



Industria de los Datos

Industrias y Servicios II
2023

PALMA, Sofía
SANCHEZ, Aldana

CONTENIDOS

01

INTRODUCCIÓN

Avance industrial

02

DATOS

Historia, obtención, web scraping, tipos, almacenamiento de datos

03

EMPRESAS Y DATOS

Procesamiento y análisis de datos, roles dentro de la empresa, confidencialidad, empresas nacionales e internacionales

04

BIG DATA

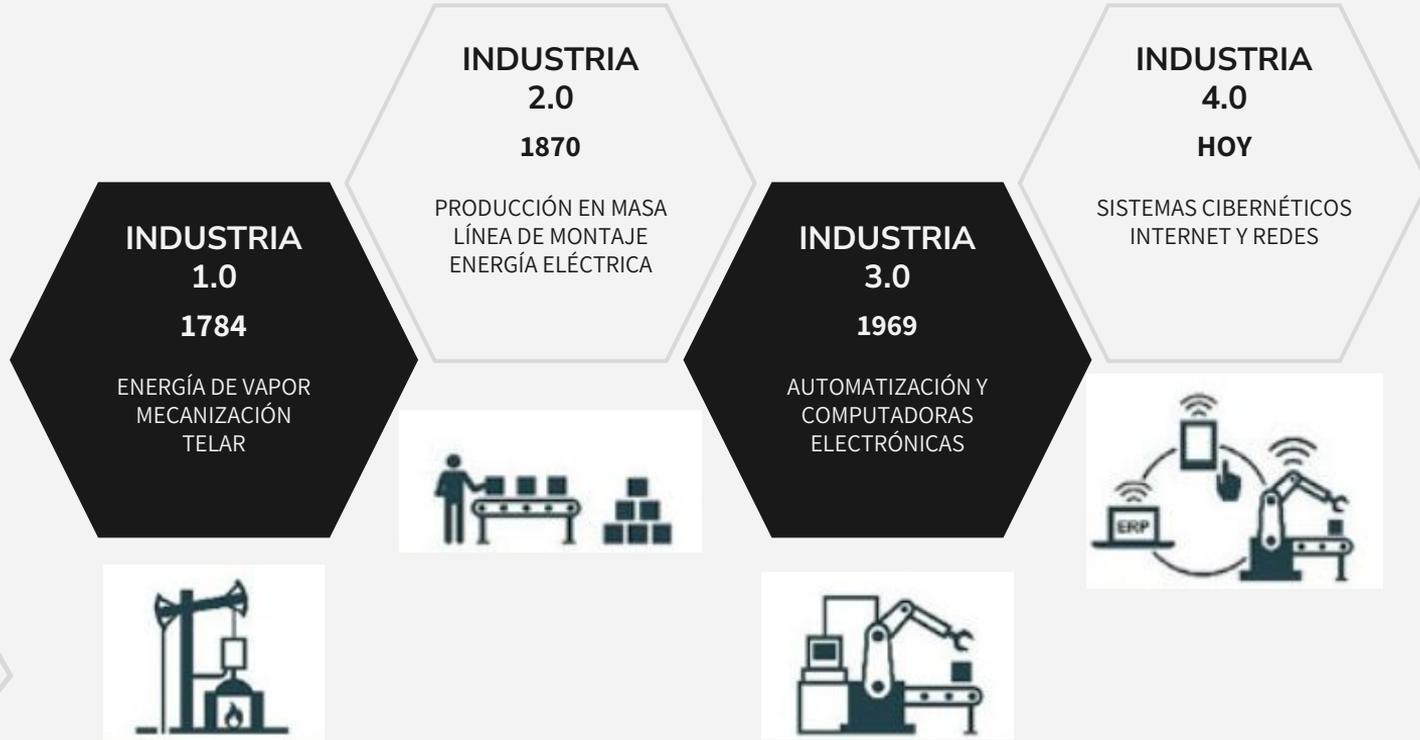
Funcionamiento, uso y aplicaciones.

A person in a dark suit is sitting at a desk, pointing with a wooden pencil at a laptop screen. The screen displays several data visualization charts, including a bar chart with blue and green bars, a line graph with blue lines, and a pie chart. The person's left hand is on the laptop keyboard, and their right hand is holding the pencil. The background is a blurred office setting with a white mug visible on the right.

01

INTRODUCCIÓN

AVANCE INDUSTRIAL



```
each: function(e, t, n) {
  var r, i = 0,
      o = e.length,
      a = N(e);
  if (n) {
    if (a) {
      for (; o > i; i++)
        if (r = t.apply(e[i], n), r === f)
          return r;
    } else
      for (i in e)
        if (r = t.apply(e[i], n), r === f)
          return r;
    } else if (a) {
    for (; o > i; i++)
      if (r = t.call(e[i], i, e[i]), r === f)
        return r;
    } else
      for (i in e)
        if (r = t.call(e[i], i, e[i]), r === f)
          return r;
  }
  return e;
},
trim: b && !b.call("\uffeff\u00a0") ? function(e) {
  return null == e ? "" : b.call(e)
} : function(e) {
  return null == e ? "" : (e + "").replace(C, "")
},
makeArray: function(e, t) {
  var n = t || [];
  return null != e && (N(Object(e)) ? x.merge(n, "st")
  );
},
isArray: function(e, t, n) {
  var r;
  if (t) {
    if (n) {
      for (i = 0; i < n; i++)
        if (!x.isArray(e[i]))
          return !1;
      return !0;
    }
    return x.isArray(e);
  }
  return !1;
}

```

02

DATOS



DATOS

- Materia prima que alimenta el desarrollo de tecnologías y servicios digitales.
- Son hechos manipulados en otras formas para que sean comprensibles y útiles.
- Se recopilan a partir de diversas fuentes y se almacenan en grandes centros de datos.

EVOLUCIÓN DE LOS DATOS



EVOLUCIÓN DE LOS DATOS

Sistemas de procesamiento de datos para manejar información en papel.

1960

Herramientas de procesamiento de texto y hojas de cálculo.

1980

Internet -> recopilar datos a través de la web -> impulsó desarrollo de Big Data.

2000

Industria en constante evolución.

HOY

1970

Sistemas de gestión de bases de datos -> almacenar y gestionar datos.

1990

Sistemas de minería de datos -> extraer información valiosa.

2010

Machine learning y AI -> herramientas clave para el análisis de datos.

OBTENCIÓN DE DATOS



SENSORES



ENCUESTAS Y CUESTIONARIOS



TRANSACCIONES Y REGISTROS



REDES SOCIALES Y SITIOS WEB

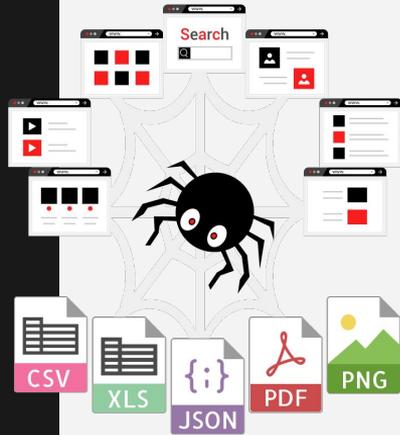


FUENTES PÚBLICAS

WEB SCRAPING



WEB SCRAPING



Raspado de precios

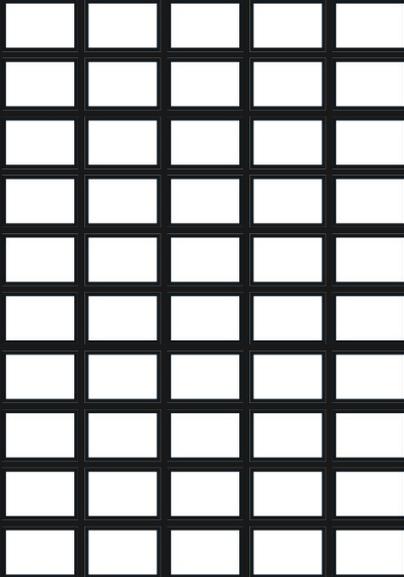
Las víctimas pueden ser agencias de viajes, vendedores de electrónica en línea, etc.

Raspado de contenido

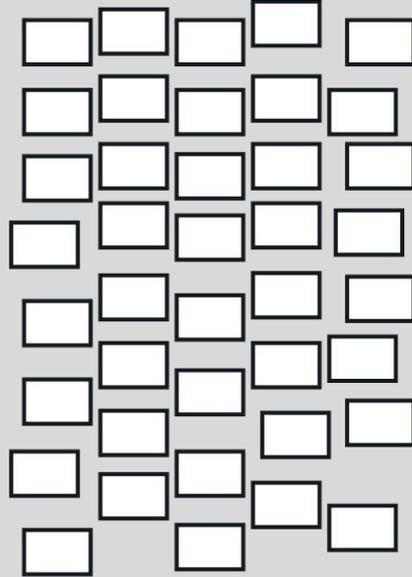
Catálogos de productos en línea, sitios web que se basan en contenido digital para impulsar el negocio

TIPOS DE DATOS

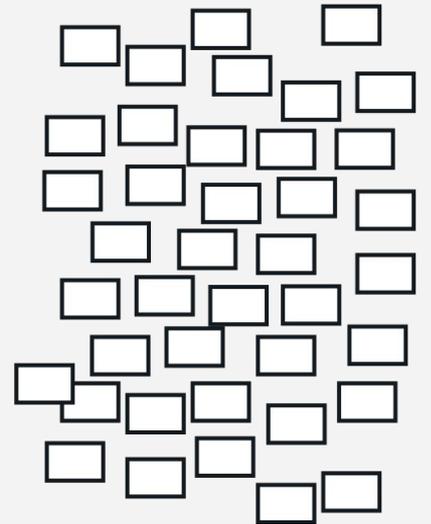
ESTRUCTURADOS



SEMIESTRUCTURADOS



NO ESTRUCTURADOS



TIPOS DE DATOS

ESTRUCTURADOS



Precios de acciones, base de datos de compras, rastreo web

SEMIESTRUCTURADOS

NO ESTRUCTURADOS



Documento de texto, video, fotografía

ALMACENAMIENTO DE DATOS

DISCOS DUROS



MEMORIA FLASH



SERVIDORES



EN LA NUBE

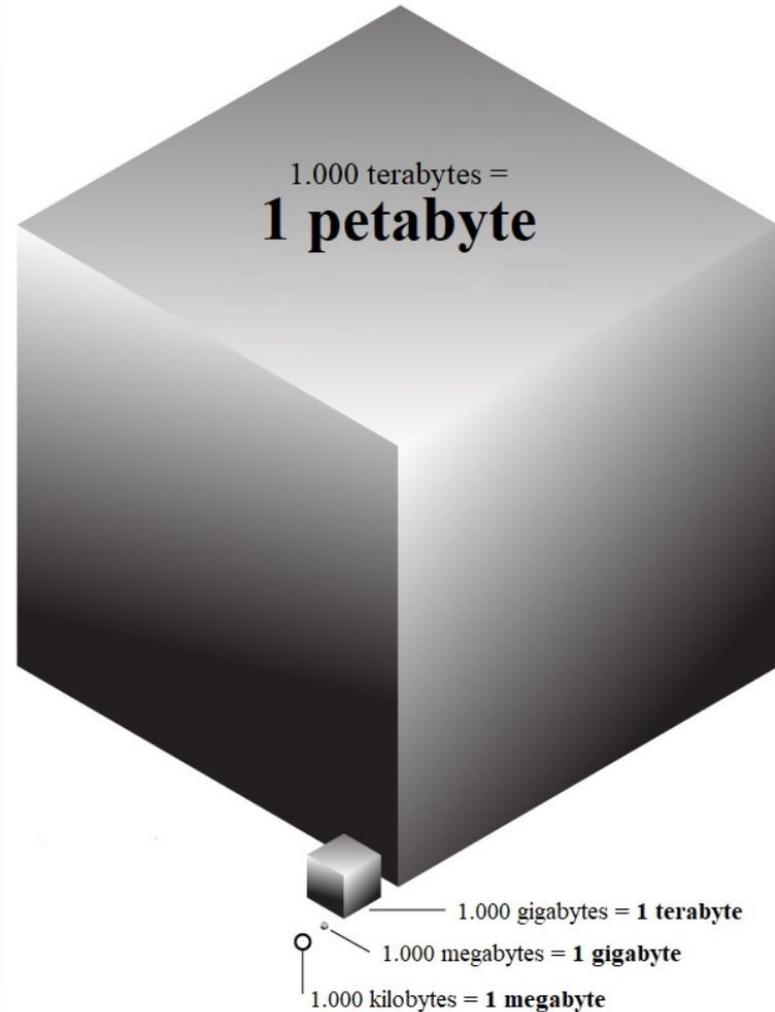




6220 TB de datos

World Data Centre for Climate (WDCC)

ALMACENAMIENTO DE DATOS



03

EMPRESAS Y DATOS



BASE DE DATOS



Cualquier organización que se precie debe contar con una **Base de Datos** con la que organizar los datos sobre su negocio.

Utilidades de la base de datos para las empresas

01

Agrupar y almacenar datos en un único lugar.

02

Facilitar que se compartan los datos entre sus miembros.

03

Evitar la redundancia y mejorar la organización de su actividad.

04

Visualizar los datos de un cliente o potencial.

05

Conectar datos de operaciones, facturación e interacciones con cada cliente o potencial

06

Activar campañas de marketing o tareas.

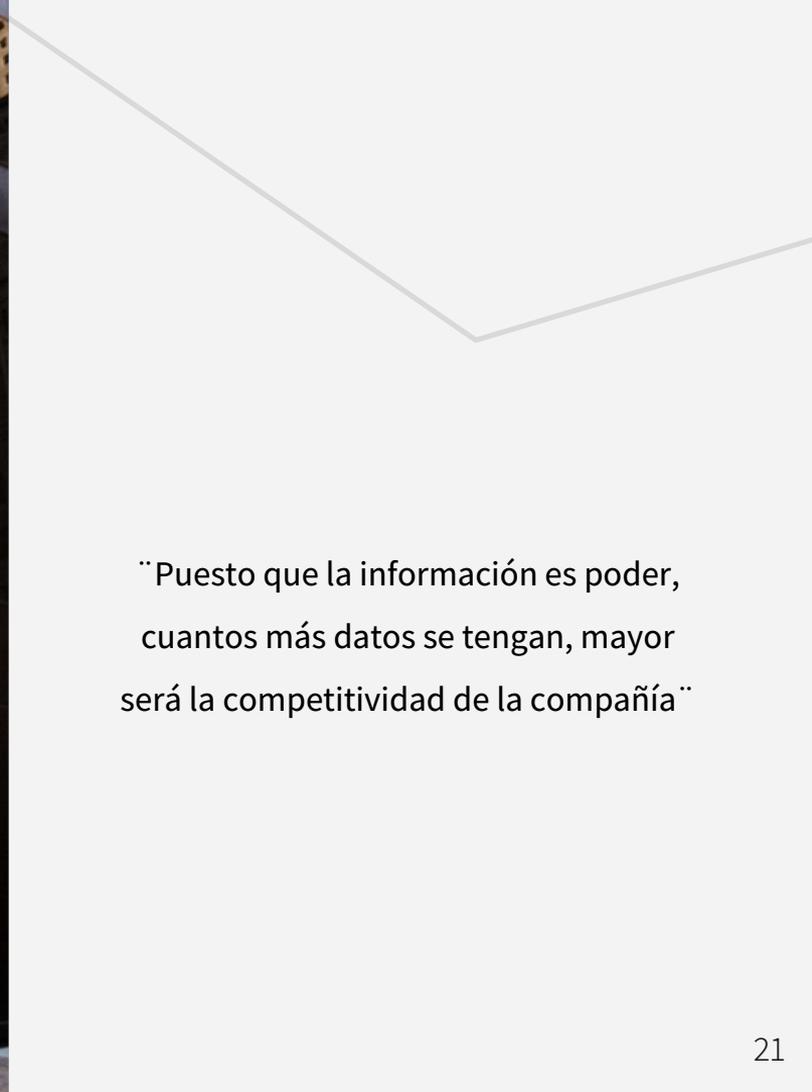
Fundamental para establecer una estrategia CRM

Utilidades de la base de datos para las empresas



A person with a backpack is seen from behind, looking towards a modern building with large glass windows at night. The building's lights are on, and the sky is dark. The person is wearing a blue patterned shirt and a black backpack. The building has a sign that says "140 O4".

“Con una base de datos
formulada correctamente,
se conseguirá que la
información y el
conocimiento sean los
mayores activos de la
compañía”

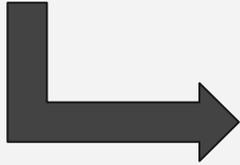
A white background with a thin grey line forming a downward-pointing V shape, centered horizontally and extending across the width of the page.

“Puesto que la información es poder,
cuantos más datos se tengan, mayor
será la competitividad de la compañía”

Google



**GENERACIÓN DE
DATOS**



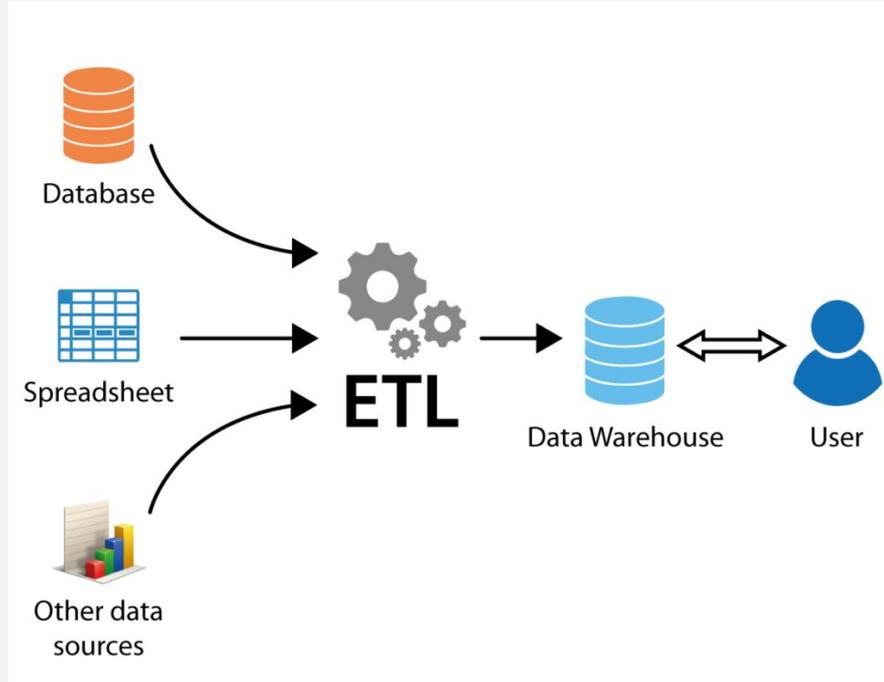
**PROCESAMIENTO DE
DATOS**



ANÁLISIS DE DATOS

PROCESAMIENTO DE DATOS

Se realiza a través de un proceso llamado “ETL” (Extract, Transform, Load)



PROCESAMIENTO DE DATOS



EXTRACT

Los datos se recopilan de fuentes de almacenamiento disponibles (formularios, sensores, etc.)



TRANSFORM

Preparación y limpieza de datos. Se descarta información repetitiva e incompleta.



LOAD

Los datos de la fase anterior son cargados en el sistema de destino.

PROCESAMIENTO DE DATOS

Herramientas que se utilizan:

SQL



Excel



Pandas (Python)



ANÁLISIS DE DATOS



OBJETIVO

Extraer conocimiento de los datos para tomar decisiones informadas.

¿CÓMO HACER UN ANÁLISIS DE DATOS?



1 Define tus preguntas, estas deben ser **medibles, claras y concisas.**



2 Decide **qué medir** y cómo medirlo.



3 **Recolecta los datos** y mantenlos bien organizados.



4 Realiza un **análisis profundo** de la información.



5 **Interpreta los resultados.**

ANÁLISIS DE DATOS

Herramientas que se utilizan:

Power BI



Tableau



Python



ANÁLISIS DE DATOS

Power BI



Link interactivo: <https://powerbi.microsoft.com/guidedtour/guidedtour/power-platform/power-bi/1/1>

ROLES DENTRO DE LA EMPRESA



Data engineer



Data analyst



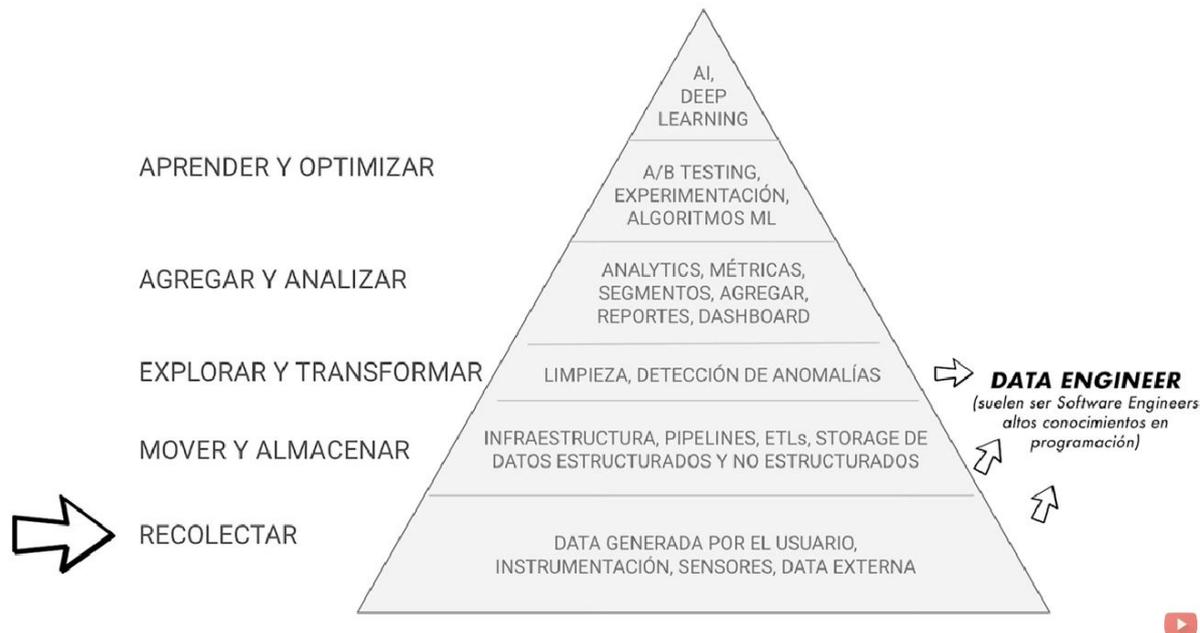
Data scientist





Data engineer

ROLES DENTRO DE LA EMPRESA





Data analyst

ROLES DENTRO DE LA EMPRESA



APRENDER Y OPTIMIZAR

AGREGAR Y ANALIZAR

EXPLORAR Y TRANSFORMAR

MOVER Y ALMACENAR

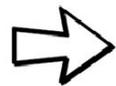
RECOLECTAR





Data scientist

ROLES DENTRO DE LA EMPRESA



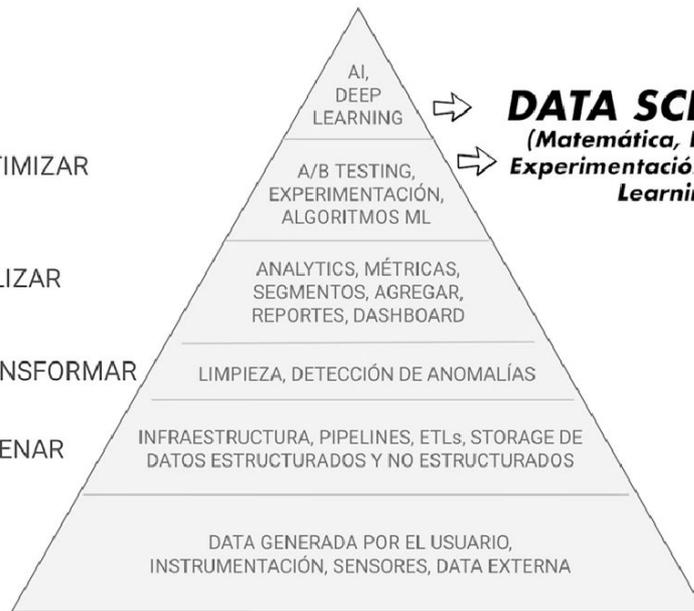
APRENDER Y OPTIMIZAR

AGREGAR Y ANALIZAR

EXPLORAR Y TRANSFORMAR

MOVER Y ALMACENAR

RECOLECTAR



DATA SCIENTIST
(Matemática, Estadística,
Experimentación y Machine
Learning)



CONTRATO DE CONFIDENCIALIDAD (NDA)



- Modelo de contrato que vincula legalmente a las partes firmantes imponiendo una relación de confidencialidad entre ellas.
- La empresa y el/la profesional acuerdan que la información privada a la que se tenga acceso **no** será puesta a disposición de terceros.
- “Cláusula de confidencialidad y no divulgación” ó NDA (non-disclosure agreement).

CONTRATO DE CONFIDENCIALIDAD (NDA)

Partes

1

Definición de
información
confidencial

2

Propósito
de la
divulgación

3

Las partes
involucradas

4

Obligaciones
de la parte
receptora de
la información

5

Período de
tiempo

6

Aviso de
inmunidad

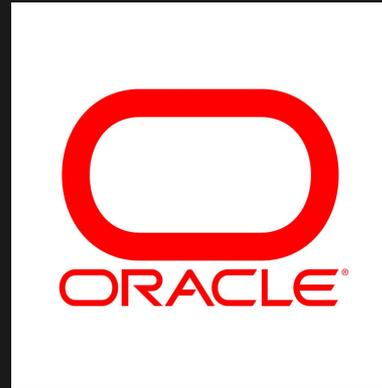
7

Penalización

PRINCIPALES EMPRESAS A
NIVEL INTERNACIONAL Y
NACIONAL



Empresas internacionales



Empresas nacionales



MEROVINGIANDATA

Creada en el año 2020 por Mario Japaz, Marcos Bruno y Giorgio Tacchini

Metodología "ABC"





04

BIG DATA

BIG DATA





VOLUMEN

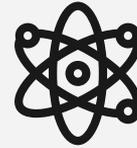
2.5 millones de gigabytes por día en el mundo



VELOCIDAD

29.2 gigabytes por segundo en el mundo

BIG DATA



VARIEDAD

Los datos se presentan en todo tipo de formatos

FUNCIONAMIENTO DEL BIG DATA

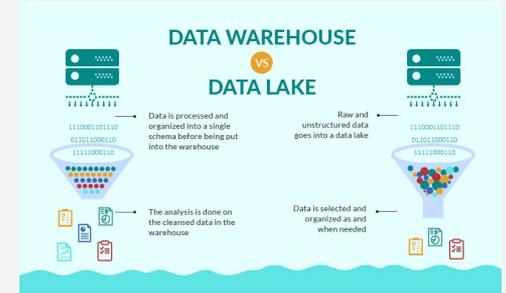
INTEGRACIÓN

Incorporar y procesar datos para que queden formateados y disponibles para su uso. (Herramientas: Apache Hadoop, NoSQL).

GESTIÓN

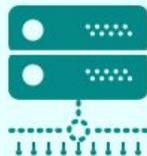
ANÁLISIS

Se analizan y utilizan los datos.



GESTIÓN

DATA WAREHOUSE VS DATA LAKE

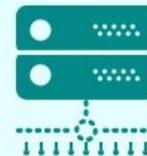


1110001101110
011011000110
11111000110

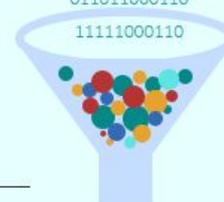


Data is processed and organized into a single schema before being put into the warehouse

The analysis is done on the cleansed data in the warehouse



1110001101110
011011000110
11111000110



Raw and unstructured data goes into a data lake

Data is selected and organized as and when needed

FUNCIONAMIENTO DEL BIG DATA

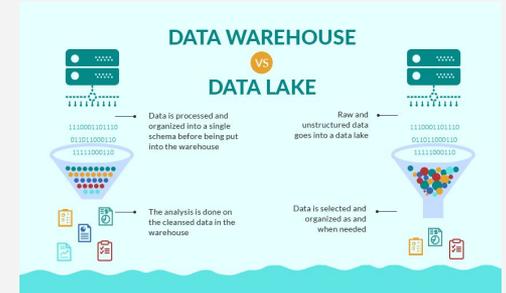
INTEGRACIÓN

Incorporar y procesar datos para que queden formateados y disponibles para su uso. (Herramientas: Apache Hadoop, NoSQL).

GESTIÓN

ANÁLISIS

Se analizan y utilizan los datos.



USO DEL BIG DATA



**DESARROLLO
DE
PRODUCTOS**

**MANTENIMIENTO
PREDICTIVO**

**EXPERIENCIA
DEL CLIENTE**

**FRAUDE Y
CONFORMIDAD**

USO DEL BIG DATA

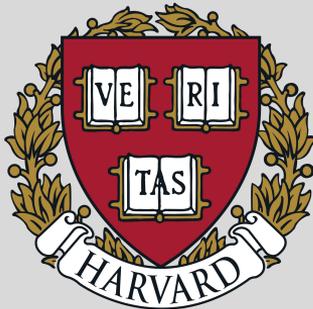


USO DEL BIG DATA

Ginger.io



Brote
Viral



USO DEL BIG DATA

Amazon

amazon

Netflix

NETFLIX

USO DEL BIG DATA

Coca-Cola

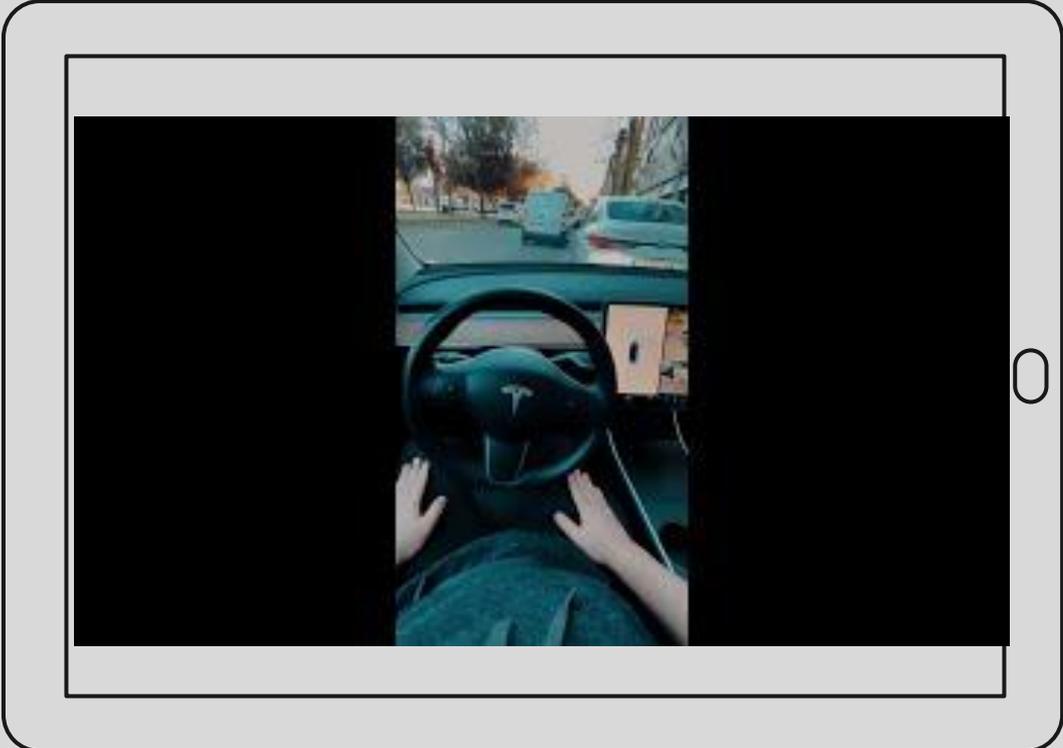
Coca-Cola

Starbucks



STARBUCKS®

USO DEL BIG DATA



- 
- 
1. ¿Sienten que big data es una invasión a la privacidad?
 2. La recopilación de datos nos genera cierta comodidad al filtrar las cosas que vemos y buscamos, ¿cómo creen que nos afecta hoy y a las próximas generaciones?
 3. ¿El uso de big data es a beneficio del que otorga la información, del que genera ganancia o de ambos?

¡MUCHAS GRACIAS!