

Estadística Descriptiva

(1) Los siguientes son los pesos, en kilogramos, y las estaturas, en centímetros, de una muestra de 14 individuos que participan de un ensayo clínico.

Pesos: 84, 99, 71, 64, 65, 80, 70, 70, 56, 66, 89, 60, 65, 79.

Estaturas: 185, 180, 173, 168, 175, 183, 184, 174, 164, 169, 205, 161, 177, 174.

- Para cada variable construya un diagrama de tallo y hojas.
- Para cada variable calcule la media, el desvío estándar, la mediana y los cuartiles.
- ¿Qué conjunto de mediciones tiene mayor variabilidad?

(2) Una empresa de informática dedicada al análisis de virus en computadoras, contabiliza los virus detectados con su software propio en 9 ordenadores de domicilios particulares. Los resultados obtenidos son los siguientes: 46, 29, 35, 61, 53, 37, 55, 47, 54.

- Obtenga la media, la mediana, el rango y el desvío estándar de las cantidades de virus detectadas en las computadoras.
- ¿Por debajo de qué cantidad de virus se encuentra el 75% de computadoras?

(3) Conocer cuál es el dato central de una distribución de datos es importante, pero también lo es el saber qué tan concentrado o extendido está nuestra información. Por ejemplo, el saber que un comercio tiene ingresos diarios de \$10.000 es interesante, pero además es importante saber si todos los días esas ventas están muy cerca de ese monto o, en realidad se alejan mucho.

A continuación, se brindan los datos de dos comercios que tienen la misma media de ventas diarias:

Comercio A: \$10.000, \$10.500, \$11.000, \$9.000, \$9.500.

Comercio B: \$10.000, \$5.000, \$15.000, \$19.000, \$1.000.

Averigüe cuál negocio tiene un desempeño de ventas diarias más consistentes.

(4) La tabla siguiente nos muestra las calificaciones de 10 estudiantes en dos evaluaciones de curso de ingreso 2022 a la LSI, un diagnóstica inicial y otra al final del curso.

Alumno	1	2	3	4	5	6	7	8	9	10
Inicio	4	5	1	5	2	3	2	1	1	3
Final	6	8	5	9	3	6	7	6	4	9

- Determine la media, la mediana, la moda, desvío estándar y varianza de las notas al inicio y al final del curso de ingreso.
- ¿Cuál conjunto de notas presenta mayor dispersión?
- Obtenga la media, mediana y desviación estándar del incremento de la calificación obtenida.

(5) Los siguientes son los puntajes que obtuvieron 40 estudiantes de una comisión en un parcial de Probabilidad y Estadística:

75, 89, 66, 52, 90, 68, 83, 94, 77, 60, 63, 77, 31, 88, 74, 37, 85, 76, 74, 63, 38, 47, 87, 65, 97, 49, 65, 72, 73, 81, 69, 72, 91, 87, 76, 58, 63, 70, 72, 65,

- Defina la variable en estudio. Realice el gráfico de tallo y hojas de los datos.
- Agrupe los datos en intervalos de clase y construya la

tabla de distribución de frecuencias completa.

c) Realice el histograma de frecuencias absolutas y el polígono de frecuencias correspondiente. Comente acerca de la simetría del gráfico construido.

d) Asumiendo que los datos están agrupados en intervalos de clase, calcule e interprete las siguientes medidas: media, mediana, cuartiles, moda, desvío estándar, rango.

e) Indique si la dispersión de los datos puede considerarse pequeña o grande aplicando el teorema de Chevishev.

f) Calcule el coeficiente de asimetría de Pearson e interprete. ¿Guarda relación con lo concluido en el inciso c?

g) Construya el histograma de frecuencias acumuladas porcentuales y la ojiva correspondiente.

h) Realice el diagrama de caja y extensión de los datos. ¿Qué puede decir acerca de la asimetría del conjunto de datos a partir del gráfico?

(6) Un estudio realizado por Weplan en 2022 recabó información sobre el consumo mensual de datos móviles. La tabla muestra la distribución de frecuencias para los datos agrupados medidos en Gb para una muestra de 60 argentinos.

Intervalo de clase	X_{PM}	fa	Fa	fr	Fr
[0, 2)		6			
[2, 4)					
[,)		23			
[,)		12			
[,)		10			

a) Complete la tabla. Calcule $P(X < 4)$, $P(X > 6)$, $P(2 \leq X < 8)$.

b) Elabore el histograma de frecuencias absolutas y el polígono de frecuencias correspondiente. Marque la moda en el histograma.

c) ¿Cuál es el consumo promedio mensual de datos móviles en Argentina? Calcule también el desvío estándar para estos datos.

d) Calcule la mediana e interprete.

e) Si los consumos mensuales de datos móviles en Uruguay tienen un promedio de 8.4 Gb con desvío estándar de 3 Gb, ¿cuál grupo de mediciones tiene mayor variabilidad: Argentina o Uruguay? ¿Cuál medida elige para responder esta pregunta? ¿Por qué?

(7) Un corrector de textos contabiliza el número de errores ortográficos que encuentra en cada página. Después de pasar este corrector por un texto de 50 páginas, se obtiene el siguiente número de erratas por página: 2, 3, 5, 0, 1, 4, 0, 6, 2, 1, 1, 0, 2, 4, 5, 3, 1, 2, 3, 2, 3, 1, 2, 4, 4, 2, 5, 4, 1, 3, 2, 6, 8, 2, 0, 1, 0, 2, 3, 1, 5, 10, 2, 1, 3, 6, 2, 0, 1, 3.

a) Identifique la variable estadística y clasifique.

b) Construya la tabla de frecuencias correspondiente.

c) Construya el diagrama de bastones para los datos. Solamente mirando este gráfico, ¿puede decir cuánto vale la moda?

d) ¿Qué porcentaje de páginas, respecto del total de las que se han corregido, tienen 2 errores?, ¿qué porcentaje de páginas respecto del total tienen menos de 6 errores?, ¿y qué porcentaje de páginas respecto del total tienen como mínimo 5 errores?

e) Calcule los siguientes estadísticos descriptivos: media aritmética, mediana, moda, varianza y cuartiles.

f) Grafique para las frecuencias acumuladas relativas.

(8) Se han comparado 5 laboratorios que analizaron varias veces con el mismo procedimiento la concentración de plomo en una misma muestra de agua de río. El objetivo del estudio es evaluar los errores obtenidos al realizar diversos análisis en cada laboratorio. Es importante que cada laboratorio analice sus muestras de manera independiente y con precisiones parecidas a las del resto de los laboratorios. La siguiente es la salida de un software estadístico del procesamiento inicial que realizó uno de los investigadores sobre los datos expresados en g/L.

Estadísticos descriptivos: Lab A; Lab B; Lab C; Lab D; Lab E									
Variable	N	Media	Desv.Est.	Varianza	Mínimo	Q1	Mediana	Q3	Máximo
Lab A	6	3,433	0,993	0,987	2,300	2,450	3,400	4,300	4,900
Lab B	5	5,080	1,215	1,477	4,000	4,100	4,400	6,400	6,500
Lab C	6	2,833	1,098	1,207	1,600	1,675	2,750	4,100	4,100
Lab D	7	3,743	1,003	1,006	2,100	2,800	3,800	4,800	4,800
Lab E	4	7,075	1,497	2,242	5,500	5,650	7,150	8,425	8,500

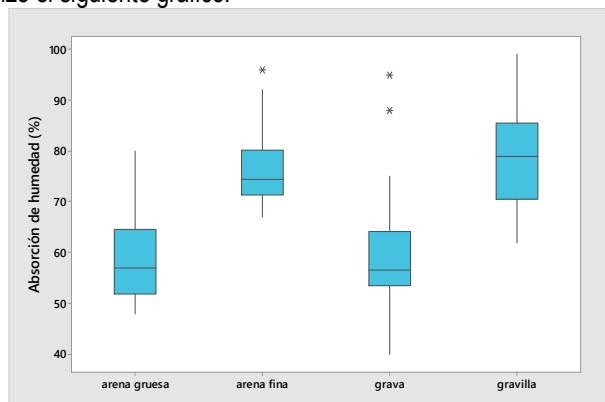
a) Complete la tabla.

Lab.	Media \pm DS	Mínimo	Cuartil 1	Mediana	Cuartil 3	Máximo	IQR	CV

b) ¿En cuál laboratorio se realizaron más cantidad de análisis?

c) ¿En cuál laboratorio se observe mayor dispersión en las mediciones? ¿Cuál medida emplea para responder? Justifique su elección.

(9) En un experimento industrial un ingeniero está interesado en la absorción de humedad para diferentes tipos de concreto en función de los distintos agregados que pueden incorporarse (arena gruesa o fina, grava y gravilla). Para ello, expuso a humedad durante 48 horas distintas muestras de concreto con los diferentes aditivos, registrando al finalizar el porcentaje de absorción de humedad de cada pieza. El ingeniero cargó los datos obtenidos en Minitab y realizó el siguiente gráfico.



a) Completa la tabla observando el gráfico.

Agregado	Mínimo	Cuartil 1	Mediana	Cuartil 3	Máximo

b) Si el objetivo es reducir la humedad absorbida, ¿cuál agregado le parece más adecuado incorporar a la mezcla de concreto? Justifique.

c) ¿Observa valores extremos en algún conjunto de observaciones? ¿en cuál? Reporte sus valores.

(10) Se ha determinado empíricamente que la llegada de automotores a un puesto de control caminero durante intervalos elegidos al azar de 10 minutos, tiene la frecuencia que se muestra en la tabla:

Nº de autos: X	0	1	2	3	4	5
frecuencia	20	40	30	10	5	1

a) Grafique la distribución de frecuencias relativas simples y las frecuencias acumuladas.

b) Calcule las medidas de tendencia central (media, mediana y moda).

c) Calcule las medidas de dispersión (variancia, desvío estándar y coeficiente de variación).

d) Calcule el coeficiente de asimetría de Pearson e indique su significado.