,9	25,4
1995	28,0	25,5
1996	28,3	25,8
1997	28,5	26,0
1998	28,9	26,3



Fuente: INE. Anuario de Demografía. Serie 1980-1988.

- a. ¿Qué conclusiones se obtienen con respecto a las edades en que se casan las personas?
- b. ¿Cómo es la distribución de las edades en que se casa la gente? ¿por qué?
2. En una provincia se desea establecer la media de los sueldos, con un 99% de confianza y con un error máximo de \$ 15.000. Si se sabe que la desviación estándar es de \$ 100.000, ¿de qué tamaño debe ser la muestra?

3. El IPC se calcula sobre la base de un promedio ponderado, de modo que cada rubro tiene distinta importancia de acuerdo a los consumos de la población.

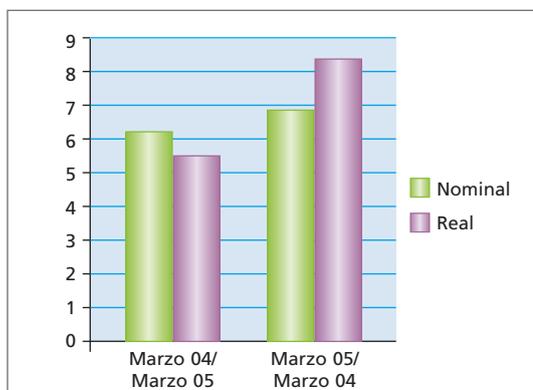
- a. Observa la tabla y determina cuál es el rubro que tiene mayor importancia.
- b. ¿Por qué crees tú que los rubros de salud y educación son los que más subieron?
- c. ¿Por qué crees tú que el rubro vestuario es el que más bajó?

IPC 2001 Variaciones e incidencias anuales		
Grupos	Variación anual %	Incidencia anual
Alimentación	2,0	0,52
Vivienda	3,5	0,73
Equipamiento de la Vivienda	0,1	0,00
Vestuario	-4,8	-0,31
Transporte	4,4	0,65
Salud	4,9	0,47
Educación y Recreación	4,9	0,56
Otros	0,0	0,00
Índice General	2,6	2,64

Indicadores del mes precios y remuneraciones del INE, Boletín N° 38, www.ine.cl, enero 2002.

4. El gráfico muestra la variación de las ventas de marzo de 2004 comparado con noviembre del 2005.

- a. ¿Por qué crees tú que se compara con el mismo mes del año anterior?
- b. Averigua qué diferencias hay entre los indicadores nominal y real.



Fuente: www.ine.cl, julio 2005.

HISTORIA



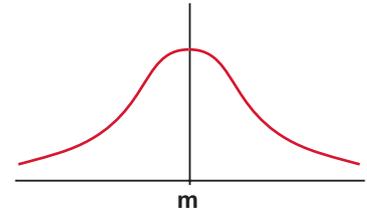
Abraham de Moivre
(1667-1754)

Matemático francés exiliado en Londres, donde publicó en 1733 una obra en la que aparece por primera vez la curva de distribución de los errores, que con el tiempo conocemos como distribución normal de Gauss.

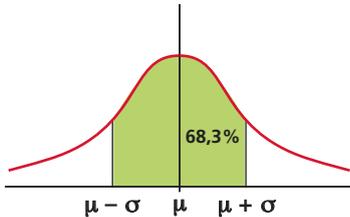
Distribución normal

→ La **distribución normal**, una de las más importantes, recibe su nombre debido a que en cierto momento se pensó que la mayoría de los fenómenos estaban distribuidos de dicha manera. Esta distribución nos permite representar fenómenos estadísticos de manera probabilística.

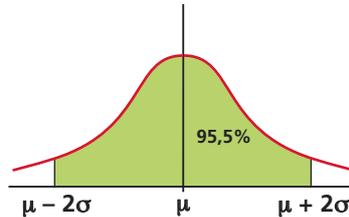
Cuando una variable continua tiene distribución normal, su gráfico es similar al indicado. Como se observa, tiene forma de campana (conocida como **campana de Gauss**) y es simétrico con respecto a la media, además presenta pocos valores extremos. Además, se sabe que si una población tiene media μ y desviación típica σ , se cumple lo siguiente:



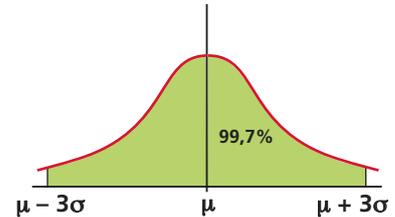
a. El **68,3%** de los individuos se encuentran en el intervalo $[\mu - \sigma, \mu + \sigma]$.



b. El **95,5%** de los individuos se encuentran en el intervalo $[\mu - 2\sigma, \mu + 2\sigma]$.



c. El **99,7%** de los individuos se encuentran en el intervalo $[\mu - 3\sigma, \mu + 3\sigma]$.



AYUDA

Si una población tiene distribución normal con media μ y desviación típica σ , anotamos que ella distribuye $N(\mu, \sigma)$.



PARA ARCHIVAR

La **distribución normal** describe la distribución de datos, que en general se relacionan con mediciones relacionadas con variables, tales como, el tamaño de las especies, rendimiento intelectual, variables sociales, etc.

EJERCICIOS

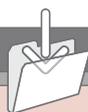
1. Determina en cuáles de los siguientes casos se trata de una población con distribución normal.

- a. Sueldos que se pagan en una empresa.
- b. Edad a la que una persona muere.

2. De un colegio mixto egresaron 210 varones y 225 damas. Las edades de los varones se dis-

tribuyen $N(18,8; 0,4)$ y las de las damas, $N(18,2; 0,6)$.

- a. ¿Cuántos varones tenían más de 18 años?
- b. ¿Cuántas damas tenían más de 17 años?
- c. Si se selecciona un alumno al azar, ¿cuál es la probabilidad de que tenga a lo menos 18,8 años?



PARA ARCHIVAR

Se puede demostrar que si x es una variable que se distribuye $N(\mu, \sigma)$, utilizando la variable $Z = \frac{x - \mu}{\sigma}$, distribuirá $N(0, 1)$.

A este procedimiento se le conoce como **tipificación**.

La ventaja de la variable Z es que existen valores tabulados para ella (ver ayuda), de modo que se pueden hacer cálculos de probabilidades y tamaños de grupos de población con solo usar correctamente la tabla, y luego hacer los cálculos correspondientes.

HISTORIA



C. Friedrich Gauss
(1777–1855)

Matemático alemán llamado el "príncipe de las matemáticas". Entre sus contribuciones destacan la demostración del teorema fundamental de álgebra y el descubrimiento de la distribución normal.

Ejemplo

El resultado de una prueba de cuarto medio, tiene una distribución $N(5,3 ; 0,6)$. El total de estudiantes que rindió la prueba es de **150**. ¿Cuál es la probabilidad de que al escoger un estudiante al azar este haya obtenido al menos un **6,0**?

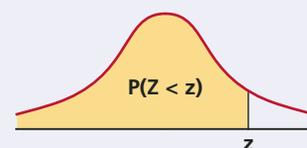
Calcularemos la probabilidad de que un alumno tenga menos de un **6,0**; para facilitar el uso de la tabla, el complemento será lo buscado.

$$z = \frac{6,0 - 5,3}{0,6} = \frac{0,7}{0,6} = 1,16 \approx 1,2$$

En la tabla, **1,2** corresponde a **0,8849**; por lo tanto, $1 - 0,8849 = 0,1151$ (probabilidad de obtener un alumno con nota igual o superior a **6,0**, o bien el **11,51%** de los alumnos obtuvieron una nota perteneciente a ese intervalo).

AYUDA

z	P(Z < z)	z	P(Z < z)	z	P(Z < z)	z	P(Z < z)
-3,0	0,0013	-1,4	0,0808	0,2	0,5793	1,8	0,9641
-2,9	0,0019	-1,3	0,0968	0,3	0,6179	1,9	0,9713
-2,8	0,0026	-1,2	0,1151	0,4	0,6554	2,0	0,9772
-2,7	0,0035	-1,1	0,1357	0,5	0,6915	2,1	0,9821
-2,6	0,0047	-1,0	0,1587	0,6	0,7257	2,2	0,9861
-2,5	0,0062	-0,9	0,1841	0,7	0,7580	2,3	0,9893
-2,4	0,0082	-0,8	0,2119	0,8	0,7881	2,4	0,9918
-2,3	0,0107	-0,7	0,2420	0,9	0,8159	2,5	0,9938
-2,2	0,0139	-0,6	0,2743	1,0	0,8413	2,6	0,9953
-2,1	0,0179	-0,5	0,3085	1,1	0,8643	2,7	0,9965
-2,0	0,0228	-0,4	0,3446	1,2	0,8849	2,8	0,9974
-1,9	0,0287	-0,3	0,3821	1,3	0,9032	2,9	0,9981
-1,8	0,0359	-0,2	0,4207	1,4	0,9192	3,0	0,9987
-1,7	0,0446	-0,1	0,4602	1,5	0,9332	3,1	0,9990
-1,6	0,0548	0,0	0,5000	1,6	0,9452	3,2	0,9993
-1,5	0,0668	0,1	0,5398	1,7	0,9554	3,3	0,9995



EJERCICIOS

- Utilizando los datos dados en el ejemplo anterior, determina cuántos alumnos reprobaron.
- Las estaturas de los recién nacidos en un hospital distribuyen $N(46, 2)$, en cm. Calcula la probabilidad de que:
 - Un bebé mida menos de **44 cm**.
 - Un bebé mida más de **50 cm**.
- Se ha calculado que los gastos de los jóvenes en un fin de semana tienen una distribución $N(8.500, 5.700)$, en pesos.
 - ¿Cuál es la probabilidad de que un joven gaste más de \$ **20.000**?
 - ¿Cuál es la probabilidad de que un joven gaste entre \$ **5.000** y \$ **10.000**?
 - Si un joven invita a su pareja, ¿cuál es la probabilidad de que gaste menos de \$ **25.000**?

Ejercicio 1

Jorge obtuvo un **5,4** en biología y un **5,7** en física. Si los promedios en ambas asignaturas fueron **4,8** y **5,0** y las desviaciones estándar **0,6** y **0,8**, respectivamente, ¿en qué asignatura obtuvo un lugar relativo mejor?

Solución

Los datos entregados, por cada asignatura son:
 nota cualesquiera: **5,4** y **5,7**
 media aritmética: **4,8** y **5,0**
 desviación estándar: **0,6** y **0,8**

El valor tipificado se encuentra a través de la expresión: $z = \frac{x - \mu}{\sigma}$

Como se quiere conocer en cuál de las asignaturas Jorge tuvo un rendimiento relativamente mejor, obtendremos los puntajes tipificados de cada asignatura (o puntajes z).

Biología

$$z = \frac{5,4 - 4,8}{0,6} \Rightarrow z = 1 \quad \Rightarrow \text{valor tabulado correspondiente: } \mathbf{0,8413}$$

$$\Rightarrow \text{equivale aproximadamente a } \mathbf{84\%}$$

Física

$$z = \frac{5,7 - 5,0}{0,8} \Rightarrow z = \mathbf{0,875} \quad \Rightarrow \text{valor tabulado correspondiente: } \mathbf{0,7881}$$

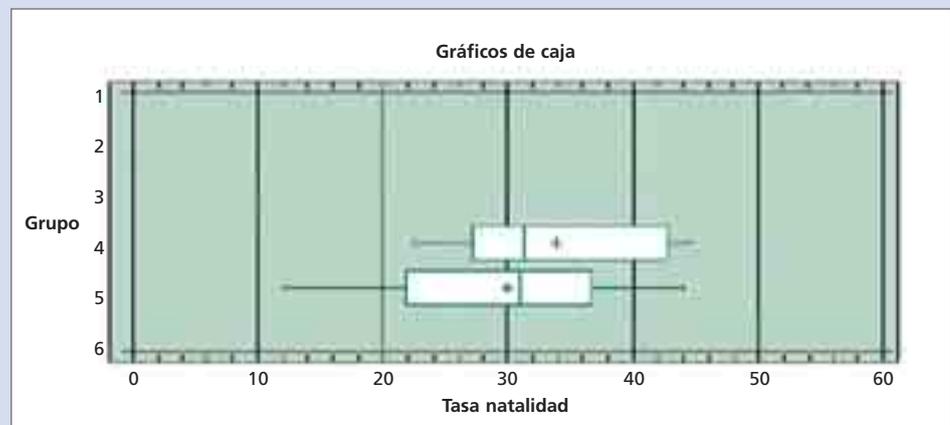
$$\Rightarrow \text{equivale aproximadamente a } \mathbf{78\%}$$

Dicho de otra forma: Jorge obtuvo un mejor rendimiento en biología ya que $1 \cdot s$ está por encima de \bar{x} , mientras que en física está solamente $0,875 \cdot s$ sobre \bar{x} , aún cuando obtuvo nota más alta en física.

En relación a la media \bar{x} , Jorge obtuvo un mejor rendimiento en biología que en física, ya que, su puntaje z, está por encima de la media \bar{x} .

Ejercicio 2

El siguiente gráfico corresponde a las tasas de natalidad de ciertos países de Oriente Medio (grupo 4) y Asia (grupo 5).



Los datos, de **tasa de natalidad**, utilizados para construir el gráfico anterior corresponden a:

grupo Oriente Medio:	42,6	22,3	26,8	29,2	38,9	42,1
	28,4	42,5	22,8	45,6	31,7	
grupo Asia:	28,6	21,2	30,5	21,3	11,7	30,3
	42,2	39,6	41,4	31,6	17,8	31,8
	33,2	23,5	36,1	40,4	22,3	

La tasa de natalidad corresponde a niños nacidos vivos en el año por cada **1.000** habitantes.

- Obtén los cuartiles de cada grupo de datos.
- Encuentra el promedio de cada grupo de datos.
- Encuentra la mediana de cada grupo de datos.
- Obtén el valor máximo y mínimo de cada grupo de datos.
- Con los datos obtenidos y la gráfica dada, ¿qué puedes concluir, en relación a la tasa de natalidad de cada país?

Solución

a.

	Q ₁	Q ₂	Q ₃
Oriente Medio	27,6	31,7	42,3
Asia	22,3	30,5	36,1

El cálculo de los cuartiles se pueden obtener en el programa EXCEL. Recuerda que se debe anotar, = **CUARTIL()**.
Ejemplo:
= **CUARTIL(A1: A6; 2)**.
A1: fila en la cual se ubica el primer dato ordenado.
A6: fila en la cual se ubica el último dato ordenado.
2: segundo cuartil.

b. $\bar{x}_{\text{Oriente Medio}} = \frac{373,9}{11} \approx 33,99$; $\bar{x}_{\text{Asia}} = \frac{503,5}{17} \approx 29,62$

- c. La mediana de cada grupo corresponde al valor del cuartil 2 (Q₂).

Oriente medio: **31,7**
Asia: **30,5**

Para obtener el valor de la mediana hay que ordenar los datos entregados. El valor debe coincidir con el valor del **Q₂**.

d.

	Valor mínimo	Valor máximo	Rango
Oriente Medio	22,3	45,6	23,3
Asia	11,7	42,2	30,5

- e. Como el valor mínimo de Oriente Medio coincide con el valor del Q₁ de Asia, podemos decir que la tasa de natalidad de los países de este último grupo es más baja.

Las tasas de natalidad de los países de Asia se encuentran más dispersas, ya que la distancia del primer al segundo cuartil es bastante más amplia que del segundo al tercero. En cambio la mayor dispersión, en el otro grupo, se encuentra entre el segundo y tercer cuartil, confirmando de esta manera que los países de Oriente Medio tienen una menor tasa de natalidad.

1. (Ensayo PSU, 2004) Alberto, Sebastián y Carlos juegan a lanzar un dado 2 veces y gana el que obtiene una suma par. En el primer lanzamiento Alberto obtiene un 2, Sebastián un 3 y Carlos un 6. ¿Cuál de las siguientes afirmaciones es verdadera?

- A. Todos tienen probabilidad $\frac{1}{2}$ de ganar.
- B. Todos tienen probabilidad $\frac{1}{3}$ de ganar.
- C. El que tiene más probabilidad de ganar es Carlos.
- D. Carlos tiene más probabilidad de ganar que Alberto.
- E. Sebastián tiene menos probabilidad de ganar que Alberto y Carlos.

2. (Ensayo PSU, 2004) La tabla muestra las edades de 220 alumnos de un colegio. ¿Cuál(es) de las siguientes afirmaciones es(son) verdadera(s)?

- I) La moda es 17 años.
- II) La mediana es mayor que la media (promedio).
- III) La mitad de los alumnos del colegio tiene 17 o 19 años.

Edad(en años)	15	16	17	18	19
Alumnos	50	40	60	50	20

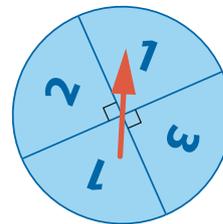
- A. Solo I
- B. I y II
- C. I y III
- D. II y III
- E. I, II y III

3. (Facsimil PSU, Demre, 2004) En un supermercado el precio de costo de un kilogramo de pan es de \$ 600 y lo venden en \$ 820; las conservas de mariscos tienen un costo de \$ 800 y las venden en \$ 1.060. Si la política de asignación de precios del supermercado es lineal, ¿cuál es el precio de venta de un kilogramo de arroz cuyo costo es \$ 400?

- A. \$ 600
- B. \$ 580
- C. \$ 547
- D. \$ 537
- E. \$ 530

4. (Pisa, 2003) En la figura, se tiene una ruleta en que la flecha puede indicar cualesquiera de los 4 sectores y ella nunca cae en los límites de dichos sectores. ¿Cuál(es) de las siguientes proposiciones es(son) verdadera(s)?

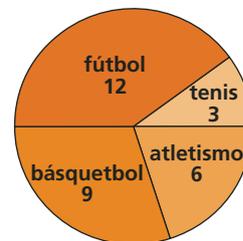
- I) La probabilidad de que la flecha apunte al número 1 es de $\frac{1}{2}$.
- II) La probabilidad de que la flecha apunte al número 2 es de $\frac{1}{4}$.
- III) La probabilidad de que la flecha apunte al número 2 o al 3 es de $\frac{2}{3}$.



- A. Solo I
- B. Solo II
- C. Solo III
- D. I y II
- E. Todas

5. (Facsimil PSU, Demre, 2004) El gráfico de la figura muestra las preferencias de 30 personas en actividades deportivas. ¿Cuál(es) de las siguientes afirmaciones es(son) correctas(s)?

- I) La frecuencia relativa del grupo de fútbol es de 40%.
- II) La frecuencia relativa del grupo de básquetbol es de 30%.
- III) La mitad del grupo no prefirió fútbol ni tenis.



- A. Solo I
- B. Solo II
- C. Solo III
- D. I y II
- E. Todas

¿Cuántas personas tendrán un accidente mañana?

A partir de la información diaria relativa a los accidentes del tránsito, se puede construir estadísticas que permitan inferir de manera aproximada la cantidad de muertos en futuros accidentes. Por ejemplo, se espera que en las épocas de fiestas patrias mueran, aunque no quisieramos, cierta cantidad de personas.

Carabineros de Chile, en su sitio web pública a diario estadísticas sobre accidentes en la que también se indican las causantes de ellos. (ver www.carabinerosdechile.cl)



1. Construye en una planilla de cálculo distintos tipos de gráficos asociados a estos datos.
2. Actualiza estos datos según el día actual.
3. ¿En qué medida ayuda conocer las estadísticas de accidentes del tránsito?
4. Quizás muy pronto tendrás la oportunidad de obtener tu licencia de conducir, por lo que sería muy bueno averiguar las causas más importantes detectadas en los accidentes. Construye un gráfico con la información que obtengas.



Mapa conceptual

Conceptos clave:

Media

Moda

Mediana

Rango

Desviación media

Desviación estándar

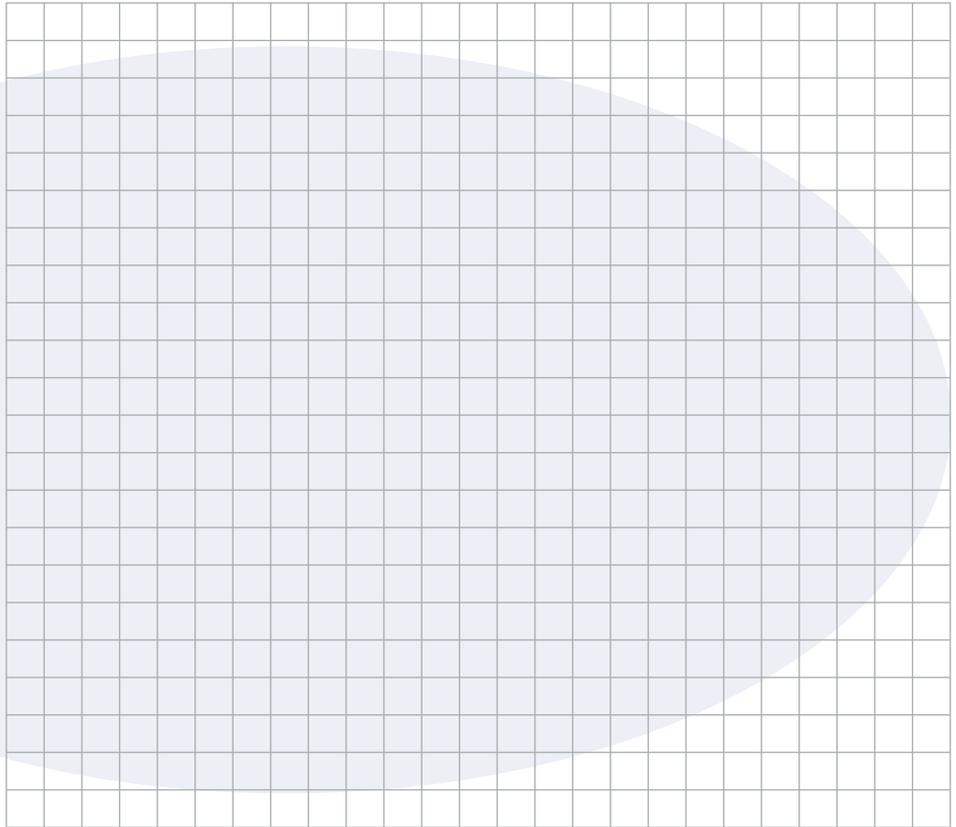
Correlación

Muestra

Error porcentual

Intervalo de confianza

Construye tu mapa conceptual que relacione **al menos** los conceptos clave dados.



Resumen

- 1 Media aritmética:** cociente entre la suma de todos los valores de un conjunto de datos y la frecuencia total de estos. Está dada por la expresión:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + x_4 + \dots + x_n}{n} = \frac{\sum_{k=1}^{k=n} x_k}{n}$$

- 2 Mediana:** dato que ocupa el valor central de un conjunto de datos ordenados según magnitud (decreciente o creciente).
- 3 Moda:** la moda de un conjunto de datos es el valor que presenta mayor frecuencia.
- 4 Rango:** diferencia entre el mayor valor y menor valor de una distribución de datos.
- 5 Desviación:** representa el mayor o menor alejamiento de un dato respecto a la media aritmética, para calcularla, utilizamos la fórmula:

$$d = x - \bar{x}$$

- 6 Desviación respecto a la media:** media aritmética de las desviaciones absolutas respecto de la media. Se calcula utilizando la fórmula:

$$DM = \frac{\sum_{k=1}^n |x_k - \bar{x}|}{n}$$

(para datos no agrupados)

o

$$DM = \frac{\sum f \cdot |x - \bar{x}|}{\sum f}$$

(para datos agrupados)

- 7 Desviación estándar o típica:** representa el grado de dispersión de los datos respecto de la media, la calculamos utilizando la expresión:

$$s = \sqrt{\frac{\sum_{k=1}^n (x_k - \bar{x})^2}{n}}$$

(para datos no agrupados)

o

$$s = \sqrt{\frac{\sum f \cdot (x - \bar{x})^2}{n}}$$

(para datos agrupados)

- 8 Correlación:** indica el grado de asociación de dos variables, esta puede ser positiva, negativa o nula.

- 9 Tamaño de la población:** número de individuos que pertenecen a una cierta población, se calcula mediante la proporción: $\frac{m}{n_2} = \frac{n_1}{N}$,

(n_1 : tamaño de la primera muestra; n_2 : tamaño de la segunda muestra; m : número de marcados en segunda muestra; N : tamaño de la población).

- 10 Tamaño muestral:** se calcula utilizando la expresión: $n = \left(\frac{k \cdot \sigma}{E} \right)^2$

(k : nivel de confianza; σ : desviación estándar de la población; E : margen de error).

- 11 Intervalo de confianza:** intervalo que con cierto nivel de confianza, nos asegura que dentro de él se encuentra la media poblacional.

- 12 Margen de error:** está dado por $[\bar{x} - E, \bar{x} + E]$ (con $E = k \cdot \frac{s}{\sqrt{n}}$;

s : desviación estándar; k : coeficiente asociado al nivel de confianza; n : número de elementos de la muestra).

El error puede ser expresado de manera porcentual dado por $\frac{E}{\bar{x}} \cdot 100$.

- 13 Medidas de localización:** dividen a una distribución de datos en una cierta cantidad de partes iguales, los más conocidos son cuartiles (cuatro partes iguales), deciles (diez partes iguales) y percentiles (cien partes iguales).

1. Las edades de los jóvenes de un grupo musical son **15, 14, 13, 15, 14 y 13** años. Entonces, es **verdadero** que:

- I) la media es **14** años.
- II) la mediana es **15** años.
- III) la desviación típica es $\sqrt{\frac{2}{3}}$ años.

- A. Solo I
- B. Solo II
- C. Solo III
- D. I y II
- E. I y III

2. La tabla muestra las edades de los jóvenes de un grupo de una parroquia. Con respecto a la información de la tabla, es **falso** que:

Edad	fi
14	6
15	8
16	12
17	6
Total	32

- A. el **25%** tiene **15** años.
- B. la moda es **16** años.
- C. la media es alrededor de **15** años.
- D. el **35,7%** tiene **16** años.
- E. la mediana es **16** años.

3. Las notas de Claudia en Física son: **3,5; 4,2; 5,3; 2,8; 5,6 y 5,6**. Con respecto a esta situación, es **verdadero** que:

- I) su media es **4,5**.
- II) la moda es un **5,6**.
- III) si Claudia obtiene en un trabajo un **6,5** y lo reemplaza por su peor nota, su media ahora es un **5,1**.

- A. Solo I
- B. Solo II
- C. II y III
- D. I y II
- E. Todas.

4. Antonia lleva un **5,5** de promedio con **4** notas en Física y debe rendir la Prueba Global. Con respecto a esta situación, es **verdadero** que:

- I) si se saca un **7,0** en la prueba global su promedio sube a **6,0**.
- II) si se saca un **4,0** en la prueba global su promedio baja a **5,0**.
- III) No se puede sacar el promedio si no se conocen las otras notas.

- A. Solo I
- B. Solo II
- C. Solo III
- D. I y II
- E. Ninguna.

5. En la selección de voleibol de un colegio **A**, la media de las estaturas es **183 cm** y la desviación típica **3,5 cm**. En otro colegio **B**, la media es **174 cm** y la desviación típica es **5 cm**. Entonces:

- I) los seleccionados de **B** tienen una estatura más pareja que en **A**.
- II) los seleccionados más altos están en **A**.
- III) los seleccionados más bajos están en **B**.

- A. Solo I
- B. Solo II
- C. Solo III
- D. I y II
- E. II y III

6. Con respecto al coeficiente de correlación de Pearson es **verdadero** que:

- I) Cuando su valor es cercano a **1**, hay correlación positiva.
- II) Cuando su valor es cercano a **0,5**, la correlación es nula.
- III) Cuando su valor es **0**, la correlación es negativa.

- A. Solo I
- B. Solo II
- C. Solo III
- D. II y III
- E. Todas

7. En un zoológico desean saber cuántos loros hay. Escogen una muestra de **50** loros y los marcan; al día siguiente toman una muestra de **40** y observan que **5** de ellos están marcados. El total aproximado de loros del zoológico es:
- A. 100 C. 400 E. 350
B. 250 D. 200
8. Cecilia, en su preparación para la PSU de lenguaje, realizó **10** ensayos y su tiempo promedio fue de una hora y media. Ella sabe que su desviación típica es de **20** minutos. Si se asume un nivel de confianza del **95%**, el error máximo en tiempo, el día que rinda la prueba, será aproximadamente:
- A. 0,14 minutos.
B. 12 minutos.
C. 13 minutos.
D. 3,92 minutos.
E. Ninguna de las anteriores.
9. En un colegio de **4.000** alumnos, las notas en matemáticas se distribuyen $N(5,2; 0,6)$. ¿Alrededor de cuántos alumnos tienen promedio sobre **6,0**?
- A. 0,9032 C. 10% E. 500
B. 0,0968 D. 390
10. Un consultorio realizó un estudio para determinar la masa de la población femenina de su comuna obteniendo una distribución $N(62, 5)$. ¿Alrededor de qué porcentaje de la cantidad de mujeres de la comuna tienen una masa entre **57** y **62** kilogramos?
- A. 99% C. 68% E. 24%
B. 95% D. 34%
11. Se desea saber las preferencias musicales de la juventud chilena y para ello se decide hacer una encuesta. ¿Cuál de los siguientes procedimientos asegura una muestra representativa?
- A. Se encuesta a **100** jóvenes en el centro de las principales ciudades.
B. Se encuesta a **2.000** jóvenes a la salida de los liceos de acuerdo a la cantidad de alumnos de cada liceo.
C. Se consigue en el registro civil una lista de todos los jóvenes del país y se seleccionan **2.500** al azar.
D. Se pide a los jóvenes que den su opinión en una radio de alcance nacional.
E. Se invita a los jóvenes a participar en sus comunas habilitando formularios y buzones.
12. La vida media de una pila (en horas) tiene una distribución $N(150, 50)$. ¿Cuál es la probabilidad (en porcentaje) de que dure menos de **50** horas?
- A. 2% C. 16% E. 68%
B. 4% D. 32%
13. En la selección de personal para un museo de historia, se realizará una prueba de conocimientos básicos de Historia de Chile. Se sabe que los puntajes distribuyen $N(132, 18)$ y tan solo el **10%** de los puntajes más altos será seleccionado. Aproximadamente, ¿desde qué puntaje se aceptará a los candidatos?
- A. 109
B. 155
C. 190
D. No se puede determinar.
E. Ninguna de las anteriores.

- En un curso universitario se sabe que la moda de las edades es de **20** años, la mediana es **21** años, el menor de los alumnos tiene **19** años y el mayor **23** años. Si hubiera uno más que tuviera **22** años, la moda sería ésta. Si en total hay **9** alumnos, construye una tabla de frecuencias con sus edades y calcula la media, la desviación media y la desviación estándar.
- Calcula la media del tiempo de espera en un consultorio de acuerdo a la siguiente tabla.

Tiempo (min)	f_i
[0 – 10[2
[10 – 20[12
[20 – 30[15
[30 – 40[10
[40 – 50[8
[50 – 60[7

- En un curso hay **24** hombres y **16** mujeres. En la tabla se muestra la estatura y la masa promedio.

	Estatura (m)	Masa (kg)
Hombres	1,78	74
Mujeres	1,59	56

Calcula la media de la estatura y de la masa del total del curso.

- Los números que aparecen a continuación corresponden a la cantidad de preguntas omitidas en un ensayo de PSU de un cuarto medio:

6 - 0 - 7 - 15 - 2 - 5 - 36 - 18 - 9 - 3 - 2 - 0 - 1 - 4 -
4 - 6 - 7 - 5 - 8 - 10 - 9 - 0 - 3 - 0 - 2 - 0 - 8 - 9 -
22 - 16 - 0 - 4 - 7 - 0 - 12 - 11 - 0 - 6 - 8 - 0 - 0 - 9

- Calcula la media, la mediana, moda, rango y desviación estándar.
- ¿Qué valores distorsionan la media y no son representativos del curso?

- En las siguientes situaciones indica si la correlación es positiva, negativa o nula. Fundamenta tu respuesta.
 - Sueldo de una persona comparado con el dinero que destina a recreación.
 - Estatura de una persona comparado con el número de calzado que usa.
 - Estatura de una persona comparado con el número de cabezazos que se da con lámparas colgantes.
 - Edad de una persona comparada con la cantidad de veces que ha salido de vacaciones en su vida.
 - Notas promedio de una persona en la enseñanza media comparada con puntaje en la PSU.
 - Número de cesáreas comparada con número de partos normales.
 - Peso de una persona comparado con su cantidad de dientes.
 - Lugar de un tenista en el ranking mundial comparado con su número de derrotas.

- En una feria ganadera se remataron **9** terneros, de acuerdo con el siguiente cuadro:

Cabezas	Peso (kg)	Precio (\$ x kg)
3	204	496
2	148	488
4	196	482

Calcula:

- El peso promedio ponderado.
- El precio promedio ponderado por kilogramo.

7. La siguiente tabla presenta los puntajes obtenidos por los jóvenes que rindieron la PSU Matemática en el año 2003.

Intervalo	Frecuencia
[100,149]	193
[150,199]	251
[200,249]	1.156
[250,299]	2.747
[300,349]	9.152
[350,399]	10.718
[400,449]	24.176
[450,499]	27.609
[500,549]	28.480
[550,599]	22.830
[600,649]	14.183
[650,699]	6.223
[700,749]	2.721
[750,799]	1.822
[800,849]	1.209

- a. ¿Qué porcentaje de alumnos obtuvo puntajes iguales o superiores a los del intervalo **[700,749]**?
- b. Calcula los percentiles **10, 30, 40, 60, 70, 80** y **90**.
- c. Calcula la mediana y la desviación estándar de esta distribución de puntajes.
- d. Calcula Q_1 , Q_2 y Q_3 .
- e. Calcula el segundo decil (D_2) y el quinto decil (D_5).
- f. ¿A qué percentil corresponde, aproximadamente, el puntaje **628**?
8. Un fabricante asegura que el contenido promedio de nicotina de sus cigarrillos es de **2 mg**. Para verificar esto se realizó un estudio con una muestra aleatoria de **45** cigarrillos, obteniéndose un promedio de **3 mg** de nicotina. Se sabe que el contenido de nicotina de un cigarrillo sigue una distribución normal con desviación estándar de **0,5 mg**.

- a. Obtén e interpreta un intervalo con un **95%** de confianza para el verdadero promedio.
- b. Obtén el intervalo con un **80%** de confianza para la media.
- c. ¿Qué puedes concluir en relación con lo que dice el fabricante?

9. En una misma prueba de Inglés dos cursos, **C** y **D**, obtuvieron resultados cuyos datos estadísticos son los siguientes:

	Curso C	Curso D
\bar{x}	5,0	5,1
s	0,6	0,5

De acuerdo con estos datos:

- a. Compara el resultado de ambos cursos.
- b. Un alumno del curso **C** obtuvo un **6,0** y uno del curso **D**, un **6,2**. ¿A cuál de los dos alumnos les fue mejor en la prueba, en relación a su curso?
- c. Suponiendo que las notas se distribuyeron en forma normal, ¿entre qué notas por debajo y por encima del promedio se encuentra el **68,3%** central de los alumnos en cada curso?
10. Un apicultor desea conocer, con fines industriales, la cantidad de miel producida por las abejas de colmenas. Estudios anteriores indican que la desviación estándar es de **10 kg** anuales.
- a. ¿Cuántas colmenas debe incluir en su estudio, si admite un error máximo de **2 kg** y un **98%** de confianza?
- b. Si desea disminuir su error en un **50%**, ¿cuántas colmenas más debe incluir en el estudio?
- c. ¿Qué relación hay entre los resultados obtenidos?