

Taller de manejo de datos con herramientas libres



Primera Clase: a conocer los datos...



Instituto Nacional de
Tecnología Agropecuaria



Taller de manejo de datos con herramientas libres



Primera Clase: a conocer los datos...



Instituto Nacional de
Tecnología Agropecuaria





Por qué un taller de datos?

Porque los datos y la
forma en que se
gestionan, están
cambiando nuestra vida



Una persona promedio actual procesa más datos en un día



Fuente: PISA hasta finales 2016 (http://www.oecd.org)

mas celulares que
cepillos de dientes
millones de personas tienen
celulares

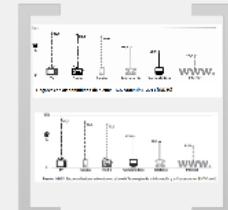


4.200 millones tienen cepillos de dientes

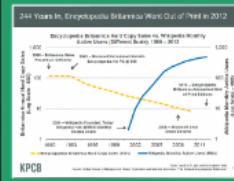
Hay más celulares que cepillos de dientes
4.800 millones de personas tienen celulares

1

三



Acceso al conocimiento



Redes Sociales
en
Argentina

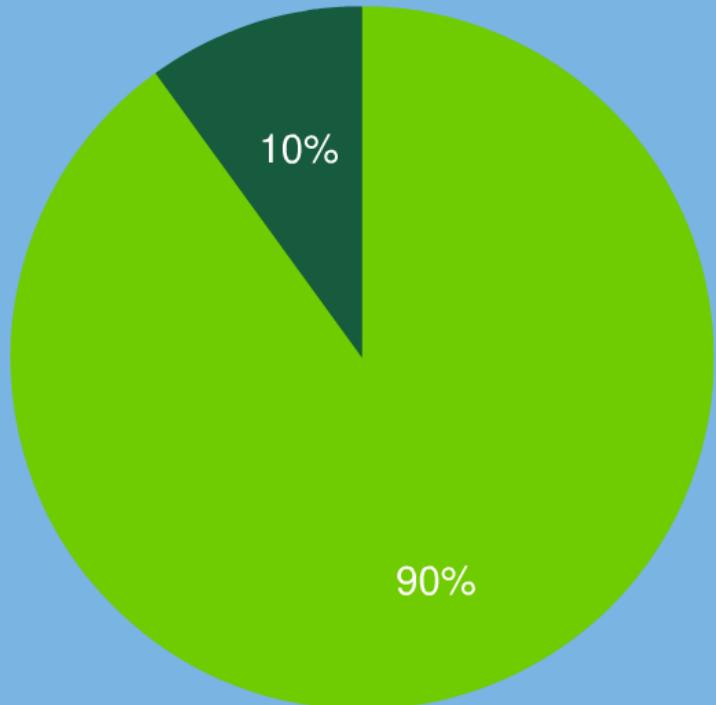
La humanidad siempre usó datos para avanzar:



6000 A.C.: datos de rendimiento y barbechos para producir mas y alimentar más personas.

siglo XV: se utilizaron los datos del firmamento para navegar por el mundo y generar nuevas rutas comerciales.





● 2014-2015 ● Antes del 2014

90% de los datos en el
mundo fueron creados en
los últimos dos años

Fuente: IBM <http://www.ibm.com/software/data/bigd>

Una persona promedio actual,
procesa más datos en un día

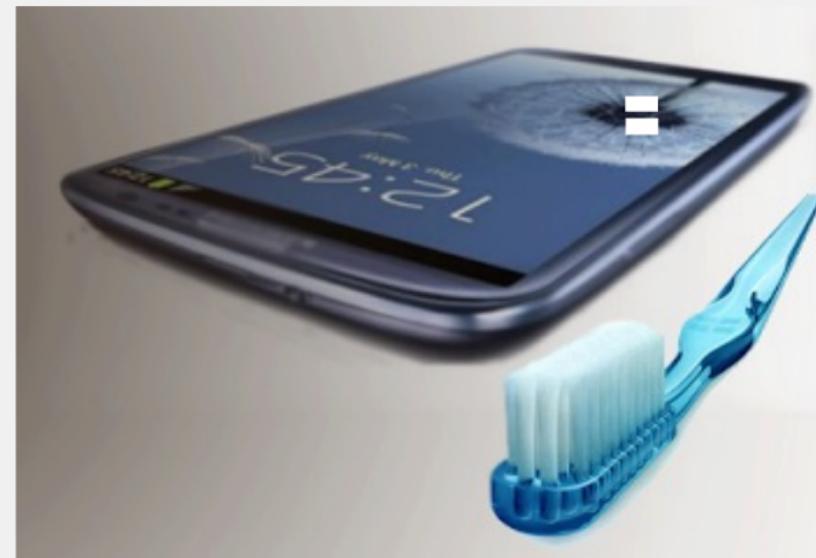


que una persona del siglo XVI
durante toda su vida

Fuente: BSA Data Study 2015 (<http://data.bsa.org/>)

Hay mas celulares que
cepillos de dientes

4.800 millones de personas tienen
celulares



4.200 millones tienen cepillos de
dientes

A ver nuestra encuesta

2005

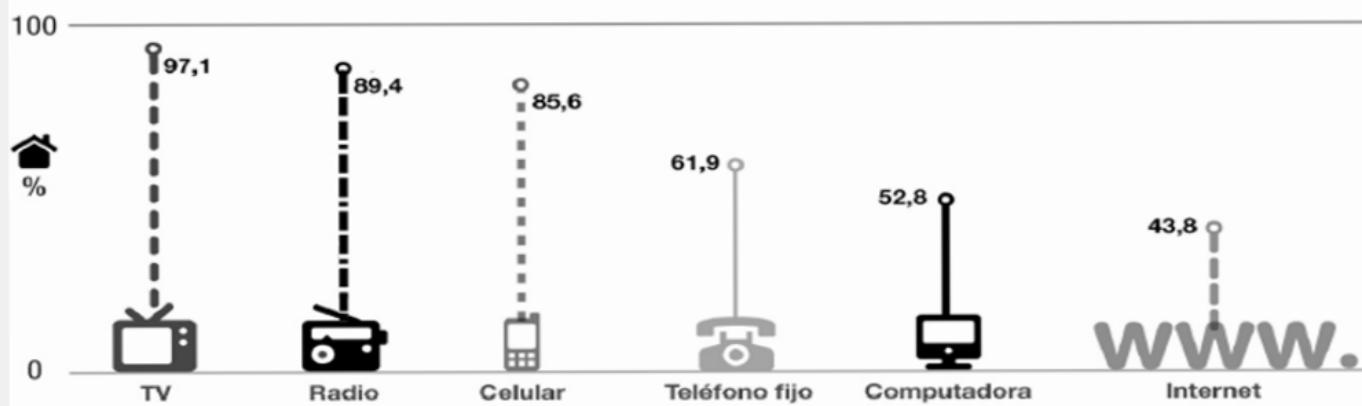




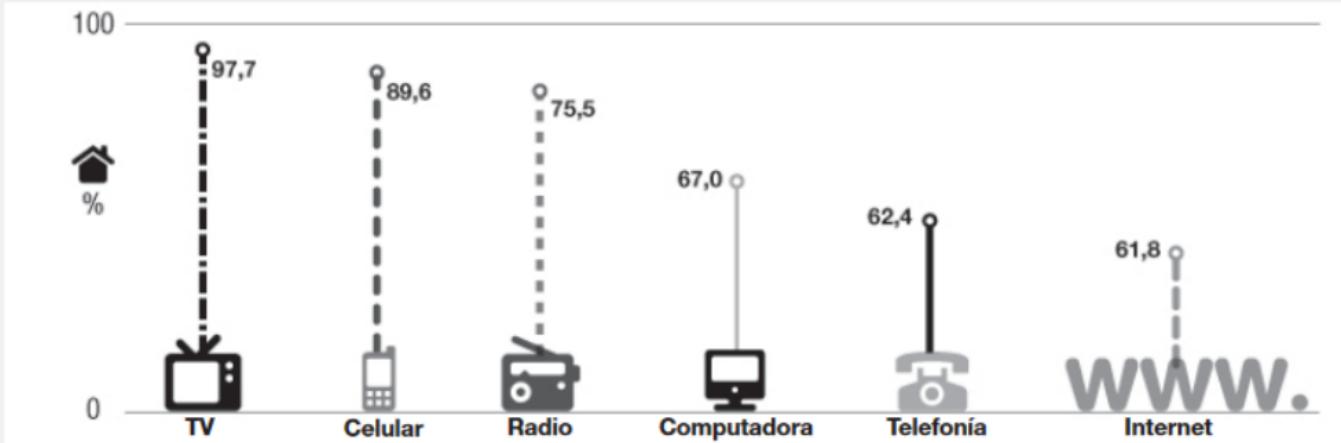
2013



Michael Sohn / AP



Hogares con disponibilidad de bienes TICs. Argentina. 2011 (INDEC)



Fuente: INDEC. Encuesta Nacional sobre Acceso y Uso de Tecnologías de la Información y la Comunicación (ENTIC) 2015.



Redes Sociales en Argentina

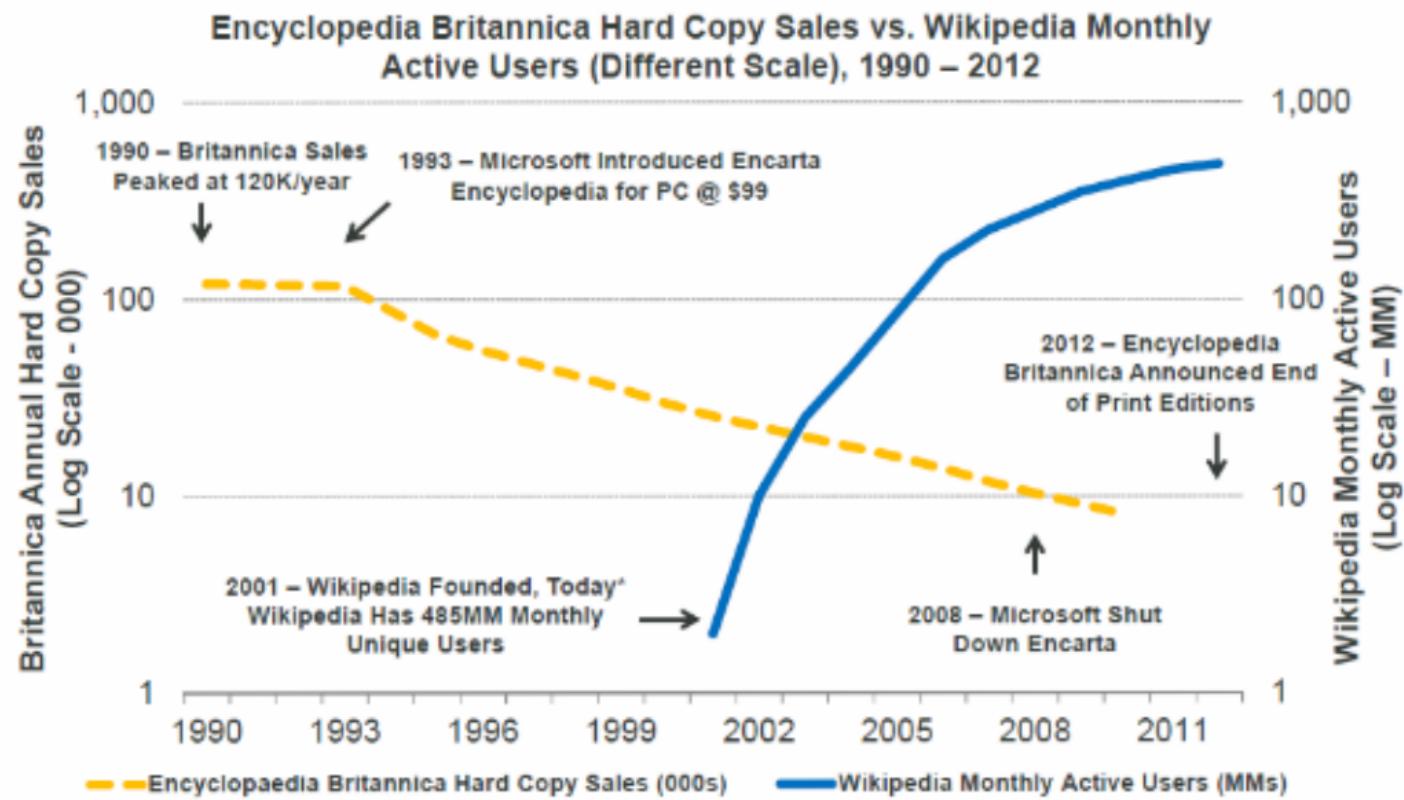
- + de 28.000.000 de usuarios
- + de 9,1 horas mensuales de uso
- + Facebook y Twitter las dos más utilizadas
- + 15 millones acceden por medio del celular

1 minuto en Internet



Acceso al conocimiento

244 Years In, Encyclopedia Britannica Went Out of Print in 2012



KPCB

Note: ^aas of 9/12, per comScore global data.
Source: Kellogg School of Management, Shane Greenstein and Michelle Devereux, "The Crisis at Encyclopedia Britannica."

que son
exactamente
los "datos"?

Quién o qué
loS genera?

Que
posibilidades
tienen de
mejorarnos la
vida?

Cómo hay que
usarlos para
obtener un
máximo
provecho?

Cómo loS usamoS de
maneras que Sean
compatibles con
nuestros valores e
inquietudes

que
hacemoS
con tantos
datos?

qué Son
exactamente
los “datos”?

Quién o qué
los genera?

Qué
posibilidades
tienen de
mejorarnos la
vida?

Cómo hay que
usarlos para
obtener un
máximo
provecho?

Cómo los usamos de
maneras que Sean
Compatibles con
nuestros valores e
inquietudes

Big Data (Datos
Masivos o Grandes
Datos)



que
hacemos
Con tantos
datos?

Big Data (Datos Masivos o Grandes Datos)



You**T**ube



You Tube

que son
exactamente
los "datos"?

Quién o qué
loS genera?

Que
posibilidades
tienen de
mejorarnos la
vida?

Cómo hay que
usarlos para
obtener un
máximo
provecho?

Cómo los usamoS de
maneras que Sean
compatibles con
nuestros valores e
inquietudes

que
hacemoS
con tantos
datos?

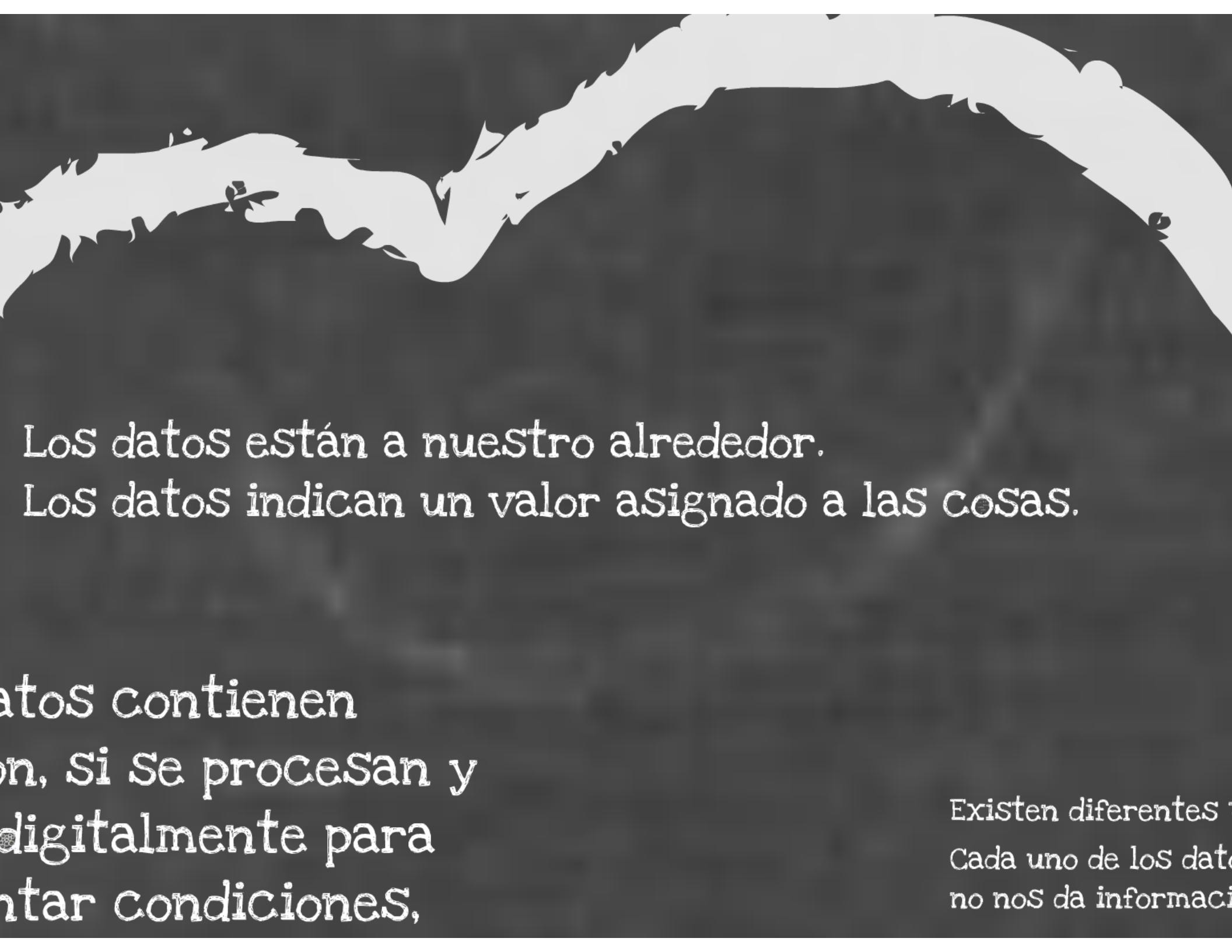
Los datos están a nuestro alrededor.
Los datos indican un valor asignado a las cosas.

Los datos contienen
información, si se procesan y
ordenan digitalmente para
representar condiciones,
objetos o ideas.

Existen diferentes tipos de datos

Cada uno de los datos por sí mismo
no nos da información.

Hay que interpretar los datos para crear información, hay que
comprender, aplicar y aprender la información para generar
conocimiento.



Los datos están a nuestro alrededor.
Los datos indican un valor asignado a las cosas.

los datos contienen
información, si se procesan y
se analizan digitalmente para
determinar condiciones,

Existen diferentes tipos de datos.
Cada uno de los datos que nos rodean
no nos da información completa.

3,7 kg

51 cm

16.30 hs

03/12/2013

Hombre

Un varoncito, llamado Juan Bautista,
nació el 3/12/2013 a las 16.30 horas
pesó 3,7 kilos y midió 51 cm.

Hoy tiene 3 años y medio
y los profes del curso
Son los babosos padres.



Los datos indican un va.

Los datos contienen información, si se procesan y ordenan digitalmente para representar condiciones, objetos o ideas.

1 BIT

Contiene solo dos valores

0 1

Sin energía Con energía

1 PETABYTE

1024 TB

2 Petabytes almacenan toda la biblioteca de investigación académica de los EEUU.

1 EXABYTE

1024 Petabytes

5 Exabytes pueden guardar todas las palabras pronunciadas por la humanidad

1 ZETTABYTE

1024 Exabyte

El tráfico de Internet alcanzó los 1.3 ZB durante el 2016.

1 BYTE

Son 8 bits.

Puede contener un carácter (una letra)

1 TERABYTE

1024 GB

5 TB contienen la información sin procesar del radar meteorológico de Argus desde el año 2009

1 YOTTABYTE

1024 Zettabytes

Nombrado por Yoda
Nos tomarían 11 billones de años descargar 1 YT de información

1 KILOBYTE

Son 1024 bytes

Puede contener una historia corta, un resumen de un paper

1 MEGABYTE

1024 Kbytes

Puede contener una novela

1 GIGABYTE

1024 MB

Puede almacenar una película en formato estándar de televisión

1 BRONTOBYTE

1024 Yottabytes

1 000 000 000 000
000 000 000 000 000

27 ceros!!!

Fuente: www.whatsabyte.com

pcbunn.cithec.edu/presentations/giod_status_sep97/tsl012.htm

1

BIT

Contiene Solo dos
valores

0

1

Sin energía

Con energía

Pue

co

1

BYTE

Son 8 bits.

Puede contener un
caracter (una letra)

1

KILOBYTE

Son 1024 bytes

Puede contener una historia

corta, un resumen de un

paper

1

1

MEGABYTE

1024 Kbytes

Puede Contener
una novela

1

GIGABYTE

1024 MB

Puede almacenar una película
en formato estándar de
televisión

1

TERABYTE

1024 GB

5 TB contienen la información
sin procesar del radar
meterológico de Anguil desde
el año 2009

1

PETABYTE

1024 TB

2 Petabytes almacenan
toda la biblioteca de
investigación académica
de los EE.UU.

1

EXABYTE

1024 Petabytes

5 Exabytes pueden
guardar todas las
palabras pronunciadas
por la humanidad

1

ZETTABYTE

1024 Exabyte

El tráfico de
Internet alcanzó
los 1.3 ZB
durante el 2016.

1 YOTTABYTE

1024 Zettabytes

Nombrado por Yoda
NoS tomarían 11 billones de
años deScargar 1 YT de
información

1

BRONTOBYTE

1024 Yottabytes

1 000 000 000 000

000 000 000 000 000

27 ceros!!!

Existen diferentes tipos de datos
Cada uno de los datos por sí mismo
no nos da información.

Hay que interpretar los datos para crear información, hay que comprender, aplicar y aprender la información para generar conocimiento.



Según Como Se ordenan

Datos
estructurados:
sumamente
organizados



Datos **no estructurados**: no tienen una
estructura predefinida. Representan el 80% de
toda la información.



Por la forma en que Se generan

Generados por computadoras



En tiempo real



Transaccionales



Por lo que contienen

Cuantitativos: contestan la pregunta "¿Cuál?"

Cualitativos: son los datos que se refieren a números.



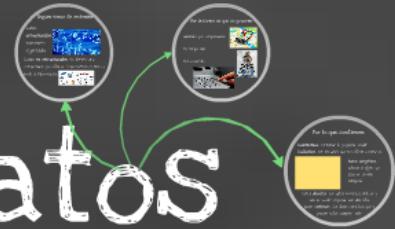
Datos **categóricos**: colocan el objeto por describir en una categoría.

Datos **discretos**: son datos numéricos enteros y sólo se pueden expresar con una cifra.

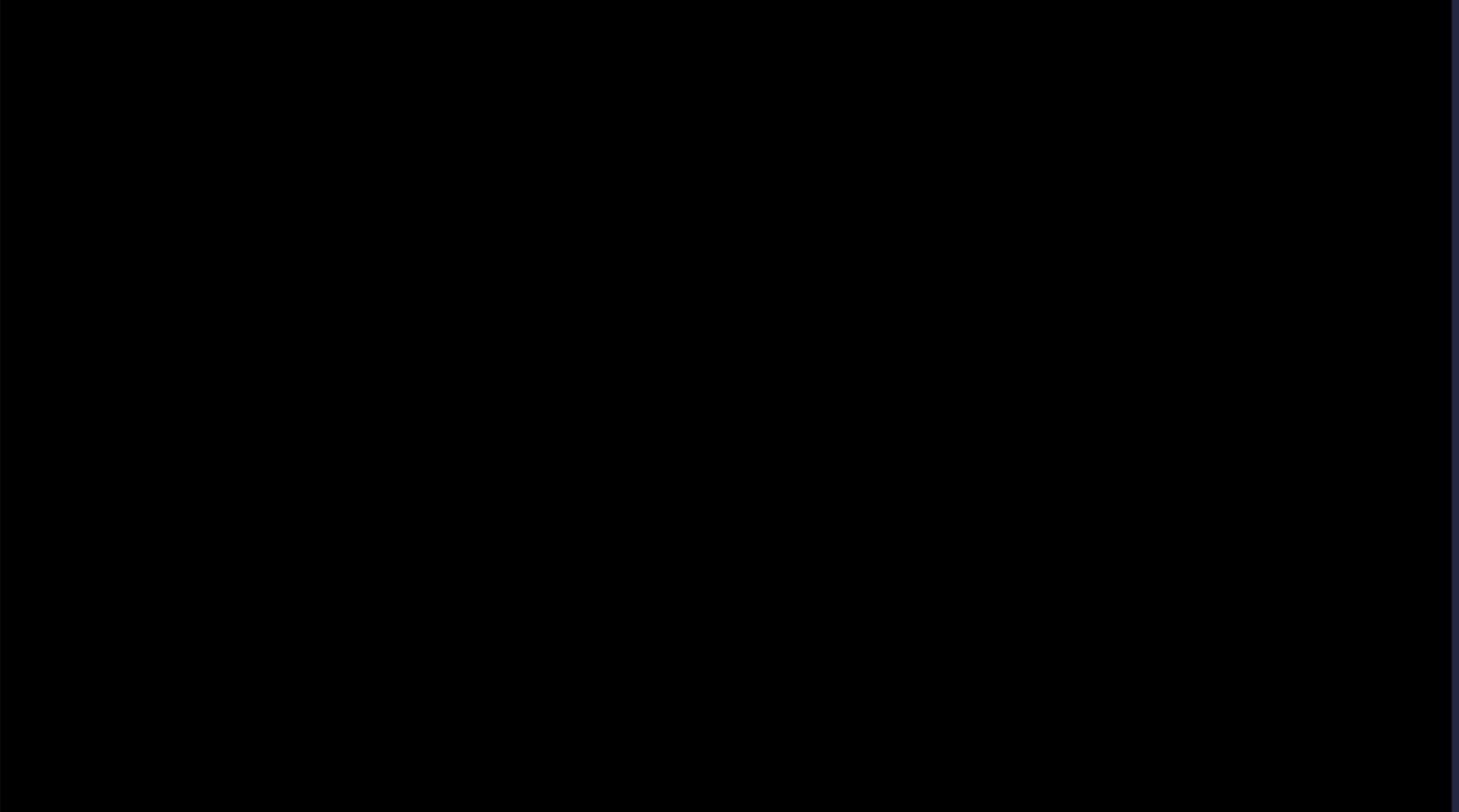
Datos **continuos**: Son datos numéricos que sí pueden recibir cualquier valor.

Existen diferentes tipos de datos
Cada uno de los datos por sí mismo
no nos da información.

Hay que interpretar los datos para crear información, hay que comprender, aplicar y aprender la información para generar conocimiento.







The Open Data Institute: <http://theodi.org/>

Y por casa como andamoS???

En 2016 el congresO Nacional aprobó la Ley de
Acceso a la Información



YouT^{ube}

Qué nota obtuvieron las provincias en 2010 y 2016
Según el índice que evalúa su información
pública?

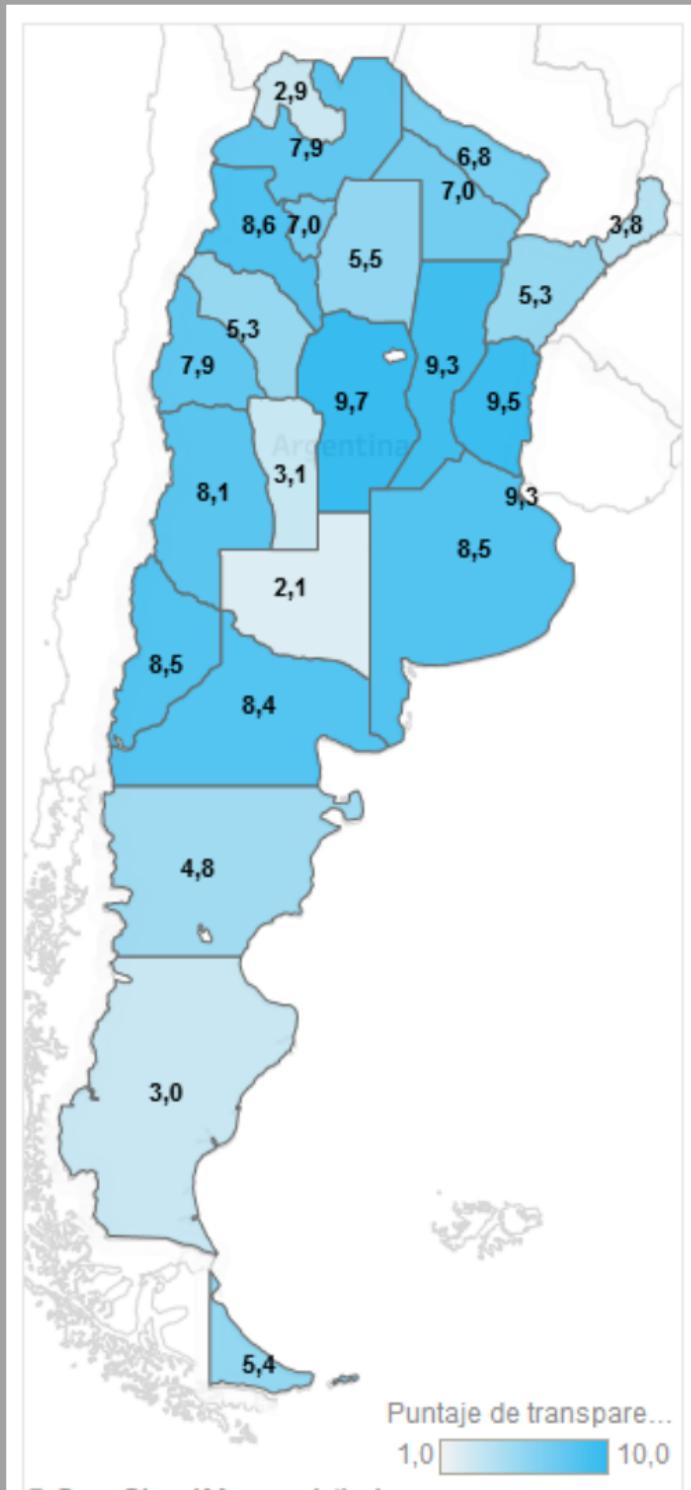
La nota de La Pampa....

Cha...chan...cha...chaaannn

2016: 2,1

últimos del país....





El podio de arriba:

- 1° Córdoba: 9,7
- 2° Entre Ríos: 9,5
- 3° Santa Fé y CABA: 9,3

El podio de abajo:

- 1° La Pampa: 2,1
- 2° Jujuy: 2,9
- 3° Santa Cruz: 3,0

Fuente: <http://www.lanacion.com.ar/1974611-presupuesto-provincial>

Tipos de licencia



LAS 5
Estrellas

Open Data Commons

- Public Domain Dedication and License
- Open Database License



Práctica



con mate

a las compu !!!

Taller de manejo de datos con herramientas libres



Primera Clase: a conocer los datos...



Instituto Nacional de
Tecnología Agropecuaria

