

# Act.UT01. Sistemas de almacenamiento de la información

Bases de datos

Francisco José García Cutillas | 1FPGS\_DAM



## Índice

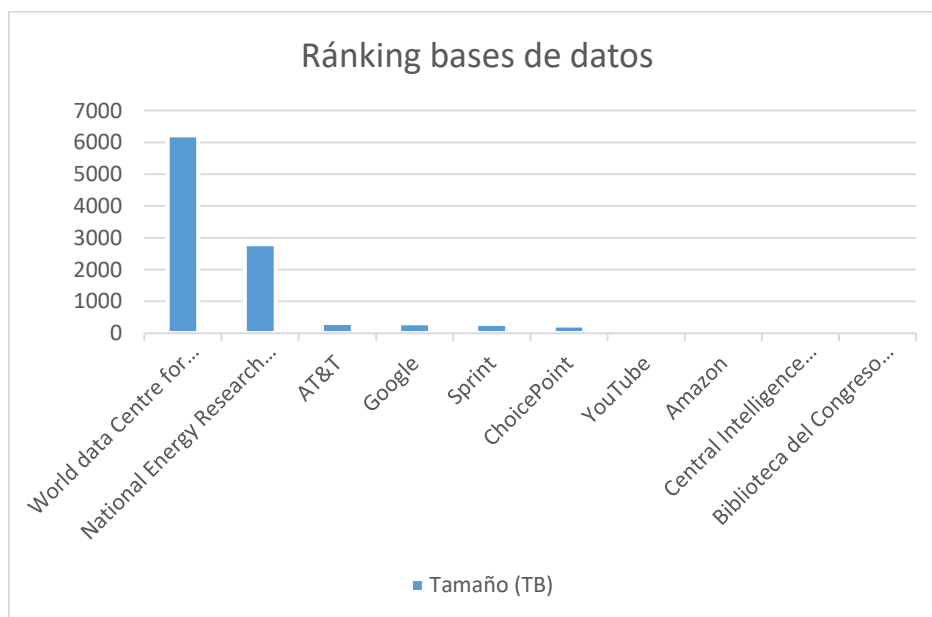
Ejercicio 1 .....	3
Ejercicio 2 .....	4
Ejercicio 3 .....	4
Ejercicio 4 .....	5
Ejercicio 5 .....	6
Ejercicio 6 .....	6
Ejercicio 7 .....	7

## Ejercicio 1

**Busca en Internet las 10 bases de datos más grandes del mundo. Anota su nombre y su tamaño y trata de generar en una hoja de cálculo (o similar) un gráfico que muestre la comparativa de tamaño de estas bases de datos. Indica el origen de la información mostrada.**

Las 10 bases de datos más grandes del mundo ordenadas de mayor a menor son:

1. World Data Centre for Climate (Centro Mundial de datos para el clima). Almacena unos 220 terabytes de información y 6 petabytes de información adicional, incluyendo datos sobre el clima, predicciones y simulaciones.
2. National Energy Research Scientific Computing Center. Dedicado a investigar sobre distintos tipos de energía. Su base de datos tiene 2,8 petabytes.
3. AT&T. se trata de una compañía de telecomunicaciones que almacena 323 terabytes.
4. Google. Se desconoce el verdadero tamaño de su base de datos, pero la compañía recibe unos 91 millones de consultas al día, las cuales son almacenadas.
5. Sprint. Es una de las mayores compañías de telecomunicaciones del mundo. Guarda los detalles de 365 millones de llamadas al día. Se estima que tiene una base de datos de unos 300 terabytes.
6. ChoicePoint. Esta compañía almacena los datos personales de la población de EEUU. Almacena unos 250 terabytes.
7. YouTube. Ocupa el séptimo lugar, aunque se desconoce el volumen de su base de datos actual, se estima en unos 45 terabytes.
8. Amazon. Cuenta con más de 42 terabytes de datos de sus 59 millones de clientes.
9. Central Intelligence Agency(CIA). Tampoco se conoce el tamaño de su base de datos.
10. Biblioteca del Congreso de EEUU. Posee más de 5 millones de documentos digitales. Su base de datos está en torno a los 20 terabytes.



Para la realización de la gráfica, tanto a la base de datos de Google como a la de la CIA, se les ha dado un valor estimado a su capacidad teniendo en cuenta la posición que ocupan en el ránking.

Fuentes:

<https://www.20minutos.es/noticia/203609/0/bases/datos/grandes/>

[https://www.sinnexus.com/blog/bases\\_datos\\_top.aspx](https://www.sinnexus.com/blog/bases_datos_top.aspx)

## Ejercicio 2

**Realiza un esquema o infografía sobre la historia de las bases de datos (puedes servirte de aplicaciones tipo genial.ly para ello).**

<https://view.genial.ly/6334290781746f00176f2670/interactive-content-timeline-corporativo-ilustrado>

Fuente:

<https://conocelahistoria.com/historia-de-la-base-de-datos/>

## Ejercicio 3

**Define los siguientes conceptos e indica algún ejemplo**

- a) **Dato.** Es un trozo de información concreta sobre algún concepto o suceso. Por ejemplo, 7 es la cantidad de días que tiene la semana.
- b) **Tipo de dato.** Indica la naturaleza del dato. Es decir, se puede tener datos numéricos, que son aquellos con los que se pueden realizar cálculos aritméticos y datos alfanuméricos, que son los que contienen caracteres alfabéticos y numéricos. Ambos se pueden combinar para obtener tipos de datos más elaborados. Por ejemplo, el tipo de dato semana que contiene siete datos alfanuméricos: Lunes, Martes, Miércoles, Jueves, Viernes, Sábado, Domingo.
- c) **Campo.** Es un identificador para una parte de los datos de una entidad. Cada campo pertenece a un tipo de dato. También se le llama columna. Por ejemplo, Curso representa al conjunto de los alumnos que están en el mismo nivel académico.
- d) **Registro.** Es una colección de datos referentes a un mismo concepto o suceso. También conocido como tupla o fila. Por ejemplo, los datos de las notas de un mismo alumno de cada asignatura.
- e) **Tabla.** Es un conjunto de registros bajo el mismo nombre que representa el conjunto de todos ellos. Por ejemplo, podemos tener una tabla que contenga para un mismo curso, las diferentes notas de todas las asignaturas de todos los alumnos.

## Ejercicio 4

**Indica al menos 3 ventajas e inconvenientes de usar bases de datos frente a los tradicionales sistemas de ficheros.**

### Ventajas:

- Independencia de los datos respecto a los tratamientos y viceversa. No es necesario cambiar la forma lógica de la base de datos si se cambia de aplicación al mismo tiempo que tampoco es necesario volver a reprogramar la aplicación si se producen cambios en los datos.
- Coherencia de los resultados. Los datos sólo se guardan una vez y no existe redundancia de los mismos.
- Mejor disponibilidad de los datos para el conjunto de los usuarios.
- Reducción del espacio de almacenamiento. Esto se consigue al eliminar la redundancia.

### Inconvenientes:

- Instalación costosa. Tanto en equipo físico (nuevas instalaciones, ampliaciones), como en el lógico (sistemas operativos, programas, compiladores...).
- Personal especializado. Implican una necesidad de personal especializado que resulta difícil de encontrar, y de formar.
- Implantación larga y difícil. Las dificultades que van apareciendo a lo largo de su desarrollo llevan en general a que se superen los plazos inicialmente previstos.
- Falta de rentabilidad a corto plazo. En instalaciones grandes el plazo puede llegar a ser de años.

Fuente:

<https://es.slideshare.net/lsa06t/ventajas-y-desventajas-de-las-bases-de-datos-frente-a-los-archivos>

## Ejercicio 5

Imagina que quieres almacenar los trabajadores de tu empresa y además quieres almacenar los distintos teléfonos que pueden tener. ¿Qué tabla o tablas crearías para ello? Imagina los campos que pueden tener dicha o dichas tablas e introduce los datos de 2 trabajadores cada uno de ellos que tenga dos teléfonos (piensa en el concepto redundancia y consistencia de los datos)

Teléfono_1	Teléfono_2	Trabajador
666666666	666666667	Juan García Palazón
666666668	666666669	Evaristo Ayala Mondéjar

Se podría tomar como principal el Teléfono\_1, ya que Teléfono\_2 podría ser opcional, y también podría haber dos trabajadores con el mismo nombre y apellidos. En este caso el campo que no se podría repetir nunca sería el del Teléfono\_1. Se podría utilizar como un identificador único de cada trabajador.

## Ejercicio 6

**Consultar en Internet información sobre las prestaciones de los SGBD comerciales y libres más utilizados en la actualidad.**

Entre los Sistemas Gestores de Bases de Datos comerciales más utilizados podemos encontrar:

- Oracle. Es el más utilizado, ya que es uno de los más fiables. Está construido alrededor de un marco en el que se puede acceder directamente a los objetos a partir del lenguaje de consulta SQL. Tiene su propio componente de red para que pueda haber comunicación a través de las redes. Se ejecuta en casi todas las plataformas. Su principal característica es que su arquitectura se divide entre la lógica y la física. Esto permite, a grandes rasgos, una mayor flexibilidad en las redes de datos y, a la vez, robustez en la estructura de los mismos. Oracle es muy utilizada por las empresas.
- SQL Server. Junto a Oracle son los que más parte de la cuota de mercado ocupan. SQL comparte muchas características con Oracle, aunque también existen diferencias. SQL Server se ejecuta en Transact-SQL, un conjunto de programas que añaden características, como pueden ser el tratamiento de errores y excepciones, procesamiento de datos, extracción de datos directos de Web, uno de distintos lenguajes de programación, etc. Además, destaca por su carácter administrativo tanto en sus funciones y seguridad como en la flexibilidad de las bases de datos.

Por otro lado, podemos encontrar los SGBD de acceso libre, los cuáles son gratuitos. Entre ellos podemos destacar:

- MySQL. Actúa del lado del cliente (servidor) y es de código abierto con licencia comercial disponible. Gestiona bases de datos relacionales, es multiusuario y es el más utilizado dentro del software libre. Destaca por requerir poca memoria y procesador para funcionar. Es usado principalmente para el desarrollo web.
- FireBird. Potente, y a la vez, sencillo SGBD relacional SQL. Es uno de los mejores gestores de código abierto compatible con Windows y Linux. Entre otras funciones, da soporte completo para procedimientos almacenados, transacciones compatibles con las características ACID (Atomicity, Consistency, Isolation, Durability) y métodos de acceso múltiple (nativo, Python, .NET, etc.).

Fuente:

<https://www.kyoceradocumentsolutions.es/es/smarter-workspaces/insights-hub/articles/los-mejores-gestores-de-base-de-datos-del-mercado.html>

## Ejercicio 7

### **¿En qué consiste y para qué se utiliza la minería de datos?**

La minería de datos permite encontrar información escondida en los datos que no siempre resulta aparente, debido al gran volumen de datos existentes. Se utiliza con el objetivo de descubrir patrones y tendencias estructurando la información obtenida de un modo comprensible para su posterior utilización.

Fuente:

<https://www.esic.edu/rethink/tecnologia/mineria-datos-proceso-areas-se-puede-aplica>