



FHyCS

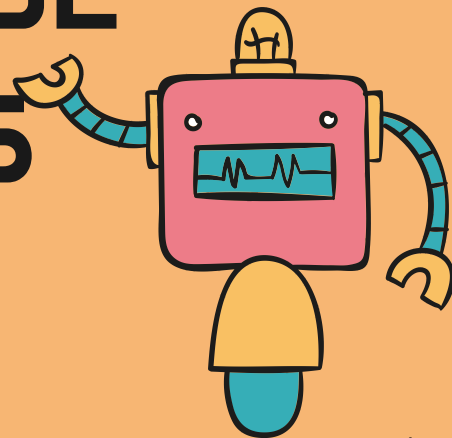
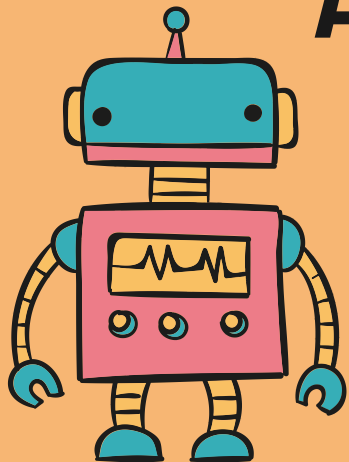
Facultad de Humanidades
y Ciencias Sociales

Secretaría de Posgrado

CURSO DE POSGRADO

ARMADO Y USO DE BASE DE DATOS

PARA PROYECTOS DE
INVESTIGACIÓN



Docentes



Javier Altamirano

Dr. en Cs. Biológicas
Prof. de Epidemiología y Estadística en FHyCS

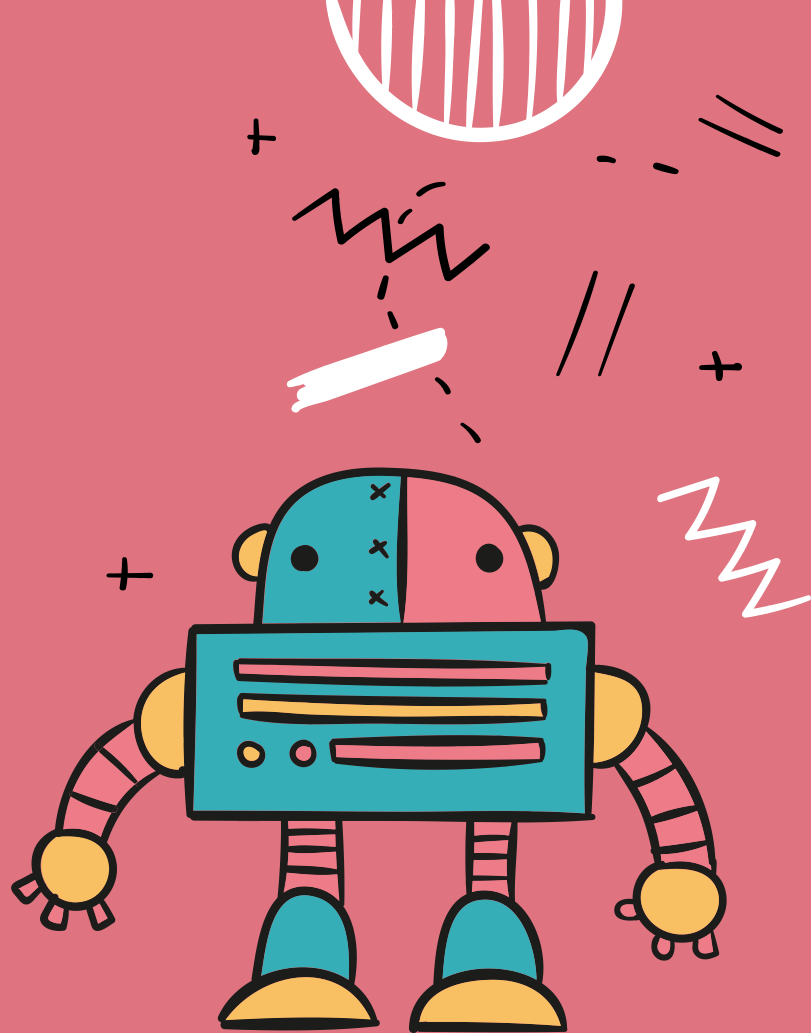


Ana María Chalabe

Bioquímica. Magister.
Especialista en Salud Pública
Prof. Titular de Epidemiología y
Estadística en FHyCs

Objetivo General

Este curso tiene como propósito brindar a los participantes las herramientas teóricas y prácticas para el armado y uso de bases de datos en proyectos de investigación científica.



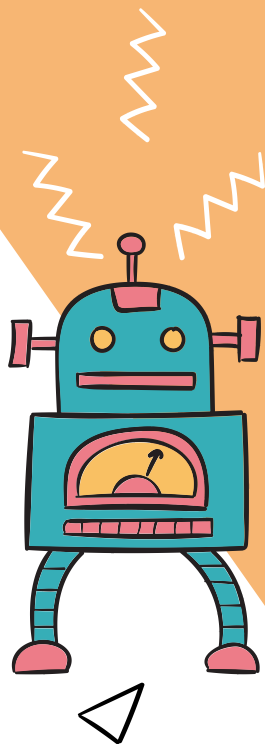


Objetivos Específicos



1. Comprender la importancia del uso de bases de datos en proyectos de investigación científica.
2. Identificar los diferentes tipos de bases de datos y sus características.
3. Diseñar y estructurar bases de datos adecuadas para proyectos de investigación.
4. Construir y gestionar bases de datos eficientes.
5. Realizar análisis exploratorios de datos utilizando técnicas y herramientas adecuadas.
6. Visualizar y comunicar los resultados del análisis de datos de manera efectiva.
7. Interpretar los resultados obtenidos a partir del análisis de datos.
8. Aplicar herramientas de inteligencia artificial para optimizar procesos.





01

Introducción a BBDD

y su importancia en la
investigación científica

02

Diseño y estructura

de bases de datos
para proyectos de
investigación

03

Construcción y gestión

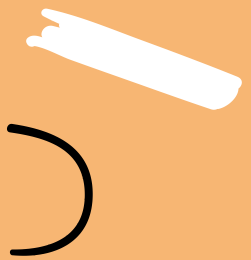
de bases de datos

04

Análisis y visualización

de datos en proyectos de
investigación científica





Cronograma de Actividades



**1er
Encuentro**



**Jueves 4 de junio
16:00 hs**

**2do
Encuentro**



**Jueves 11 de junio
16:00 hs**

**3er
Encuentro**



**Jueves 18 de junio
16:00 hs**

**4to
Encuentro**



**Jueves 25 de junio
16:00 hs**



**Entrega de
proyecto final**



**Fecha a coordinar
(sem. 6 al 12 de julio?)**





SISTEMA DE EVALUACIÓN

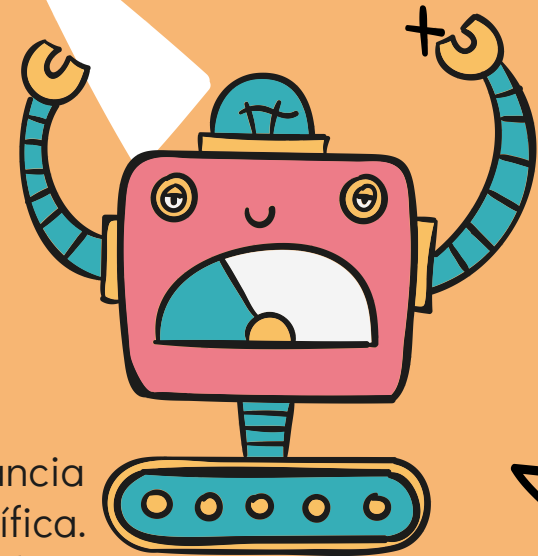


1. Participación en clase (10%)
 2. Trabajos prácticos individuales (30%)
 3. Proyecto final (60%): Desarrollo de un proyecto aplicado donde los estudiantes diseñen, construyan y gestionen una base de datos relacionada con su área de investigación específica.
- 

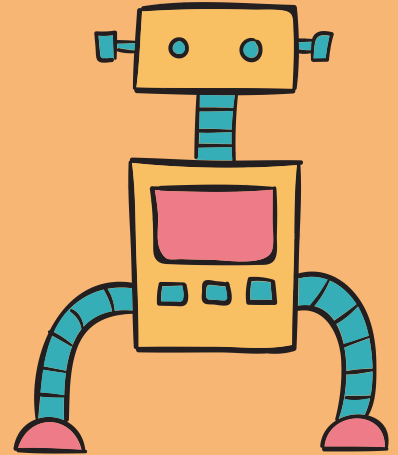
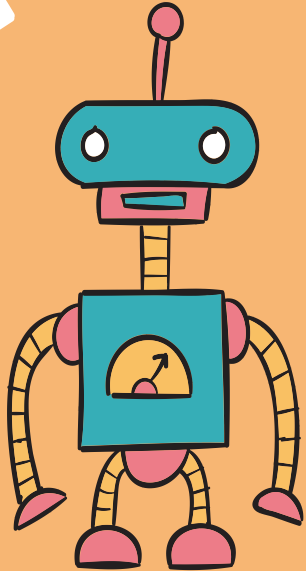
01.

Introducción a Bases de Datos

Definición y conceptos básicos de las bases de datos. Importancia del uso de bases de datos en proyectos de investigación científica.
Tipos de bases de datos más utilizadas en investigación científica.
Buenas prácticas en la nomenclatura y estructura de datos.
Ejemplos prácticos



¿Qué es una Base de Datos?

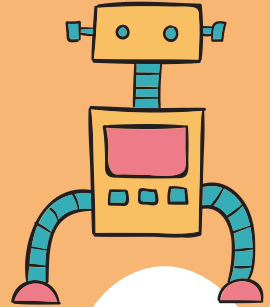
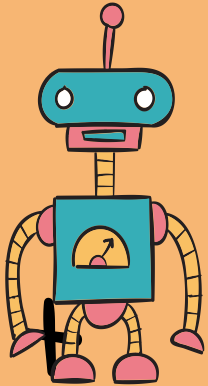


B2 X ✓ fx =BUSCARV(A2,Generales!\$B\$2:\$G\$17)

	A	B	C	D	E	F	G
1							
2		Hotel	Precio	Región	Habitaciones	Fundación	Gerente
3		Villa Purificación	BUSCARV(valor_buscado, matriz_buscar_en, indicador_columnas, [ordenado])				
4		Quitupan	\$500	Oeste	36	27/12/2013	Humberto Macias
5		El Salto	\$1,500	Este	26	28/09/2013	Ramon Hinojosa
6		San Cristóbal	\$900	Norte	30	01/11/2013	Erivan Calderón
7		San Diego de Alejandria	\$1,500	Norte	24	31/10/2013	Roberto Cortes
8		San Juan de los Lagos	\$400	Sur	33	16/06/2014	Carlos Cedeño
9		San Julián	\$9,000	Este	21	21/11/2013	Eduardo Mancillas
10		San Marcos	\$1,200	Sur	20	23/11/2014	Mariana Ramirez
11		San Martín de Bolaños	\$500	Norte	39	25/01/2013	Alejandra Sanchez
12		San Martín Hidalgo	\$1,200	Norte	43	25/04/2013	Sofia Larios
13		San Miguel el Alto	\$1,200	Norte	23	09/11/2014	Eduardo Corona
14		Gómez Farías	\$1,200	Oeste	25	22/11/2013	Luis Lopez
15		San Sebastián del Oeste	\$1,500	Oeste	23	05/12/2013	Ana Gonzalez
16		Santa María de los Ángeles	\$1,300	Este	22	16/11/2014	Mario Avila
17		Sayula	\$900	Este	36	26/03/2014	Nayeli Parraga

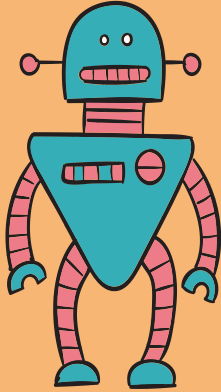
¿Qué es una Base de Datos?

Una base de datos es una **colección organizada** de información estructurada, almacenada y accesible de manera **sistemática**, que permite registrar, gestionar y recuperar datos de forma eficiente.



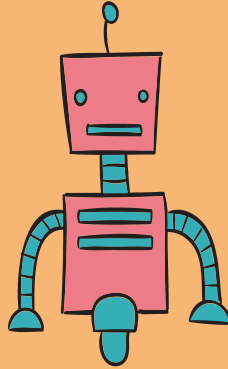
Funciones

+



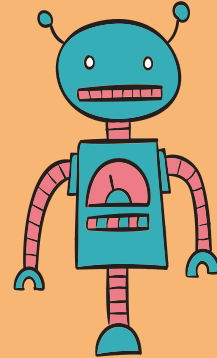
Almacenar

Guardar grandes volúmenes de datos en un lugar centralizado y organizado, facilitando su acceso posterior.



Ordenar

Estructurar los datos de manera lógica y coherente, permitiendo encontrar y relacionar información rápidamente.



Búscar y Analizar

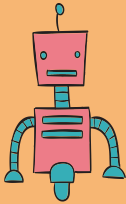
Recuperar datos específicos y aplicar técnicas de análisis para extraer conclusiones y generar conocimiento.

Características de una base de datos

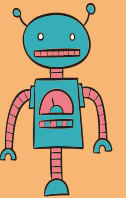
+



Organización: Los datos se almacenan en una estructura organizada: tablas, filas y columnas con roles definidos.



Redundancia mínima: Los datos se registran una sola vez, reduciendo la posibilidad de inconsistencias y errores.

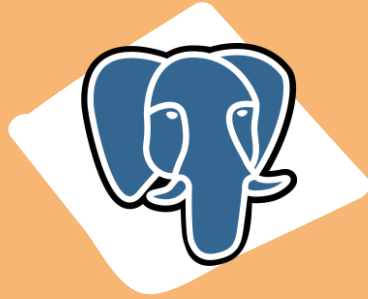


Consistencia: Los datos se mantienen coherentes en toda la base: los mismos criterios aplican a toda la información.

Acceso: Los datos pueden recuperarse de forma rápida usando filtros, búsquedas y consultas.

Seguridad: Los datos se protegen mediante controles de acceso: quién puede ver, editar o eliminar registros.

Sistemas de gestión de bases de datos Relacionales



PostgreSQL

Base de datos relacional de código abierto, avanzada y muy robusta. Ampliamente usada en investigación científica y análisis de datos.



MySQL

Base de datos relacional de código abierto. La más utilizada en aplicaciones web y proyectos de investigación con datos estructurados.



SQLite

Base de datos relacional liviana, sin servidor. Ideal para proyectos individuales de investigación y prototipos.

Sistemas de gestión de bases de datos No relacionales (NoSQL)



MongoDB

Base de datos documental de código abierto. Almacena registros con estructuras variables, útil para datos cualitativos heterogéneos.



Redis

Base de datos clave-valor de código abierto. Extremadamente rápida, usada para almacenar configuraciones y datos de acceso frecuente..

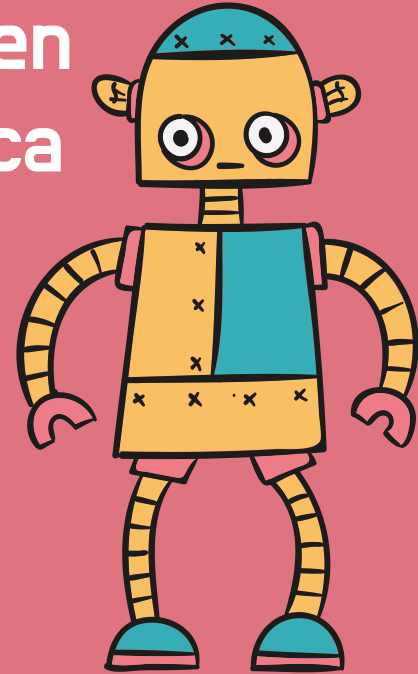


Neo4j

Base de datos de grafos. Especializada en modelar relaciones complejas entre actores, ideal para análisis de redes sociales e institucionales.

Importancia de las bases de datos en proyectos de investigación científica

Saber armar y gestionar bases de datos es una habilidad esencial para cualquier investigador. Permite **organizar, almacenar, analizar y recuperar** datos de manera eficiente, aumentando la calidad y la confiabilidad de los resultados.



1. Eficiencia en la gestión de datos +

Almacenamiento



Las BBDD permiten almacenar grandes cantidades de datos en un solo lugar, facilitando acceso y gestión

Organización

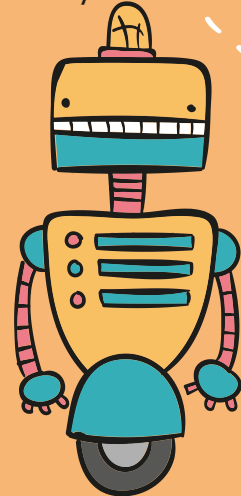


Los datos se pueden organizar de forma lógica y estructurada, lo que facilita su búsqueda y análisis

Automatización



Se pueden automatizar tareas repetitivas, como la limpieza y el análisis de datos, lo que ahorra tiempo y esfuerzo.



2. Mejora en la calidad de la investigación+

Análisis



Las bases de datos permiten realizar análisis de datos más complejos y sofisticados

Proyección

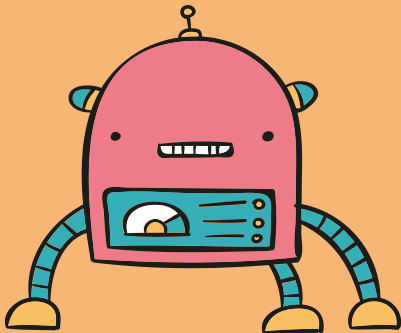


Se pueden identificar patrones y tendencias en los datos que serían difíciles de detectar manualmente.

Test Estadísticos



Se pueden realizar pruebas de hipótesis y generar nuevos conocimientos.



3. Facilita la colaboración entre investigadores

Compartir



Las bases de datos permiten compartir datos con otros investigadores de manera fácil y segura.

Trabajo en Equipo

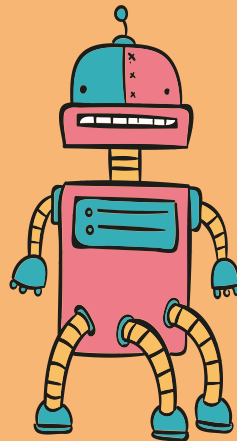


Se pueden crear proyectos de investigación colaborativos en los que los investigadores pueden trabajar simultáneamente en una sola BD.

Reproducibilidad



Se pueden fomentar la reproducibilidad de la investigación al compartir los datos y el código utilizado para analizarlos.

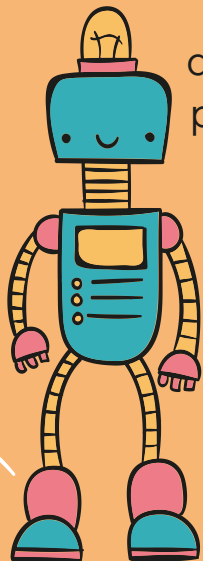


4. Aumenta la transparencia de la investigación

Claridad



Las bases de datos permiten documentar de manera clara y precisa los datos utilizados en una investigación.



Etiquetas



Se pueden crear metadatos que describan los datos, lo que facilita su comprensión y uso por parte de otros investigadores.

Confiabilidad



Se puede fomentar la confianza en la investigación al hacer que los datos sean accesibles y transparentes

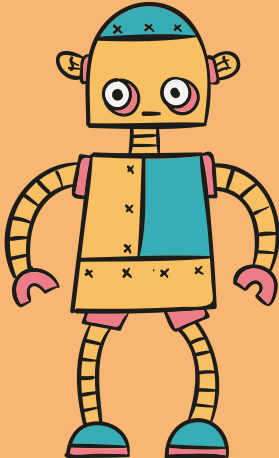


5. Preservación de los datos

Seguridad



Las bases de datos permiten almacenar los datos de forma segura y permanente.



Back Up



Se pueden crear copias de seguridad de los datos para evitar su pérdida

Restricciones



Se pueden crear políticas de acceso a los datos para asegurar que solo las personas autorizadas puedan acceder a ellos.

Ejemplos de uso de bases de datos para proyectos de investigación científica



Historia/ Comunicación

- Repositorio de fuentes primarias o entrevistas.
- Análisis de noticias periodísticas.



Medicina

- Almacenamiento de registros médicos electrónicos.
- Análisis de datos de ensayos clínicos.
- Desarrollo de nuevos medicamentos.
- Estudio de enfermedades y epidemias.



Administración/ Economía

- Seguimiento presupuestario de proyectos.
- Análisis macro y microeconómicos regionales.



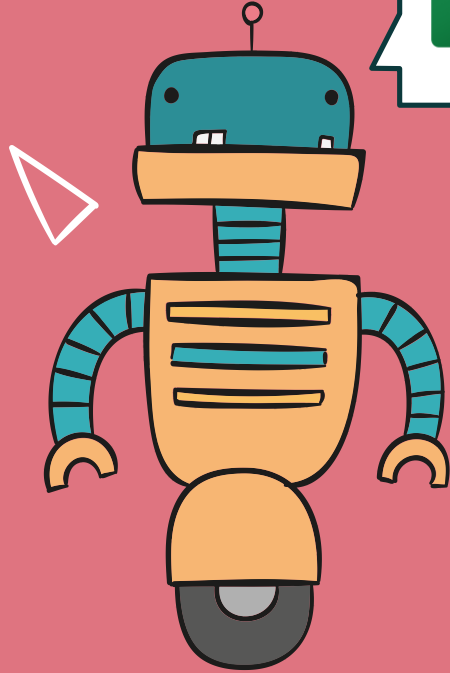
Ciencias Sociales

- Almacenamiento de datos de encuestas y censos.
- Análisis de datos de comportamiento humano.
- Estudio de tendencias sociales.
- Modelado de sistemas sociales.



Ciencias Ambientales

- Almacenamiento de datos de sensores ambientales.
- Análisis de datos del cambio climático.
- Modelado de la calidad del aire y del agua.
- Estudio de la biodiversidad.



Excel como base de datos

Excel no es una base de datos en el sentido tradicional, ya que no está diseñado específicamente para almacenar y gestionar grandes cantidades de datos. Sin embargo es la herramienta más accesible para organizar y analizar datos en proyectos de investigación de pequeña y mediana escala



Ventajas

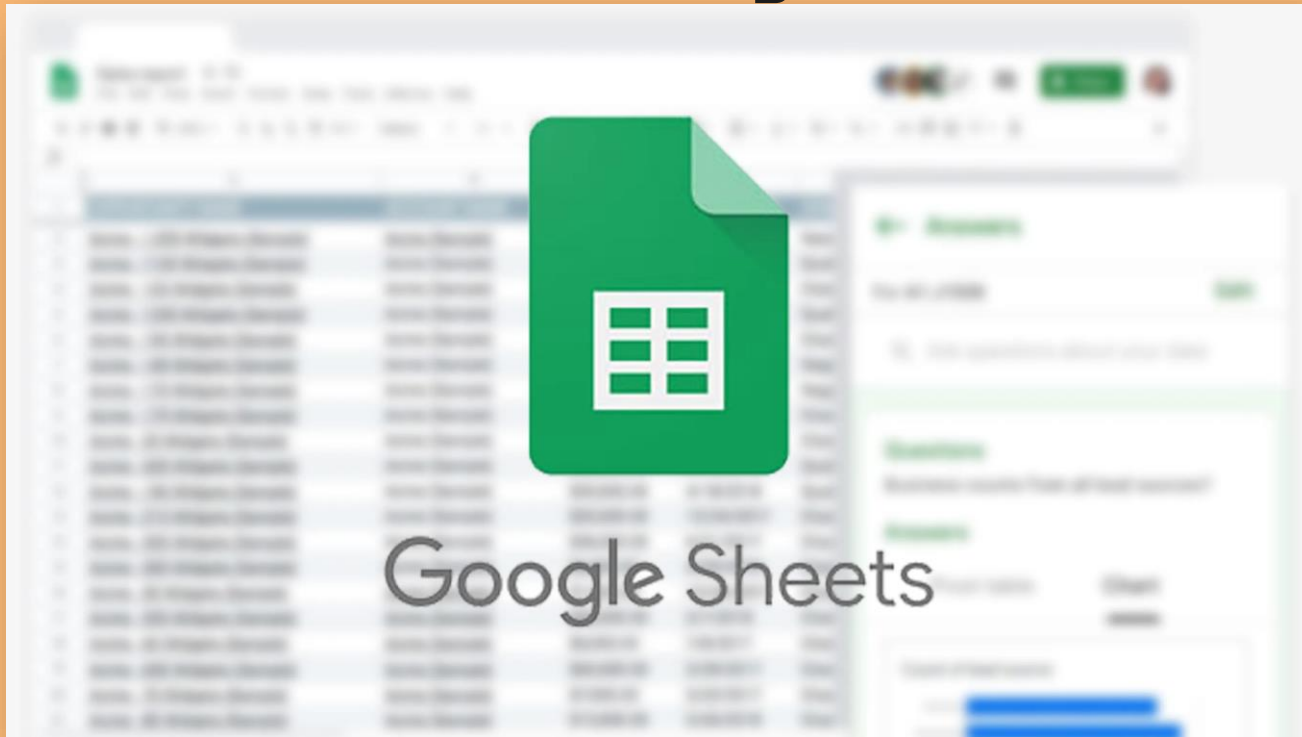
- Fácil de usar · curva de aprendizaje baja
- Flexible para diferentes tipos de datos
- Funciones potentes para análisis y visualización
- Económico (incluido en entornos académicos)
- Compatible con R, Python y otros software



X Desventajas

- Capacidad máxima: 1.048.576 filas × 16.384 columnas
- Mayor riesgo de errores en datos compartidos
- Sin control de versiones ni acceso concurrente robusto
- Falta de integridad referencial entre hojas
- No recomendado para datos sensibles sin protección adicional

Alternativa: Hojas de cálculo de Google



Característica

Excel



Hojas de cálculo
de Google



Costo

De pago

Gratuito

Instalación

Local

Basado en la nube

Capacidad de almacenamiento

7.179.869.184 celdas
(1.048.576 filas x 16.384 columnas)

5.000.000 de celdas
(400.000 filas x 12.500 columnas)

Interfaz

Compleja

Sencilla

Colaboración

Limitada

En tiempo real

Funciones básicas

Similar

Similar

Funciones avanzadas

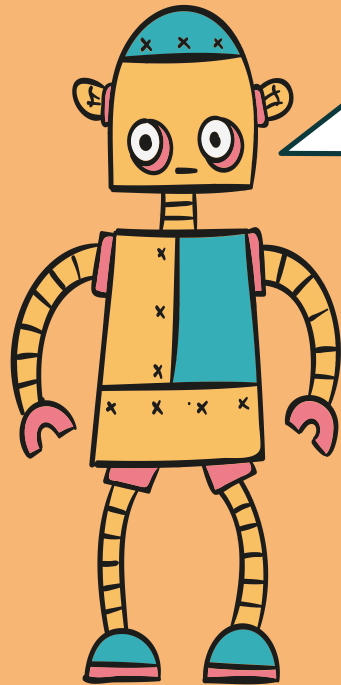
Más amplias
(Power Query, macros)

Limitadas

**Integración con otras
aplicaciones**

Limitada (ecosistema Microsoft)

Amplia



Entonces... ¿Cómo se
organizan los datos en
una Base de Datos?

Principios de Organización de Datos +

Filas = Observaciones

1
2
3
4
5
6
7
8
9
10
11
12

✓ Cada fila representa una unidad de análisis única: una persona, un caso, un evento, una fecha.

✗ No mezclar dos tipos de unidades en la misma tabla.

Columnas = Variables

A	B	C	D
---	---	---	---

✓ Cada columna contiene un único tipo de información: edad, sexo, fecha, localidad.

✗ No guardar dos variables en la misma celda (ej: 'nombre apellido' en una sola columna).

Celdas = Valores

✓ Cada celda tiene un único valor. Sin fórmulas ni cálculos dentro de la tabla de datos.

5,5
6,5
7,1
3,4

✗ Sin celdas fusionadas, sin filas de totales intercaladas.

Buenas prácticas en nomenclatura y estructura

+

Nombres de variables

Sin espacios: usar **guion_bajo** (fecha _ nac) o **camelCase** (fechaNac). Sin tildes, ñ ni caracteres especiales. Descriptivos pero cortos.

Valores consistentes

Usar siempre el mismo criterio: 'Masculino' nunca mezclar con 'M', 'masc.', 'masculino'. Definir categorías antes de cargar datos.

Encabezados

Una sola fila de encabezado, en la primera fila. Nombres únicos para cada columna. Sin columnas sin nombre.

Datos faltantes

Identificar con **NA**, 999 o el código definido en el diccionario de datos. Nunca dejar celdas vacías sin aclarar.

Estructura del archivo

Sin filas ni columnas vacías entre datos. Sin hojas ocultas con datos sin documentar. Sin formatos decorativos que oculten información.



Funciones esenciales de Excel para gestión de datos

BUSCARV

Busca un valor en una columna y devuelve un dato de otra columna. Ideal para cruzar tablas o validar registros.

`=BUSCARV(A2,TablaRef,3,0)`

CONTAR.SI

Cuenta el número de celdas que cumplen una condición. Útil para frecuencias por categoría.

`=CONTAR.SI(B:B,"Femenino")`

SI

Ejecuta una acción si se cumple una condición y otra si no. Permite clasificar o etiquetar datos automáticamente.

`=SI(C2>=18,"Adulto","Menor")`

CONCATENAR / &

Une el contenido de varias celdas en una sola. Útil para crear identificadores únicos o unir nombre y apellido.

`=A2&"_"&B2`

TEXTO A COLUMNAS

Divide el contenido de una celda en varias columnas usando un delimitador (coma, punto y coma, espacio). Se accede desde la pestaña Datos.

Datos → Texto a columnas

¡Comencemos
con la práctica!

