

# Analítica de Datos

---

Introducción al Big Data, Inteligencia de Negocios y Fuentes de Datos.





Garrincha- El Gráfico - El Gráfico n° 2233

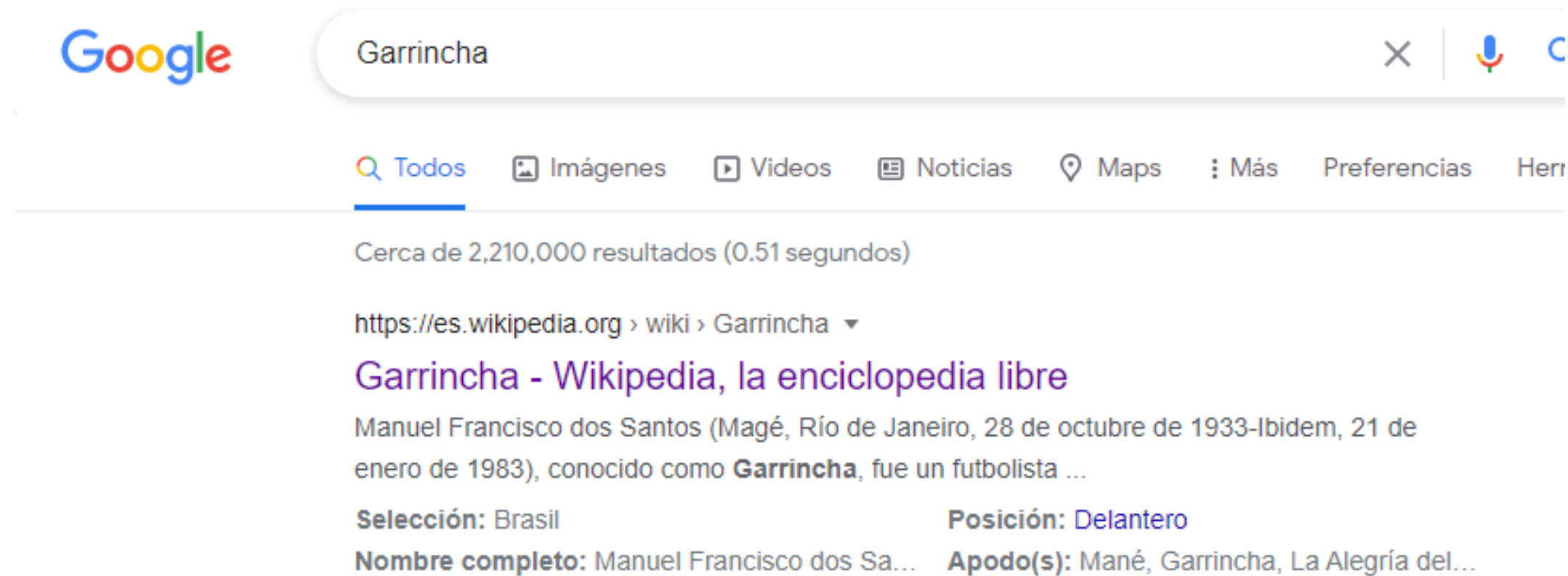
# La importancia de los datos en la toma de decisiones.

La toma de decisiones en base a datos es algo que se hace todos los días, pero se hace generalmente con uno o dos datos.

En algún país de Latinoamérica, años 50:

- Dicen que Botafogo tiene un buen puntero derecho
- No, yo lo vi entrenando, es cojo.

Hoy en día existen millones de datos disponibles para extraer información.



Garrincha– Google.com

¿De donde salen todos esos datos?  
¿Y que hacer con ellos?

# 2020 *This Is What Happens In An Internet Minute*

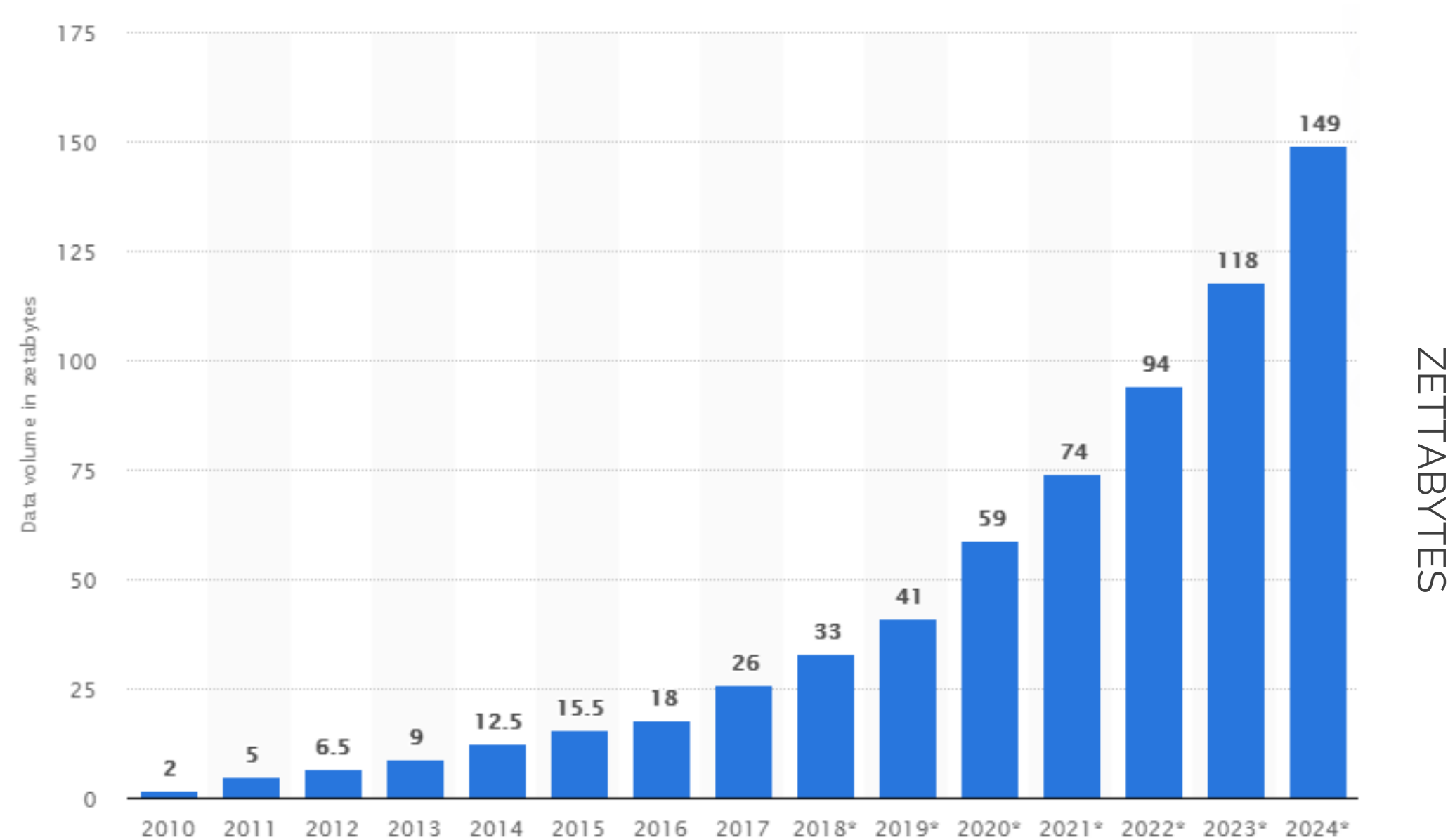


## BIG DATA

La cantidad de datos disponibles es gigantesca y se agregan nuevos datos cada instante. Y todo se guarda.

Un minuto en Internet – @LoriLewis, @OfficiallyChadd

Al año pasado, se estimaban 2.5 millones de terabytes de información nueva por día, y al finalizar el año se decía que cada persona en promedio generaba 1.7 megabytes por segundo.



Volumen de datos – por Statista.com

© Statista

# 1 Zettabyte

es una cantidad de datos  
igual a la capacidad de

## 17.200 millones de iPods

con 64 GB de memoria

Una **película HD**  
tendría una duración de  
**36 millones  
de años**



Un zettabyte - The Zettabyte Era: Trends and Analysis June 2017

En resumen:  
Son muchos datos.

Actividad: Entre a <https://quizizz.com/join> para solucionar una pequeña encuesta.

Tiempo	Datos
Menos de una hora	Máximo 6 gigas
Entre una y dos horas	Entre 6 y 12 gigas
Entre dos y tres horas	Entre 12 y 18 gigas
Mas de 3 horas	Mas de 18 gigas

Todo eso pasa a sumar al **BIG DATA**

# Las características del **BIG DATA**



Volumen



Variedad



Velocidad



Veracidad



Valor



# Volumen

La cantidad de información que se genera. Diariamente se produce mas información que todo lo que hizo la humanidad desde sus inicios hasta la década del 2000.

Genera el problema de almacenamiento optimo de cantidades ingentes de datos.

# Variedad

Cientos de formatos de datos se usan diariamente: videos, texto, imágenes, listas, georreferenciación, sin contar la meta data asociada a cada uno de los datos.

Genera el problema de analizar los datos en contexto.

# Velocidad



La velocidad a la que se genera nueva información hace que los procesos de almacenaje deban ser eficientes para mantener la información actualizada.

Genera el problema del análisis en tiempo real.

# Veracidad



Entre la cantidad de información que se genera hay mucha que solo es ruido, bien sea por falsa o por no ser de interés particular.

Genera el problema de limpiar y filtrar los datos.

# Valor

Los datos pueden aportar algo, desde patrones de conducta o tendencias, detección de problemas, irregularidades hasta ser capaces de monetizarse por si solos.

Son una fuente de oportunidades.

# El valor del **BIG DATA**

¿Le sirve a alguien la información que subimos o consumimos en internet?

- Facebook segmenta a sus usuarios para generar publicidad personalizada.
- Twitter permite perfilar a los usuarios y sus reacciones, para medir sus reacciones a determinados estímulos.
- YouTube ofrece servicios de reproducción adaptados al historial de vistas para mantener usuarios y generar puntos de impacto de manera mas efectiva.

# La Inteligencia de Negocios

La inteligencia de negocios hace referencia a las estrategias y herramientas que nos permiten tomar buenas decisiones basadas en datos



# Identificar las Preguntas

Tomar mejores decisiones de negocios

- Se consideran decisiones de negocios aquellas que se refieren a prioridades de la empresa y problemas sin solucionar.

Tomar mejores decisiones operativas

- Las decisiones operativas tienen que ver con las problemáticas internas de la empresa.

Monetizar los datos

- Además de analizar los datos, es posible comercializarlos.



¿Cómo...

...Tomar mejores decisiones de negocios...

- En una institución educativa?
- En una gran superficie?

...Tomar mejores decisiones operativas...

- En una empresa de transporte?
- En una aplicación de citas?

...Monetizar los datos...

- De una campaña electoral?
- De una plataforma de streaming?

# Recolección y Análisis

Escoger los datos necesarios

- Dependiendo de las necesidades a resolver, los datos necesarios.

Recolectar los datos

- Ubicar las fuentes de datos y métodos para su obtención.

Analizar los datos

- Utilizar diversas técnicas para limpiar, filtrar y extraer información de los datos.

# ¿Cómo...

...Escoger los datos necesarios...

- Para una institución educativa?
- Para una gran superficie?

...Recolectar Datos...

- En una empresa de transporte?
- En una aplicación de citas?

...Analizar los datos...

- De una campaña electoral?
- De una plataforma de streaming?

# Obtener Resultados

De acuerdo a los procesos particulares de la toma de decisiones basada en datos en una empresa, es necesario medir cual es el avance de la estrategia hacia los objetivos que pretende cumplir. La manera más común de realizar esta medición son los KPI (Key Performance Indicator), indicador clave de desempeño. Estos permiten medir el avance de los objetivos entre dos puntos del avance de la estrategia.

## Algunos tipos de KPI

- Aquellos que miden características de los clientes.
- Aquellos que miden el desempeño financiero.
- Aquellos que miden la calidad de los procesos internos.
- Aquellos que miden el desempeño y estado de los trabajadores.

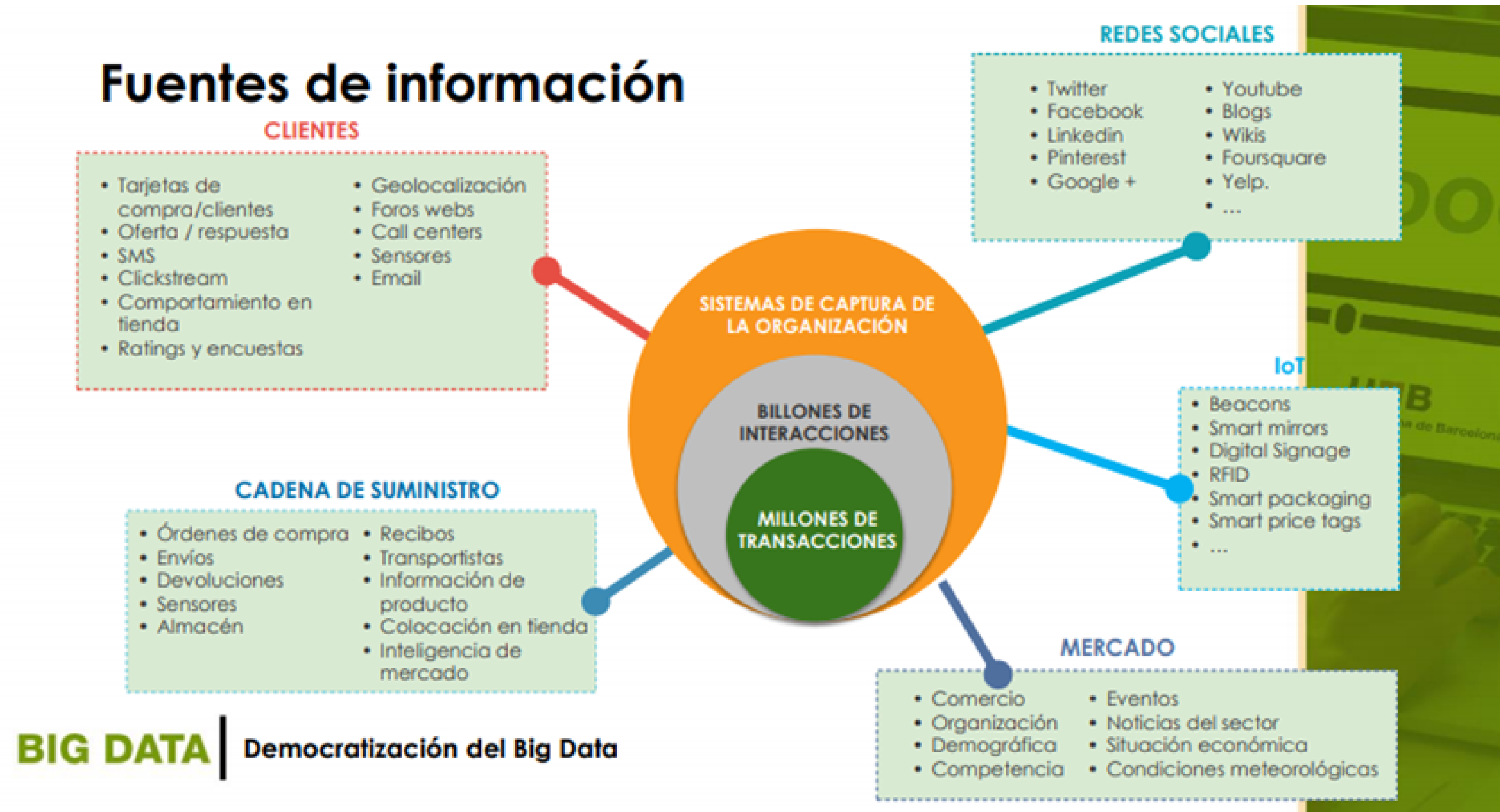
**Actividad:** En base a los escenarios planteados en la actividad anterior, ¿cuales podrían ser KPIs adecuados?



Metodología SMART – AllKPIs.com

**Actividad:** Los casos de éxito suelen ser una buena forma sacar inspiración para una estrategia.

# Fuentes de Datos



Fuentes de datos – BIG DATA

¿Preguntas, comentarios?



# Actividad

Supóngase la siguiente situación: en 2018, la autoridad ambiental de Cali decide implementar un programa de mitigación de ruido, identificando las zonas de la ciudad que mas reportes generan, y desarrollando acciones en dichas zonas.

Se recogen datos los primeros cinco meses del año, se pone en practica la estrategia entre junio y julio, y se evalúan los resultados los meses siguientes.

Usando la base <https://datos.cali.gov.co/dataset/mapa-de-calor-visitas-de-ruido-2018>, responda:

- ¿Se consideraría este proceso como una estrategia de datos?
- ¿Que KPI considera que debió haberse planteado para medir la eficacia del proceso?
- ¿Se puede medir ese KPI con los datos?
- ¿Qué conclusión se obtendría?

# Actividad

Identifique una problemática dentro de su institución, que pueda ser solucionada por medio de mejores decisiones empresariales u operacionales.

Dentro del marco de una estrategia de datos

- ¿Qué datos necesitaría para la estrategia?
- ¿Es posible encontrar datos de esa naturaleza en las bases abiertas del estado?

# Actividad

Un ejemplo de cómo usar los datos de manera efectiva es la plataforma Uber Movement (<https://movement.uber.com/?lang=es-CO>). En esta se muestran los registros de viajes en diversas ciudades del mundo, con tiempos de viaje y velocidades. Estos datos están disponibles para desarrollar mejores estrategias de movilidad.

- ¿Podrían ser usados estos datos en la generación de estrategias que mejoren la movilidad?
- ¿Cuál serían los KPIs adecuados para medir el avance de la estrategia?
- ¿Qué otros datos quisiera obtener para diseñar una mejor estrategia de movilidad?