

Sesión 2

Aprendizajes esperados

Al final de esta sesión se espera que puedas:



Identificar diferentes características de gráficas y tablas.



Identificar ventajas de presentar datos en gráficas o en tablas.



Calcular diferentes medidas de acuerdo a los datos presentados como la medida de tendencia central y una de dispersión.

Duración sugerida



40%

40%

20%

Material para la clase

- ☐ Anexo 2.1
- ☐ Anexo 2.2



Nota

Recuerda que solo podemos hacernos una idea. No podemos afirmar que el clima es diferente. Eso solo se puede hacer usando fórmulas de inferencia estadística.

Lo que sabemos,

lo que debemos saber



Esta sección corresponde al 40% de avance de la sesión

En esta sesión profundizaremos en las dos formas de presentar datos y dos características importantes para describirlos.

¿Recuerdas que dijimos que el tiempo atmosférico y por ende el clima también dependen de un lugar y un momento? En la sesión pasada vimos cómo el clima parece cambiar en los diferentes meses del año en un mismo lugar (ver nota al margen). Además, a pesar de que con tu grupo trabajaste en una sola ciudad, si miras rápidamente las gráficas de Cartagena y las de Bogotá, podrás darte cuenta de que son diferentes. Es decir que cambiando solo el lugar también parece que cambió el clima. Pero Cartagena y Bogotá están muy lejos y entre un mes y el siguiente pasa mucho tiempo. ¿Sí cambiará el clima en cuestión de horas y de pocos kilómetros? ¿Qué tanto? Vamos a mirar más datos para hacernos una idea.

En el Anexo 2.1 tenemos una gráfica del Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM). Antes de mirar los datos, vamos a prestar atención a este tipo de gráfica y a las partes que tiene. En esta parte de la sesión vas a trabajar en grupos. Reúnete en grupos de 2 o 3 personas, según las indicaciones de tu docente.

Respondan las preguntas 1 a 9 del Anexo 2.1.



¿Cuántos datos creen que hay en la gráfica?

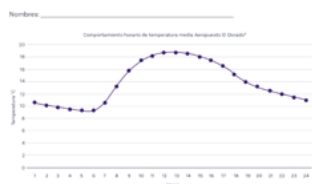
A pesar de que no lo dice directamente en la gráfica del Anexo 2.1, los cuadrados rojos en la línea son los datos y las líneas son solo para unir los puntos, para que sea más fácil de visualizar el comportamiento. Esos cuadrados se llaman marcadores de datos. ¿Ahora saben cuántos datos hay?



¿Notan alguna otra diferencia entre esta gráfica y la usada en la sesión pasada? En la Figura 1 se puede ver la información presentada en la sesión anterior.

Anexo

Anexo 2.1



Encierran en un círculo en la gráfica lo que aparece a continuación y escriban el número correspondiente al lado del círculo:

- ☐ Un dato
- ☐ Unidades
- ☐ Título de la gráfica
- ☐ Eje x que nombra o describe
- ☐ Eje y que nombra o describe
- ☐ Título del eje x
- ☐ División del eje x
- ☐ División del eje y
- ☐ Mínimo
- ☐ ¿A qué hora hace más frío en Bogotá normalmente?
- ☐ ¿Normalmente cuál es la temperatura mínima?
- ☐ Máximo
- ☐ ¿A qué hora hace más calor en Bogotá normalmente?
- ☐ ¿Normalmente cuál es la temperatura máxima?

¿Qué más podrían decir a partir de la gráfica? Comparen la velocidad con la que se sube con la que se baja.

Anexo 2.2

Nombre: _____

Tabla 2. PROMEDIOS MENSUAL Y ANUAL DE TEMPERATURA (°C)

Estación	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	Anual
Bogotá San Jorge	15.5	15.8	16.8	16.9	16.9	15.8	15.2	15.3	15.3	15.5	15.7	15.8	15.8
Medellín	16.8	16.8	17	17.3	17.3	17	16.4	16.5	16.8	16.9	16.9	16.9	16.9

Encierran en un círculo en la tabla lo que aparece a continuación y escriban el número correspondiente al lado del círculo:

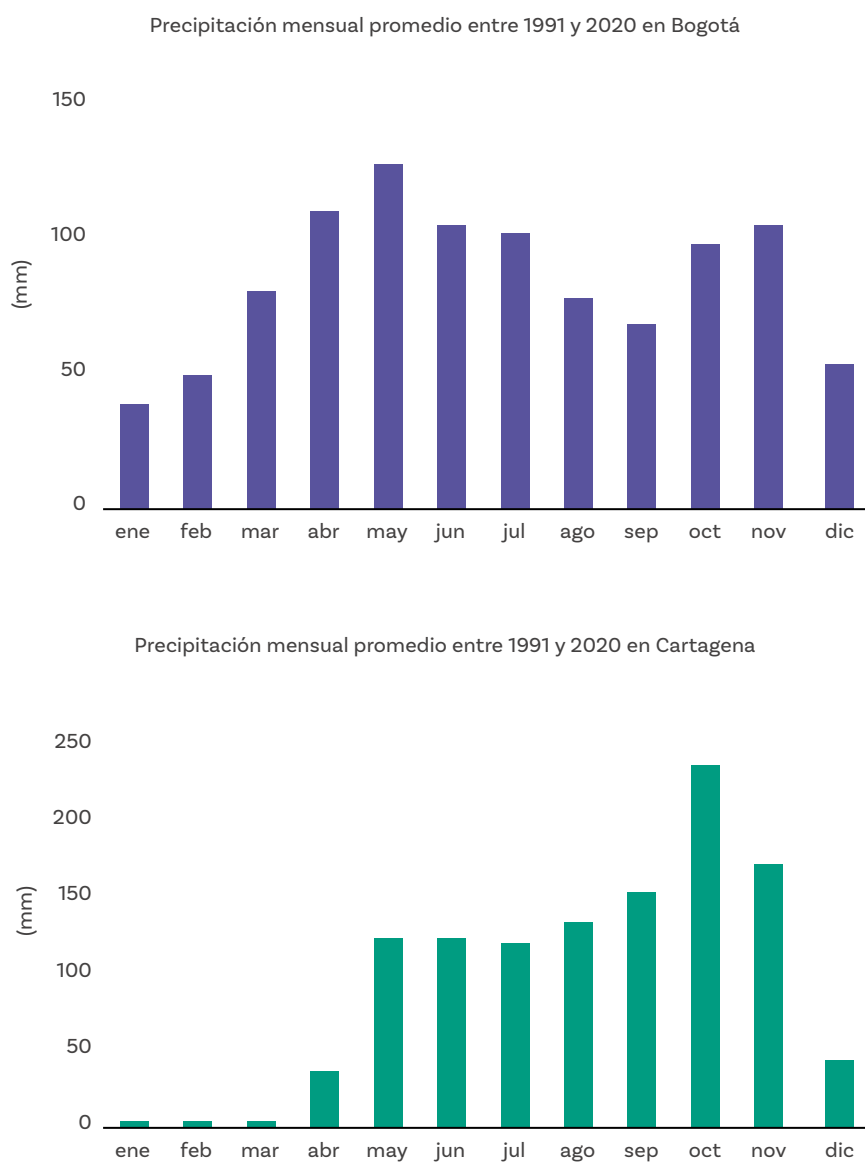
- ☐ Un dato
- ☐ Unidades
- ☐ Título de la tabla
- ☐ Encabezado (primera fila)
- ☐ Una columna
- ☐ Una serie de datos (fila fila)
- ☐ Temperatura máxima mensual promedio
- ☐ Gráfica San Jorge
- ☐ Universidad Nacional
- ☐ Temperatura máxima mensual promedio
- ☐ Gráfica San Jorge
- ☐ Universidad Nacional
- ☐ Rango de temperatura mensual promedio
- ☐ Gráfica San Jorge
- ☐ Universidad Nacional

6. Datos tomados del Estudio de la Caracterización Climática de Bogotá y Cuenca Alta del Río Tiquis por R. S. <http://www.ideam.gov.co/informacion/2002/07/05/CARACTERIZACION-CLIMATICA-BOGOTA.pdf> y <http://www.ideam.gov.co/informacion/2002/07/05/CARACTERIZACION-CLIMATICA-MEDALLA.pdf>

Las gráficas presentadas en la *Figura 1* tenían unas barras, por lo que se llaman gráficas de barras. La de esta sesión usa líneas y se llama gráfica de líneas.

Es momento de mirar los datos, pero antes, debemos asegurarnos de que sí nos sirvan para lo que estamos haciendo. Como queremos saber si el clima puede variar y qué tanto, debemos partir de datos del clima.

Figura 1. Precipitaciones mensuales promedio en Bogotá y Cartagena



Nota

Las coordenadas son un conjunto de valores numéricos que permiten localizar un punto en un espacio determinado.

Con las coordenadas se pueden conocer las distancias entre varios puntos. Si quieren saber qué tan cerca están las estaciones, se comparten las coordenadas de ambas estaciones:

San Jorge 04°30' 20,5"N,
74°11' 21,3"W.

UNAL: 04°38' 17,3"N,
74°05' 20,3"W.

Pueden buscar en internet una calculadora de distancia a partir de estas coordenadas suministradas.

No nos servirían datos del tiempo atmosférico de un día en particular.

La gráfica sí corresponde a datos de clima. En el documento de donde se obtuvo afirman que incluye datos de por lo menos diez años, tomados cada hora. Es decir, aproximadamente 3.650 días y como tenemos 24 datos por día, serían un total de 87.600 datos aproximadamente.

Ahora sí vamos a mirar algunas características de los datos. En primer lugar, en esta gráfica cada dato tiene dos componentes asociados: la hora, que sirve para clasificar, ordenar, de alguna manera los datos y el valor de la temperatura. En esta gráfica se clasifican los datos en el eje horizontal, mientras que el valor de la temperatura va en el vertical. El **mínimo** es el menor valor de un conjunto de datos, en este caso la temperatura. El **máximo** es el mayor valor.

Respondan las preguntas 10 y 11 del Anexo 2.1.



¿Fue fácil identificar la zona en la que se encontraban el máximo y el mínimo?

¿Fue fácil identificar el valor exacto del máximo y del mínimo?

Una vez más observamos que la gráfica nos ayuda a hacernos una idea del comportamiento, pero no es muy útil cuando se necesitan valores exactos.

Volvamos a las preguntas del principio.



¿Es posible que cambie el clima en un lugar en cuestión de horas?

¿Cuánto es el máximo cambio de temperatura promedio entre dos horas consecutivas?

¿Cuánto cambia la temperatura horaria promedio cada día en Bogotá? Es decir, ¿cuál es la diferencia entre el máximo y el mínimo?

Lo único que pueden decir es si les parece que hay un cambio. No pueden decir nada más sin las fórmulas de inferencia estadística. Además, ¿hay una definición única de qué es cambiar? ¿Qué tanto tiene que cambiar para que se considere que cambia?, ¿1 °C, 0,1 °C, 0,01 °C?, ¿y en cuánto tiempo? Este es el tipo de aspectos que es necesario definir para poder usar las fórmulas de la inferencia estadística.

Fíjate que, al igual que en la sesión pasada, en este caso estamos interesados en cambios en un parámetro del clima con respecto al tiempo. En la sesión pasada queríamos comparar la precipitación durante un año, dividiéndolo en meses. En este caso queremos comparar la temperatura durante un día dividiéndolo en horas.

Manos a la obra

Desconectadas



Esta sección corresponde al 80% de avance de la sesión

Continuarás trabajando en los mismos grupos en los que hiciste la actividad que acabas de terminar. En la actividad anterior nos hicimos una idea sobre si era posible que el clima de un lugar cambiara en cuestión de horas. Ahora, vamos a ver si al desplazarse unos pocos kilómetros es posible que el clima cambie.



¿Cómo creen que se obtienen los datos de temperatura de un lugar?

Las estaciones meteorológicas son lugares donde tienen instrumentos para medir diferentes parámetros del tiempo atmosférico, como la temperatura.



¿Cómo será posible dar información sobre toda una ciudad a partir de la medición en un solo lugar? ¿Cómo se elige ese lugar?

En realidad, conforme nos alejemos del lugar de la medición, la medida va perdiendo exactitud. En consecuencia, lo que se puede hacer es tomar varias medidas en diferentes lugares. Si se usan o no, y la forma en la que se usan las medidas, dependerá de lo que se quiera averiguar.

En el Anexo 2.2 aparece una tabla con algunos datos de dos estaciones meteorológicas en Bogotá. Así que esta tabla tiene dos **series, o grupos de datos**: los de la estación Granja San Jorge y los de la estación de la Universidad Nacional. Al igual que en la actividad pasada, antes de mirar los datos, vamos a prestar atención a este tipo de tabla y las partes que tiene.

Respondan las preguntas 1 a 6 del Anexo 2.2.



¿Es fácil identificar el número de datos? ¿Cuántos hay?

Antes de mirar los datos, debemos asegurarnos de que sí nos sirvan para lo que estamos haciendo. Como queremos saber si es posible que el clima varíe y qué tanto, al desplazarse unos pocos kilómetros, necesitamos datos del clima en lugares diferentes.

No nos servirían datos del tiempo atmosférico de un día en particular. En el documento de donde se obtuvo la tabla afirman que son datos de clima. Por otro lado, son dos estaciones meteorológicas en la misma ciudad, así que consideraremos que están a pocos kilómetros de distancia.

Ahora sí, veamos algunas características de los datos. En primer lugar, en esta tabla cada dato tiene tres componentes asociados. La serie que, como dijimos, es la estación; el mes es el que sirve para clasificar, ordenar, de alguna manera los datos y el valor de la temperatura. Por otro lado, vemos que la última columna tiene una etiqueta particular, que no corresponde a un mes. Esta columna tiene la temperatura promedio anual de cada una de las estaciones.



¿Cómo podríamos comparar esas dos series de datos?

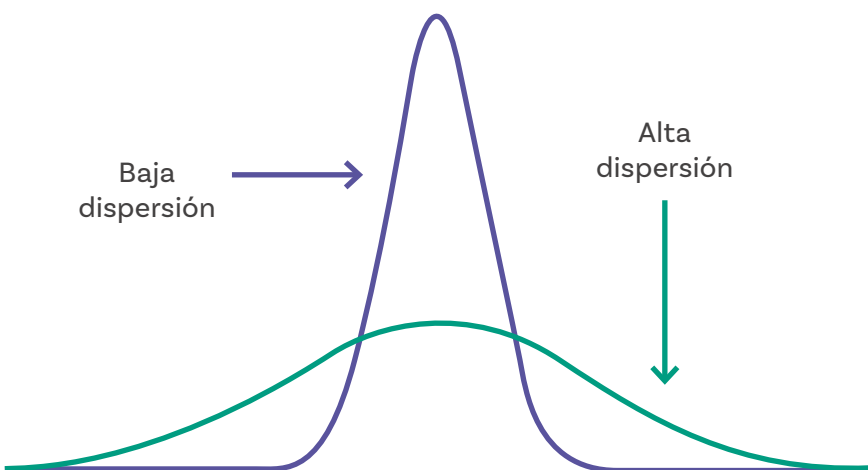
Los aspectos principales que se suelen comparar entre conjuntos de datos son medidas que dan una idea de dos cosas: 1) dónde está el centro de los datos y 2) qué tan espaciados están los datos.

Es muy importante que ambos aspectos se comparen. Para que se considere que dos conjuntos de datos se comportan de la misma manera, tanto la medida del centro de los datos como la medida de qué tan espaciados están deben considerarse iguales.

Un ejemplo de una medida del centro de los datos es el promedio. En estadística se prefiere llamarlo **media**, pero es lo mismo: la suma de los términos dividida por el número de términos.

Un ejemplo de una medida de qué tan espaciados están los datos es la diferencia entre el máximo y el mínimo. En terminología de estadística se diría que una medida de la **dispersión** de los datos es el **rango**, la diferencia entre el mínimo y el máximo. Puedes ver un ejemplo de datos con baja y alta dispersión en la *Figura 2*.

Figura 2. Representación de datos con baja y alta dispersión



En la tabla ya tenemos las medias de las dos estaciones. Recuerden que aquí lo que nos interesa es ver si hay una variación en la temperatura al cambiar de lugar. Ahora faltaría calcular los rangos. Esta es la primera vez que podemos hacer esto, ya que tenemos los datos que se usaron para calcular la media que estamos comparando.

Ahora, respondan las preguntas 7, 8 y 9 del Anexo 2.2.



¿Consideran que hay un cambio entre la temperatura anual promedio de las dos estaciones?



Una vez más, lo único que pueden decir es si les parece que hay un cambio. No pueden decir nada más sin las fórmulas de inferencia estadística. En este caso también hay ambigüedad con respecto a qué tanto tiene que cambiar la media de la temperatura y a qué distancia deben estar las estaciones para considerarlas cercanas. Además, vimos que se debe tener en cuenta otro aspecto que no habíamos considerado antes, la dispersión de los datos. Esto nos acerca un poco más a lo que hace la inferencia estadística con sus fórmulas matemáticas.

Glosario



Series de datos: en estadística, una serie de datos es una colección ordenada de valores que representan observaciones o mediciones de una variable.



Dispersión: en estadística, la dispersión mide qué tan alejados están los datos entre sí o respecto a un valor central, como la media. Indica la variabilidad o la extensión de un conjunto de datos.



Rango: en estadística, el rango es una medida de dispersión que indica la diferencia entre el valor máximo y el valor mínimo de un conjunto de datos.



Antes de irnos



Esta sección corresponde al 100% de avance de la sesión

Revisa los aprendizajes esperados de forma individual respondiendo las preguntas de forma que mejor reflejen tu progreso:

- 1 ¿Puedes identificar partes de gráficas y tablas?
 - ☐ Sí
 - ☐ Parcialmente
 - ☐ Aún no
- 2 ¿Puedes identificar ventajas de presentar datos en gráficas o en tablas?
 - ☐ Sí
 - ☐ Parcialmente
 - ☐ Aún no
- 3 ¿Puedes calcular una medida de tendencia central y una de dispersión?
 - ☐ Sí
 - ☐ Parcialmente
 - ☐ Aún no

Si tus respuestas fueron “Parcialmente” o “Aún no”, vuelve a las actividades propuestas en prácticas de datos. Luego, discute con tus compañeras y compañeros de grupo lo que se hizo en cada momento de la actividad y el rol al que correspondía. Si todavía te quedan dudas, consúltale a tu docente.

Si regresamos al reto, ¿para qué te servirá conocer las partes de una gráfica y los conceptos de estadística de esta sesión?

Haz un esquema en el que muestres las partes principales de una tabla y una gráfica. Además, completa lo que aparece abajo. Si necesitas volver a leer, no dudes en hacerlo para completar este esquema.

La _____, comúnmente conocida como _____, es una medida del _____ de los datos. El _____ es una medida de qué tan _____ están los datos, lo que en estadística se conoce como _____.

Elige 4 números y a partir de estos calcula las dos medidas que se mencionaron en el párrafo anterior.

