

MÉTODOS ESTADÍSTICOS BÁSICOS

GRADO II



CONTENIDOS

- Referentes de aprendizaje
- Población, Muestra, Variable
- Medidas de tendencia central
- Medidas de dispersión
- Distribución de frecuencias
- Datos NO agrupados y agrupados
- Probabilidad
- Combinaciones y Permutaciones

■ METODOS DE APRENDIZAJE

Estudio de un problema o pregunta de interés a la comunidad de la institución y/o ciudadanía.

■ Estándares

- Usar comprensivamente algunas medidas de centralización, localización, dispersión y correlación.
- Diseñar experimentos aleatorios para estudiar un problema o pregunta.

REFERENTES DE APRENDIZAJE

Propósitos

- Reconocer la importancia del análisis estadístico, para manejar apropiadamente los conceptos e interpretaciones de medidas y representaciones de conjuntos de datos, y así aplicar los conocimientos adquiridos para la resolución de problemas de situaciones cotidianas.

Desempeños

- Comprende la importancia del análisis estadístico para la resolución de problemas de situaciones cotidianas, aplicando en forma acertada los conocimientos estadísticos para la interpretación e inferencia de tablas, gráficos y conjuntos de datos.

MEDIDAS DE DISPERSIÓN

Dispersión

- Las medidas de dispersión son números positivos, que informan sobre cuánto se alejan los datos del centro (promedio).
- Cuando los datos son iguales este valor es cero.
- Cuando los datos se vuelven más diversos este se incrementa.
- Determinación requiere cálculo de la media (obligatorio).

Clases

- Desviación media

$$\begin{aligned} D_m &= \frac{\text{Suma valor absoluto de desviaciones}}{\text{Número total datos}} \\ &= \frac{\sum |x - \bar{X}|}{n} \end{aligned}$$

- Desviación estándar

$$\begin{aligned} \sigma &= \sqrt{\frac{\text{Suma cuadrado de desviaciones}}{\text{Número total datos}}} \\ &= \sqrt{\frac{\sum (x - \bar{X})^2}{n}} \end{aligned}$$

DISTRIBUCIÓN DE FRECUENCIA

Definición

- Método de recolección de datos según su magnitud y que indican el número de observaciones en cada una de ellas.
- Frecuencia (repetición)
- Clases:
 - Frecuencia absoluta (conteo)
 - Frecuencia relativa (porcentual respecto a total de datos)
 - Frecuencia acumulada (suma consecutiva dato anterior).
- Gráficos
 - Histograma
 - Polígono de frecuencias

Ejemplo

- Se indagó por el número de hermanos que tienen los 25 alumnos de un curso de grado sexto. Los resultados se presentan en la siguiente tabla.

No de hermanos	No de alumnos	% de alumnos	Frecuencia Acumulada
	Frecuencia absoluta	Frecuencia relativa	
0	3	12%	
1	8	32%	
2	6	24%	
3	4	16%	
4	2	8%	
5	2	8%	

ANÁLISIS ESTADÍSTICO DE DATOS


NO agrupados

- Independiente de la Población (pequeño o grande).
- La muestra es menor a 20 - 30 elementos.
- La información está en bruto (no necesita clasificación).
- El manejo o tratamiento puede ser manual (a suela!).



Agrupados

- Independiente de la Población (pequeño o grande).
- La muestra es mayor a 30 elementos.
- La información debe ser refinada: es aconsejable agrupar los datos en clases o categorías.
- El manejo o tratamiento debe ser metódico (fórmulas, tablas) y numérico (calculadora, programas informáticos).

A photograph of a rectangular piece of paper. The top half of the paper is covered with a dense, random pattern of small black dots. The bottom half of the paper is white and contains the handwritten formula $P_v = n R T$ in black ink.

- Objetivo: resumir la información en tendencias numéricas (media, desviación, gráficos, ...).

ANÁLISIS EN DATOS AGRUPADOS

Definición: objeto

- Los elementos (datos) son clasificados por grupos o intervalos según su frecuencia.
- Con datos agrupados se pierde detalle de la información original, pero se reduce la información numérica.
- La información (muestra) es resumida en una distribución de frecuencia (tablas, gráficos).

Clasificación: intervalo de clase

- En esencia, localizar y contar un dato dentro de los valores numéricos de un intervalo.

ANÁLISIS EN DATOS AGRUPADOS

Definición: objeto

- Los elementos (datos) son clasificados por grupos o intervalos según su frecuencia.
- Con datos agrupados se pierde detalle de la información original, pero se reduce la información numérica.
- La información (muestra) es resumida en una distribución de frecuencia (tablas, gráficos).

Clasificación: intervalo de clase

- En esencia, localizar y contar un dato dentro de los valores numéricos de un intervalo.

Ejemplo ilustrativo

■ Consumo bebidas en el PAE

Se necesita realizar un estudio sobre el número de bebidas (jugos, lácteos) que consume la comunidad educativa del colegio en un mes. De la población del colegio se elige una muestra de 100 personas.



ANÁLISIS EN DATOS AGRUPADOS

Distribución de frecuencia

Tabla de frecuencia		
Intervalo de clase	Frecuencia Absoluta	Frecuencia acumulada
Bebidas consumidas	Número de estudiantes	Número de estudiantes
I – 10		
11 - 20		
21 - 30		
31 - 40		
Total		

Oiga! No olvide que son 4 columnas. Si! Ud. expositor!

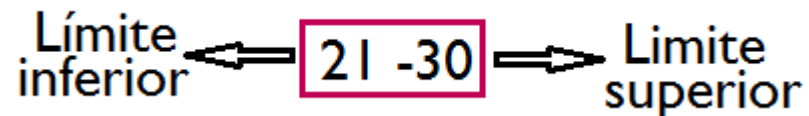
Ejemplo ilustrativo

■ Consumo bebidas en el PAE

31	18	10	27	14	31	18	10	27	14
12	24	16	28	20	12	24	16	28	20
13	30	22	9	26	13	30	22	9	26
19	11	23	15	32	19	11	23	15	32
25	17	29	21	8	25	17	29	21	8
31	18	10	27	14	31	18	10	27	14
12	24	16	28	20	12	24	16	28	20
13	30	22	9	26	13	30	22	9	26
19	11	23	15	32	19	11	23	15	32
25	17	29	21	8	25	17	29	21	8

ANÁLISIS EN DATOS AGRUPADOS

■ Intervalo de clase



■ Marca de clase (Mc)

Promedio de los límites de clase; es la medida representativa del intervalo.

■ Ancho de clase real (C)

Diferencia entre dos marcas de clase contiguas; puede ser (des)igual.

Con las marcas de clase y frecuencias relativas se calculan las medidas estadísticas.

Ejemplo ilustrativo

Tabla de frecuencia		
Intervalo de clase	Frecuencia Absoluta (f_i)	Marca de clase (Mc)
1 - 10		
11 - 20		
21 - 30		
31 - 40		
Total		

ANÁLISIS EN DATOS AGRUPADOS

Cálculo de intervalos de clase

Tabla de frecuencia		
Intervalo	f_i	Mc
I – 10		
II - 20		
2I - 30		
3I - 40		
Total		

ANÁLISIS EN DATOS AGRUPADOS

Cálculo de intervalos de clase

■ Rango (R)

Muestra el ancho numérico de variación de la muestra. Diferencia del dato máximo con el dato mínimo.

■ Número de clases (k)

Número de intervalos de agrupación; depende del total de datos (N).

- Elección empírica (justificada).
- Fórmula de Sturges (redondeada a un entero*):

$$k = 1 + \frac{\ln N}{\ln 2}$$

- Regla de Rice (de arroz!) (redondeada a un entero*):

$$k = 2\sqrt[3]{N}$$

■ Ancho de clase (C)

División del rango entre el número de clases:
 $C=R/k$.

■ Formar las clases

Poner como límite inferior del primer intervalo un valor un poco menor que el dato menor hallado en la muestra y posteriormente se suma a este valor C, para hallar el límite superior del primer intervalo.

Se procede a obtener los límites del intervalo siguiente y así sucesivamente.

ANÁLISIS EN DATOS AGRUPADOS

Cálculo de medidas estadísticas

Fórmulas y ejemplos



ACTIVIDAD I

***Welcome to reality:* el pasaje de TransMilenio costará 2400 pesos**

En los próximos días, la noticia de la firma del decreto que aumentó en \$100 pesos el pasaje será una realidad

Mientras estamos en vacaciones se nos olvidan las responsabilidades del día a día: pagar un recibo, tomar un bus, hacer vueltas en la mañana, porque si no es así, se nos va la tarde, y de ese modo nos entregamos al descanso.

Diciembre tiene ese no sé qué que nos hace relajar. Puede ser el aire de un nuevo comienzo, un cambio de página y más de uno se 'va a curar el alma' a la playa. Pero dicho esto, la verdad es otra y enero se convierte en el mes de *welcome to reality* y más si vive en Bogotá.

Toda esa palabrería, como tal vez lo podrá entender si lee estas páginas, es porque de verdad este mes viene con una dosis completa de realidad –hay que revisar la semana pasada para darse cuenta de que fue así– y la falta de

Publím metro
incluyente



Nota disponible en audio,
escaneando este código QR.
Utilice la aplicación Igetax

Por supuesto no a todos les agrada esa bienvenida. A nadie le suena que le toquen el bolsillo y menos por un sistema que tiene serias deficiencias. Su seguridad no



- Tomado del diario gratuito Publimetro Bogotá, del 24 Enero 2019, pág 6.

ACTIVIDAD I

***Welcome to reality:* el pasaje de TransMilenio costará 2400 pesos**

En los próximos días, la noticia de la firma del decreto que aumentó en \$100 pesos el pasaje será una realidad

Mientras estamos en vacaciones se nos olvidan las responsabilidades del día a día: pagar un recibo, tomar un bus, hacer vueltas en la mañana, porque si no es así, se nos va la tarde, y de ese modo nos entregamos al descanso.

Diciembre tiene ese no se qué que nos hace relajar. Puede ser el aire de un nuevo comienzo, un cambio de página y más de uno se 'va a curar el alma' a la playa. Pero dicho esto, la verdad es otra y enero se convierte en el mes

¿Un precio justo?

Esta Alcaldía ha dejado claro más de una vez que el sistema tiene un déficit grande,

Publimetro
incluyente



Nota disponible en audio,
recorriendo esta rutina CID



- Tomado del diario gratuito Publimetro Bogotá, del 24 Enero 2019, pág 6.

Sin embargo, ese no es el único pero de los ciudadanos: el descontento que hay entre los bogotanos que usan el sistema es grande y todo se debe a que no encuentran un servicio de calidad.

ACTIVIDAD I

Welcome to reality: el pasaje de Troncal 2400 pesos

En los próximos
decreto q

Mientras estas
nes se nos olvi
sabilidades de
gar un recibo
hacer vueltas
porque si no e
la tarde, y de
entregamos a
Diciembre
se qué que no
Puede ser el a
vo comienzo,
página y más
curar el alma
dicho esto, la
enero se conv
da suolomo in

¿Un pro
Esta Al
más de
ma tiene un dencn grande,



desmotivaciones.es

pagas más

Y te estripan más



- Tomado del diario gratuito Publimetro Bogotá, del 24 Enero 2019, pág 6.

Sin embargo, ese no es el único pero de los ciudadanos: el descontento que hay entre los bogotanos que usan el sistema es grande y todo se debe a que no encuentran un servicio de calidad.

ACTIVIDAD I

Aumento del pasaje de TM en los últimos 10 años

- 2019: \$2400
- 2018: \$2300
- 2017: \$2200
- 2016: \$2000
- 2015: \$1800
- 2014: \$1800 y \$1500
- 2013: \$1700 y \$1400
- 2012: \$1700
- 2011: \$1750
- 2010: \$1700



Resolver

1. Medidas básicas estadísticas.
2. Gráfico de barras: año vs. Precio.
3. Aumento porcentual entre años consecutivos (explicar).
4. Cambio porcentual entre 2010 y 2019.
5. Estimar una proyección debidamente argumentada para el 2020.

ACTIVIDAD 2



POLÍTICA ECONOMÍA INTERNACIONAL CIENCIA Y TECNOLOGÍA MEDIOAMBIENTE EDUCACIÓN SALUD ARTE Y CULTURA | OTRAS ▼



Está en: [UN Periódico Digital](#) /
[Tanto temblor: ¿qué pasa?](#)



MEDIOAMBIENTE



ENERO 30 DE 2019 | BOGOTÁ D.C.



Tanto temblor: ¿qué pasa?

GONZALO DUQUE ESCOBAR | PROFESOR EMÉRITO, UNIVERSIDAD NACIONAL DE COLOMBIA SEDE MANIZALES

Desde el sismo del sábado 26 de enero (de magnitud 5,4), asociado con el Nevado del Huila, se han presentado más de 8.500 eventos; el más fuerte, del lunes 28 en la tarde, se sintió especialmente en el centro y suroccidente del país. La vigilancia de este complejo volcánico impone desafíos como la optimización de estrategias para la gestión del riesgo.



El área glaciar forma parte de las territorialidades del resguardo Wila del pueblo Nasa y de los municipios de Páez (Cauca), Teruel (Huila) y Planadas (Tolima). Foto: Ideam.

ACTIVIDAD 2

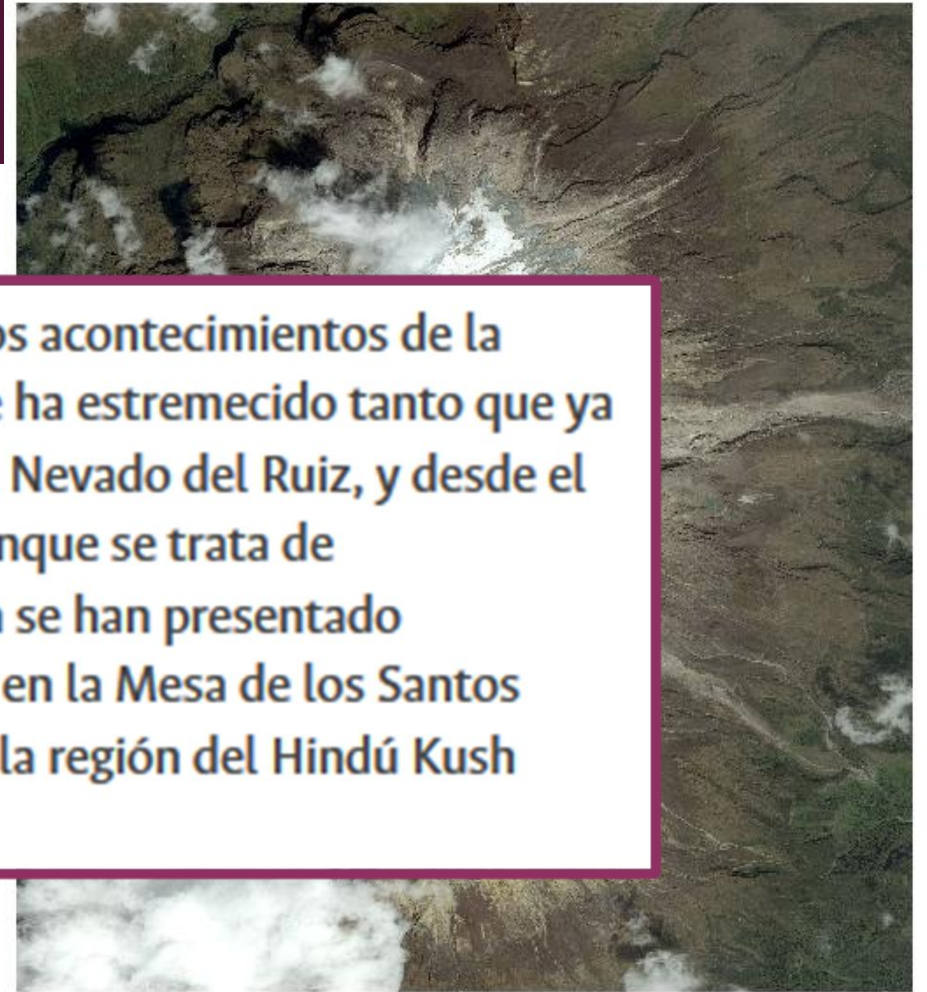


POLÍTICA ECONOMÍA INTERNACIONAL CIENCIA Y TECNOLOGÍA MEDIOAMBIENTE EDUCACIÓN SALUD ARTE Y CULTURA | OTRAS ▼



Enero ha sido un mes movido en Colombia, y no solo por los acontecimientos de la actualidad nacional e internacional, sino porque la tierra se ha estremecido tanto que ya genera preocupación. Por ejemplo el lunes 21 tembló en el Nevado del Ruiz, y desde el 26 se han presentado cientos de sismos en el del Huila. Aunque se trata de movimientos asociados con enjambres volcánicos, también se han presentado temblores asociados con fallas geológicas, como el del 22, en la Mesa de los Santos (Santander) –segundo nido sísmico del mundo después de la región del Hindú Kush (Afganistán)– y el del 28 en Casanare.

mas fuerte, del lunes 28 en la tarde, se sintio especialmente en el centro y suroccidente del país. La vigilancia de este complejo volcánico impone desafíos como la optimización de estrategias para la gestión del riesgo.



El área glaciar forma parte de las territorialidades del resguardo Wila del pueblo Nasa y de los municipios de Páez (Cauca), Teruel (Huila) y Planadas (Tolima). Foto: Ideam.

ACTIVIDAD 2

un Periódico
digital

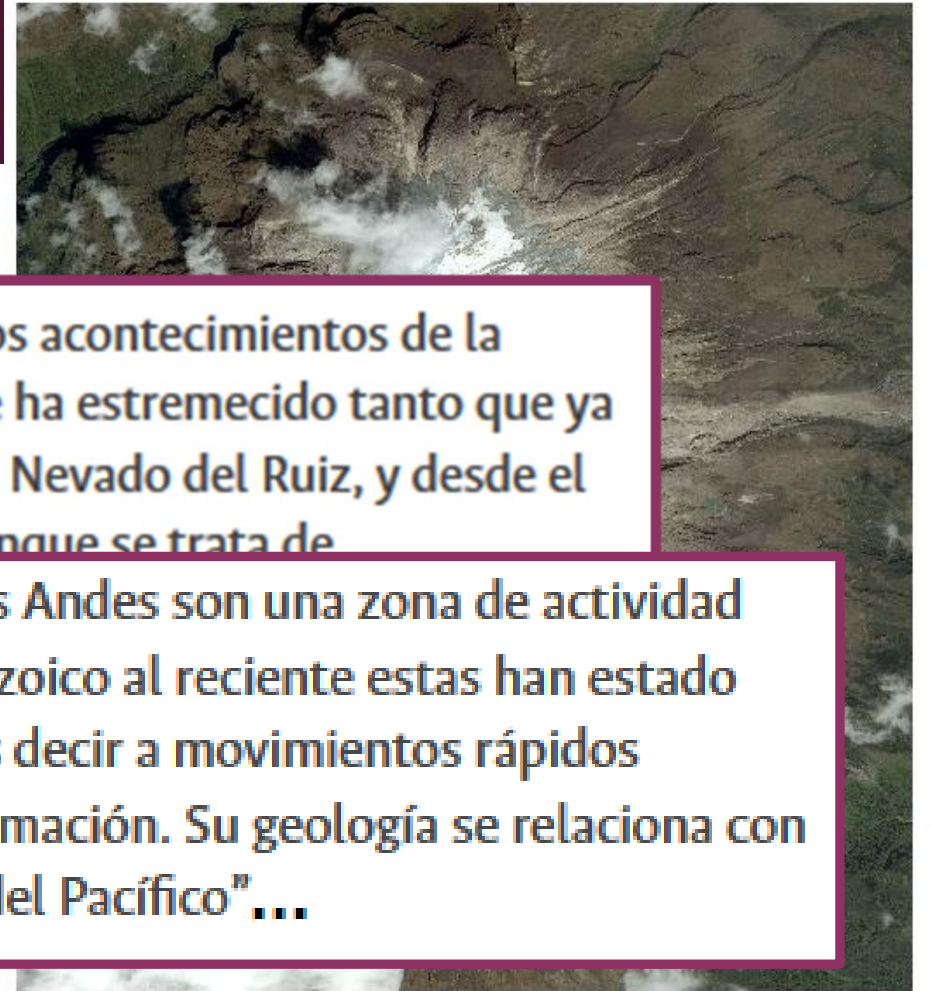
POLÍTICA ECONOMÍA INTERNACIONAL CIENCIA Y TECNOLOGÍA MEDIOAMBIENTE EDUCACIÓN SALUD ARTE Y CULTURA | OTRAS ▼



Enero ha sido un mes movido en Colombia, y no solo por los acontecimientos de la actualidad nacional e internacional, sino porque la tierra se ha estremecido tanto que ya genera preocupación. Por ejemplo el lunes 21 tembló en el Nevado del Ruiz, y desde el 26 se han presentado cientos de sismos en el del Huila. Aunque se trata de

movimientos asociados a temblores asociados a (Santander) –según (Afganistán)– y el de Por la juventud de sus montañas, los Andes son una zona de actividad **vulcano-tectónica**, ya que del Paleozoico al reciente estas han estado sometidas a procesos orogénicos, es decir a movimientos rápidos horizontales, responsables de su formación. Su geología se relaciona con el denominado “cinturón de Fuego del Pacífico”...

mas fuerte, del lunes 26 en el centro y suroccidente del volcánico impone desafíos como la optimización de estrategias para la gestión del riesgo.



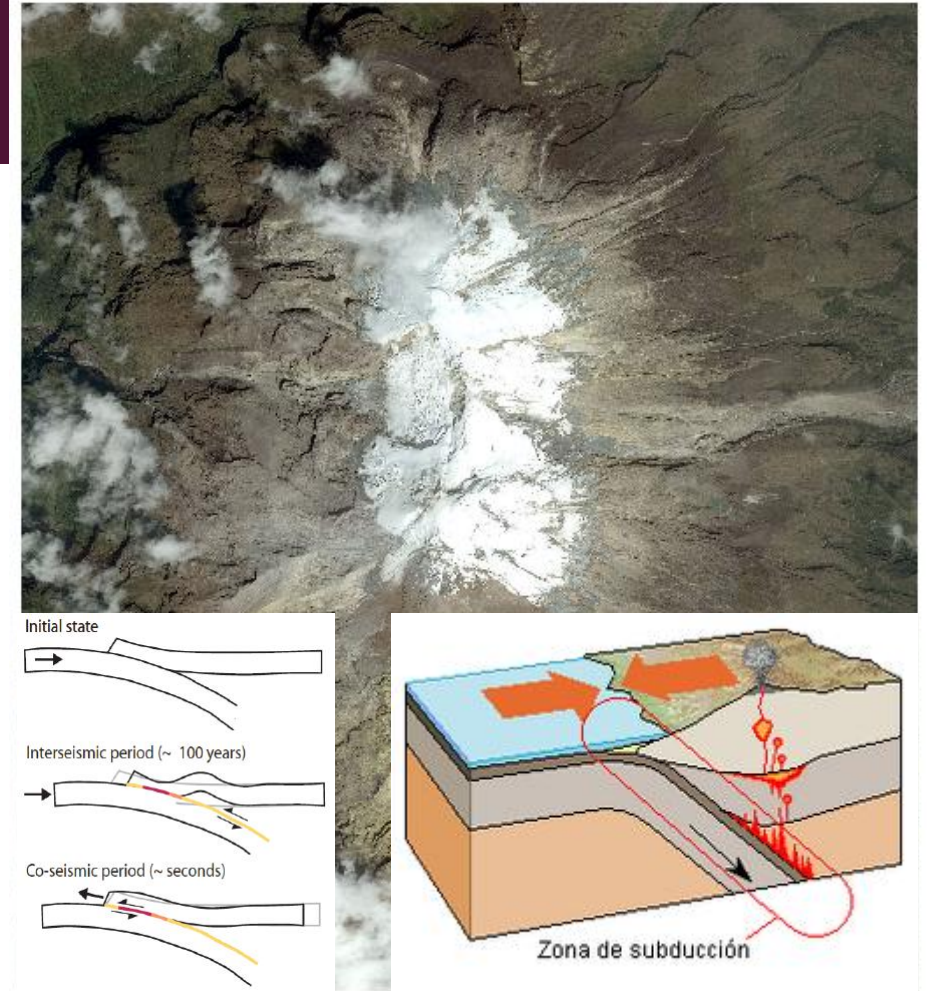
El área glaciar forma parte de las territorialidades del resguardo Wila del pueblo Nasa y de los municipios de Páez (Cauca), Teruel (Huila) y Planadas (Tolima). Foto: Ideam.

ACTIVIDAD 2

Una hoja de vida con potencial destructivo

Con 5.365 msnm, el Nevado del Huila es el volcán activo de mayor altura de los Andes colombianos. Su cubierta glaciar –en proceso de desglaciación– alcanza los 13 km². Su primera erupción histórica fue una pequeña explosión ocurrida a mediados del siglo XVI, entre los años 1550 y 1560.

A diferencia de los volcanes hawaianos, caracterizados por derrames fluidos de lava, el del Huila es explosivo y suele presentar lanzamiento violento y a gran presión de magma pulverizado y fragmentos de roca.



El área glaciar forma parte de las territorialidades del resguardo Wila del pueblo Nasa y de los municipios de Páez (Cauca), Teruel (Huila) y Planadas (Tolima). Foto: Ideam.

ACTIVIDAD 2

Resolver

1. Realizar una tabla de distribución de frecuencias para la magnitud de sismos registrados en el Volcán Nevado del Huila entre el 26 de Enero 2019 y el 30 de Enero 2019.
2. Encontrar el valor promedio y desviación típica estándar para los sismos registrados.
3. Respecto a las magnitudes mínima y máxima, hallar incremento sísmico porcentual.
4. Aunque los sismos no se pueden predecir como los eclipses, sí se pueden monitorear los eventos probables mediante un parámetro: suponga que una desviación estándar para magnitud sísmica debajo de 0.1 indica mayor probabilidad de ocurrencia de un evento. De acuerdo con los registros, ¿existe amenaza de probabilidad de ocurrencia?



El área glaciar forma parte de las territorialidades del resguardo Wila del pueblo Nasa y de los municipios de Páez (Cauca), Teruel (Huila) y Planadas (Tolima). Foto: Ideam.

ACTIVIDAD 3



Funcionarios del Centro de Operaciones de [TransMilenio](#) aseguran que **cada ocho minutos, en hora pico, son despachados buses en las 269 rutas urbanas, complementarias y especiales** que transitan todos los días en Bogotá y que cuentan con 5.970 paraderos. Todos los buses son controlados desde un centro de operación con 40 puestos de trabajo.

Las rutas del SITP que más se demoran, según usuarios

Tomado de periódico el tiempo, Por: Sara Daniela Caicedo y Sharon Fontecha; 11 de agosto 2018

<https://www.eltiempo.com/bogota/las-rutas-mas-demoradas-del-sitp-en-bogota-segun-los-usuarios-254350>

ACTIVIDAD 3



Actualmente, **1'520.000 personas usan a diario el SITP** y para la mayoría de ellas se volvió parte de su rutina salir con hora y media de anticipación; unos 30 minutos para esperar el bus y una hora para llegar a su destino.

T11 Calle 222-Alpes

Los conductores que manejan esta ruta tardan seis horas y media en realizar el recorrido –ida y vuelta– que atraviesa la ciudad de norte (calle 222) a sur (Ciudad Bolívar), convirtiendo **la ruta en una de las que más buses tiene en circulación en hora pico**, pues sus operadores programan 66 vehículos a diario aproximadamente, los cuales salen cada seis minutos. Esta es otra de las rutas que utiliza un bus patrón, Masivo Capital S.A.S. y Suma son los operadores de la T11.

Las rutas del SITP que más se demoran, según usuarios

Tomado de periódico el tiempo

<https://www.eltiempo.com/bogota/las-rutas-mas-demoradas-del-sitp-en-bogota-segun-los-usuarios-254350>

ACTIVIDAD 3

Como es mencionado, en los últimos meses los Bogotanos han expresado su inconformismo con el servicio de transporte SITP, por la demora de las rutas. Consultando a 30 usuarios, se les pregunto “¿cuánto demora su ruta de transporte?”. Ellos respondieron, dando su tiempo en minutos:

33 18 4 45 13 27 48 60 23 33
39 45 37 23 30 19 37 42 55 22
45 43 40 19 28 32 35 53 23 7

Realizar

1. Tabla de distribución de frecuencias con las respectivas marcas de clase; hallar ancho de clase real.
2. Polígono de frecuencias; estime del gráfico la demora promedio.
3. ¿Qué porcentaje de tiempos consultados superan los 25 minutos?



Tiempo (min)	fi	Mc
1 – 10		
11 – 20		
21 – 30		
31 – 40		
41 – 50		
51 – 60		

ACTIVIDAD 4

Realizar tabla de distribución de frecuencias completa y polígono de frecuencias para la situación del consumo de bebidas en el PAE

1. Usando la fórmula de Sturges.
2. Usando la regla de Rice.
3. Mencione las diferencias entre los métodos de clasificación.



31	18	10	27	14	31	18	10	27	14
12	24	16	28	20	12	24	16	28	20
13	30	22	9	26	13	30	22	9	26
19	11	23	15	32	19	11	23	15	32
25	17	29	21	8	25	17	29	21	8
31	18	10	27	14	31	18	10	27	14
12	24	16	28	20	12	24	16	28	20
13	30	22	9	26	13	30	22	9	26
19	11	23	15	32	19	11	23	15	32
25	17	29	21	8	25	17	29	21	8

TAREAS

- Tarea I
- ¿Qué es la frecuencia acumulada relativa?
- ¿Qué es el límite real superior/inferior de una clase?