Graphical user interface, application

Description automatically generated

Apuntes Unidad 1

Tipos de gráficos estadísticos



Shape, arrow

Description automatically generated

**EXPLORAR Y ANALIZAR LOS DATOS**

Para entender un fenómeno, tal como la megasequía, puede resultar útil buscar datos como las precipitaciones anuales.  Además, debemos recordar que el **promedio** **o media** de un conjunto de datos se obtiene de **calcular la suma de ellos y luego dividir esto por el número de datos.**

Para explorar y analizar los datos buscados, estos se organizan en tablas y gráficos. En los **gráficos de barras** existe una correspondencia **entre la longitud de la barra y el valor del dato.** 

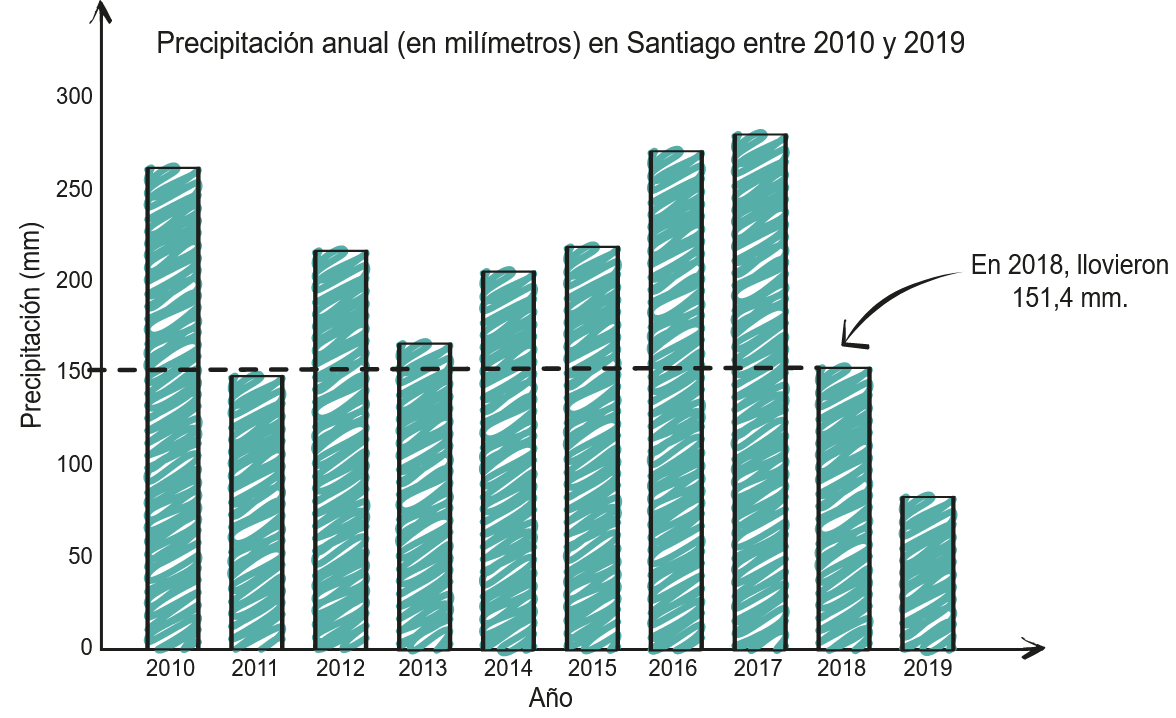


Figura 1: Gráfico de barras correspondiente a la precipitación anual en Santiago entre 2010 y 2019

Esto permite visualizar y comparar los datos. Por ejemplo, para observar en qué año hubo menos lluvia solo hay que ver la barra de menor altura.

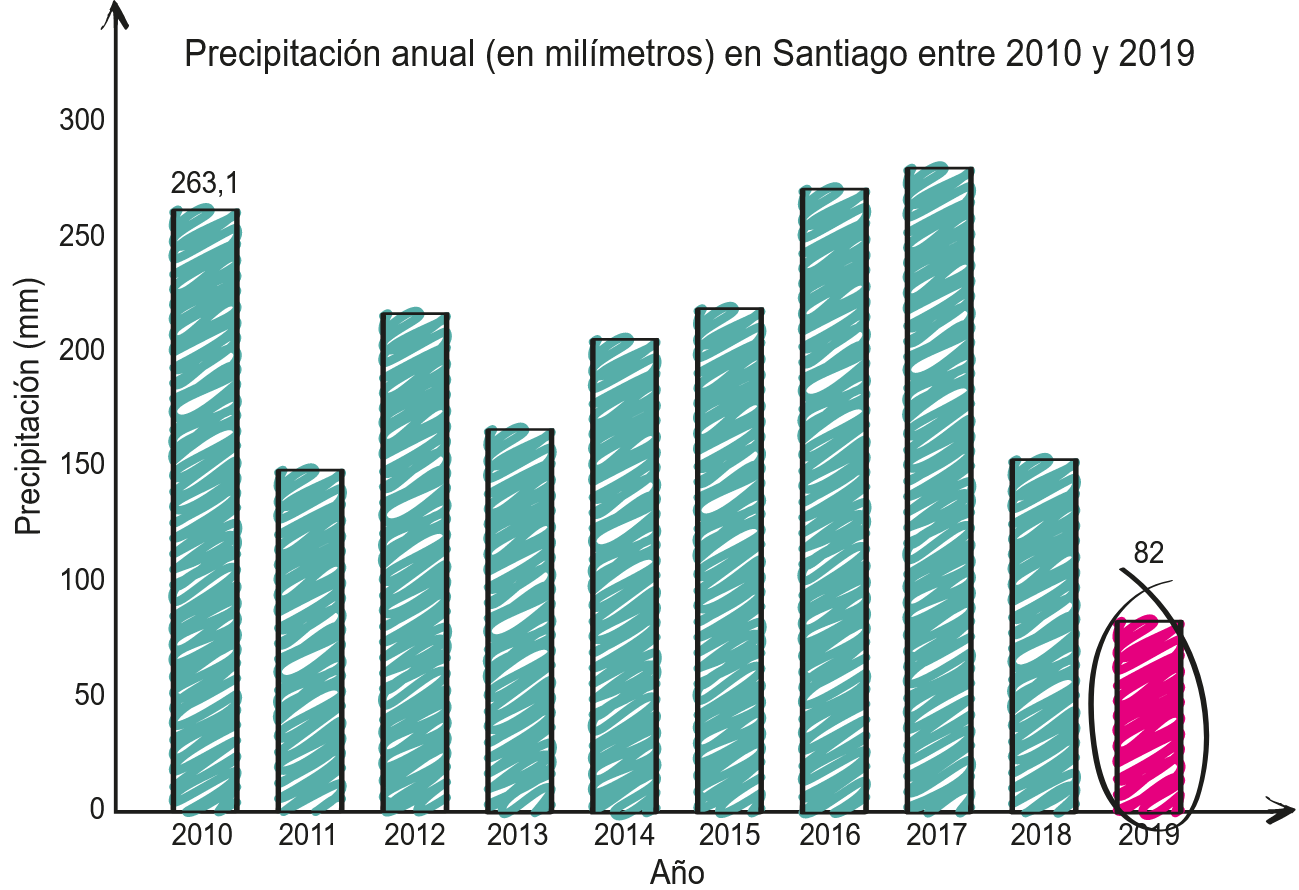


Figura 2: Gráfico de barras idéntico al de Figura 1 donde se resalta la barra asociada al año que presentó menos lluvia

.

Observación: En la notación de números decimales se puede utilizar tanto la coma como el punto. Aunque lo usual es que los escribamos con coma, en las calculadoras y en programas computacionales se suelen expresar con punto. En este curso, ocuparemos recursos digitales que, por defecto, denotan los decimales con punto.

**GRÁFICO DE LÍNEAS**

Los **gráficos de líneas** se suelen emplear para **ver tendencias, particularmente, a lo largo del tiempo.**

**Ejemplo**

El Producto Interno Bruto o PIB es un indicador macroeconómico que permite medir el valor monetario de la producción de bienes y servicios de consumo final de un país o región durante un período en particular.

Para observar cómo ha evolucionado el Producto Interno Bruto (PIB) per cápita de Chile a lo largo de la última década. A continuación podemos ver el gráfico de esta situación.

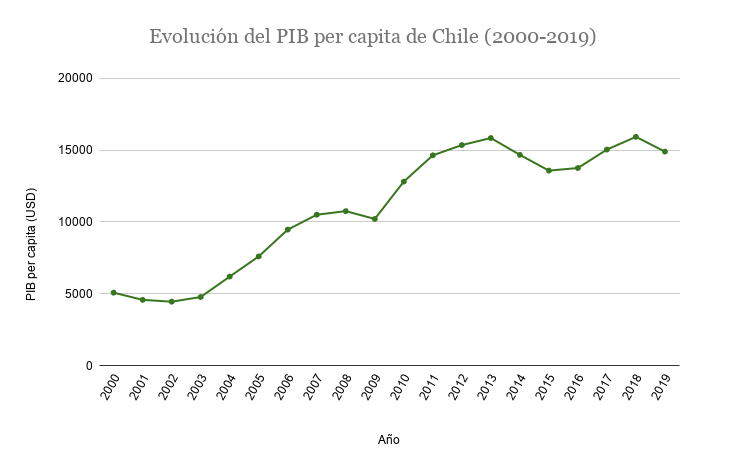


Figura 1: Evolución del PIB per cápita de Chile, extraído de Banco Mundial

**¿PARA QUÉ SIRVEN LOS GRÁFICOS DE BARRAS?**

Los **gráficos de barras** se suelen emplear cuando queremos **comparar cantidades,** considerando que las **barras son proporcionales a la magnitud.**

**Ejemplo**

Se puede utilizar para comparar el PIB per cápita de los países sudamericanos. A continuación, podemos ver la representación gráfica de lo antes mencionado, donde en este caso cada barra representa la magnitud del PIB per cápita de cada país. Por esto último, es importante ser cuidadoso si se modifica la escala del eje vertical, pues se puede distorsionar la información e inducir al error. Por ejemplo, si observamos el siguiente gráfico cuya escala vertical fue ajustada, se podría pensar que el PIB per cápita de Colombia es alrededor de la mitad del de Brasil, lo que no es cierto.

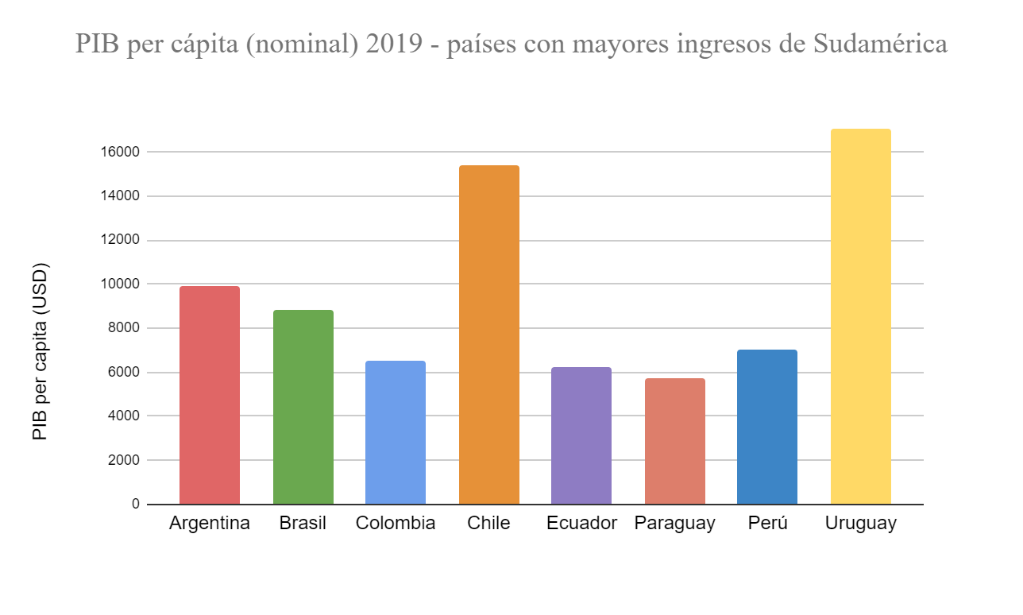


Figura 2: PIB per cápita 2019 de países con mayores ingresos en Sudamérica, sin ajustar.

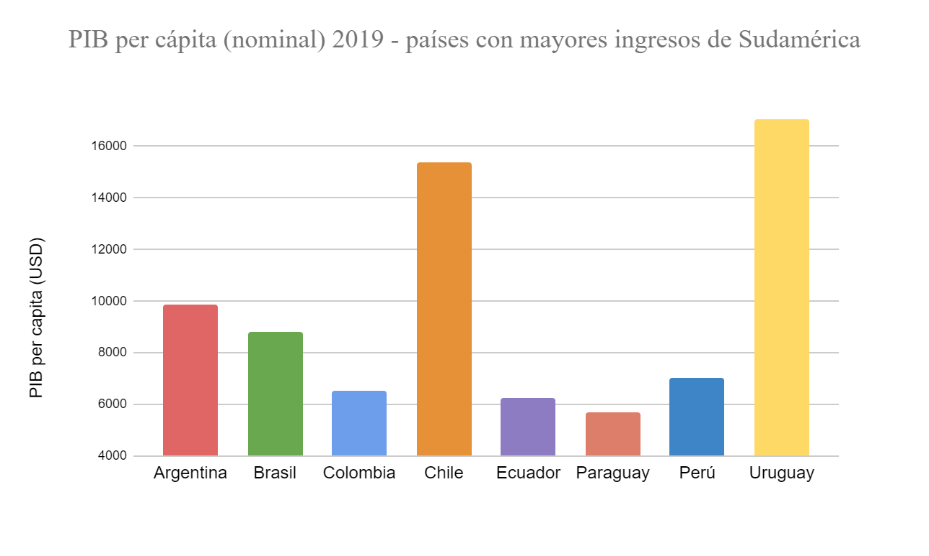
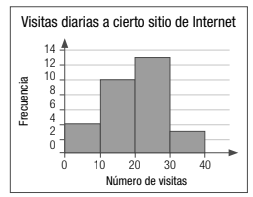


Figura 3: PIB per cápita 2019 de países con mayores ingresos de Sudamérica, con escala vertical ajustada.

Entonces, puede resultar conveniente **modificar la escala del eje vertical**. Esto ocurre, por ejemplo, **si se busca hacer más notorias las diferencias** entre los datos. Es importante tener en cuenta que **si modificamos la escala** de alguno de los ejes, esto debe aparecer **claramente indicado** en el gráfico.

**Histograma**  
  
Los histogramas se usan exclusivamente para variables cuantitativas y permiten observar la distribución de datos que se encuentran agrupados en intervalos.

  
Figura 4: Histograma

**FORMAS DE DISTRIBUCIONES**

Tal y como hemos visto en lecciones anteriores, los histogramas permiten mostrar la distribución que sigue cierta muestra de datos. Por ello, revisaremos algunas formas de distribuciones y la relación que tienen estas con las medidas de tendencia central.

**Distribución simétrica**

Una distribución es simétrica si la distribución de los datos al lado derecho de la mediana es igual a la del lado izquierdo, pero reflejada con respecto a un eje que pasa por la mediana, denotado en la figura 1 con una línea punteada.

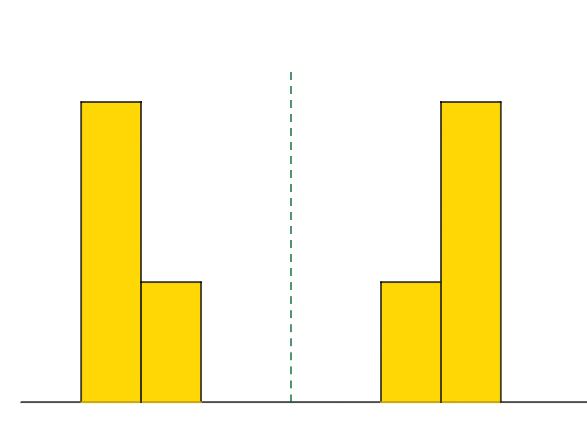
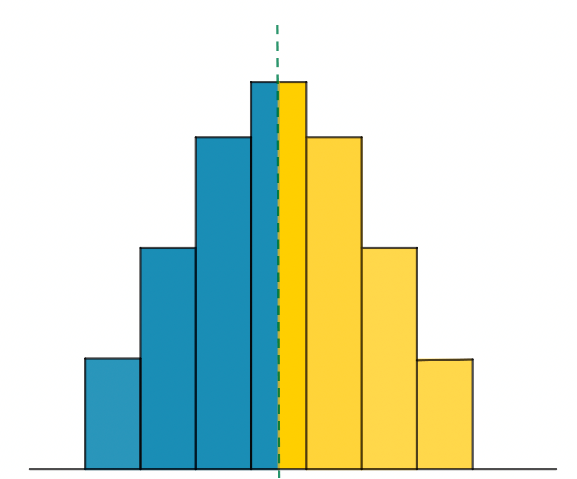
.

Figura 1: Distribuciones simétricas

**Distribución simétrica con los datos agrupados al centro**

Los datos se concentran en torno a un valor central, es decir, las mayores frecuencias se encuentran cerca de ese valor.

Las distribuciones de datos reales que son aproximadamente simétricas también las consideraremos como simétricas.

Por ejemplo, el peso de los estudiantes de un curso expresado en kilogramos.

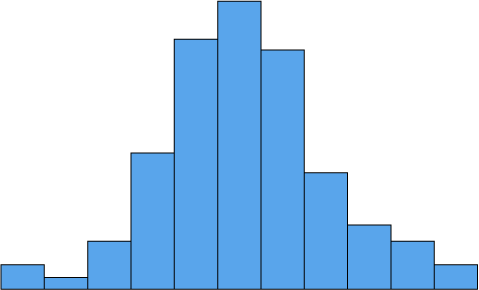


Figura 2: Distribución de datos simétrica con datos concentrados en el centro.

**PIRÁMIDE DE POBLACIÓN**

La **pirámide de población** es una forma gráfica de visualizar la distribución de las edades de una población. Cada mitad de la pirámide corresponde a un histograma que muestra la distribución de las edades de los hombres y las mujeres de dicha población.

Lo usual es que los largos de los intervalos correspondan a períodos de 5 años. Además, es común que la frecuencia esté dada en porcentajes, de forma de facilitar la comparación entre distintas poblaciones.

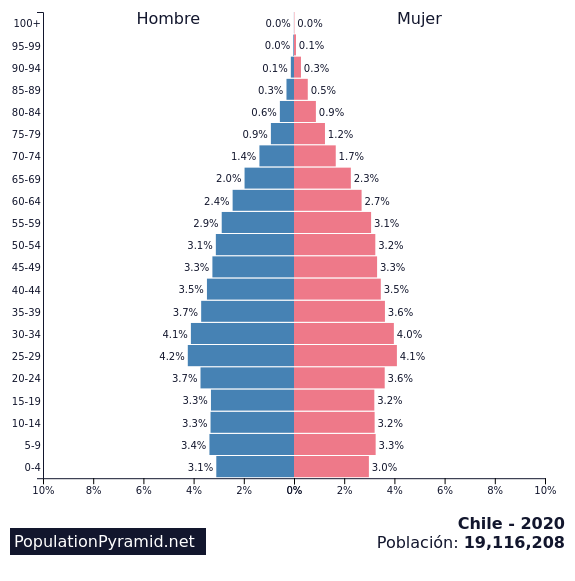


Figura 2: Pirámide de población chilena en 2020.

Los porcentajes asociados a cada barra son tanto para hombres como para mujeres respecto a la **población total**.

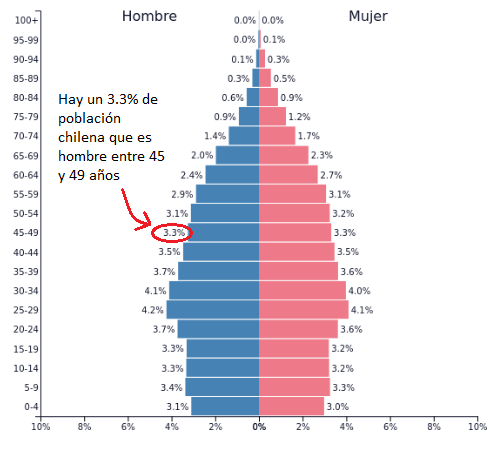


Figura 4: Pirámide de población resaltando al grupo de hombres entre 45 y 49 años, junto a su frecuencia relativa

Las pirámides poblacionales ayudan a estudiar muchos fenómenos, tales como el envejecimiento o que tan equilibrada es la población con respecto a hombres y mujeres. 

**SÍNTESIS**

* Para estudiar un fenómeno se debe buscar información y evaluar si los datos que tenemos son **suficientes** para entenderlo. Si no es suficiente, se deberán buscar o recolectar más datos que nos permitan comprender mejor la situación.
* Para explorar y analizar los datos, resulta conveniente organizarlos en **tablas** y **gráficos.**
* En los gráficos de barras hay una correspondencia entre **la longitud** de la barra y **el valor del dato.**
* Los gráficos de **líneas** se suelen emplear para ver **tendencias,** particularmente, **a lo largo del tiempo.**
* Los **gráficos de barras** son **útiles para comparar los valores** correspondientes a las **distintas categorías** estudiadas.
* Para comparar las **frecuencias de datos agrupados en intervalos**, conviene usar un **histograma**.
* Al construir gráficos es fundamental elegir bien el tipo de gráfico, usar títulos, etiquetas y leyendas simples e informativas, y ser consistente en el uso de ejes y escalas.
* La **pirámide de población** es una forma gráfica de visualizar la **distribución de las edades de una población tanto para hombres como para mujeres.** Donde una mitad de la pirámide corresponde a un histograma que muestra la distribución de las edades de los hombres y la otra a la de las mujeres **respecto a la población total.**