CENTRO EDUCACIONAL
SANTA ROSA
Más que un colegio, una tradición

DIRECCIÓN ACADÉMICA
DEPARTAMENTO DE MATEMÁTICA

Respeto – Responsabilidad – Resiliencia – Tolerancia

GUÍA N°3

Medidas de Dispersión de Datos

Nombre: ____Curso 3°___ Fecha: ____/____/2020

Objetivos: 1) Calcular Rango, Desviación Media, Varianza, y Desviación Estándar. 2) Calcular Coeficiente de Variación. 3) Comparar conjuntos de datos utilizando medidas de tendencia central y de dispersión.

Estimado(a) Estudiante: Para apoyar tu estudio desde casa, tus profesores(as) de Matemática han preparado guías de apoyo, buscando links con videos en YouTube y un correo electrónico para atender consultas.

INSTRUCCIONES:

1. **Lee la información que contiene la guía y de ser necesario observa el material de apoyo.**
2. **Imprime y desarrolla, si no puedes imprimirla responde cada pregunta en tu cuaderno.**
3. **Ante cualquier consulta, enviar un correo a matematicacestarosa@gmail.com indicando nombre, curso y la consulta.**
4. **Videos sugeridos: https://www.youtube.com/watch?v=InXJIzJ44bc**

COMPARACION DE CONJUNTOS DE DATOS

Si se desea **comparar dos o más conjuntos de datos**, se pueden utilizar medidas de tendencia central, como el promedio y la mediana; medidas de dispersión como el rango, varianza, desviación estándar. Así podemos juzgar cuál de ellos tiene un **promedio más representativo**, es decir, **aquel conjunto cuyos valores son más cercanos al promedio**.

Veamos un ejemplo:

Un equipo de fútbol femenino necesita una delantera, para lo cual tiene dos candidatas. En los últimos 10 partidos del campeonato, las delanteras registraron las siguientes cantidades de goles:

Navas: 1, 0, 3, 0, 4, 1, 0, 0, 0, 3 **PROMEDIO** $(\overline{X}) = \frac{1+0+3+0+4+1+0+0+0+3}{10} = \frac{12}{10} =$ **1,2**

Flores: 1, 1, 2, 0, 1, 1, 2, 1, 1, 2 **PROMEDIO** $(\overline{X}) = \frac{1+1+2+0+1+1+2+1+1+2}{10} = \frac{12}{10} =$ **1,2**

La DT observa que ambas marcaron 12 goles en 10 partidos, con un promedio de 1,2 goles por partido. Entonces decide usar otros indicadores:

- $\text{Rango}_{\textbf{Navas}} = 4 - 0 = 4$
- $\text{Rango}_{\textbf{Flores}} = 2 - 0 = 2$

El **mayor rango** que presenta **Navas** puede indicar que en algunos partidos anota muchos goles, pero en otros, no anota, mientras que **los de Flores están más repartidos**.

Calculemos ahora la **varianza** y **desviación estándar** para cada jugadora:

NAVAS			
Nº Goles	$X - \bar{X}$	$\lvert X - \bar{X}\rvert$	$(X - \bar{X})^2$
0	0 – 1,2 = -1,2	1,2	$1{,}2^2 = 1{,}44$
0	0 – 1,2 = -1,2	1,2	$1{,}2^2 = 1{,}44$
0	0 – 1,2 = -1,2	1,2	$1{,}2^2 = 1{,}44$
0	0 – 1,2 = -1,2	1,2	$1{,}2^2 = 1{,}44$
0	0 – 1,2 = -1,2	1,2	$1{,}2^2 = 1{,}44$
1	1 – 1,2 = -0,2	0,2	$0{,}2^2 = 0{,}04$
1	1 – 1,2 = -0,2	0,2	$0{,}2^2 = 0{,}04$
3	3 – 1,2 = 1,8	1,8	$1{,}8^2 = 3{,}24$
3	3 – 1,2 = 1,8	1,8	$1{,}8^2 = 3{,}24$
4	4 – 1,2 = 2,8	2,8	$2{,}8^2 = 7{,}84$
			Σ = 21,6

La varianza será: $\sigma^2 = \frac{21{,}6}{10}$ **= 2,16**

La desviación estándar se obtiene calculando la raíz cuadrada del valor anterior:

$\sigma = \sqrt{2{,}16}$ **= 1,47**

FLORES			
Nº Goles	$X - \bar{X}$	$\lvert X - \bar{X}\rvert$	$(X - \bar{X})^2$
0	0 – 1,2 = -1,2	1,2	$1{,}2^2 = 1{,}44$
1	1 – 1,2 = -0,2	0,2	$0{,}2^2 = 0{,}04$
1	1 – 1,2 = -0,2	0,2	$0{,}2^2 = 0{,}04$
1	1 – 1,2 = -0,2	0,2	$0{,}2^2 = 0{,}04$
1	1 – 1,2 = -0,2	0,2	$0{,}2^2 = 0{,}04$
1	1 – 1,2 = -0,2	0,2	$0{,}2^2 = 0{,}04$
1	1 – 1,2 = -0,2	0,2	$0{,}2^2 = 0{,}04$
2	2 – 1,2 = 0,8	0,8	$0{,}8^2 = 0{,}64$
2	2 – 1,2 = 0,8	0,8	$0{,}8^2 = 0{,}64$
2	2 – 1,2 = 0,8	0,8	$0{,}8^2 = 0{,}64$
			Σ = 3,6

La varianza será: $\sigma^2 = \frac{3{,}6}{10}$ **= 0,36**

La desviación estándar se obtiene calculando la raíz cuadrada del valor anterior:

$\sigma = \sqrt{0{,}36}$ **= 0,6**

Estos indicadores confirman que los goles de Flores presentan menor dispersión, lo que se refleja en que cada partido marca una cantidad de goles similares, lo que no ocurre con Navas.

La jugadora elegida por la DT dependerá de lo que busque. Si se consideran los promedios de goles por partido, en ambos casos es el mismo, pero el promedio de Flores resulta mucho más **representativo**, ya que presenta una cantidad de goles por partido más homogénea (parecida, similar).

COEFICIENTE DE VARIACION

- **El coeficiente de variación (CV)** permite realizar comparaciones entre conjuntos con respecto a la dispersión de sus datos, e **incluso entre variables que se miden con diferentes unidades de medida**. Matemáticamente, corresponde al cociente (división) entre la desviación estándar y la media aritmética (promedio). Esto es:

$$CV = \frac{\sigma}{\lvert\bar{X}\rvert}$$

Para expresar el CV en porcentaje, basta con multiplicar el coeficiente obtenido por 100.

- Mientras **menor** sea el coeficiente de variación, el conjunto es más **homogéneo** (los datos son más parecidos entre sí).
- Mientras **mayor** sea el coeficiente de variación, el conjunto es más **heterogéneo** (los datos se diferencian más entre sí).

Sigamos con Navas y Flores……..

NAVAS: σ **= 1,47** $\bar{X}$ **= 1,2**	**FLORES:** σ **= 0,6** $\bar{X}$ **= 1,2**
$CV = \frac{1,47}{1,2} = 1,225$	$CV = \frac{0,6}{1,2} = 0,5$

Estos valores indican que los resultados de **FLORES** son más parecidos entre sí **(más homogéneos)**, comparados con los de **NAVAS**.

ACTIVIDADES: COMPARACION DE CONJUNTOS DE DATOS

OBSERVACION: EXPRESE SUS RESULTADOS CON DOS DECIMALES.

1. Dos candidatos, Elvira y Juan, han rendido 7 pruebas de selección para una empresa. Los puntajes que obtuvieron cada uno fueron los siguientes:

   Elvira: 80, 40, 62, 72, 46, 80, 40
   Juan: 57, 55, 54, 52, 62, 55, 59

Si el director de la empresa debe decidir por aquel que tuvo mejor rendimiento, ¿a quién contratará? Aplica los indicadores de dispersión, **rango, varianza, desviación estándar y coeficiente de variación** para determinar quién será el elegido.

a) CALCULOS

b) RESPUESTA

Para resolver la siguiente situación, considera que **el coeficiente de variación (CV) permite** realizar comparaciones **entre variables que se miden con diferentes unidades de medida**.

2. Paulina trabaja en una ferretería y ha recibido de una distribuidora dos tipos de muestras; una muestra de 9 clavos (medidos en pulgadas) y de la otra distribuidora, una muestra de 9 varas de madera (medidas en metros):

Clavos	2,0	2,5	3,4	2,6	3,3	3,5	2,1	2,3	2,1

Varas	3,3	3,0	3,5	3,2	3,5	3,6	2,7	3,5	3,5

a) Calcule el coeficiente de variación para cada conjunto de datos (para clavos y varas):

b) Lee lo que dice Paulina y responde:

¿Por cuál distribuidora optará Paulina? Justifica:

R E C O R D E M O S

MEDIDAS DE DISPERSION

- **El rango (R)** corresponde a la **diferencia (resta)** entre el **mayor** y el **menor** de los **datos** de la distribución. Esta medida indica de alguna manera cuán dispersos están los datos de la distribución.

$$\boldsymbol{R = Xmáx - Xmin}$$

- **La desviación de una variable x con respecto a su media aritmética está dada por:**

$$D = Xi - \bar{X}$$

- **La desviación media (DX)** corresponde a la **media aritmética** de los valores absolutos de las desviaciones $D = Xi - \bar{X}$ de los n datos, esto es:

$$\mathbf{D_X} = \frac{|X1-\bar{X}|+|X2-\bar{X}|+|X3-\bar{X}|+\cdots+|Xn-\bar{X}|}{n}$$

$\boldsymbol{\bar{X}}$ = promedio

$|\boldsymbol{X1} - \boldsymbol{\bar{X}}|$ = valor absoluto de la diferencia (resta) entre el dato y su promedio. Es siempre un valor positivo.

- **La varianza** y **la desviación estándar** permiten cuantificar la dispersión dada por la desviación media.

- **La varianza ($\boldsymbol{\sigma^2}$)** corresponde a la media aritmética de los cuadrados de las desviaciones de los n datos. Se expresa en unidades cuadradas.

$$\boldsymbol{\sigma^2} = \frac{(X1-\bar{X})^2+(X2-\bar{X})^2+(X3-\bar{X})^2+\cdots+(Xn-\bar{X})^2}{n}$$

- **La desviación estándar ($\boldsymbol{\sigma}$)** se obtiene extrayendo la raíz cuadrada de la varianza. Se expresa en la misma unidad que la variable, por lo que puede dar una idea más cercana de lo disperso que es el conjunto.

$$\boldsymbol{\sigma} = \sqrt{\boldsymbol{\sigma^2}}$$

- **A mayor dispersión, mayor valor de la varianza; a menor dispersión, menor valor de la varianza.**

SOLUCIONARIO

PREGUNTA 1

a)

	PROMEDIO $\overline{X}$	**VARIANZA** σ^2	**DESVIACION ESTANDAR (σ)**	**COEFICIENTE DE VARIACION (CV)**
ELVIRA	60	277,71	16,66	0,28
JUAN	56,29	9,63	3,1	0,06

b) Juan tiene un coeficiente de variación menor, por lo que su rendimiento es más homogéneo. Se debería contratar a Juan.

PREGUNTA 2

a) Coeficiente de Variación de los clavos: 21,59%
Coeficiente de Variación de las varas: 8,46%

b) Los clavos son los datos con mayor coeficiente de variación, por lo tanto, más heterogéneos. Debe pedir que manden una muestra más homogénea de clavos.