

# Data Visualization and Data Analysis 2026-1s

Area Académica de Ciencias Básicas y Modelado

Hugo Escobar

# Introducción al Curso

## Parte 1

1. Presentación del docente y estudiantes.
2. Objetivos y Resultados de Aprendizaje.
3. Contenido temático.
4. Lineamientos del curso.
5. Cronograma.
6. Asistencia.
7. Evaluación.
8. Bibliografía.
9. Recursos de Apoyo.

## SÍLABUS DE LA ASIGNATURA

(Data Visualization & Data Analysis)  
010867 y 01133

INFORMACIÓN DE LA ASIGNATURA					
Facultad	Ciencias Naturales e Ingeniería				
Área o Unidad Académica responsable	Ciencias Básicas y Modelado				
Nombre de la asignatura	Data Visualization and Data Analysis				
Código de la asignatura	010867 y 011339	Grupo	1, 2 y 5	Periodo académico	2025-25
Lugar (aulas y/o enlace sala sincrónica)	Aula de Clase	Intensidad horaria de trabajo autónomo semanal	4		
Intensidad horaria de trabajo presencial semanal	2	Intensidad horaria de trabajo presencial total del periodo	32		
Horario y lugar de atención de monitorías (si aplica)	Módulo 6 – Oficina 501				
Actividad académica	Clase teórico-práctica				



El curso es de modalidad **teórico/práctica**: los estudiantes **deberán asistir a por lo menos el 70% de las sesiones**

La **no presentación** de actividades de evaluación deben ser justificados por el estudiante ante el profesor dentro de un término no superior a cinco (5) días hábiles siguientes. El profesor dentro de su autonomía aceptará o no la justificación.



## Asistencia

**ARTÍCULO 42º. Asistencia a clases.** La asistencia a las actividades académicas programadas es un derecho del estudiante y se tomará con fines de acompañamiento académico. La asistencia a evaluaciones es de carácter obligatorio. Tratándose de talleres, laboratorios, salidas de campo y en general, actividades de orden práctico, el estudiante deberá asistir por lo menos al 70 % de dicha programación.

**ARTÍCULO 43º. Justificación de la inasistencia.** Los casos de inasistencia a evaluaciones, talleres, laboratorios, salidas de campo y en general actividades de orden práctico, deben ser justificados por el estudiante ante el profesor de la asignatura, dentro de un término no superior a cinco (5) días hábiles siguientes a la fecha de la inasistencia. El profesor dentro de su autonomía aceptará o no la justificación, la cual, si es aceptada, dará la oportunidad al estudiante de presentar los trabajos o evaluaciones que se hayan realizado durante su ausencia y adoptar las medidas que estime pertinentes en relación con la programación de orden práctico. La justificación **no aceptada** dará lugar a una calificación de cero punto cero (0.0) en la evaluación o actividad desarrollada.

### **1. Evaluación de contenidos y competencias**

Quices (en clase o en extraclase por medio de AVATA), tareas y participación.

### **2. Desarrollo de un proyecto de datos**

En equipos de trabajo los estudiantes definirán un tema de interés y desarrollarán un proyecto de visualización y análisis de datos y la presentación de resultados. El proyecto se desarrollará en tres fases.

## Fase 1

El equipo definirá un tema de interés y realizará una **consulta de fuentes de información** que contengan **visualizaciones** y análisis de datos relacionados con el tema escogido. Además, **descargará** una (o varias) base(s) de datos relacionada(s) con el tema escogido. Se realizará una entrega inicial en la semana 5 (última del corte 1).

## Fase 2

El equipo identificará el tipo de datos (cuantitativos, cuantitativos). Adelantará un **análisis exploratorio de datos**: primeras visualizaciones y estadísticas, con sus respectivos análisis. Empezará a plantearse un conjunto de preguntas y propósitos que buscará resolver con sus datos. Entregará un informe escrito en la semana 10 (última del corte 2).

# Evaluación

## Fase 3

El equipo consolidará sus visualizaciones, estadísticas y el análisis de sus datos, concluyendo adecuadamente y resolviendo las preguntas planteadas. Además: presentará sus resultados siguiendo una metodología de **storytelling** o narración de datos y en un **póster** (datafolio) en las semanas 15 y 16.



Los tres cortes tendrán igual ponderación para la nota definitiva y se evaluarán de la siguiente manera:

<b>Actividad</b>	<b>Corte 1</b>	<b>Corte 2</b>	<b>Corte 3</b>
<i>Evaluación de contenidos y competencias</i>	60%	50%	40%
<i>Proyecto de datos</i>	40%	50%	60%

### Libros guía

- Cairo, A. (2019). *How Charts Lie: Getting Smarter about Visual Information*. W. W. Norton & Company.
- Knafllic, C. N. (2015). *Storytelling with Data: A Data Visualization Guide for Business Professionals*. John Wiley & Sons.
- Knafllic, C. N. (2020). *Storytelling with Data: Let's practice!*. John Wiley & Sons.
- González, L. (2018). *Análisis exploratorio de datos: Una introducción a la estadística descriptiva y probabilidad*. Universidad Jorge Tadeo Lozano.
- Tufte, Edward R. (1990). *Envisioning Information*. Graphics Press.
- Yau, N. (2011). *Visualize This: The FlowingData Guide to Design, Visualization, and Statistics*.

## Software y recursos online

- Excel, PowerBI ([Licencia Office 365 desde la cuenta de UTADEO](#)) 
- Hojas de cálculo de Google
- [CODAP](#)
- [Tableau Public](#)

## Otros recursos

- [Tableau: Visualization of the day](#)
- [Gapminder](#)
- [Information is beautiful](#)
- [The who of data visualization](#)
- [Flowing data](#)
- [Dataviz project](#)

Aspirantes | Estudiantes | Egresados

Correo | Biblioteca | Utadeo Virtual | ENTRA | Portal de Servicios en línea | Idiomas | Publicaciones | AVATA | Uconsultoría | Unitadeístas | HJUT 106.9 FM | CrossmediaLab



f X Instagram YouTube LinkedIn Dailymotion

Líneas de emergencia



Servicios

Utadeo | Admisiones | Sedes | Programas | Educación Continua | Facultades | Internacionalización | Investigación



## AVATA

Ambiente Virtual de Aprendizaje Tadeista

SOY PROFESOR



SOY ESTUDIANTE



TIC PARA LA EDUCACIÓN



CONDICIONES SERVICIO AVATA



CONTÁCTANOS



Conoce los servicios que te ofrece la **asesoría** para trabajos de grado.

**ESTADO DEL ARTE**  
Estructura la base teórica de tu investigación

**ARTÍCULO ACADÉMICO**  
Conoce la estructura y forma básica de un artículo académico

**MENDELEY**  
GESTORES BIBLIOGRÁFICOS  
Uso y gestión de referencias con Mendeley y Zotero

# Introducción al Curso

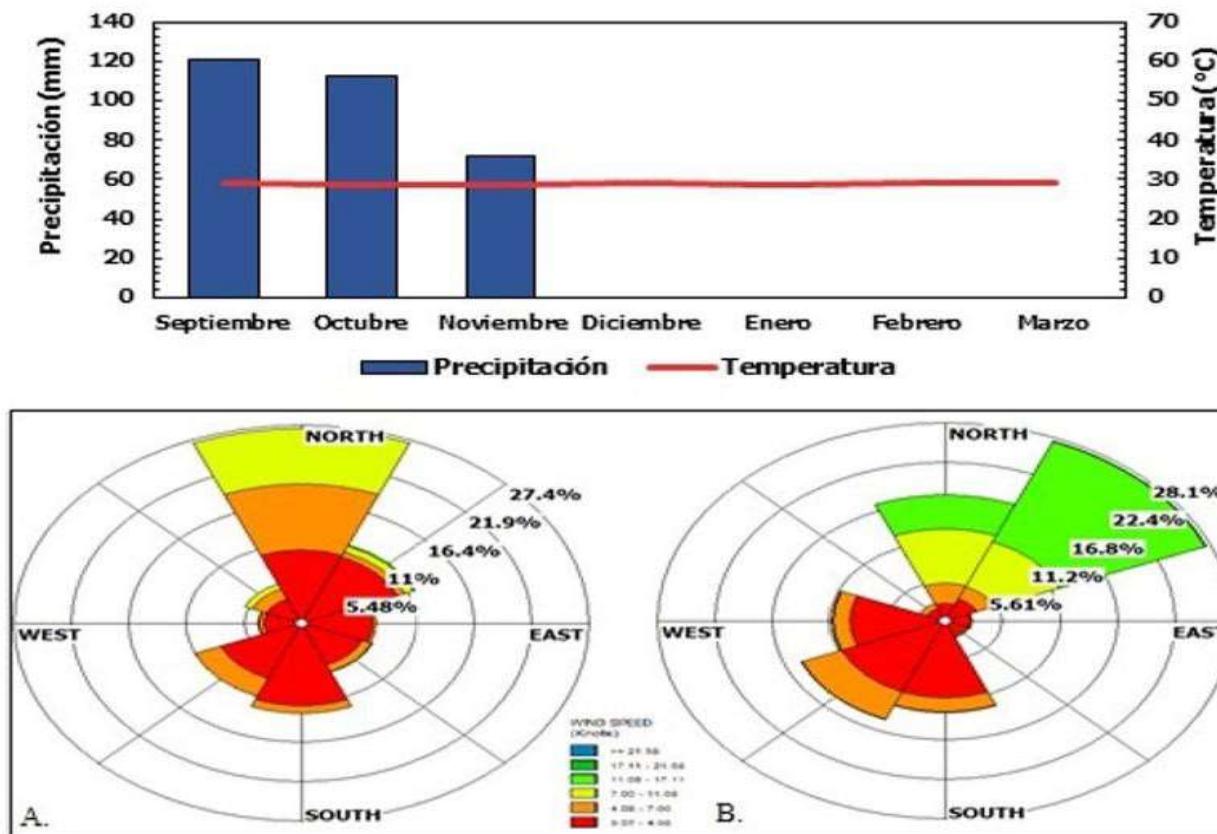
## Parte 2

# ¿Qué es la visualización y el análisis de datos?

¿Qué piensan que es la visualización y el análisis de datos?



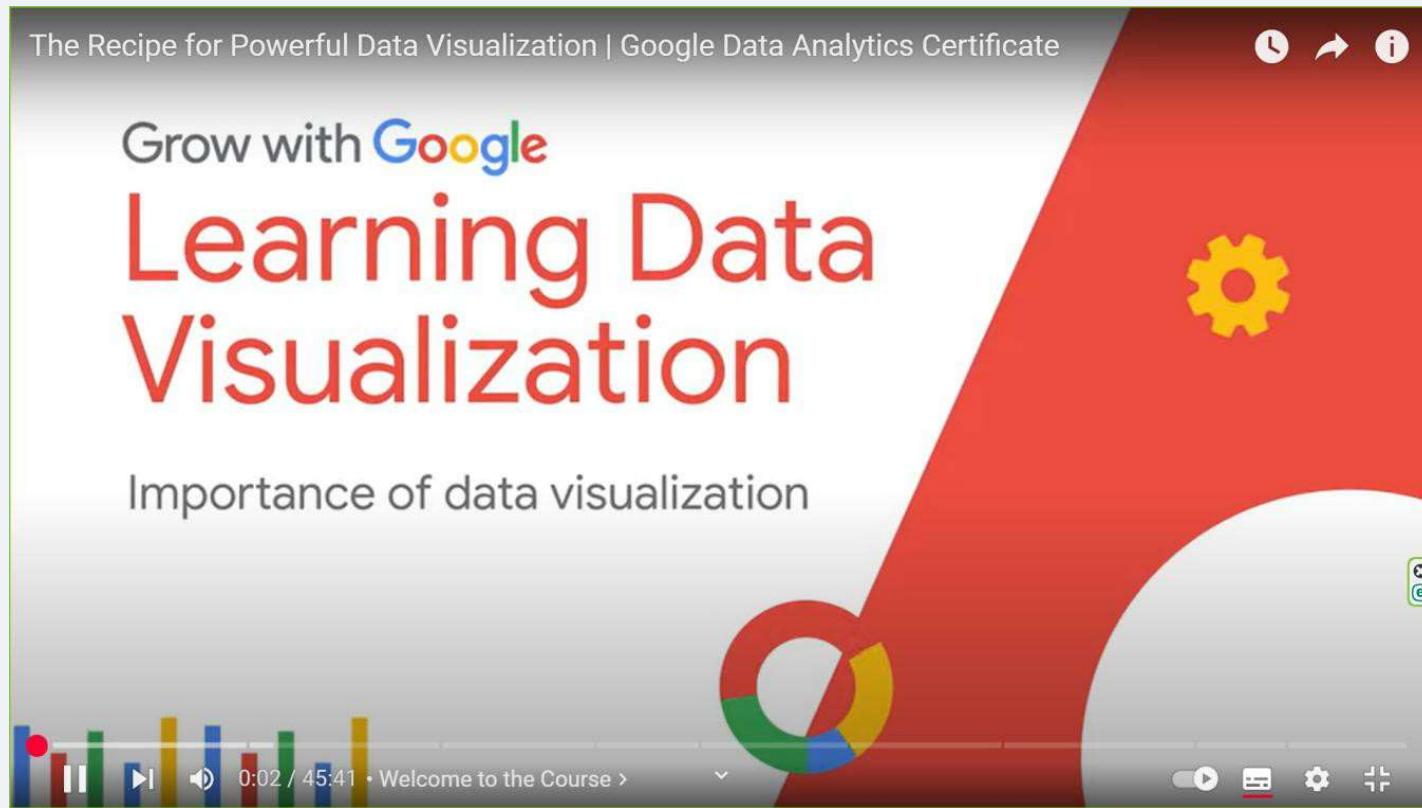
# ¿Qué es la visualización y el análisis de datos?



**Figura 3.** Climograma para la zona costera del departamento del Magdalena en el periodo entre septiembre del 2019 y marzo del 2020 y campos de vientos para la estación meteorológica del aeropuerto Simón Bolívar de la ciudad de Santa Marta en (A) noviembre del 2019 y (B) febrero y marzo del 2020

# ¿Qué es la visualización y el análisis de datos?

Video: The Recipe for Powerful Data Visualization (0:00-6:35)



t: 15 min; S: 70 min

# Factores de éxito en la Visualización de Datos

1. Visualización de Datos es la representación gráfica y presentación de los datos.
2. Es una imagen que sea fácil de entender para los demás.
3. Una visualización bien hecha tiene el poder de cambiar las mentes de las personas.
4. Regla al crear una visualización:

**5 s**

La audiencia debe saber exactamente lo que estas mostrando

**10 s**

La audiencia debe comprender la conclusión a la que llega la visualización.

5. Elementos claves en una visualización: Datos +Objetivos + Historia + Forma visual.

# ¿Qué es la visualización y el análisis de datos?

En este curso nos enfocaremos en la **visualización de datos** y el **análisis visual y exploratorio de datos**. No profundizaremos en todo el campo de la analítica de datos.

# ¿Qué es la visualización y el análisis de datos?



Rev. Acad. Colomb. Cienc. Ex. Fis. Nat. 46(179):406-425, abril-junio de 2022  
doi: <https://doi.org/10.18257/raccefyn.1578>

## Ciencias Naturales

Artículo original

### Distribución espacial y temporal de microplásticos flotantes en aguas del Caribe central colombiano

Spatial and temporal distribution of floating microplastics in waters of the Colombian central Caribbean region

 **Andrés Franco-Herrera\***,  **Paula Irene Polanía-Zenner**,  **Carlos David Otálora-Rincón**,  
 **Paulo César Tigreros-Benavides**

Grupo de Investigación Dinámica y Manejo de Ecosistemas Marino Costeros DIMARCO, Universidad de Bogotá “Jorge Tadeo Lozano”, Santa Marta, Colombia

Artículo de posesión para admisión de Andrés Franco-Herrera, como miembro correspondiente a la Academia Colombiana de Ciencias Exactas, Físicas y Naturales

Texto simple  
**91%**

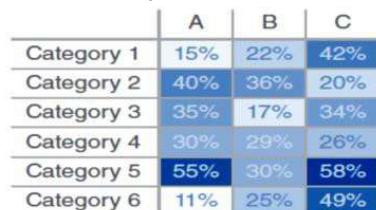
Simple text

Tabla

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

Table

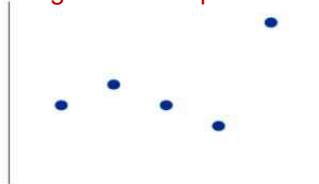
Mapa de calor



Heatmap

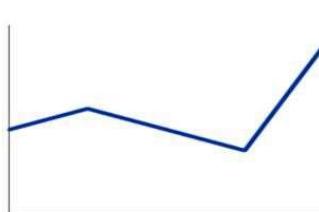
# Ejemplos de visualización de datos

Diagrama de dispersión



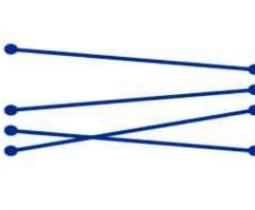
Scatterplot

Línea



Line

Gráfico de pendientes



Slopegraph

Barras verticales



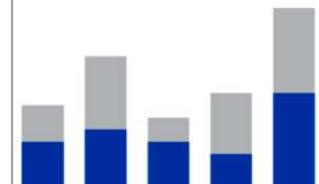
Vertical bar

Barras horizontales



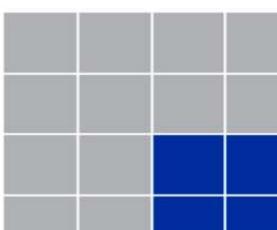
Horizontal bar

Barras verticales apiladas



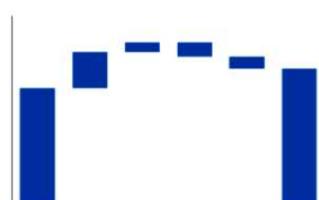
Stacked vertical bar

Gráfico de área



Square area

Cascada

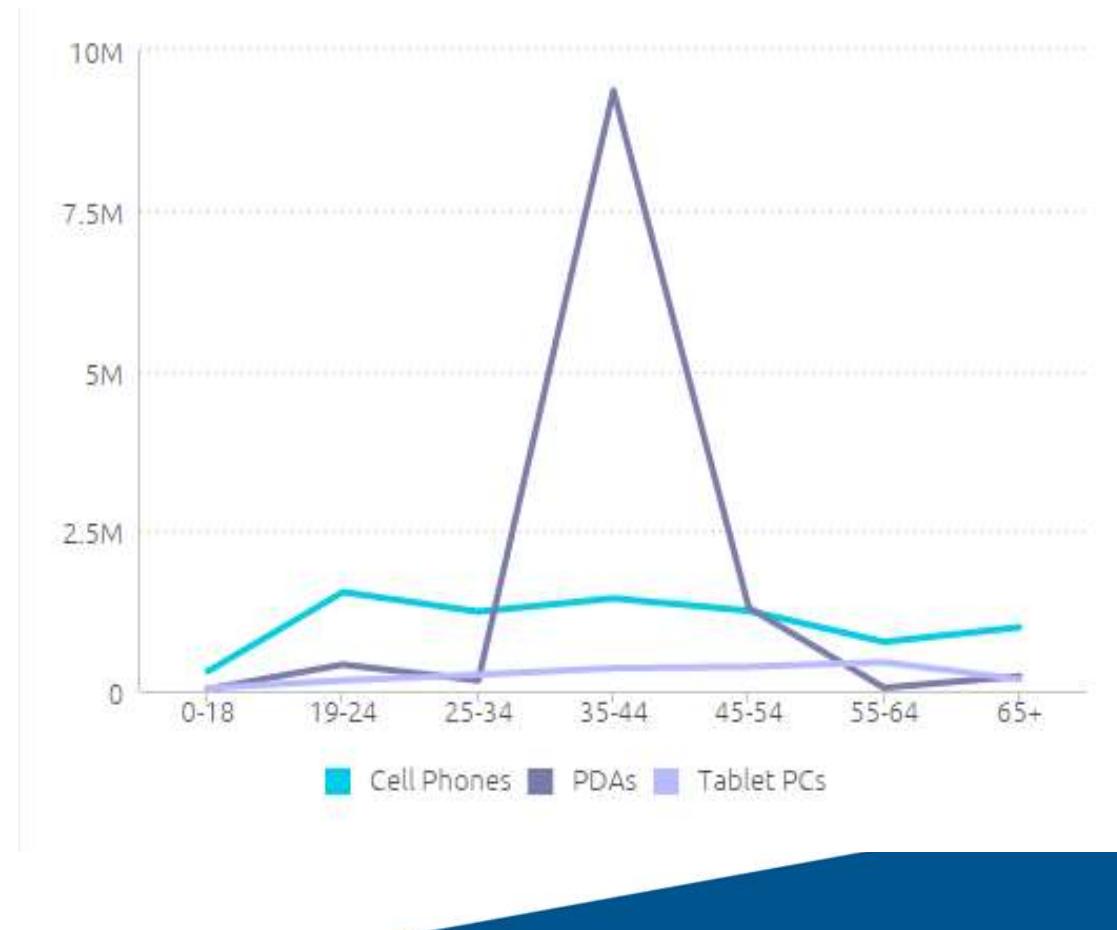


Waterfall

# 1. Los gráficos de líneas

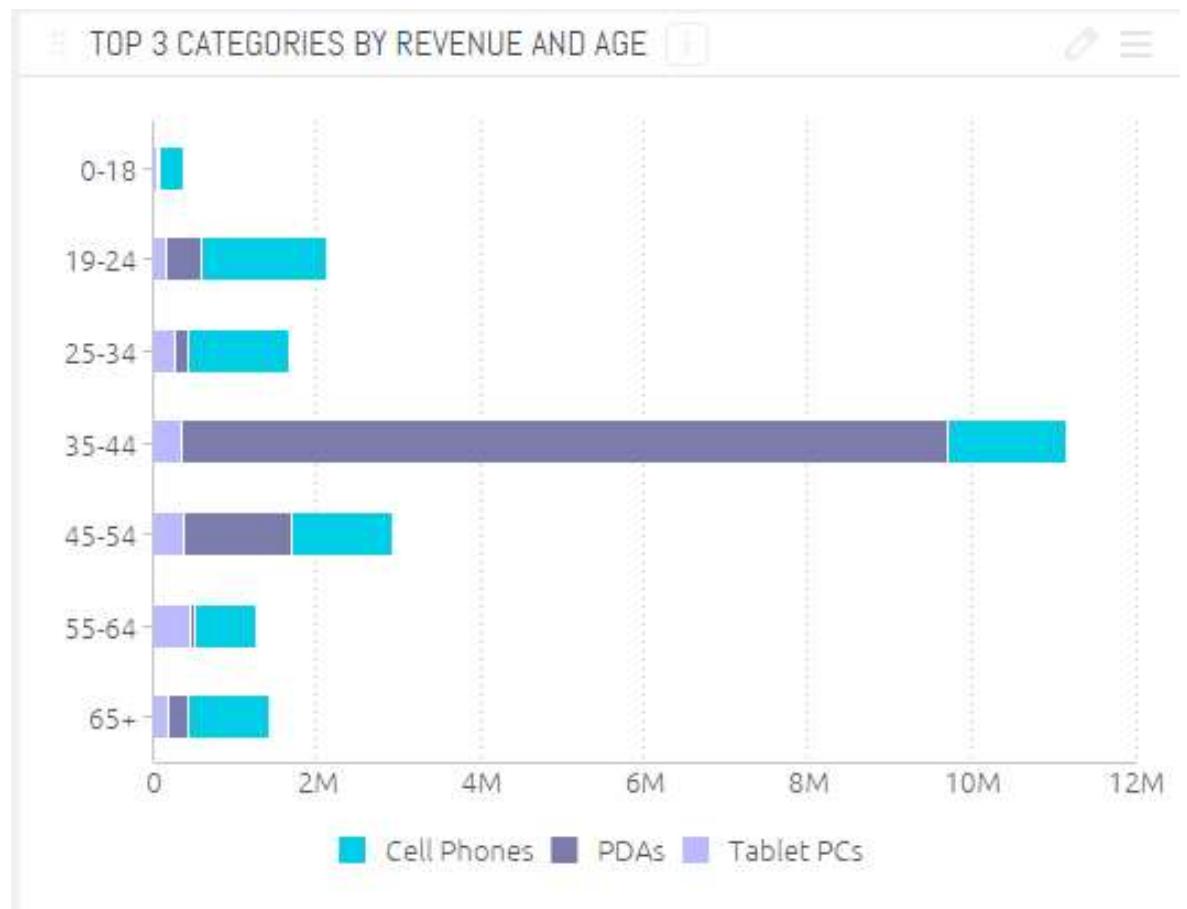
Los gráficos de líneas son muy populares en diversos casos, ya que muestran una tendencia general de forma rápida y concisa, siendo difícil malinterpretar la información. En particular, son útiles para representar tendencias de diferentes categorías durante el mismo período, lo que facilita la comparación.

Por ejemplo, este gráfico visualiza las cifras de ventas por grupo de edad para tres líneas de productos diferentes:



## 2. Los gráficos de barras

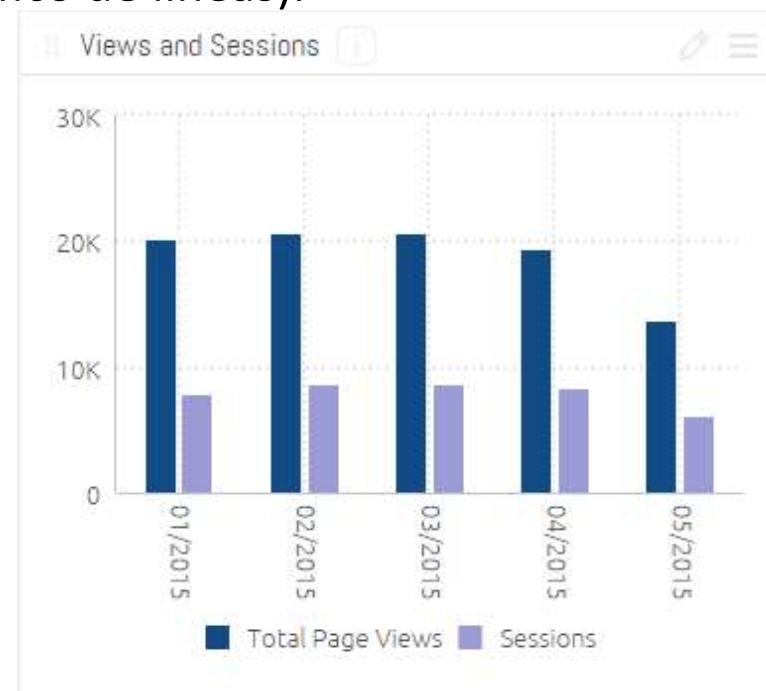
Son excelentes para comparar varios valores diferentes, especialmente cuando algunos de ellos se dividen en categorías con códigos de colores. Para ilustrar la diferencia entre esto y un gráfico de líneas, tomemos la misma información anterior y visualicémosla como un gráfico de barras:



### 3. Los gráficos de columnas

Conviene usar gráficos de columnas para comparar valores en paralelo (Uno al lado del otro). También se pueden usar para mostrar cambios a lo largo del tiempo, aunque es más conveniente hacerlo cuando se desea destacar las cifras totales en lugar de la forma de la tendencia (lo cual es más efectivo con un gráfico de líneas).

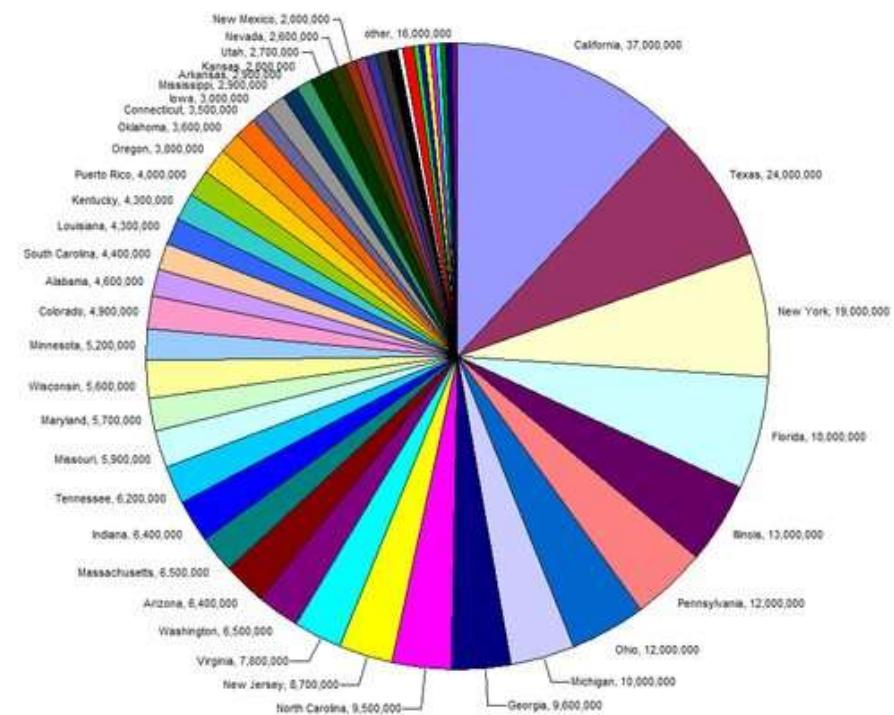
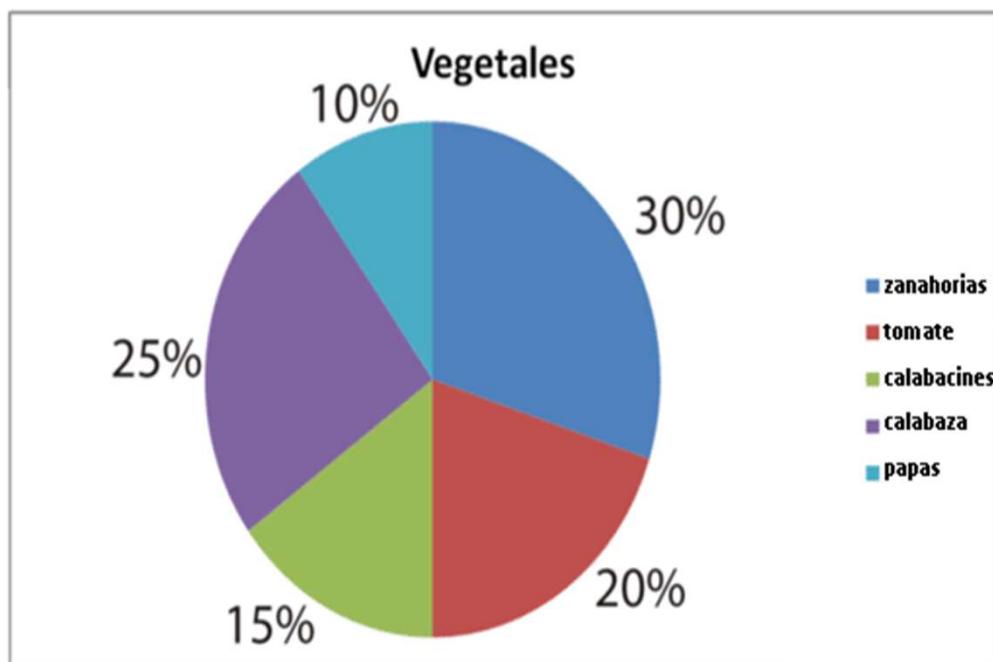
*Total de visitas a páginas web frente a las sesiones en una serie de fechas. Las cifras no varían mucho de un día para otro, por lo que un gráfico lineal no revelaría nada revelador en términos de tendencias; más bien, la información relevante aquí es el número concreto de visitantes al sitio web cada día.*



## 4. Los gráficos circulares

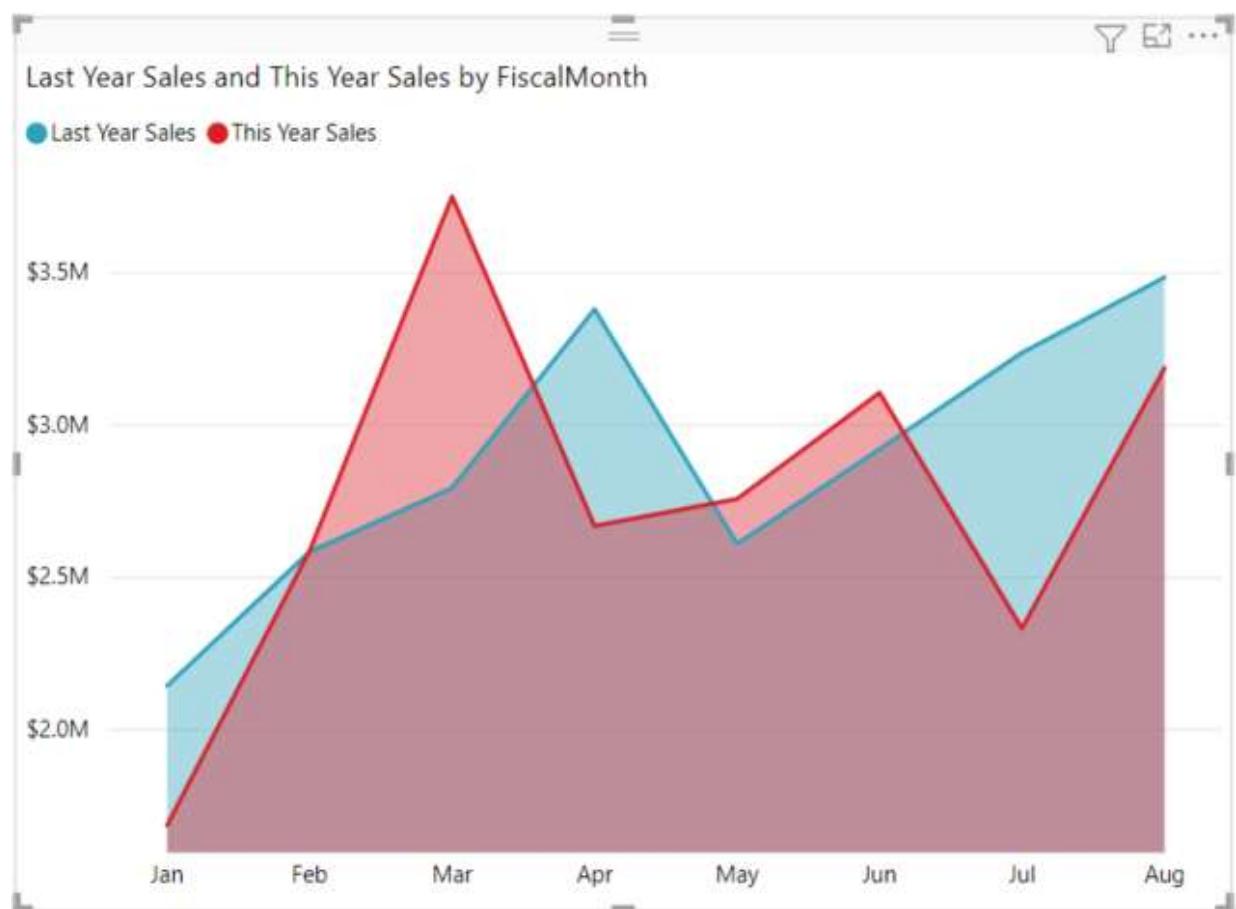
Muestran fácilmente la proporción que cada valor representa del total. Son mucho más intuitivos que simplemente mostrar porcentajes que suman el 100 %.

Para que un gráfico circular sea efectivo, se necesitan **seis categorías o menos**. Si se superan, el gráfico estará demasiado saturado y los valores serán demasiado imprecisos como para obtener información útil.



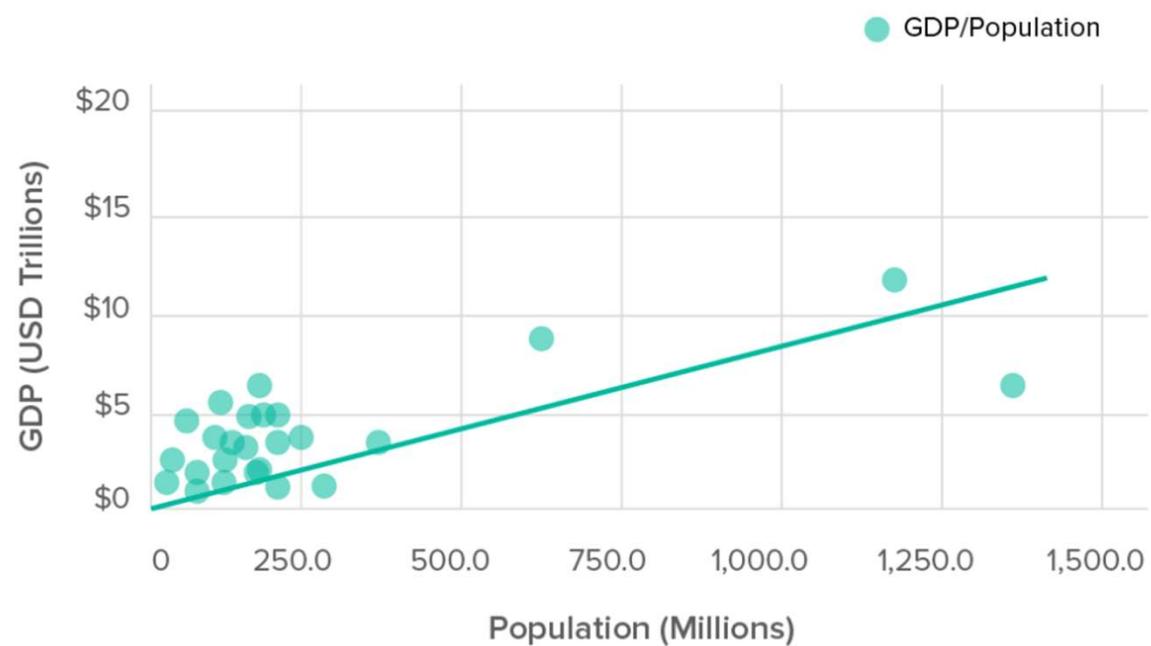
## 5. Los gráficos de área

Los gráficos de área comparan proporciones y son útiles porque dan una idea del volumen general, así como de la proporción del mismo que ocupa cada categoría. Ayudan a visualizar tendencias de datos y a resaltar los cambios de volumen a lo largo del tiempo. Los gráficos de áreas enfatizan la magnitud del cambio a lo largo del tiempo y hacen hincapié en el valor total en una tendencia. Se basa en el gráfico de líneas, con el área entre el eje y la línea llena de colores para indicar el volumen.



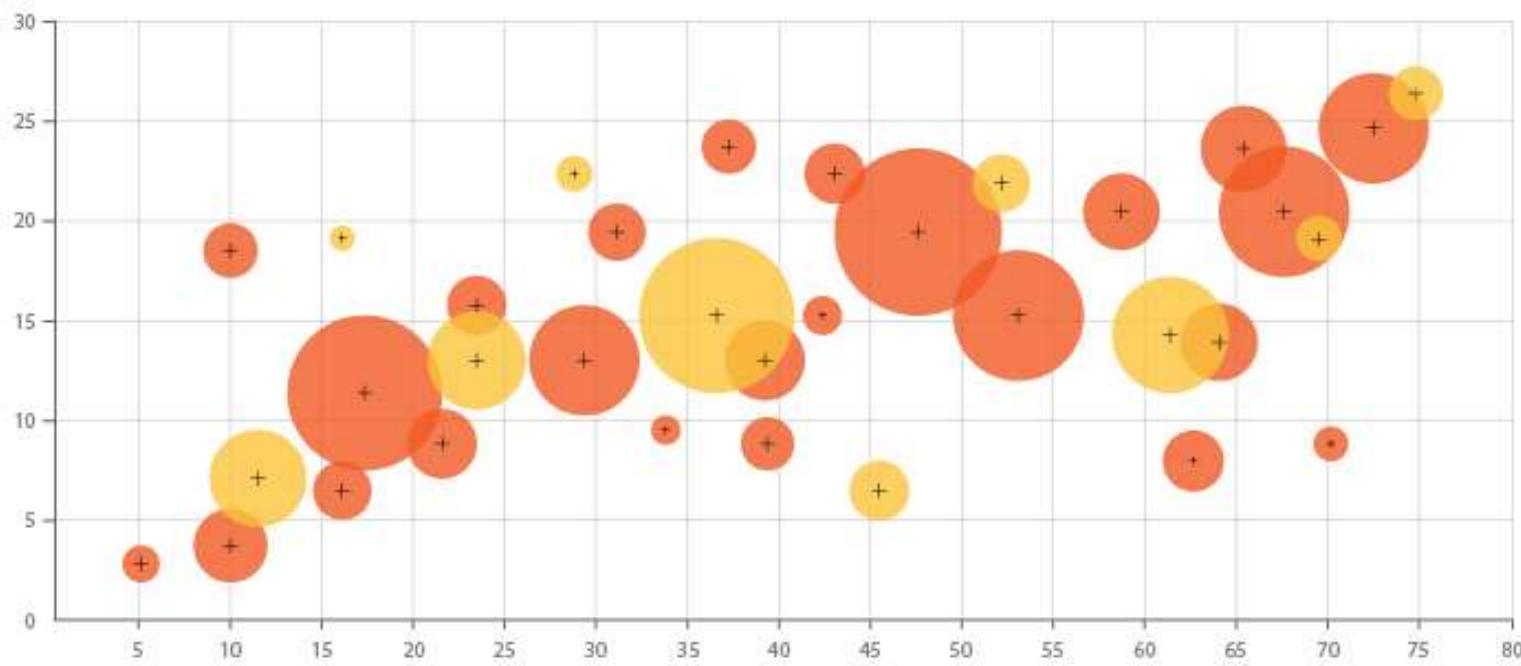
## 6. Gráficos de dispersión

Los gráficos de dispersión presentan categorías de datos por color del círculo y el volumen de los datos por tamaño del círculo; se utilizan para visualizar la distribución y la relación entre dos variables.



## 7. Gráficos de burbujas

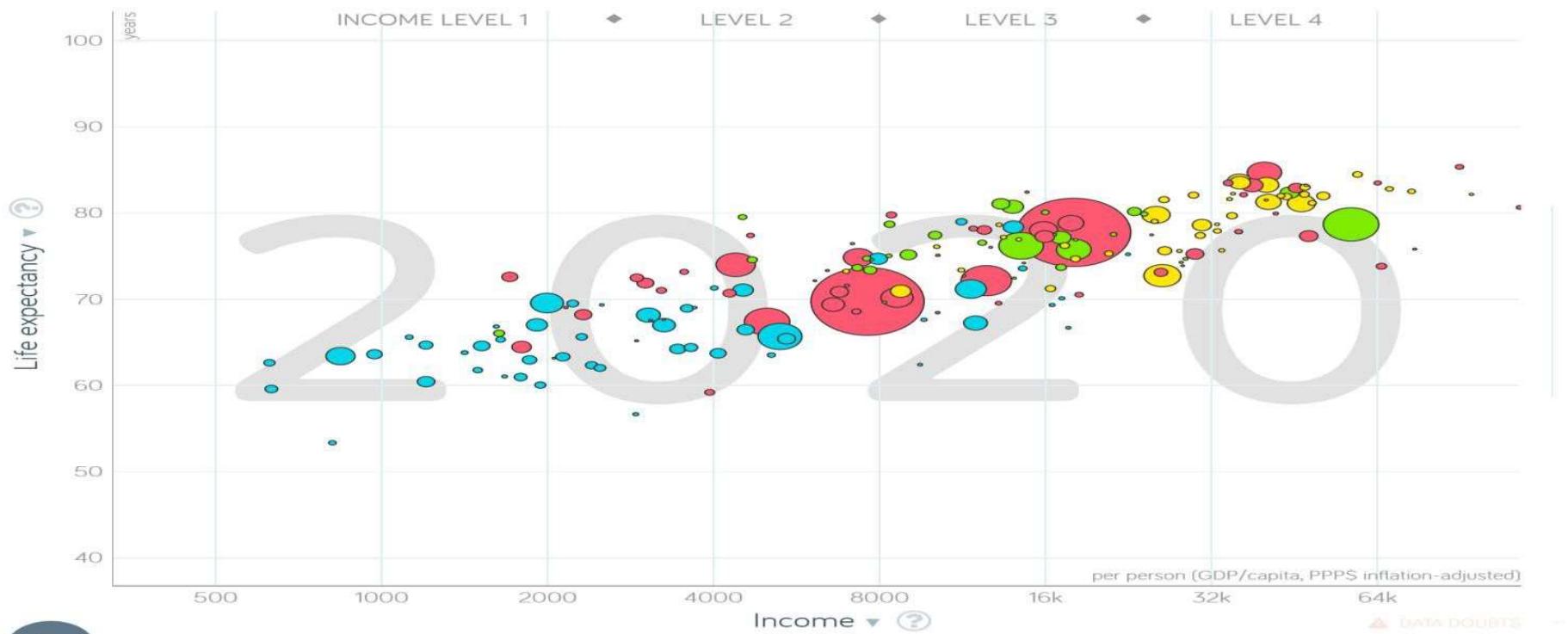
Los gráficos de burbujas representan la ponderación de los valores según el tamaño de la circunferencia del círculo. Sin embargo, se diferencian en que concentran muchos valores diferentes en un espacio reducido y representan una sola medida por categoría. Son útiles para demostrar la gran importancia de unas pocas categorías en comparación con un mar de categorías insignificantes.



*Al igual que un diagrama de dispersión, los gráficos de burbujas utilizan un sistema de coordenadas cartesianas para trazar puntos a lo largo de una cuadrícula donde los ejes X e Y son variables separadas*

# Ejemplos de visualización de datos

Gráfico de burbujas animado



Esperanza de vida VS Ingreso per cápita  
[Gapminder](#)



## 8. Las tablas dinámicas

Las tablas dinámicas no son las formas más hermosas o intuitivas de visualizar datos, pero son útiles cuando desea extraer rápidamente cifras clave mientras ve números exactos

The screenshot shows the Microsoft Excel ribbon with the 'Tablas dinámicas' tab selected. In the main workspace, a dynamic table is displayed with columns for Product, Quantity, and Sales Amount. The sales amount for the row 'Clavos' is highlighted. Below the table, a 'Suma de Monto venta por Vendedor' table is shown, listing sales amounts for Fernando, Juan, Laura, and Pedro, with a total of 29613000.

Producto	Cantidad	Monto venta
Martillo	32	\$ 995.000
Destornillador	32	\$ 1.720.000
Tubos	50	\$ 1.252.000
Sierra eléctrica	36	\$ 1.685.000
Clavos	6	\$ 1.663.000
Tornillos	49	\$ 1.014.000
Lentes de seguridad	10	\$ 420.000
Pintura	3	\$ 1.318.000
Brochas	19	\$ 1.401.000
Cinta	20	\$ 982.000
Tablas	39	\$ 1.735.000
Pintura	40	\$ 1.542.000
Lentes de seguridad	19	\$ 1.178.000
Destornillador	31	\$ 1.747.000
Tubos	22	\$ 1.174.000
Cinta	48	\$ 335.000
Tablas	27	\$ 1.347.000
Clavos	15	\$ 253.000
Tornillos	6	\$ 839.000
Martillo	1	\$ 713.000

Suma de Monto venta por Vendedor

Etiquetas de fila	Suma de Monto venta
Fernando	8123000
Juan	7899000
Laura	6511000
Pedro	7080000
Total general	29613000

Suma de Costo por Vendedor

Etiquetas de fila	Suma de Costo
Fernando	4754000
Juan	4202000
Laura	3557000
Pedro	4499000
Total general	17012000

Suma de Cantidad por Vendedor

Etiquetas de fila	Suma de Cantidad
Fernando	231
Juan	166
Laura	174
Pedro	153
Total general	724

Suma de Región por Vendedor

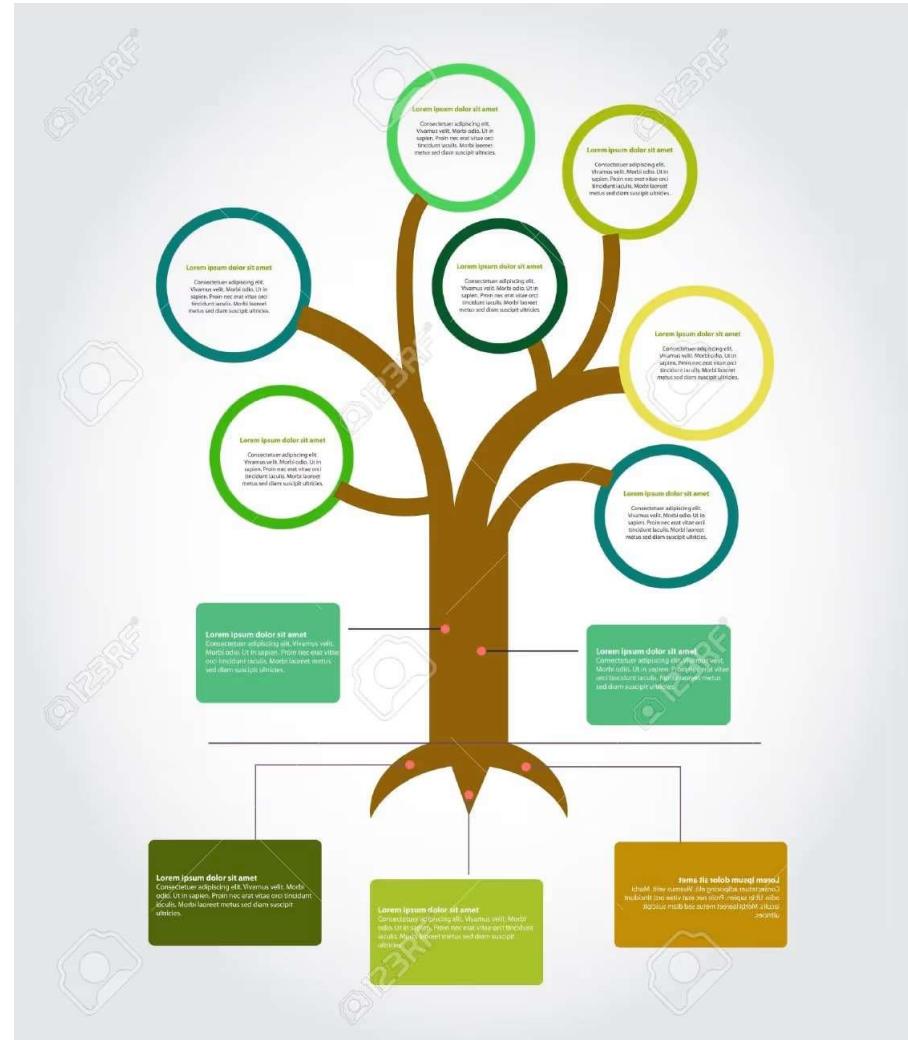
Etiquetas de fila	Suma de Monto venta
Fernando	8123000
Juan	7899000
Laura	6511000
Pedro	7080000
Total general	29613000

Tabla dinámica en blanco Cambiar datos de origen... Aceptar Cancelar

## 9. Los mapas de árbol

Los mapas de árbol son útiles para mostrar jerarquías y valores comparativos entre categorías y subcategorías, además de permitirle conservar los detalles mientras proyecta una idea instantánea de qué áreas son las más importantes en general.

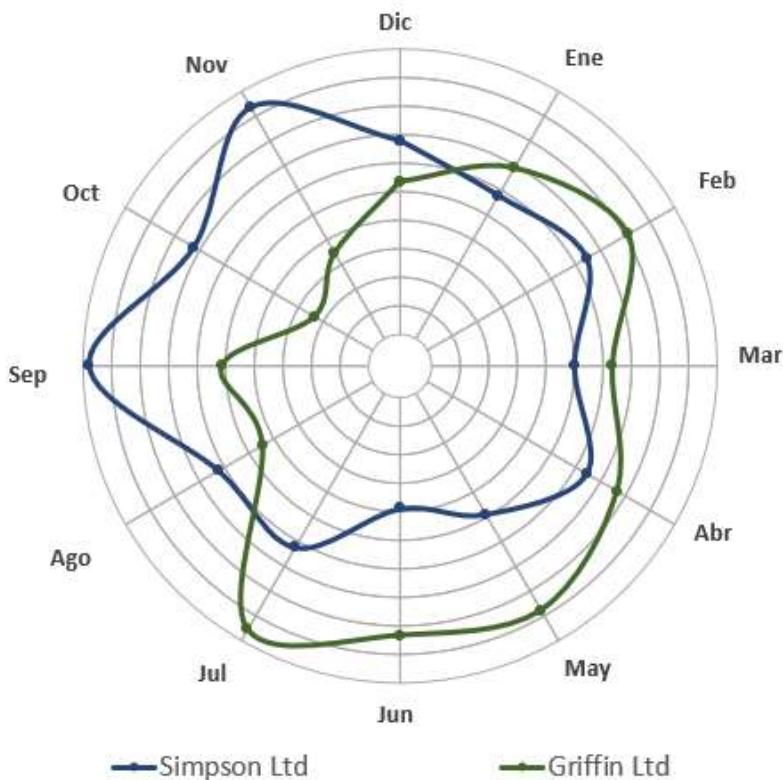
Esto se consigue anidando rectángulos con códigos de colores, ponderados para reflejar su participación en el total.



# 10. Los gráficos polares

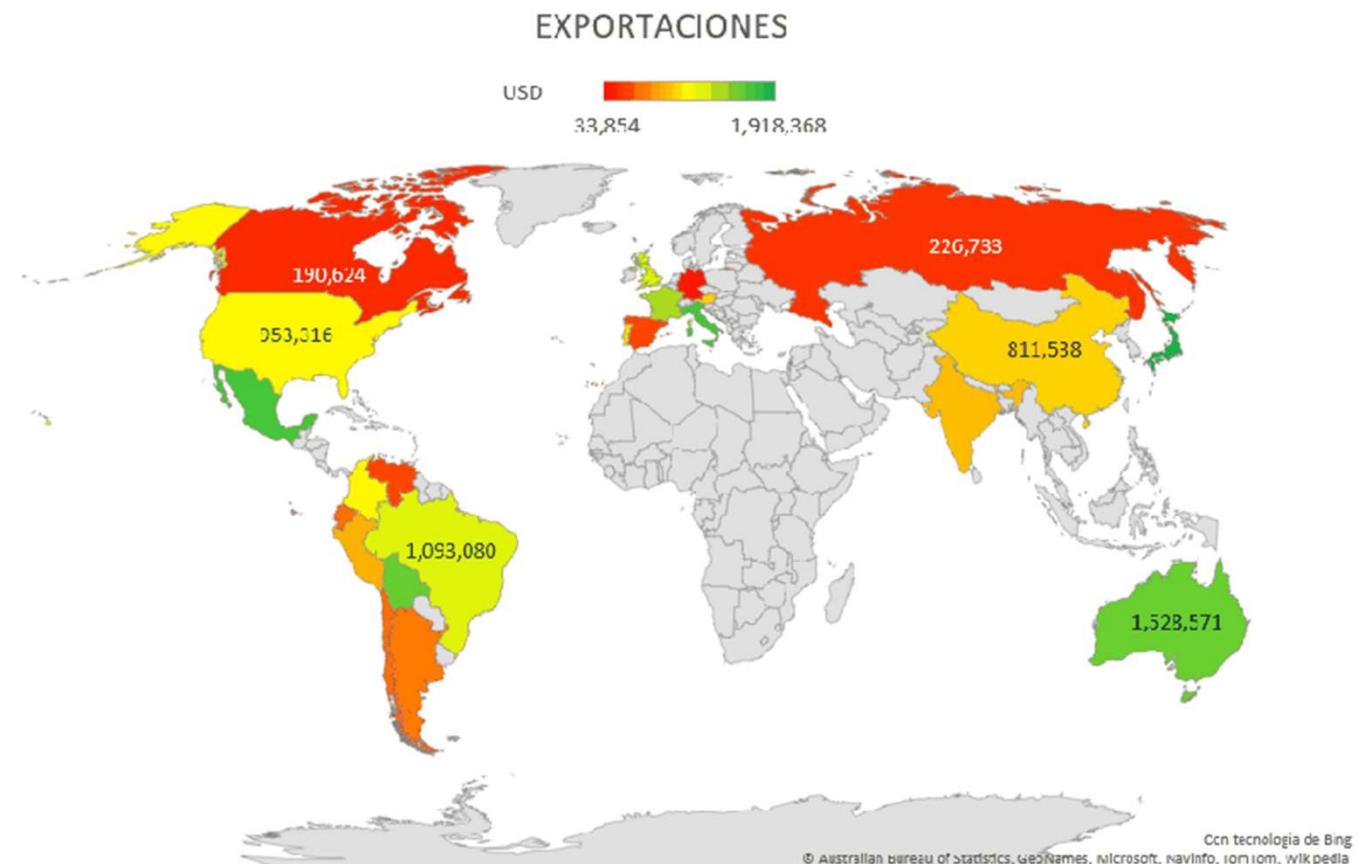
Un gráfico polar (o diagrama de área polar) es un tipo de gráfico circular. Sin embargo, en lugar de representar la proporción de cada valor en el total mediante el tamaño del ángulo, todos los sectores tienen ángulos iguales y el valor se muestra según su distancia desde el centro del círculo.

Puntuaciones de Satisfacción del Cliente en 2019: Simpson Ltd y Griffin Ltd



## 11. Los mapas de área

Estas visualizaciones de datos le permiten ver inmediatamente qué ubicaciones geográficas son más significativas para una determinada variable. Los datos se visualizan como puntos de color en un mapa; los valores se representan mediante el tamaño de un círculo.

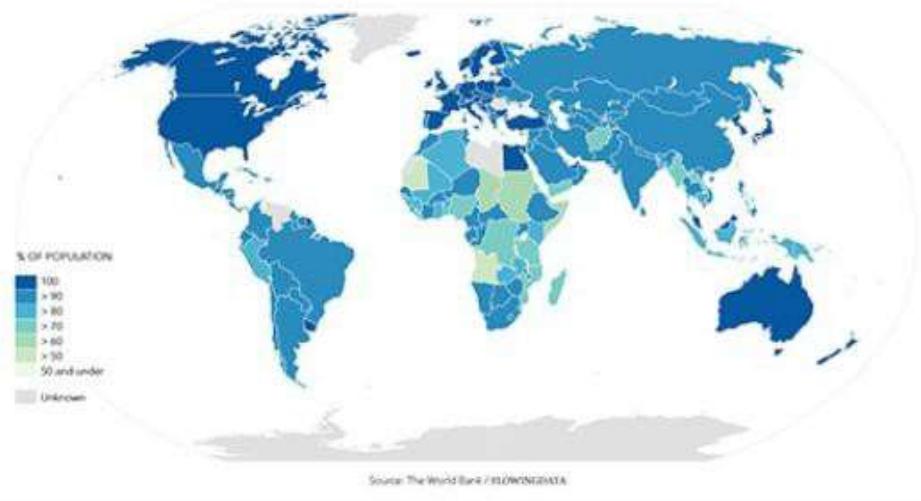


# Ejemplos de visualización de datos

## Mapas

### ACCESS TO IMPROVED WATER SOURCE

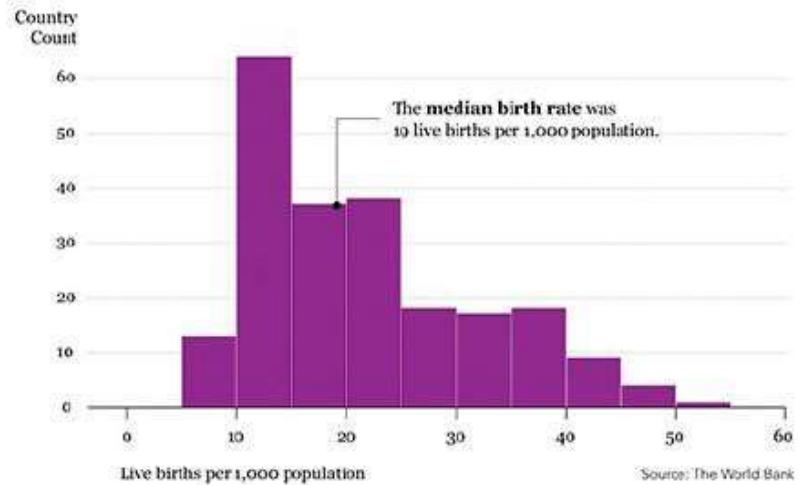
Most people have access to clean, potable water, but not everyone is so lucky. This map shows the percentage of urban population in each country that had "reasonable access" to an improved water source in 2009.



## Histogramas y distribuciones

### GLOBAL DISTRIBUTION OF BIRTH RATES

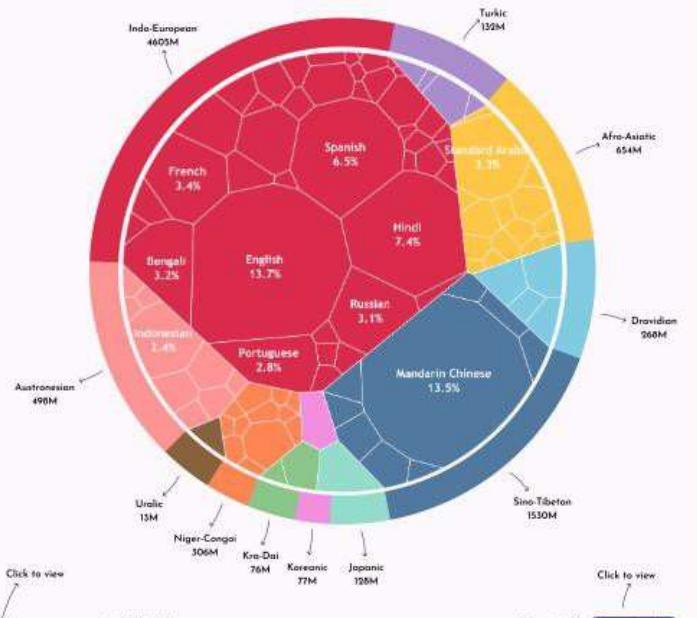
In 2008, most countries had birth rates less than 25 live births per 1,000 population. There are, however, many developing countries where women tend to bear more children.



## 100 Most Spoken Languages Around The World

Around the world, there are more than 7,000 languages in existence. Here's the 100 most spoken languages in the world, the number of total speakers for each language, and the origin tree that each language has branched out from.

The data comes from the 22nd edition of Ethnologue, the database of living languages of the world's population, detailing approximately 7,000 living languages in existence today. The data was published on February 15, 2020.



Click to view  
It comes as no surprise that English reigns supreme, with over 1 billion Total Speakers—or roughly 15% of the global population. Mandarin Chinese, Hindi, Spanish, and French round out the top five.

One reason these languages are popular is that they're widely and consistently used. Unfortunately, nearly 3,000 (about 40%) of all languages are at risk of being lost, or are already in the process of dying out today.

Indo-European languages have the widest spread worldwide. According to Ethnologue, the language family contains over 3 billion speakers in total. Interestingly, there are actually 1,396 Niger-Congo languages altogether, though only 12 are represented here.

The United Nations declared 2019 the International Year of Indigenous Languages (IYIL), with a resolution to continue fostering these languages and pass on their knowledge for future generations.

Data Source: ethnologue.com

#DiversityInData

Design: Nandini Pundit | @nandinipundit

View on Tableau Public



### Details

☆ 489 ○ 41,204

#DiversityInData | The visualization shows the 100 most spoken languages in the world, the number of total/native speakers for each language, and the origin tree that each language has branched out from. Data Published: Feb 2020.

Tableau: viz of the day



The world cannot be understood without numbers.  
And it cannot be understood with number alone.

—HANS ROSLING. *FACTFULNESS* 2018

[Tableau: viz of the day](#)



## Actividades Extraclase

Videos:

[Crime spotting \(6:45\)](#)

[Florence Nightingale \(3:41\)](#)

[200 countries, 200 years, in 4 minutes \(4:56\)](#)



**Evaluación**

[Preguntas estadísticas \(11:31\)](#)

[Preguntas estadísticas y no estadísticas \(11:37\)](#)



**Próxima  
sesión**

