	INSTITUCIÓN EDUCATIVA LA PRESENTACIÓN					
	NOMBRE ALUMNA:					
	ÁREA / ASIGNATURA: Estadística					
	DOCENTE: ÉDISON MEJÍA MONSALVE					
	PERIODO	TIPO GUÍA	GRADO	Nº	FECHA	DURACIÓN
	II	APRENDIZAJE	11°		3/05/2022	

INDICADOR DE DESEMPEÑO

Calcula e interpreta las medidas de tendencia central en datos agrupados y no agrupados.

MEDIDAS DE TENDENCIA CENTRAL

Los parámetros más útiles son las medidas de Tendencia Central, las cuales ubican el valor alrededor del cual se concentra un conjunto de datos y las Medidas de Dispersión que describen la variabilidad o dispersión de los mismos.
Las tres medidas de tendencia central o de centralización más importantes son la moda, la mediana y la media.

- ***Moda.***

La *moda* se define como el dato con la frecuencia más alta, es decir, el que más se repite. No siempre existe una moda y en ocasiones puede haber más de una. Además, es la única medida de tendencia central que se puede calcular para variables nominales.

Ejemplos:

En el conjunto de datos: {2, 3, 3, 4, 4, 4, 5, 5, 8, 8, 12, 13} la moda es 4.
En la distribución {2, 2, 3, 3, 5, 5, 8, 8, 12, 12, 13, 13} no hay moda.
Para el conjunto de datos ordinales: {pequeña, pequeña, mediana, mediana, mediana, *grande*, grande, grande, extra-grande, extra-grande}, hay dos modas: “mediana” y “grande”, porque ambos se repiten el mismo número de veces.

- ***Mediana.***

La *mediana* se define como el dato central de la distribución, es decir el dato que queda justo en el medio, cuando el conjunto de datos se encuentra ordenado.

La mediana se puede utilizar con variables ordinales (además de la moda). Si el número de datos es impar, entonces la mediana corresponde al valor que se encuentra en el medio. Pero si el número de observaciones es par, entonces se toman los dos valores que se hallan en el medio de la distribución y se dice que la

mediana se encuentra entre esos dos valores, (en el caso de variables numéricas se suman esos valores y se divide entre dos)
Ejemplos: En el conjunto de datos: {a, b, b, c, c, *c, d*, d, g, g, k, m} la mediana esta entre c y d.
Para el conjunto de datos {2, 2, 3, 3, 5, *5*, 8, 8, 12, 12, 13} la mediana es 5
En el conjunto de datos: {2, 3, 3, 4, 4, *4, 5*, 5, 8, 8, 12, 13} la mediana es 4.5
En el siguiente conjunto de datos ordinales {pequeña, pequeña, mediana, mediana, mediana, *grande*, grande, grande, grande, grande, grande, extragrande, extragrande}, la mediana es "grande" La mediana divide al conjunto de datos justo a la mitad por lo que nos proporciona información del estilo: "El 50% de los datos esta por debajo de la mediana y el otro 50% por arriba de ella"

- ***Media.***

Si los datos son numéricos (en escala intervalar o de razón), entonces es posible calcular una tercera medida de tendencia central: la *media aritmética,* la cual consiste en la suma de todos los valores dividida por el número de ellos.
Se denota con $\overline{x}$
y queda expresada como:

$$Media(X) = \overline{x} = \frac{\sum_{i=1}^{N} X_i}{N}$$

siendo $(X_1, X_2, ..., X_N)$ el conjunto de observaciones

La media aritmética es lo que usualmente conocemos como "promedio", y se interpreta como tal. Una característica de la media es que resulta sensible a datos extremos, lo que no sucede con la mediana ni con la moda.

Ejemplos

En el conjunto de datos: {2, 3, 3, 4, 4, 4, 5, 5, 8, 8, 12, 13}, la moda es 4, la mediana es 4.5 y la media es 6.45.
Para el conjunto de datos {2, 3, 3, 4, 4, 4, 5, 5, 8, 8, 12, 93}, la moda es 4, la mediana es 4.5 y la media resulta 13.72.

Un ejemplo más:

En un grupo de Estadística I de la U de A, se observó la estatura de 16 alumnos y se obtuvieron los siguientes datos (ya ordenados):
1.52 1.52 1.53 1.53 1.57 1.58 1.58 1.60 1.64
1.64 1.64 1.66 1.66 1.74 1.76 1.79

Calculemos las Medidas de Tendencia Central:

moda = 1.64

mediana = $\frac{1.60+1.64}{2}$ = 1.62

$$\text{media} = \bar{x} = \frac{\sum_{i=1}^{16} x_i}{16} = \frac{25.96}{16} = 1.6225$$

Información proporcionada:

moda: “La estatura más frecuente entre los estudiantes es de 1.64 m”

mediana: “El 50% de los estudiantes miden menos de 1.62 m y el otro 50% mide más de 1.62m”

moda: “Los estudiantes tienen una estatura promedio de 1.6225 m ”

ACTIVIDAD # 1

1. La cuenta de la luz (en pesos) del mes de marzo de 30 familias escogidas Aleatoriamente se muestra a continuación.

250	560	340	780	890	960	470	340	540	440	120	340	340	550	440
450	450	670	860	430	330	230	810	70	970	360	560	1120	370	840

Calcula las tres medidas de tendencia central y escribe la información que proporcionan.

2. Realizar el anexo # 2. Puntos: 1 al 11.

MEDIDAS DE DISPERSIÓN

A las Medidas de Dispersión también se les llama Medidas de Variación. La variación es la cantidad de dispersión, o “separación”, que presentan los datos.

- ***Rango***

El rango de un conjunto de números es la diferencia entre el mayor y el menor de todos ellos. Se denota por *R* y se tiene que $R = x_n - x_1$

- ***Varianza***

La varianza es la suma de los cuadrados de las diferencias de los datos con relación a su media aritmética, dividida entre el tamaño de la muestra menos 1.

Se denota por S^2, y se tiene $S^2 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$

Si se dispone de una tabla de distribución de frecuencias el calculo varía, utilizando la expresión :

$S^2 = \dfrac{\sum_{i=n}^{k}(x_i - \bar{x})^2 * f_i}{n-1}$ en la cual, k es el número de datos distintos en la muestra.

- ***Desviación Estándar.***

Un inconveniente de la varianza es que sus unidades de medición se encuentran al cuadrado, por lo que no se puede comparar con la media aritmética. Debido a esto, se define la Desviación Estándar como la raíz cuadrada de la varianza.

Se denota por S, y se tiene $S = \sqrt{\dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$

De igual manera, existe una expresión equivalente: $S = \sqrt{\dfrac{\sum_{i=n}^{k}(x_i - \bar{x})^2 * f_i}{n-1}}$

- ***Coeficiente de variación.***

El coeficiente de variación es una medida relativa de la variación. Mide la dispersión de los datos con respecto de su media.

Se denota por CV y se expresa en porcentaje: $CV = \left(\dfrac{S}{\bar{x}}\right) \cdot 100\%$

El coeficiente de variación se utiliza principalmente cuando se desea comparar dos distribuciones de frecuencia que tienen diferente unidad de medida.

Ejemplo:

En un grupo de Estadística I del Cch Sur, se observó la estatura de 16 alumnos y se obtuvieron los siguientes datos (ya ordenados):

1.52 1.52 1.53 1.53 1.57 1.58 1.58 1.60 1.64 1.64 1.64 1.66 1.66 1.74 1.76 1.79

Calculemos las Medidas de Dispersión

Rango R = 1.79 – 1.52 = 0.27

Para realizar los cálculos de la varianza "a mano", resulta conveniente construir una tabla como la siguiente

Estatura x_i	Frecuencia f	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	$(x_i - \overline{x})^2 * f_i$
1.52	2	-0.1025	0.01051	0.02101
1.53	2	-0.0925	0.00856	0.01711
1.57	1	-0.0525	0.00276	0.00276
1.58	2	-0.0425	0.00181	0.00361
1.6	1	-0.0225	0.00051	0.00051
1.64	3	0.0175	0.00031	0.00092
1.66	2	0.0375	0.00141	0.00281
1.74	1	0.1175	0.01381	0.01381
1.76	1	0.1375	0.01891	0.01891
1.79	1	0.1675	0.02806	0.02806

$\overline{x} = 1.6225$ $\Sigma = 0.1095$

Varianza $S^2 = \frac{0.1095}{15}$ = 0.0073

Desviación Estándar $S = \sqrt{0.0073}$ = 0.08544

Coeficiente de Variación CV = $\frac{0.08544}{1.6225}$ 100 % = 5.266%

Démosle sentido a estos números:

R “La máxima diferencia de estaturas entre los estudiantes es de 27 cm.”

S “Las estaturas de los estudiantes se desvían en promedio 8.54 cm. de su media.” (equivalente a 0.08544 m.)

CV “Las estaturas varían 5.266% con respecto a su media”

ACTIVIDAD # 2

La cuenta de la luz (en pesos) del mes de marzo de 30 familias escogidas aleatoriamente se muestra a continuación.

250	560	340	780	890	960	470	340	540	440	120	340	340	550	440
450	450	670	860	430	330	230	810	70	970	360	560	1120	370	840

Calcula las medidas dispersión y escribe la información que proporciona cada una .

El pensamiento estadístico será un día tan necesario para el ciudadano eficiente como la capacidad de leer y escribir.
Herbert George Wells (1866 -1946)