

# Introducción a la visualización de datos

Maria Isabel Serrano G.  
maria-serrano@javeriana.edu.co

# Agenda



Que es visualización de datos



Presentación de la asignatura



Visualización y ciencia de datos



Visualización en proyectos de BI



Taller



¿POR QUÉ ESTÁS AQUÍ?

# Composición de algunas bebidas de café

- **Expreso:** la medida básica incluye 30 ml (una onza) de café concentrado.
- **Café americano:** una medida de expreso y tres medidas de agua.
- **Café late:** una medida de expreso, dos medidas de leche al vapor y una medida de espuma.
- **Capuchino:** una medida de expreso, 1.5 medidas de leche al vapor y 1.5 medidas de espuma de leche.

¿Cómo mostramos esto?  
¿Es atractivo? ¿Vende?

# Coffee drinks illustrated | Lokesh Dhakar



*“The ability to take data — to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it — that’s going to be a hugely important skill in the next decades”*

Hal Varian - Google’s chief economist

Hal Varian on how the Web challenges managers | McKinsey





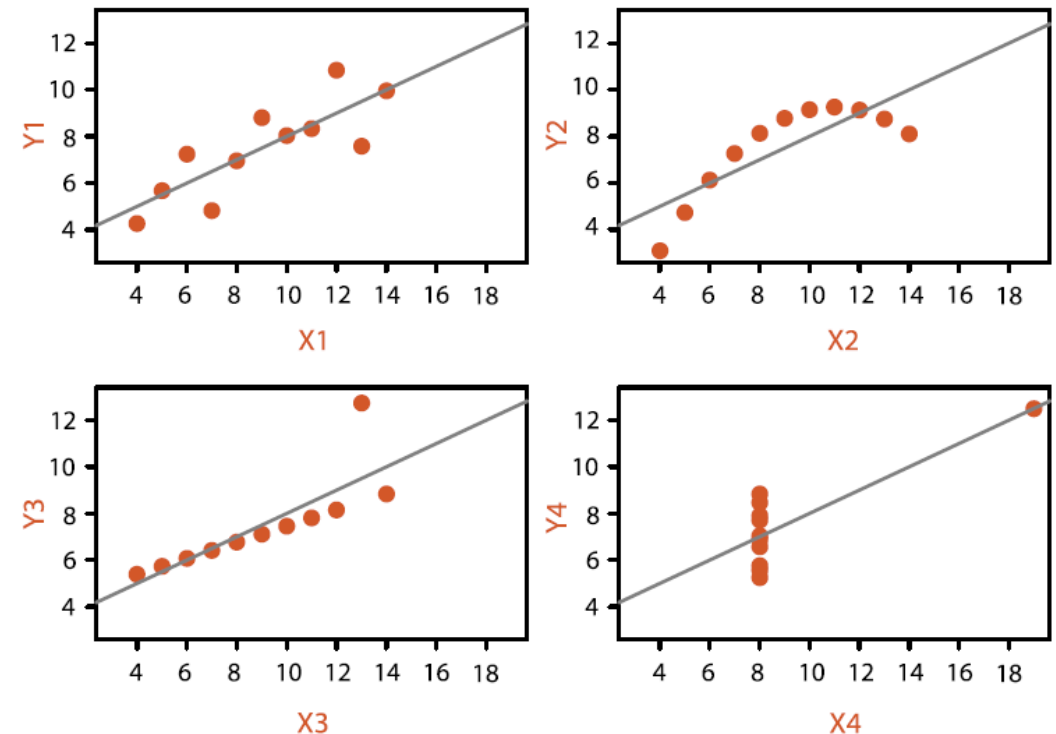
*“The greatest value of a picture is when it forces us to notice what we never expected to see”*

John W Tukey – Exploratory Data Analysis

Anscombe's Quartet: Raw Data

	1		2		3		4	
	X	Y	X	Y	X	Y	X	Y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
Correlation	0.816		0.816		0.816		0.816	

Regresión lineal:  $y = 3.0 + 0.5 x$





# Visualización de datos

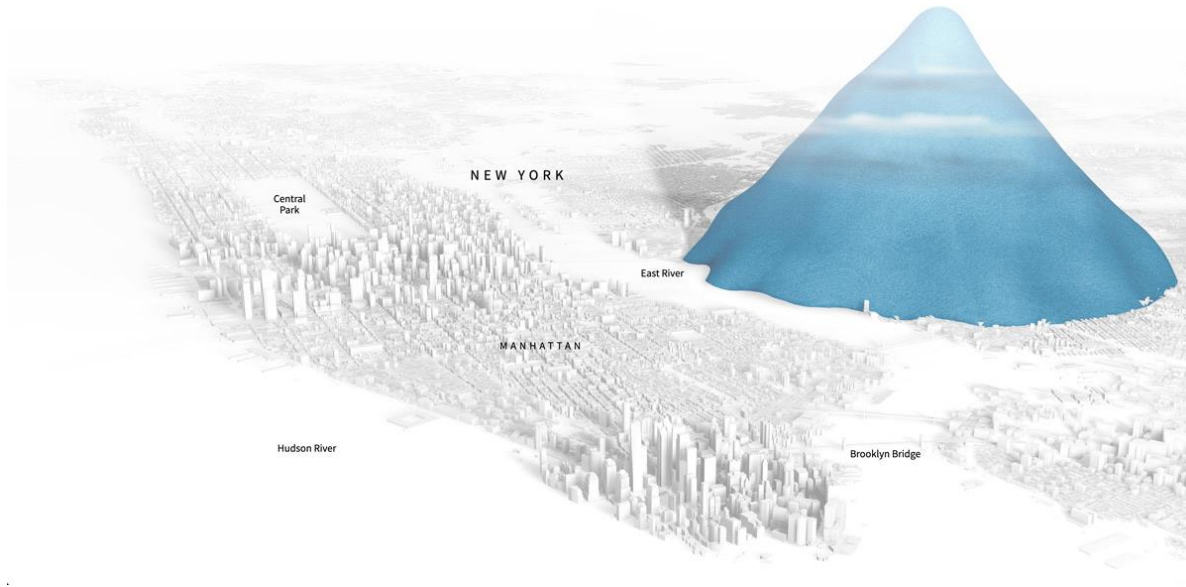
Es contar una historia



*Visualization gives you  
answers to questions you  
didn't know you had.*

Ben Shneiderman  
Profesor de computer science  
Universidad Maryland

# Importancia de la visualización de datos



Por Reuters Graphics

¿Qué es la visualización de datos?

- Arte visual que capta nuestro interés y mantiene nuestros ojos en el mensaje.
- Combina arte, creatividad, matemáticas, ciencias de la computación, percepción visual.

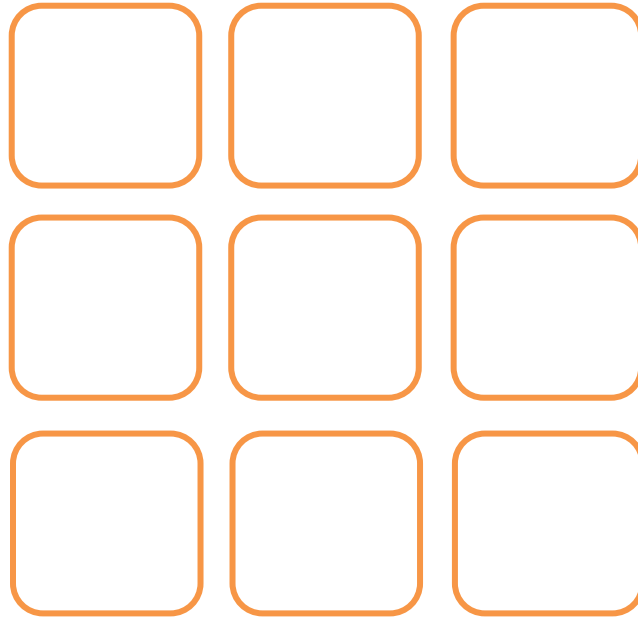
¿Por qué es importante?

- Analizar grandes cantidades de información y tomar decisiones basadas en los datos
- Crea formas visibles para representar el significado de datos abstractos

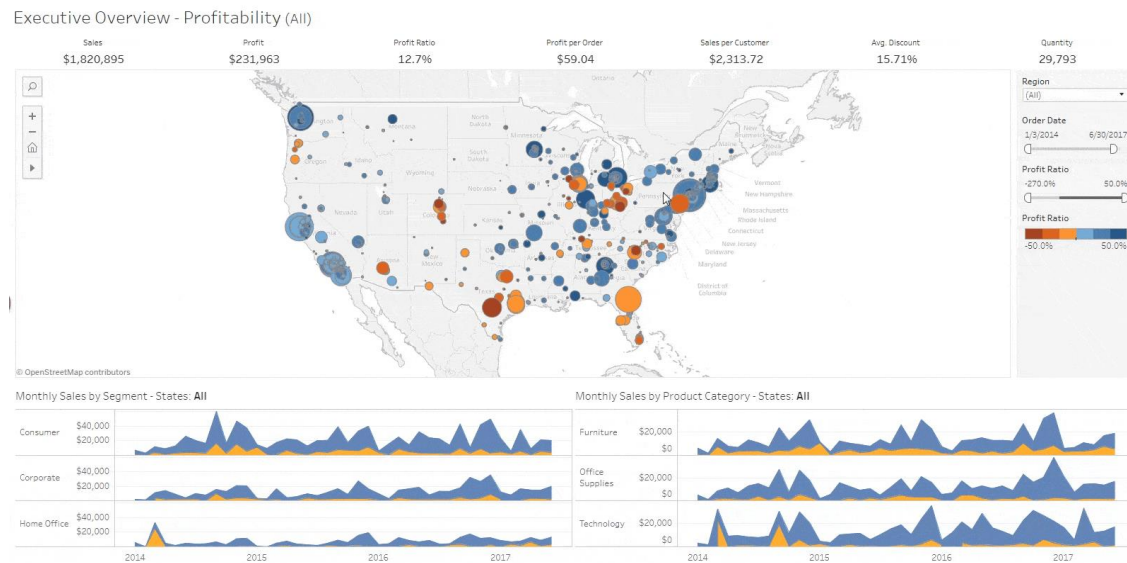
# “Juego del 15”



# Triqui



# Beneficios de la visualización de datos

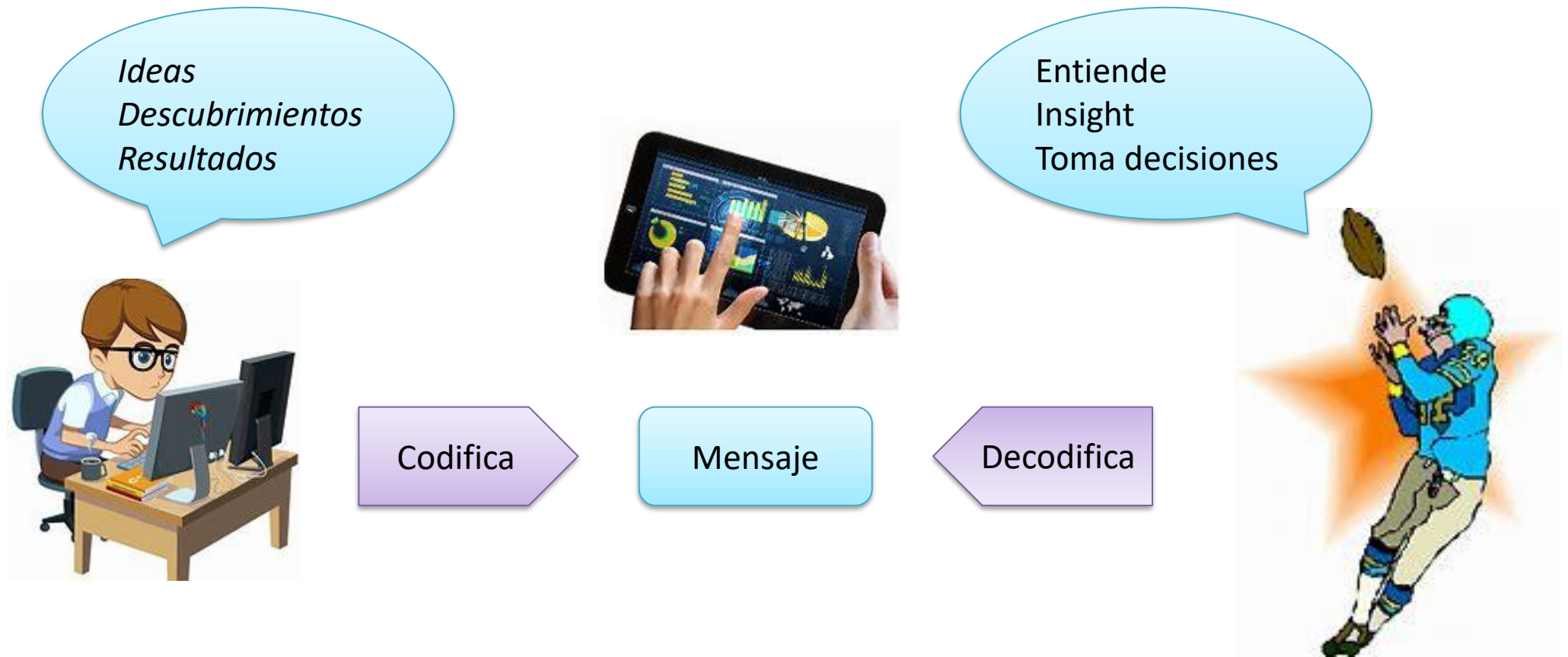


- Las visualizaciones de datos **muestran los datos** de una manera más **accesible y comprensible**.
- Un dashboard puede **contar** rápidamente una **historia** y **resaltar tendencias o patrones** que no se podrían descubrir fácilmente si se analizaran manualmente los datos sin procesar.
- Permite tener más conversaciones en torno a los datos y por ende, un **mayor impacto en el negocio**.
- Resalta la información útil y **elimina el ruido**.
- Ojo:
  - Equilibrio entre forma y función.
  - Combinar un gran análisis con una gran narración.

[Inteligencia de negocios: qué es y por qué es importante. \(tableau.com\)](https://www.tableau.com)



# Actores del intercambio de información





# Programa

---

## Tema

- Importancia de la visualización de datos
  - Metodología para la visualización de datos
  - Entender el contexto
  - Diseñar el modelo de analítica
  - Escoger la visualización
  - Eliminar el desorden
  - Enfocar la atención
  - Diseñar
  - Contar la historia
  - Errores comunes
  - Otras herramientas
  - Gráficos de redes sociales
-

# Evaluaciones

- Pruebas cortas 30%
- Talleres 35%
- Proyecto 35% - Semana 16



# Visualizaciones y Ciencia de Datos

## Visualizaciones y BI

# Ciencia de Datos

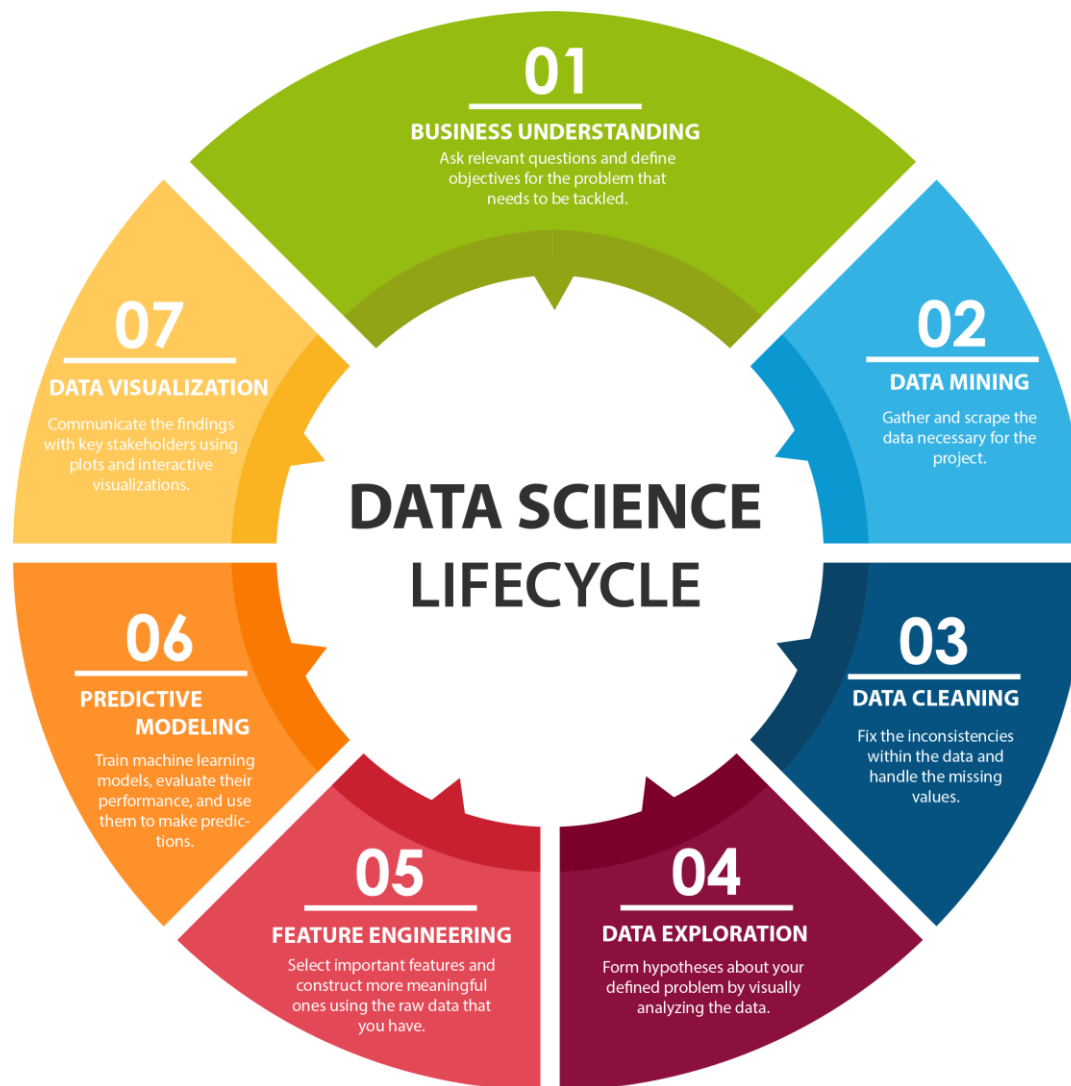
Qué es ciencia de datos:

*Data science* is the computational science of extracting meaningful insights from raw data and then effectively communicating those insights to generate value.

Lillian Pierson - Data Science for Dummies, 2017

¿Para que sirve?

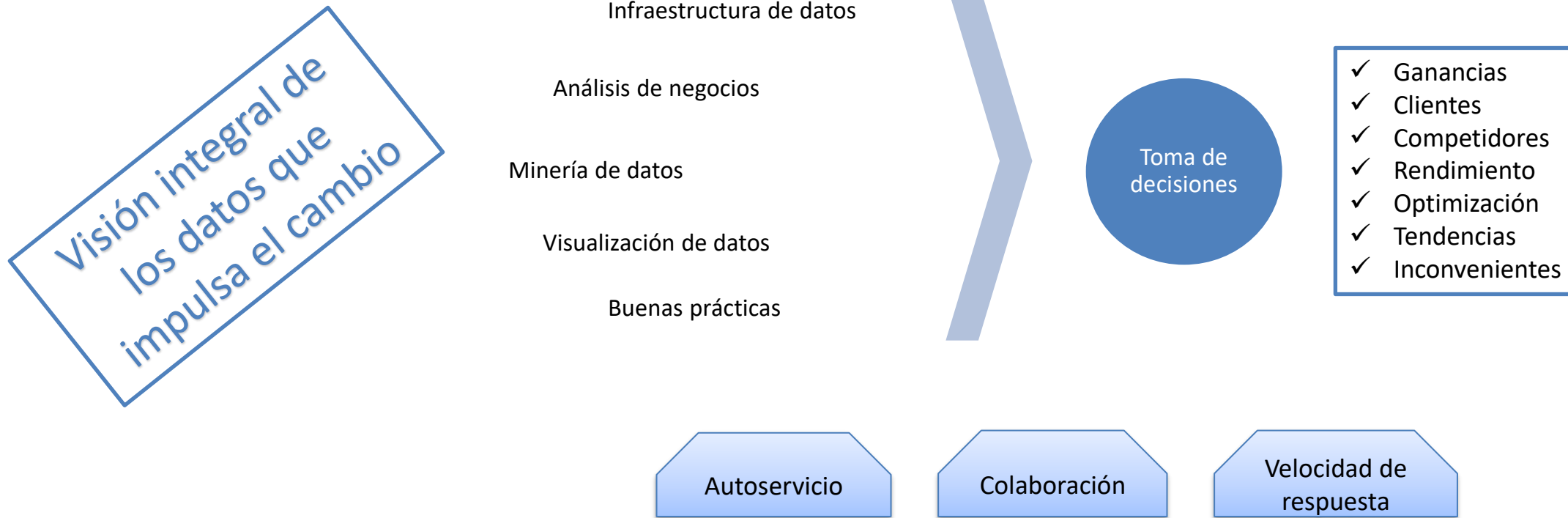
- Optimizar y mejorar negocios
- To give insights:
  - Un pasajero de Uber que espera que su conductor lo recoja exactamente a la hora y ubicación predicha por la aplicación
  - El comprador online que espera que Amazon le recomiende las mejores alternativas para el producto que busca, de forma que pueda compararlas y tomar la mejor decisión de compra.



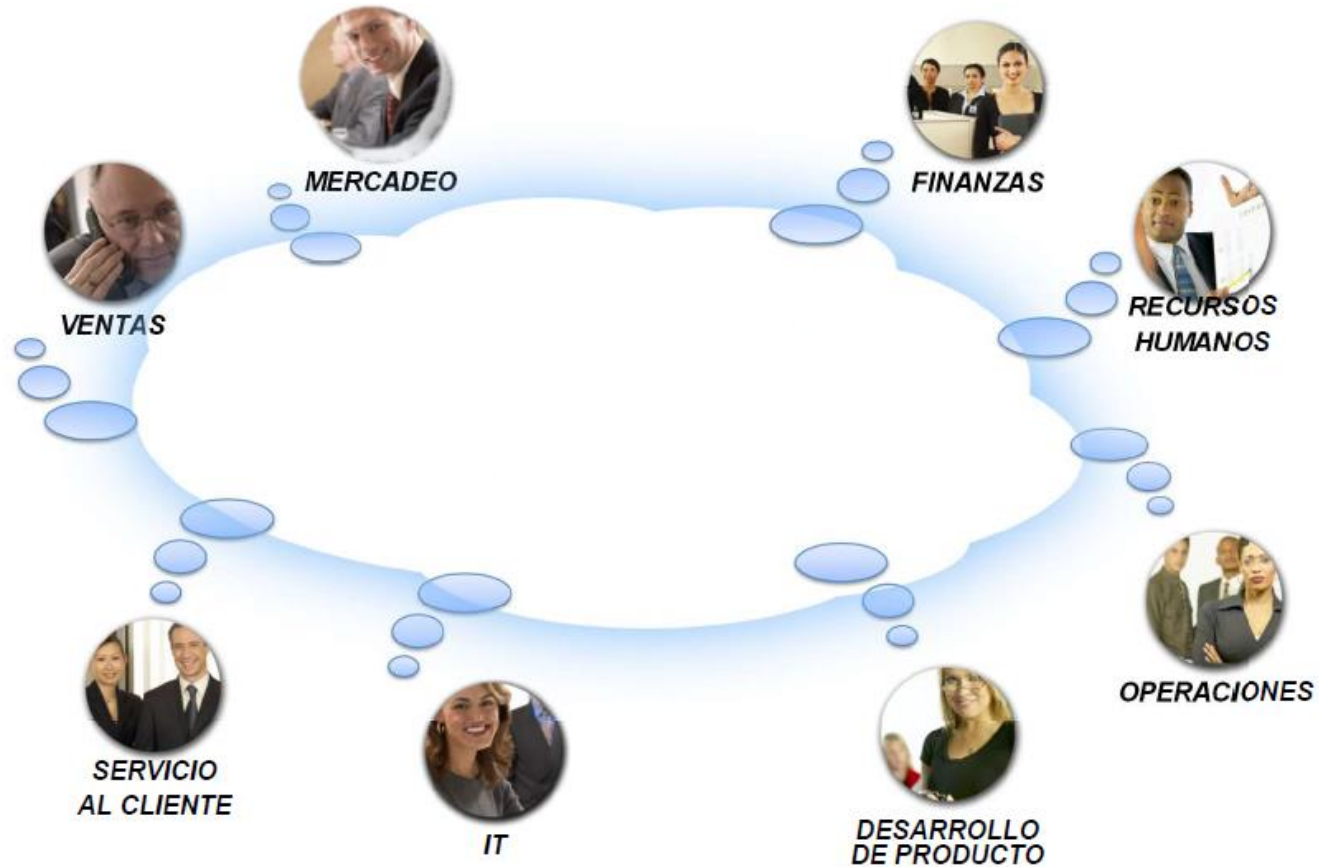
## La visualización en la Ciencia de datos

[Hans Rosling's 200 Countries, 200 Years, 4 Minutes - The Joy of Stats - BBC Four - YouTube](#)

# ¿Qué es inteligencia de negocios?



# Usuarios de la BI

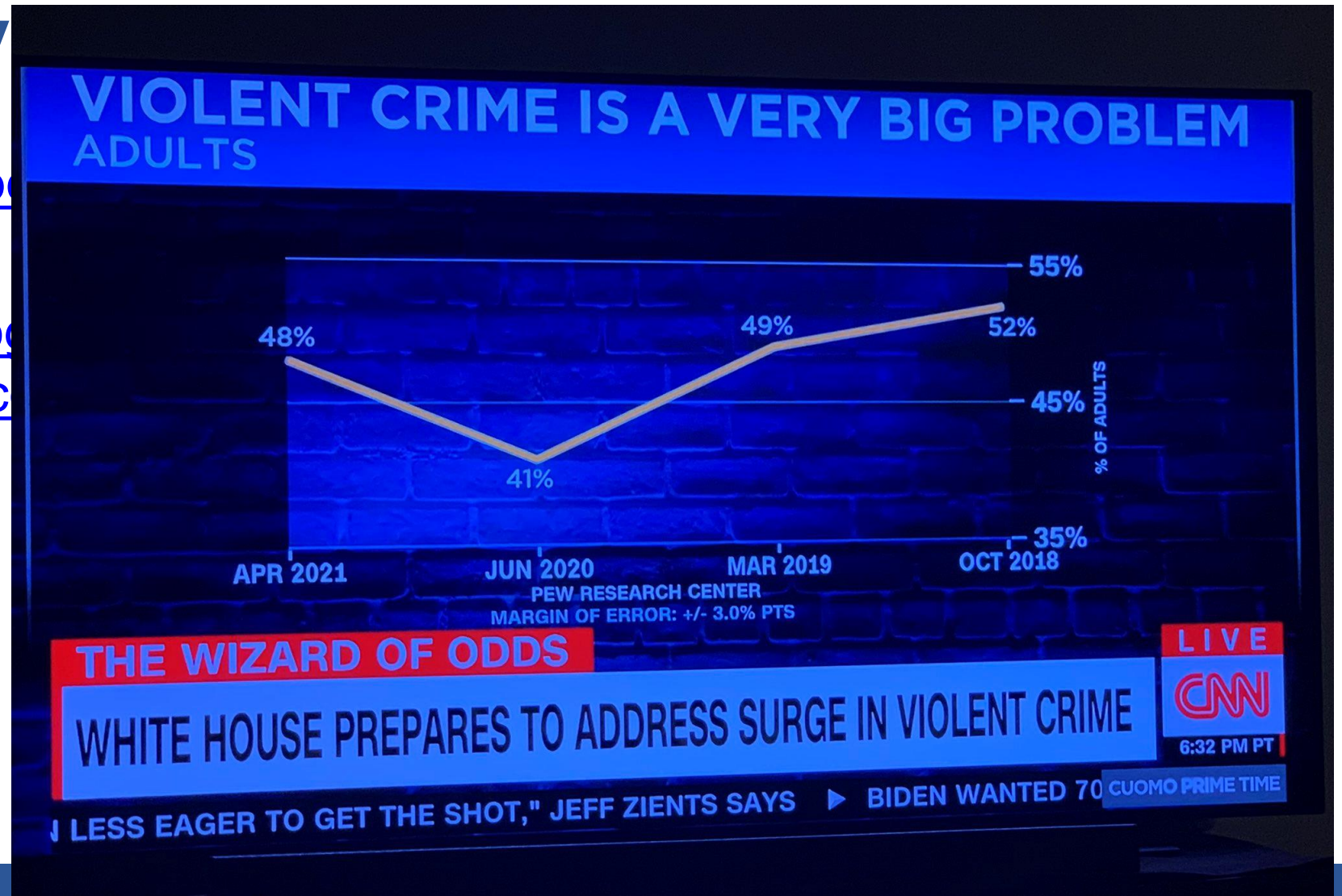




# Ejemplos de v

World Population Glob

Understanding our pro  
COVID-19 (microsoft.c



# En resumen

Por la forma en que el cerebro humano procesa la información, el uso de diagramas o gráficas para visualizar grandes cantidades de datos complejos es más fácil que vaciarlos en hojas de cálculo o reportes.

La visualización de datos es una forma fácil y rápida de transmitir conceptos de manera universal y puede experimentar con diferentes escenarios haciendo ajustes ligeros.

La visualización de datos puede:

- Identificar áreas que necesitan atención o mejoras.
- Esclarecer qué factores influyen el comportamiento de los clientes.
- Ayudarle a entender qué productos colocar en qué lugar.
- Predecir volúmenes de ventas.

# AMPLIANDO EL CONOCIMIENTO

# Lecturas

- Primer capítulo del libro “Visualization Analysis and Design” de Tamara Munzner.  
<https://dokumen.tips/documents/visualization-analysis-and-designpdf.html>

Uso de bicicletas compartidas

# TALLER



# Bicicletas compartidas

A partir de la información de renta de bicicletas que se encuentra en:

[UCI Machine Learning Repository: Bike Sharing Dataset Data Set](#)

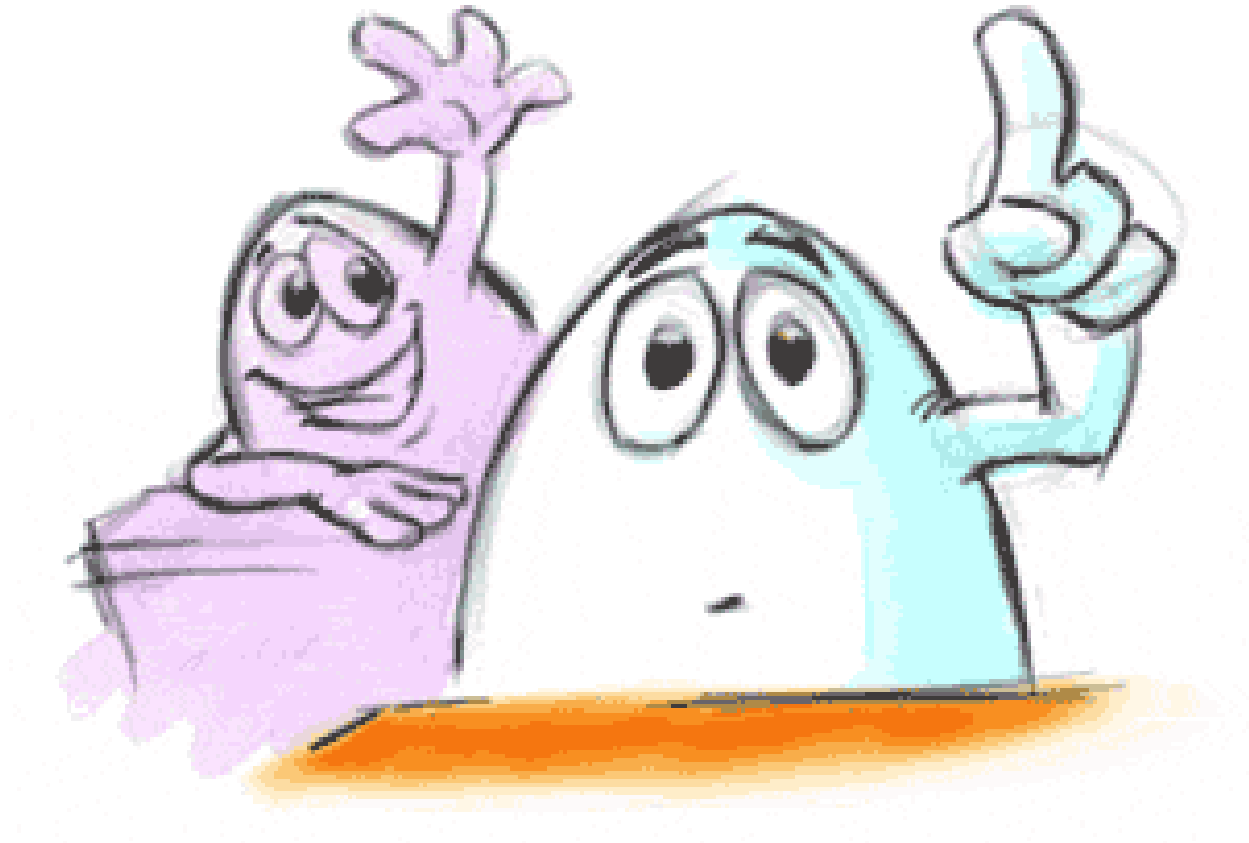
Busque responder las siguientes preguntas y cuéntenos que otra información importante encontró:

- ¿En qué época del año debo asegurarme de tener el mayor número de bicicletas disponibles?
- ¿Cuáles son las horas del día y los días de la semana en donde voy a tener mayor demanda por tipo de usuario?
- ¿Es posible inferir el uso que le dan a las bicicletas mis usuarios?
- ¿Cómo afectan las condiciones climáticas en la renta de bicicletas?
- Genere y responda su propia pregunta

# Otros datasets

- Estadísticas Eurocopa / Copa América
- Juegos Olímpicos
- Precio de aguacates
- Ventas de videojuegos
- Tik Tok





# ¿Preguntas?

# Introducción a la visualización de datos

Maria Isabel Serrano G.  
maria-serrano@javeriana.edu.co